NOTES DE COURS DE MDD251

Christian Gérard
christian.gerard@math.u-psud.fr

Table des matières

1. Introduction

Le but de ces notes est de servir de *support* au cours en amphi et non de le remplacer, leur style volontairement laconique les rendant peu appropriées à cet usage. Les étudiants trouveront ici les énoncés stables et les démonstrations des résultats décrits en amphi.

Il est probable qu' une partie de ces notes ne sera d' ailleurs pas traitée en amphi, et ne figurera donc pas dans les contrôles de connaissance.

2. Les espaces $\mathbb{R}^d$ et $\mathbb{C}^d$

2.1. **L' espace $\mathbb{R}^d$.** On rappelle que $\mathbb{R}^d$ pour $d \geq 1$ est l' espace

$$\mathbb{R}^d := \{X = (x_1, \dots, x_d) : x_i \in \mathbb{R}\}.$$

Pour $d = 2, 3$ on note parfois par (x, y) ou (x, y, z) les points de $\mathbb{R}^d$. Les réels $x_1, \dots, x_d$ sont les *coordonnées cartésiennes* du point X. Néanmoins, il est important de penser aux points de $\mathbb{R}^d$ sans faire forcément référence aux coordonnées cartésiennes. En d' autres termes, il faut penser à $\mathbb{R}^2$ comme à une *feuille de papier blanche*, et non comme à une feuille de papier millimétré. En effet d' autres systèmes de coordonnées sont souvent utiles :

– dans $\mathbb{R}^2$ on peut utiliser les *coordonnées polaires* $(r, \theta) \in \mathbb{R}^+ \times]-\pi, \pi]$ définies par

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta, \end{cases}$$

$r = (x^2 + y^2)^{\frac{1}{2}}$.

– dans $\mathbb{R}^3$ on peut utiliser les *coordonnées sphériques* $(r, \theta, \varphi) \in \mathbb{R}^+ \times [-\pi/2, \pi/2] \times]-\pi, \pi]$ définies par

$$\begin{cases} x = r\sin\theta\cos\varphi, \\ y = r\sin\theta\sin\varphi, \\ z = r\cos\theta, \end{cases}$$

et là encore $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$.

2.1.1. $\mathbb{R}^d$ *comme espace vectoriel.* On rappelle que $\mathbb{R}^d$ est un espace vectoriel sur $\mathbb{R}$. On définit

(2.1)

$$X + Y = (x_1 + y_1, \dots, x_d + y_d), \text{ pour } X = (x_1, \dots, x_d), \ Y = (y_1, \dots y_d) \in \mathbb{R}^d,$$

$$\lambda X = (\lambda x_1, \dots, \lambda x_d) \text{ pour } X = (x_1, \dots, x_d) \in \mathbb{R}^d, \lambda \in \mathbb{R},$$

$$\vec{0} = 0_d = (0, \dots, 0) \text{ (vecteur nul).}$$

2.2. **Norme et produit scalaire sur $\mathbb{R}^d$.** On rappelle que l' on définit le *produit scalaire* de deux vecteurs de $\mathbb{R}^d$:

$$X\cdot Y := \sum_{i=1}^{d} x_i y_i,$$

qui a les propriétés suivantes :

$$\begin{aligned} &i) \quad \begin{aligned} &(X+Y)\cdot Z = X\cdot Z + Y\cdot Z,\\ &Z\cdot(X+Y) = Z\cdot X + Z\cdot Y,\\ &(\lambda X)\cdot Y = X\cdot(\lambda Y) = \lambda(X\cdot Y), \end{aligned}\\ &ii) \quad X\cdot Y = Y\cdot X,\\ &iii) \quad X\cdot X \geq 0,\ X\cdot X = 0 \ \Leftrightarrow X = \vec{0}. \end{aligned} \tag{2.2}$$

On dit que le produit scalaire $X\cdot Y$ est *bilinéaire* (propriété *i)*), *symétrique* (propriété *ii)*) et *défini positif* (propriété *iii)*).

L' intérêt des coordonnées cartésiennes est que le produit scalaire s' exprime de manière simple dans ces coordonnées. Néanmoins il est souvent inutile d' utiliser les coordonnées cartésiennes dans des démonstrations.

L' inégalité suivante, appelée *inégalité de Cauchy-Schwarz*, est fondamentale.

Proposition 2.1 (Inégalité de Cauchy-Schwarz)**.** *On a*

$$|X\cdot Y| \leq (X\cdot X)^{\frac{1}{2}}(Y\cdot Y)^{\frac{1}{2}},\ \forall X, Y\in\mathbb{R}^d. \tag{2.3}$$

Démonstration. Fixons $X, Y\in\mathbb{R}^d$ et posons pour $t\in\mathbb{R}$

$$P(t) = (X+tY)\cdot(X+tY) = t^2 Y\cdot Y + 2tX\cdot Y + X\cdot X,$$

en utilisant (2.2). $P(t)$ est un polynôme en t de degré 2 avec $P(t)\geq 0$ pour tout $t\in\mathbb{R}$. Il en suit que $P(t)$ ne peut pas avoir deux racines distinctes, donc son discriminant

$$\Delta = 4(X\cdot Y)^2 - 4(X\cdot X)(Y\cdot Y),$$

est négatif ou nul, ce qui entraine (2.3).

2.2.1. *Norme associée au produit scalaire.* On associe à ce produit scalaire la *norme euclidienne* sur $\mathbb{R}^d$ définie par :

$$\|X\| = (X\cdot X)^{\frac{1}{2}} = (x_1^2+\cdots+x_d^2)^{\frac{1}{2}},$$

qui s' interprête comme la longueur du vecteur X. Pour des raisons trop compliquées à expliquer ici, on la note parfois aussi par $\|X\|_2$. Elle possède les propriétés suivantes :

Lemme 2.2. *Pour $X, Y\in\mathbb{R}^d$ et $\lambda\in\mathbb{R}$ on a :*

$$\begin{aligned} &i) \quad \|\lambda X\| = |\lambda|\|X\|,\\ &ii) \quad \|X+Y\| \leq \|X\| + \|Y\|, \textit{(inégalité triangulaire)}\\ &iii) \quad \|X\|\geq 0,\ \|X\| = 0 \ \Leftrightarrow\ X = \vec{0}. \end{aligned} \tag{2.4}$$

Démonstration. Les propriétés *i)* et *iii)* sont immédiates. Montrons *ii)*. On a en utilisant (2.2), (2.3) :

$$\begin{aligned} & \|X+Y\|^2 = (X+Y)\cdot(X+Y) \\ (2.5) \qquad = \ & X\cdot X + 2X\cdot Y + Y\cdot Y = \|X\|^2 + 2X\cdot Y + \|Y\|^2 \\ \leq \ & \|X\|^2 + 2\|X\|\|Y\| + \|Y\|^2 = (\|X\|+\|Y\|)^2, \end{aligned}$$

ce qui montre *ii)* en passant aux racines carrées. □

Le produit scalaire permet donc de mesurer la longueur d' un vecteur. Il permet aussi de mesurer les *angles* entre deux vecteurs : si θ est l' angle entre les deux vecteurs X et Y, on a :

$$|\cos\theta| = \frac{X\cdot Y}{\|X\|\|Y\|}.$$

On définit plus généralement une *norme* sur $\mathbb{R}^d$ comme une fonction

$$\mathbb{R}^d \ni X \mapsto N(X) \in \mathbb{R}$$

qui vérifie les propriétés (2.4), c'est à dire :

$$\begin{aligned} i) \quad & N(\lambda X) = |\lambda| N(X), \\ ii) \quad & N(X+Y) \leq N(X) + N(Y), \text{(inégalité triangulaire)} \\ iii) \quad & N(X) \geq 0,\ N(X) = 0 \ \Leftrightarrow \ X = \vec{0}. \end{aligned}$$

Les fonctions

$$(2.6) \qquad \begin{aligned} & \|X\|_1 := \textstyle\sum_{i=1}^d |x_i|, \\ & \|X\|_\infty = \max_{1\leq i\leq d} |x_i| \end{aligned}$$

sont des normes sur $\mathbb{R}^d$. Notons que

$$\begin{aligned} & \max_{1\leq i\leq d} |x_i| \leq \textstyle\sum_{i=1}^d |x_i| \leq d \max_{1\leq i\leq d} |x_i|, \\ & \max_{1\leq i\leq d} |x_i| \leq (\textstyle\sum_{i=1}^d x_i^2)^{\frac{1}{2}} \leq d \max_{1\leq i\leq d} |x_i|, \end{aligned}$$

et donc

$$(2.7) \qquad \|X\|_\infty \leq \|X\|_1 \leq d\|X\|_\infty,\ \forall X \in \mathbb{R}^d,$$

et de même

$$(2.8) \qquad \|X\|_\infty \leq \|X\| \leq \sqrt{d}\|X\|_\infty,\ \forall X \in \mathbb{R}^d.$$

On dit que les normes $\|\cdot\|, \|\cdot\|_1, \|\cdot\|_\infty$ sont *équivalentes*, voir Subsect. 6.4.

2.3. **L' espace $\mathbb{C}^d$.** On rappelle que $\mathbb{C}^d$ pour $d \geq 1$ est l' espace

$$\mathbb{C}^d := \{X = (x_1, \dots, x_d) : x_i \in \mathbb{C}\}.$$

On pose

$$\mathrm{Re}X = (\mathrm{Re}x_1, \dots, \mathrm{Re}x_d),\ \mathrm{Im}X = (\mathrm{Im}x_1, \dots, \mathrm{Im}x_d) \in \mathbb{R}^d$$

et on a

$$X = \mathrm{Re}X + \mathrm{i}\mathrm{Im}X.$$

Comme $\mathbb{R}^d$, $\mathbb{C}^d$ est un espace vectoriel, avec les règles données dans (2.1) mais le corps des scalaires est maintenant $\mathbb{C}$.

Remarque 2.3. *$\mathbb{C}^d$ est aussi un espace vectoriel sur $\mathbb{R}$. Par exemple $\mathbb{C}$ s' identifie à $\mathbb{R}^2$, en identifiant $z = x + \mathrm{i}y$ au couple (x, y). Comme ensemble et comme espace vectoriel sur $\mathbb{R}$, on peut identifier $\mathbb{C}^d$ avec $\mathbb{R}^{2d}$, en introduisant les parties réelles et imaginaires de chaque coordonnée complexe de X.*

2.4. **Norme et produit scalaire sur $\mathbb{C}^d$.** On définit le *produit scalaire* de deux vecteurs $X, Y \in \mathbb{C}^d$ par

$$(Y|X) := \sum_{i=1}^{d} \overline{y}_i x_i \in \mathbb{C}.$$

Le produit scalaire sur $\mathbb{C}^d$ possède les propriétés suivantes (on dit que $(\cdot|\cdot)$ est une *forme hermitienne*) :

$$i) \quad \begin{array}{l} (Z|X+Y) = (Z|X) + (Z|Y), \ (Y|\lambda X) = \lambda(Y|X), \\ (X+Y|Z) = (X|Z) + (Y|Z), \ (\lambda Y|X) = \overline{\lambda}(Y|X), \end{array}$$

$$ii) \quad (Y|X) = \overline{(X|Y)},$$

$$iii) \quad (X|X) \geq 0, \ (X|X) = 0 \Leftrightarrow X = 0_d.$$

On dit que le produit scalaire $(Y|X)$ est *sesquilinéaire* (propriété *i)*), *hermitien* (propriété *ii)*) et *défini positif* (propriété *iii)*).

Proposition 2.4 (Inégalité de Cauchy-Schwarz)**.** *On a*

$$|(X \cdot Y)| \leq (X \cdot X)^{\frac{1}{2}} (Y \cdot Y)^{\frac{1}{2}}, \ \forall X, Y \in \mathbb{C}^d. \tag{2.9}$$

Démonstration. On commence par une remarque sur le module d' un nombre complexe z. En écrivant $z = \rho \mathrm{e}^{\mathrm{i}\theta}$ on a $\rho = |z| = z\mathrm{e}^{-\mathrm{i}\theta}$ donc comme $|z|$ est réel on a $|z| = \mathrm{Re}(z\mathrm{e}^{-\mathrm{i}\theta})$.

On applique cette remarque à $z = (X|Y)$ pour $X, Y \in \mathbb{C}^d$ et on obtient que

$$|(X|Y)| = \mathrm{Re}((X|Y)\mathrm{e}^{-\mathrm{i}\theta}) = \mathrm{Re}(X|\mathrm{e}^{-\mathrm{i}\theta}Y) = \mathrm{Re}(X|\tilde{Y}),$$

pour $\theta \in \mathbb{R}$ et $\tilde{Y} = \mathrm{e}^{-\mathrm{i}\theta}Y$. Comme $\|\tilde{Y}\| = \|Y\|$, on voit qu' il suffit de montrer l' inégalité :

$$|\mathrm{Re}(X|Y)| \leq \|vX\| \|Y\|, \ X, Y \in \mathbb{C}^d.$$

Cette inégalité se démontre comme dans le cas réel, en considérant le polynôme $P(t) = (X + tY|X + tY)$ pour $t \in \mathbb{R}$ et en utilisant les propriétés vues plus haut du produit scalaire.

La norme associée à $(\cdot|\cdot)$ est la norme euclidienne, déjà introduite dans 6.2.2 :

$$\|X\| = (X|X)^{\frac{1}{2}} = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}},$$

où on rappelle que $|x|$ désigne le *module* du nombre complexe x, défini par $|x|^2 = \overline{x}x$. On a

$$\|X\|^2 = \|\mathrm{Re}X\|^2 + \|\mathrm{Im}X\|^2.$$

Lemme 2.5. *On a*

$$\|X\| = \sup_{\|Y\| \leq 1} |(Y|X)|. \tag{2.10}$$

Démonstration. par l' inégalité de Cauchy-Schwarz on a $\sup_{\|Y\|\leq 1}|(Y|X)| \leq \|X\|$. Inversement en prenant $Y = \|X\|^{-1}X$ on a $(Y|X) = \|X\|$, donc $\sup_{\|Y\|\leq 1}|(Y|X)| \geq \|X\|$. □

Comme dans le cas de $\mathbb{R}^d$ on peut munir $\mathbb{C}^d$ d' autres normes, comme les normes $\|\cdot\|_1$, $\|\cdot\|_\infty$ définies comme dans (2.6).

2.5. **Distance sur $\mathbb{R}^d$ et $\mathbb{C}^d$.** Très souvent la structure d' espace vectoriel de $\mathbb{R}^d$ ou $\mathbb{C}^d$ ne joue aucun rôle. C'est le cas par exemple si l' on considère des fonctions $f : \mathbb{R}^d \to \mathbb{R}$ qui ne sont pas linéaires, comme par exemple $f(x_1, x_2) = \cos x_1 + x_2^2$. C'est aussi le cas si l' on restreint des fonctions linéaires à des *sous ensembles* de $\mathbb{R}^d$ ou $\mathbb{C}^d$ qui ne sont pas des sous-espaces vectoriels. Dans ce cas on identifie $\mathbb{C}^d$ avec $\mathbb{R}^{2d}$ et on ne considèrera dans la suite que l' espace $\mathbb{R}^d$.

On definit la *distance euclidienne* entre deux points X, Y par

$$d(X,Y) := \|Y - X\| = \Big(\sum_{i=1}^{d}(x_i - y_i)^2\Big)^{\frac{1}{2}}. \tag{2.11}$$

On déduit immédiatement du Lemme 2.2 le lemme suivant.

Lemme 2.6.

i) $d(X,Y) \geq 0$ *(positivité)*,

ii) $d(X,Y) = d(Y,X)$, *(symétrie)*;

iii) $d(X,Y) \leq d(X,Z) + d(Y,Z)$ *(inégalité triangulaire)*,

iv) $d(X,Y) = 0 \Leftrightarrow X = Y$ *(axiome de séparation)*.

Remarque 2.7. *Si E est un ensemble, une fonction $d : E \times E \to \mathbb{R}^+$ satisfaisant les propriétés du Lemme 2.6 s' appelle une* distance *sur E, voir Section 3.*

2.5.1. *Exemples.*

(1) en posant $d_{(1,\infty)}(X,Y) = \|X - Y\|_{(1,\infty)}$, où les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sont définies dans (2.6) on obtient deux distances d_1, d_∞ sur $\mathbb{R}^d$.

(2) La fonction $d_{\log}(x,y) = |\log_{10}(y/x)|$, (où $\log_{10}(s) = \log(s)/\log(10)$ est le logarithme en base 10) est une distance sur $]0, +\infty[$. Cette distance correspond à la notion intuitive d' *ordre de grandeur* : par exemple $d_{\log}(10^4, 10^3) = 1$, alors que $d(10^4, 10^3) = 9000$.

(3) si $d(X,Y)$ est la distance usuelle dans $\mathbb{R}^2$ on pose

$$\delta(X,Y) = \begin{cases} d(X,Y) \text{ si } 0, X, Y \text{ alignés,} \\ d(0,X) + d(0,Y) \text{ sinon.} \end{cases}$$

On vérifie que δ est une distance sur $\mathbb{R}^2$.

3. Espaces métriques

Définition 3.1. *Soit E un ensemble. Une fonction $d : E \times E \to \mathbb{R}^+$ telle que :*

$$\begin{aligned} i) &\quad d(x,y) \geq 0 \textit{ (positivité)},\\ ii) &\quad d(x,y) = d(y,x), \textit{ (symétrie)}\\ iii) &\quad d(x,y) \leq d(x,z) + d(z,y) \textit{ (inégalité triangulaire)},\\ iv) &\quad d(x,y) = 0 \Leftrightarrow x = y \textit{ (axiome de séparation)}, \ \forall x,y,z \in E \end{aligned}$$

est appelée une distance *sur E. E muni de la distance d se note (E,d) et s' appelle un* espace métrique.

Il est utile de retenir la version suivante de l' inégalité triangulaire :

$$|d(x,z) - d(y,z)| \leq d(x,y), \ x,y,z \in E. \tag{3.1}$$

3.0.2. *Distance induite.* Soit $U \subset E$ une partie de E. Si d est une distance sur E la restriction de d à $U \times U$ fait de (U,d) un espace métrique.

3.1. **Boules dans un espace métrique.**

Définition 3.2. *Soit (E,d) un espace métrique, $x_0 \in E$ et $r \geq 0$.*

(1) *L' ensemble*

$$\{x \in E : d(x_0,x) < r\}, \ \textit{noté } B(x_0,R)$$

s' appelle la boule ouverte *de centre x_0 et de rayon r.*

(2) *L' ensemble*

$$\{x \in E : d(x_0,x) \leq r\}, \ \textit{noté } B_{\mathrm{f}}(x_0,R)$$

s' appelle la boule fermée *de centre x_0 et de rayon r.*

(3) *L' ensemble*

$$\{x \in E : d(x_0,x) = r\}, \ \textit{noté } S_d(x_0,R)$$

s' appelle la sphère *de centre x_0 et de rayon r.*

Parfois on notera $B(x_0,r)$ par $B_d(x_0,r)$ pour rappeler que les boules dépendent du choix de la distance. Commençons par quelques remarques faciles.

Lemme 3.3. *On a*

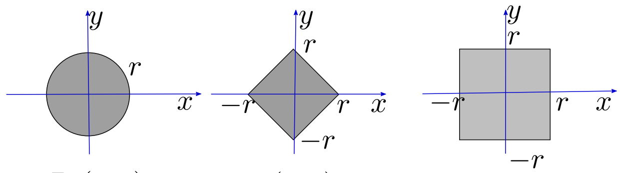
$$\begin{aligned} i) &\quad B(x_0,0) = \emptyset, \ B_{\mathrm{f}}(x_0,0) = \{x_0\},\\ ii) &\quad B(x_0,r_1) \subset B_{\mathrm{f}}(x_0,r_1) \subset B(x_0,r_2) \textit{ si } 0 \leq r_1 < r_2,\\ iii) &\quad B(x_1,r_1) \subset B(x_0,r) \textit{ si } d(x_0,x_1) + r_1 \leq r. \end{aligned}$$

Démonstration. *i)* et *ii)* sont évidents. Montrons *iii)* : si $x \in B(x_1,r_1)$ on a par l' inégalité triangulaire $d(x_0,x) \leq d(x_0,x_1) + d(x_1,x)$ et donc $d(x_0,x) < d(x_0,x_1) + r_1 < r$, donc $x \in B(x_0,R)$. □

3.1.1. *Exemples.* Dans $\mathbb{R}$ muni de $d(x,y) = |x-y|$ on a $B(x_0,r) =]x_0 - r, x_0+r[$, $B_{\mathrm{f}}(x_0,r) = [x_0-r, x_0+r]$.

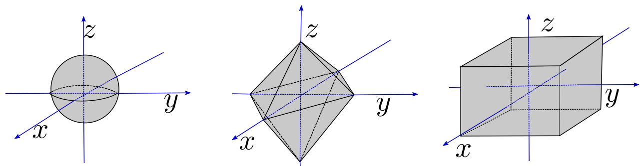
On peut remplacer la distance euclidienne dans $\mathbb{R}^d$ par une autre distance $\tilde{d}$ et définir les boules $B_{\tilde{d}}(x_0,r)$ pour cette nouvelle distance. Voici la forme des boules dans $\mathbb{R}^2$, $\mathbb{R}^3$ pour les distances d, d_1, d_∞.

$d = 2$



$B_2(0,r)$ $B_1(0,r)$ $B_\infty(0,r)$

$d = 3$



$B_2(0,r)$ $B_1(0,r)$ $B_\infty(0,r)$

3.2. **Parties bornées d' un espace métrique.**

Définition 3.4. *Une partie $A \subset E$ est* bornée *si il existe $x_0 \in E$ et $r > 0$ tels que $A \subset B(x_0,r)$.*

Lemme 3.5. *Les propriétés suivantes sont équivalentes :*
(1) *A est bornée.*
(2) *pour tout $x_0 \in E$, il existe $r > 0$ tel que $A \subset B(x_0,r)$;*
(3) *il existe $r > 0$ tel que $d(x,y) \le r$ pour tous $x, y \in A$.*

Démonstration. laissée en exercice, utiliser l' inégalité triangulaire. □

On a les propriétés élémentaires suivantes.

Lemme 3.6. (1) *toute partie finie est bornée.*
(2) *si $B \subset A$ et A borné alors B est borné.*
(3) *une union finie de parties bornées est bornée.*

Démonstration. laissée en exercice. □

Définition 3.7. *Soit $A \subset E$ bornée et non vide. Le nombre*

$$\operatorname{diam}(A) := \sup_{x,y \in A} d(x,y)$$

s' appelle le diamètre *de A.*

3.2.1. *Fonctions bornées.*

Définition 3.8. *Soit B un ensemble (par exemple $B = \mathbb{R}^n$). Une fonction $F : B \to E$ est* bornée *si*

$$f(B) = \{f(b) : b \in B\} \subset E$$

est une partie bornée de E.

3.2.2. *Distance entre deux ensembles.*

Définition 3.9. *Soit $A, B \subset E$. On pose*

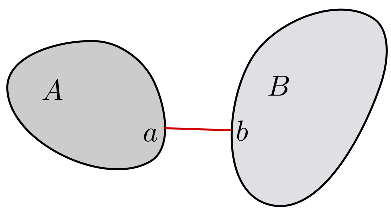
$$d(A,B) := \inf_{x \in A, y \in B} d(x,y)$$

appelé la distance *entre les deux ensembles A et B.*

De même si $x \in E$ et $A \subset E$

$$d(x,A) := \inf_{y \in A} d(x,y)$$

est appelé la distance entre le point x et l' ensemble A.



3.3. **Topologie des espaces métriques.**

Définition 3.10. (1) *Un ensemble $U \subset E$ est* ouvert *si $\forall x \in U$ $\exists \delta > 0$ tel que $B(x,\delta) \subset U$.*

(2) *Un ensemble $F \subset E$ est* fermé *si $E \setminus F$ est ouvert.*

On convient que $\emptyset$ est ouvert. Comme E est évidemment ouvert, E et $\emptyset$ sont à la fois ouverts et fermés.

Remarque 3.11. *Dans $\mathbb{R}$ les intervalles ouverts resp. fermés sont des ensembles ouverts resp. fermés. Un intervalle semi-ouvert n'est ni ouvert ni fermé, comme l'immense majorité des parties de $\mathbb{R}$.*

Lemme 3.12. *$B(x_0, r)$ est ouvert, $B_{\mathrm{f}}(x_0, r)$ est fermé.*

Démonstration. Soit $x_1 \in B(x_0, r)$ et donc $d(x_0, x_1) < r$. Par le Lemme 3.3 *iii)* si δ est assez petit pour que $d(x_0, x_1) + \delta \leq r$ on a $B(x_1, \delta) \subset B(x_0, r)$. $B(x_0, r)$ est donc bien ouvert.

Soit $x_1 \in E \setminus B_{\mathrm{f}}(x_{0,r})$, c'est à dire tel que $d(x_0, x_1) > r$. Soit $\delta > 0$ tel que $r < d(x_0, x_1) - \delta$. Si $x \in B(x_1, \delta)$ on a $d(x_0, x_1) \leq d(x_0, x) + d(x_1, x) \leq d(x_0, x) + \delta$ par l 'inégalité triangulaire et donc $r < d(x_0, x_1) - \delta \leq d(x_0, x)$, donc $x \notin B_{\mathrm{f}}(x_0, r)$. On a montré que $B(x_1, \delta) \subset E \setminus B_{\mathrm{f}}(x_0, r)$ et donc $E \setminus B_{\mathrm{f}}(x_0, r)$ est bien ouvert. $\square$

Proposition 3.13. (1) *si* $U_i, i \in I$ *sont ouverts alors* $\bigcup_{i\in I} U_i$ *est ouvert (une union quelconque d' ouverts est un ouvert).*

(2) *si* $U_1, \dots U_n$ *sont ouverts, alors* $\bigcap_{1\leq i\leq n} U_i$ *est ouvert (une intersection finie d' ouverts est un ouvert).*

(3) *si* $F_i, i \in I$ *sont fermés alors* $\bigcap_{i\in I} F_i$ *est fermé (une intersection quelconque de fermés est un fermé).*

(4) *si* $F_1, \dots F_n$ *sont fermés, alors* $\bigcup_{1\leq i\leq n} F_i$ *est fermé (une union finie de fermés est un fermé).*

Une intersection infinie d' ouverts n'est pas nécessairement un ouvert, par exemple $[0,1] = \bigcap_{n\geq 1} I_n$ pour $I_n =]-n^{-1}, 1+n^{-1}[$. De même une union infinie de fermés n'est pas nécessairement un fermé, par exemple $]0,1[= \bigcup_{n\geq 1} I_n$, pour $I_n = [n^{-1}, 1-n^{-1}]$.

Démonstration. les affirmations sur les fermés se déduisent de celles sur les ouverts par passage aux complémentaires.

(1) : si $x \in \bigcup_{i\in I} U_i$ alors il existe i_0 tel que $x \in U_{i_0}$. Comme U_{i_0} est ouvert, il existe $\delta > 0$ tel que $B(x,\delta) \subset U_{i_0} \subset \bigcup_{i\in I} U_i$ et donc $\bigcup_{i\in I} U_i$ est ouvert.

(2) : si $x \in \bigcap_{1\leq i\leq n} U_i$, alors $x \in U_i$ pour tout $1 \leq i \leq n$ et donc il existe $\delta_i > 0$ tel que $B(x,\delta_i) \subset U_i$. En prenant $\delta = \min_{1\leq i\leq n} \delta_i > 0$ on a $B(x,\delta) \subset B(x,\delta_i) \subset U_i$ pour $1 \leq i \leq n$ et donc $B(x,\delta) \subset \bigcap_{1\leq i\leq n} U_i$. $\bigcap_{1\leq i\leq n} U_i$ est donc ouvert. □

3.4. **Intérieur, adhérence, frontière.**

Définition 3.14. *Soit* $A \subset E$.

(1) *un point* $x \in E$ *est* intérieur *à* A *si il existe* $\delta > 0$ *tel que* $B(x,\delta) \subset A$. *L' ensemble des points intérieurs à* A *se note* $Int(A)$ *ou* $\mathring{A}$ *et s' appelle* l' intérieur *de* A.

Proposition 3.15. $Int(A)$ *est le plus grand ouvert inclus dans* A*, ou de manière équivalente la réunion de tous les ouverts inclus dans* A.

Démonstration. $Int(A)$ est évidemment inclus dans A. Montrons d' abord que $Int(A)$ est ouvert. Si $x_1 \in Int(A)$, il existe $\delta > 0$ tel que $B(x_1,\delta) \subset A$. Soit $x \in B(x_1, \delta/2)$. On a $B(x,\delta/2) \subset B(x_1,\delta)$ par l' inégalité triangulaire et donc $B(x,\delta/2) \subset A$, c'est à dire que $x \in Int(A)$. Ceci signifie que $B(x_1,\delta/2) \subset Int(A)$ donc $Int(A)$ est ouvert.

Soit maintenant $B \subset A$ un ouvert. Si $x_0 \in B$, il existe $\delta > 0$ tel que $B(x_0,\delta) \subset B$ comme B est ouvert et donc $B(x_0,\delta) \subset A$ car $B \subset A$. On en déduit que $x_0 \in Int(A)$ et donc que $B \subset Int(A)$. $Int(A)$ est donc bien le plus grand ouvert inclus dans A. □

Définition 3.16. *Soit* $A \subset E$.

(1) *un point* $x \in E$ *est* adhérent *à* A *si* $B(x_0, r)$ *intersecte* A *pour tout* $r > 0$. *L' ensemble des points adhérents à* A *se note* $Adh(A)$ *ou* $\overline{A}$ *et s' appelle l'*adhérence *ou la* fermeture *de* A.

Proposition 3.17. $Adh(A)$ *est le plus petit fermé contenant* A*, ou de manière équivalente l' intersection de tous les fermés contenant* A.

Démonstration. A est évidemment inclus dans $Adh(A)$. Montrons que $Adh(A)$ est fermé, c'est à dire que $E \setminus Adh(A)$ est ouvert. Soit $x_0 \notin Adh(A)$. Par définition il existe $\delta > 0$ tel que $B(x_0, \delta) \cap A = \emptyset$. Si $x \in B(x_0, \delta/2)$ alors $B(x, \delta/2) \subset B(x_0, \delta)$ par l 'inégalité triangulaire et donc $B(x, \delta/2) \cap A = \emptyset$ c'est à dire que $B(x, \delta/2) \subset E \setminus Adh(A)$. L' ensemble $E \setminus Adh(A)$ est donc bien ouvert.

Soit F un fermé contenant A. Montrons que $Adh(A) \subset F$ ou de manière équivalente que $E \setminus F \subset E \setminus Adh(A)$. Si $x_0 \notin F$ alors il existe $\delta > 0$ tel que $B(x_0, \delta) \subset E \setminus F$ car F est fermé. On a donc $B(x_0, \delta) \cap F = \emptyset$ et donc $B(x_0, \delta) \cap A = \emptyset$ comme $A \subset F$. Donc $x_0 \notin Adh(A)$.

On a donc montré que $Adh(A) \subset F$ si F est un fermé contenant A, donc $Adh(A)$ est le plus petit fermé contenant A. □

Définition 3.18. *Soit $A \subset B \subset E$. On dit que A est* dense *dans B si $B \subset Adh(A)$.*

Définition 3.19. *Soit $A \subset E$. L' ensemble $Adh(A) \cap Adh(E \setminus A)$ se note $Fr(A)$ ou ∂A et s' appelle la* frontière *ou le* bord *de A.*

Un point x_0 appartient à $Fr(A)$ si pour tout $\delta > 0$ $B(x_0, \delta)$ intersecte à la fois A et $E \setminus A$.

3.4.1. *Exemples.* Dans $\mathbb{R}$ muni de la distance usuelle on a

$$Int(\mathbb{Q}) = Int(\mathbb{R} \setminus \mathbb{Q}) = \emptyset, \ Adh(\mathbb{Q}) = Adh(\mathbb{R} \setminus \mathbb{Q}) = \mathbb{R}.$$

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

Dans $\mathbb{R}^d$ (et dans tous les espaces vectoriels normés, voir Section 6) on a

$$\begin{aligned} Int(B(x_0, r)) &= Int(B_{\mathrm{f}}(x_0, r)) = B(x_0, r), \\ Adh(B(x_0, r)) &= Adh(B_{\mathrm{f}}(x_0, r)) = B_{\mathrm{f}}(x_0, r), \\ Fr(B(x_0, r)) &= Fr(B_{\mathrm{f}}(x_0, r)) = S(x_0, r). \end{aligned}$$

Remarque 3.20. *Dans un espace métrique l' adhérence de la boule ouverte $B(x_0, r)$ n' est* pas toujours égale *à la boule fermée $B_{\mathrm{f}}(x_0, r)$. Prenons par exemple $E = \{a, b, c\}$ et posons*

$$\begin{aligned} d(a, a) &= d(b, b) = d(c, c) = 0, \\ d(a, b) &= d(b, a) = d(b, c) = d(c, b) = 1, \\ d(a, c) &= d(c, a) = 2. \end{aligned}$$

On vérifie que d est une distance sur E. On a

$$B(a, 2) = \{a, b\} = Adh(B(a, 2)), \ B_{\mathrm{f}}(a, 2) = \{a, b, c\}.$$

3.4.2. *Propriétés de base.*

Proposition 3.21. (1) *$Int(A) \subset A \subset Adh(A)$.*
(2) *$E = Int(E \setminus A) \cup Fr(A) \cup Int(A)$, les trois ensembles étant disjoints.*
(3) *$E \setminus Int(A) = Adh(E \setminus A)$.*
(4) *$E \setminus Adh(A) = Int(E \setminus A)$.*
(5) *$Fr(A) = Adh(A) \setminus Int(A)$.*

Démonstration. laissée au lecteur. □

Proposition 3.22. (1) *A est ouvert si et seulement si $A = Int(A)$.*
(2) *A est fermé si et seulement si $A = Adh(A)$.*
(3) *$x \in Adh(A)$ si et seulement si $d(x, A) = 0$.*
(4) *$x \in Int(A)$ si et seulement si $d(x, E \setminus A) > 0$.*

Démonstration. laissé au lecteur. □

3.5. **Suites dans un espace métrique.**

Définition 3.23. *Une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans E est une fonction $\mathbb{N} \ni n \mapsto x_n \in E$. Pour n fixé, l' élément x_n est appelé le* n-ième terme *de la suite $(x_n)_{n\in\mathbb{N}}$.*

Si $E = \mathbb{R}^d$n a évidemment $x_n = (x_{1,n}, \ldots x_{d,n})$, où les $(x_{i,n})_{n\in\mathbb{N}}$ sont des suites réelles. Se donner une suite $(x_n)_{n\in\mathbb{N}}$ dans $\mathbb{R}^d$ revient donc, après le choix des coordonnées canoniques, à se donner d suites réelles $(x_{i,n})_{n\in\mathbb{N}}$.

Définition 3.24. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite dans E et $x \in E$. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x et on écrit $\lim_{n\to\infty} x_n = x$, ou $\lim x_n = x$ ou encore $x_n \to x$ si pour tout $\epsilon > 0$ il existe $N \in \mathbb{N}$ tel que $d(x_n, x) \leq \epsilon$ pour $n \geq N$.*

De manière équivalente $\lim_{n\to\infty} x_n = x$ si et seulement si $\lim_{n\to\infty} d(x_n, x) = 0$.

Une suite $(x_n)_{n\in\mathbb{N}}$ est bornée *si l' ensemble de ses termes $\{x_n : n \in \mathbb{N}\}$ est borné.*

Si $E = \mathbb{R}^d$ d' après 2.8, on a

$$\sup_{1\leq i\leq d} |x_i - x_{i,n}| \leq d(x_n, x) \leq d \sup_{1\leq i\leq d} |x_i - x_{i,n}|,$$

et donc, après avoir choisi les coordonnées canoniques, $\lim_{n\to\infty} x_n = X$ si et seulement si $\lim_{n\to\infty} x_{i,n} = x_i$ pour tout $1 \leq i \leq d$.

Toute suite convergente est bornée.

Lemme 3.25. *La limite d' une suite convergente est unique.*

Démonstration. Supposons que $x_n \to x$ et $x_n \to x'$. Alors $d(x, x') \leq d(x, x_n) + d(x_n, x')$ et donc $d(x, x') = 0$ en faisant tendre n vers l' infini, ce qui montre que $x = x'$. □

Proposition 3.26. *Soit $A \subset E$.*
(1) *$x \in Adh(A)$ si et seulement si il existe une suite $(x_n)_{n\in\mathbb{N}}$* d' éléments de A *telle que $x = \lim_{n\to\infty} x_n$.*
(2) *A est fermé si et seulement si la limite de toute suite convergente $(x_n)_{n\in\mathbb{N}}$* d' éléments de A *appartient aussi à A.*

Démonstration. (1) : Soit $x \in Adh(A)$. Par définition $B(x, n^{-1}) \cap A \neq \emptyset$ pour tout $n \in \mathbb{N}$. Il existe donc un élément $x_n \in B(x, n^{-1}) \cap A$. On a $x_n \in A$ et comme $d(x_n, x) \leq n^{-1}$ $x_n \to x$. Inversement s' il existe une suite $(x_n)_{n\in\mathbb{N}}$ d' éléments de A telle que $\lim x_n = x$, donc $\lim d(x, x_n) = 0$. Pour tout $\delta > 0$ il existe donc un $N \in \mathbb{N}$ tel que $d(x, x_N) < \delta$, c'est à dire que $x_N \in B(x, \delta)$. Comme $x_N \in A$, on en déduit que $x \in Adh(A)$.

(2) : on a vu dans (1) que $x \in Adh(A)$ ssi il existe une suite $(x_n)_{n\in\mathbb{N}}$ d' éléments de A avec $\lim x_n = x$. Comme A est fermé ssi $Adh(A) = A$ on obtient 2). □

3.5.1. *Suites de Cauchy.* La notion de suite de Cauchy est absolument fondamentale. Elle permet de montrer qu' une suite est convergente sans avoir besoin de connaitre sa limite.

Définition 3.27. *Une suite $(x_n)_{n\in\mathbb{N}}$ dans E est une* suite de Cauchy *si pour tout $\epsilon > 0$ il existe $N \in \mathbb{N}$ tel que $d(x_p, x_n) \leq \epsilon$ pour tous $n, p \geq N$.*

Proposition 3.28. (1) *toute suite convergente est de Cauchy.*
(2) *toute suite de Cauchy est bornée.*

Démonstration. (1) : supposons que $\lim_n x_n = x$. Pour $\epsilon > 0$ on peut trouver $N \in \mathbb{N}$ tel que $d(x, x_n) \leq \epsilon$ pour tout $n \geq N$ et donc $d(x_n, x_p) \leq d(x_n, x) = d(x, x_p) \leq 2\epsilon$ pour tout $n, np \geq N$. Remplaçant ϵ par $\epsilon/2$ on obtient que $(x_n)_{n\in\mathbb{N}}$ est une suite de Cauchy.

(2) : Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy. Par définition (prendre $\epsilon = 1$) il existe N tel que $d(x_n, x_p) \leq 1$ pour $n;, p \geq N$ et donc $d(x_n, x_N) \leq 1$ pour $n \geq N$. On a donc

$$d(x_n, x_N) \leq 1 + \sup_{1\leq i\leq N} d(x_n, x_N) =: r_0, \ \forall n \in \mathbb{N},$$

et donc $x_n \in B(x_N, r_0)$ pour tout $n \in \mathbb{N}$. □

Définition 3.29. *Un espace métrique (E, d) est dit* complet *si toute suite de Cauchy dans E est convergente.*

Théorème 3.30. *$\mathbb{R}^d$ muni de la distance canonique est complet.*

Démonstration. on sait que $\mathbb{R}$ est complet (rappelons que cette propriété découle de la propriété de la borne supérieure). Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $\mathbb{R}^d$ et $(x_{i,n})_{n\in\mathbb{N}}$ ses suites de coordonnées. En utilisant (2.8) chaque suite réelle $(x_{i,n})_{n\in\mathbb{N}}$ est une suite de Cauchy, et donc converge vers $x_i \in \mathbb{R}$. On en déduit que $\lim X_n = (x_1, \ldots, x_d)$. □

3.5.2. *Sous-suites.*

Définition 3.31. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite dans E. Une suite $y_n = x_{\varphi(n)}$ où $\varphi : \mathbb{N} \to \mathbb{N}$ est une application* strictement croissante *est appelée une* sous-suite *ou* suite extraite *de la suite $(x_n)_{n\in\mathbb{N}}$.*

Il est utile de remarquer que si $\varphi : \mathbb{N} \to \mathbb{N}$ est strictement croissante, on a $\varphi(n) \geq \varphi(0) + n$ et donc $\lim \varphi(n) = +\infty$.

Proposition 3.32. (1) *Toute sous-suite d' une suite convergente converge vers la même limite.*
(2) *une suite de Cauchy possèdant une sous-suite qui converge vers un point x converge aussi vers x.*

Démonstration. (1) : supposons que $\lim x_n = x$ et soit $y_n = x_{\varphi(n)}$ uen sous-suite de (x_n). Pour $\epsilon > 0$ il existe $N \in \mathbb{N}$ tel que $d(x_p, x) \leq \epsilon$ pour $p \geq N$. Soit $M \in \mathbb{N}$ tel que $\varphi(M) \geq N$. Si $n \geq M$ on a $\varphi(n) \geq \varphi(M) \geq N$ et donc $d(y_n, x) = f(x_{\varphi(n)}, x) \leq \epsilon$. On a donc bien $\lim y_n = x$.

(2) : soit (x_n) une suite de Cauchy et $y_n = x_{\varphi(n)}$ une sous-suite avec $\lim y_n = x$. Soit $\epsilon > 0$. Il existe N tel que $d(x_n, x_p) \leq \epsilon$ pour $n, p \geq N$. Il existe aussi M tel que $d(x_{\varphi(n)}, x) \leq \epsilon$ si $n \geq M$. Choisissons $M_0 \geq M$ tel que $\varphi(M_0) \geq N$, ce qui est possible car $\lim \varphi(n) = +\infty$. Si $n \geq \varphi(M_0)$ on a donc $d(x_n, x) \leq d(x_n, x_{\varphi(M_0)}) + d(x_{\varphi(M_0)}, x) \leq 2\epsilon$. En remplaçant ϵ par $\epsilon/2$ on obtient bien que $\lim x_n = x$. □

3.6. **Comment déterminer l' adhérence ou l' intérieur d' un ensemble.** Un exercice classique consiste à se donner un ensemble $A \subset \mathbb{R}$, $\mathbb{R}^2$ ou $\mathbb{R}^3$ et à déterminer son adhérence ou son intérieur.

Tout d' abord on peut remarquer qu' il s' agit en fait du même problème. En effet on a

$$\begin{aligned} &i) \quad \mathrm{Int}(A) = E \setminus \mathrm{Adh}(E \setminus A), \\ &ii) \quad \mathrm{Adh}(A) = E \setminus \mathrm{Int}(E \setminus A), \end{aligned} \tag{3.2}$$

($E = \mathbb{R}^d$, $d = 1, 2, 3$ ici) et connaître un ensemble est équivalent à connaître son complémentaire. Néanmoins il vaut mieux raisonner directement pour éviter de s' embrouiller dans les complémentaires. Voici la marche à suivre :

1) on commence par dessiner sur sa feuille l' ensemble A. C'est parfois difficile, par exemple on ne peut pas dessiner $\mathbb{Q}$ dans $\mathbb{R}$, mais pour exprimer le fait que $\mathbb{Q}$ est dense dans $\mathbb{R}$, on peut représenter $\mathbb{Q}$ comme un nuage de points sur la droite réelle.

2) le dessin permet de deviner ce que sont $\mathrm{Int}(A)$, noté B et $\mathrm{Adh}(A)$, noté C.

3) Comment justifier que notre candidat B est bien égal à $\mathrm{Int}(A)$?

- on montre que B est ouvert et $B \subset A$. Comme $\mathrm{Int}(A)$ est le plus grand ouvert inclus dans A, on a donc montré que $B \subset \mathrm{Int}(A)$.

- il reste à montrer que $\mathrm{Int}(A)$ ne contient pas d' autres points que ceux de B, c'est à dire que les points de A qui ne sont pas dans B ne sont pas dans $\mathrm{Int}(A)$, c'est à dire que si $X \in A \setminus B$ alors $X \in \mathrm{Adh}(E \setminus A)$ (voir (3.2) *i)*). Pour celà pour chaque $X \in A \setminus B$, il faut trouver une suite X_n avec $X_n \notin A$ et $\lim X_n = X$.

4) Comment justifier que notre candidat C est bien egal à $\mathrm{Adh}(A)$?

- on montre que C est fermé et $A \subset C$. Comme $\mathrm{Adh}(A)$ est le plus petit fermé contenant A on a donc montré que $\mathrm{Adh}(A) \subset C$.

- on montre que $C \subset \mathrm{Adh}(A)$ en trouvant pour chaque $X \in C$ une suite X_n avec $X_n \in A$ et $\lim X_n = X$. Il suffit de considérer les $X \in C \setminus A$ (si $X \in A$ on peut prendre la suite constante $X_n = X$).

3.7. **Compacité.**

Définition 3.33. *Soit $F \subset E$. Un* recouvrement ouvert *de F est une famille $(U_i)_{i \in I}$ d' ouverts de E telle que $F \subset \bigcup_{i \in I} U_i$.*

F est compact *si tout recouvrement ouvert de F admet un sous recouvrement fini, c'est à dire que si $(U_i)_{i \in I}$ est une famille d' ouverts telle que*

$F \subset \bigcup_{i\in I} U_i$ alors il existe $i_1, \ldots, i_n \in I$ tels que

$$F \subset U_{i_1} \cup U_{i_2} \cdots \cup U_{i_n}.$$

3.7.1. *Compacité et sous-suites.* Voici une autre caractérisation de la compacité, connue sous le nom de propriété de *Bolzano-Weierstrass* ou de *compacité séquentielle.*

Théorème 3.34. *$K \subset E$ est compact si et seulement si toute suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans K possède une sous-suite qui converge vers un élément x de K.*

La preuve du Théorème 3.34 est hors programme. Une preuve est donnée dans l' appendice A.1.

3.7.2. *Propriétés de base.*

Proposition 3.35. (1) *tout compact $K \subset E$ est fermé et borné.*
(2) *si K est compact et F fermé alors $K \cap F$ est compact.*
(3) *Si K est compact, toute suite de Cauchy dans K converge vers un élément de K, c'est à dire que (K, d) est complet.*

Démonstration. (1) : soit K compact et $x \in Adh(K)$. Il existe une suite (x_n) d' éléments de K telle que $x_n \to x$. Comme K est compact, il existe une sous-suite $x_{\varphi(n)}$ telle que $\lim x_{\varphi(n)} = y \in K$. Par unicité de la limite $x = y$ et donc $x \in K$. On a donc $Adh(K) \subset K$ c'est à dire que K est fermé.

Montrons maintenant que K est borné. En posant $U_x = B(x, 1)$ on obtien un recouvrement ouvert $\bigcup_{x\in K} U_x$. Comme K est compact, il existe $x_1, \ldots, x_n \in K$ tels que $K \subset \bigcup_{i=1}^n B(x_i, 1)$ et donc K est borné.

(2) : Soit (x_n) une suite dans $K \cap F$. Comme K est compact il existe une sous-suite $(x_{\varphi(n)})$ qui converge vers $x \in K$. Comme $(x_{\varphi(n)})$ est une suite dans le fermé F on a $x \in F$ et donc $x \in K \cap F$. $F \cap K$ est séquentiellement compact, donc compact.

(3) est laissé en exercice. □

3.7.3. *Theorème de Borel-Lebesgue.*

Théorème 3.36. *[Borel-Lebesgue] $K \subset \mathbb{R}^d$ est compact si et seulement si K est fermé et borné.*

Pour montrer le Théorème 3.36 on utilise le lemme suivant.

Lemme 3.37. *Les boules fermées $B_{\mathrm{f}}(x_0, r)$ sont compactes pour tout $x_0 \in \mathbb{R}^d$.*

Démonstration. commençons par considérer le cas $d = 1$ et par montrer que les intervalles fermés bornés $[a, b]$ sont compacts. Soit $(U_i)_{i\in I}$ un recouvrement ouvert fixé de $[a, b]$. Soit F l' ensemble des $x \in [a, b]$ tels que $[a, x] \subset \bigcup_{i\in J} U_i$, où $J \subset I$ est une partie finie (dépendant a priori de x). Pour montrer que $[a, b]$ est compact, il suffit de montrer que $b \in F$.

F est non vide : en effet $a \in F$ car $[a, a] = \{a\}$ est inclus dans un des U_i. F est majoré par b, donc possède une borne supérieure c.

Montrons que $c = b$. Supposons que $c < b$ et soit $i_0 \in I$ tel que $c \in U_{i_0}$ et $\alpha > 0$ tel que $]c - \alpha, c + \alpha[\subset U_{i_0}$. Un tel α existe comme U_{i_0} est ouvert.

Soit maintenant $x \in F \cap]c-\alpha, c]$ (un tel x existe comme $c = \sup F$). Comme $x \in F$ il existe $J \subset I$ finie tel que $[a, x] \subset \bigcup_{i \in J} U_i$. Comme $]c-\alpha, c+\alpha[\subset U_{i_0}$ et $c - \alpha < x$ on a $[a, c+\alpha/2] \subset [a, c+\alpha[\subset \bigcup_{i\in J} U_i \cup U_{i_0}$. Ceci signifie que $c+\alpha/2 \in F$, ce qui contredit que c est un majorant de F. On a donc $c = b$, c'est à dire $\sup F = b$.

Il reste à montrer que $b \in F$. On choisit à nouveau i_1 tel que $b \in U_{i_1}$ et $\alpha > 0$ tel que $]b-\alpha, b+\alpha[\subset U_{i_1}$. Comme plus haut on choisit $x \in]b-\alpha, b] \cap F$, l' intervalle $[a, x]$ est recouvert par un nombre fini de U_i, et en rajoutant U_{i_1} à ce recouvrement, on obtient un recouvrement fini de $[a, b]$, c'est à dire que $b \in F$. $\square$

Considérons maintenant le cas d quelconque et montrons que $B_{\mathrm{f}}(X_0, r)$ est séquentiellement compact. Soit (X_n) une suite dans $B_{\mathrm{f}}(X_0, r)$. En posant $X_0 = (a_1, \dots, a_d)$ et $X_n = (x_{1,n}, \dots, x_{d,n})$ les suites réelles $(x_{i,n})$ sont à valeurs dans les compacts $[a_i - r, a_i + r]$. Pour $i = 1$, il existe une sous-suite $(x_{1,\varphi_1(n)})$ qui converge vers b_1. La suite $x_{2,\varphi_1(n)}$ est à valeurs dans le compact $[a_2 - r, a_2 + r]$, on peut donc en extraire une sous-suite $(x_{2,\varphi_1(\varphi_2(n))})$ qui converge vers un réel b_2. La suite $(x_{1,\varphi_1(\varphi_2(n))})$ est extraite de la suite $(x_{1,\varphi_1(n)})$ et donc converge vers b_1.

En répétant l' argument, on obtient $\varphi = \varphi_1 \circ \cdots \varphi_d : \mathbb{N} \to \mathbb{N}$ tel que $(x_{i,\varphi(n)})$ converge vers b_i pour tout $1 \leq i \leq d$. Ceci entraine que $(X_{\varphi(n)})$ converge vers $B = (b_1, \dots, b_d)$. La boule $B_{\mathrm{f}}(X_0, r)$ est donc séquentiellement compacte, donc compacte. $\square$

Preuve du Théorème 3.36. On a déjà vu qu' un compact était fermé et borné. Soit maintenant K fermé et borné. Comme K est fermé, K est inclus dans le $B_{\mathrm{f}}(0, r)$ pour r assez grand, et donc $K = B_{\mathrm{f}}(0, r) \cap K$ est compact comme intersection d' un compact et d' un fermé. $\square$

3.8. **Limites et applications continues.**

Définition 3.38. *Soit (E_1, d_1) (E_2, d_2) deux espaces métriques, $x_0 \in E_1$, $l \in E_2$ et $F : E_1 \to E_2$ une application.*

1. *on dit que $\lim_{x \to x_0} F(x) = l$ pour tout $\epsilon > 0$, il existe $\alpha > 0$ tel que pour tout $x \in E_1$ tel que $d_1(x, x_0) \leq \alpha$ on a $d_2(f(x), l) \leq \epsilon$.*
2. *on dit que F est continue en x_0 si $\lim_{x\to x_0} f(x) = f(x_0)$.*
3. *on dit que F est continue si F est continue en tout point $x_0 \in E_1$.*

Proposition 3.39. *Soit (E_1, d_1) (E_2, d_2) deux espaces métriques et $F : E_1 \to E_2$ une application. Les propriétés suivantes sont équivalentes :*

1. *F est continue.*
2. *pour tout ouvert $U \subset E_2$, $F^{-1}(U)$ est ouvert dans E_1.*
3. *pour tout fermé $F \subset E_2$, $F^{-1}(F)$ est fermé dans E_1.*
4. *pour toute suite $(x_n)_{n\in\mathbb{N}}$ dans E_1 avec $\lim x_n = x \in E_1$ on a $\lim F(x_n) = F(x)$.*

Démonstration. (1) $\Rightarrow$ (2) : Soit $U \subset E_2$ ouvert et $x_0 \in F^{-1}(U)$, c'est à dire $y_0 = F(x_0) \in U$. Comme U est ouvert, il existe $\epsilon > 0$ tel que $B(y_0, \epsilon) \subset U$. Comme F est continue au point x_0 il existe $\alpha > 0$ tel que

$$x \in E_1,\ d(x, x_0) < \alpha \Rightarrow d(F(x), y_0) < \epsilon.$$

De manière équivalente si $x \in B(x_0, \alpha)$ alors $F(x) \in B(y_0, \epsilon) \subset U$. On a donc montré que pour tout $x_0 \in F^{-1}(U)$, il existe $\alpha(x_0)$ tel que $B(x_0, \alpha(x_0)) \subset F^{-1}(U)$, c'est à dire que $F^{-1}(U)$ est ouvert.

(2) $\Rightarrow$ (3) : passer aux complémentaires.

(3) $\Rightarrow$ (4) : soit (x_n) une suite dans E_1 avec $\lim x_n = x \in E_1$. Supposons que la suite $(F(x_n))$ ne converge pas vers $F(x)$. Alors il existe $\epsilon_0 > 0$ tel que $\forall n \in \mathbb{N}$, il existe $p(n) \geq n$ avec $d(F(x_{p(n)}), F(x)) \geq \epsilon_0$. Soit $y_n = x_{p(n)}$ et $U = E_2 \setminus B(F(x), \epsilon_0)$. U est fermé, $F(y_n) \in U$, donc $y_n \in F^{-1}(U)$, qui est fermé par la propriété (3). Comme (y_n) est une sous-suite de la suite (x_n) qui converge vers x, on a $\lim y_n = x$. Comme $F^{-1}(U)$ est fermé, on en déduit que $x \in F^{-1}(U)$, c'est à dire que $F(x) \in U$ ou encore que $d(F(x), F(x)) \geq \epsilon_0$ ce qui est faux.

(4) $\Rightarrow$ (1) : supposons que F ne soit pas continue en un point x_0 de E_1. Alors il existe $\epsilon_0 > 0$ tel que $\forall \delta > 0$ il existe $x_\delta \in E_1$ avec $d_1(x_\delta, x_0) < \delta$ et $d_2(F(x_\delta), F(x)) \geq \epsilon_0$. En prenant $\delta = n^{-1}$ on obtient une suite $(x_n)_{n \geq 1}$ telle que $\lim x_n = x_0$ et $d_2(F(x_n), F(x_0)) \geq \epsilon_0$, ce qui contredit (4). $\square$

Voici l' exemple canonique d' une application continue entre espaces métriques.

Lemme 3.40. *Soit (E, d) un espace métrique et $x_0 \in E$. Alors*

$$E \ni x \mapsto d(x_0, x) \in \mathbb{R}$$

est continue.

Démonstration. L' inégalité triangulaire montre que $|d(x, x_0) - d(x', x_0)| \leq d(x, x')$ ce qui entraine la continuité. $\square$

4. Limites et continuité des fonctions de plusieurs variables

Soit $D \subset \mathbb{R}^n$ et $F : D \to \mathbb{R}^p$ une application. L' ensemble D est donc le *domaine de définition* de F et en introduisant les coordonnées cartésiennes on a

$$F(x_1, \dots, x_n) = (F_1(x_1, \dots, x_n), \dots F_p(x_1, \dots, x_n)),$$

où $f_i : D \to \mathbb{R}$ sont des fonctions de D dans $\mathbb{R}$. Se donner une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$ revient donc à se donner p fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. Dans la pratique le domaine D sera souvent égal à $\mathbb{R}^n$ ou à un ouvert de $\mathbb{R}^n$.

On peut directement appliquer la Sous-section 3.8 aux deux espaces métriques (D, d) et $(\mathbb{R}^p, d)$ où d est la distance euclidienne sur $\mathbb{R}^n$ ou $\mathbb{R}^p$. Ceci conduit aux résultats suivants.

Définition 4.1. *Soit $D \subset \mathbb{R}^n$, $F : D \to \mathbb{R}^p$ et $X_0 \in D$.*

(1) *on dit que $\lim_{X \to X_0} F(X) = L$ pour $L \in \mathbb{R}^p$ si pour tout $\epsilon > 0$, il existe $\alpha > 0$ tel que pour tout $X \in D$ tel que $d(X, X_0) \leq \alpha$ on a $d(f(X), L) \leq \epsilon$.*

(2) *on dit que F est continue en X_0 si $\lim_{X \to X_0} f(X) = f(X_0)$.*

(3) *on dit que F est continue sur D si F est continue en tout point $X_0 \in D$.*

Proposition 4.2. *Soit $D \subset \mathbb{R}^n$ et $F : D \to \mathbb{R}^p$ une application. Les propriétés suivantes sont équivalentes :*

(1) *F est continue sur D.*

(2) *pour tout ouvert* $U \subset \mathbb{R}^p$ $F^{-1}(U) = V \cap D$ *où* V *est ouvert dans* $\mathbb{R}^n$.
(3) *pour tout fermé* $U \subset \mathbb{R}^p$ $F^{-1}(U) = V \cap D$ *où* V *est fermé dans* $\mathbb{R}^n$.
(4) *pour toute suite* $(X_n)_{n\in\mathbb{N}}$ *dans* $\mathbb{R}^n$ *avec* $\lim X_n = X \in D$ *on a* $\lim F(X_n) = F(X)$.

4.1. **Comment construire des applications continues.**

Lemme 4.3. *Soit* $D \subset \mathbb{R}^n$ *et* $F : D \to \mathbb{R}^p$ *une application. Alors* F *est continue sur* D *si et seulement si les composantes* F_i *de* F *définies par*

$$F(X) = (F_1(X), \dots, F_p(X))$$

sont continues sur D.

Démonstration. Le lemme suit directement de la Proposition 4.2 (4). □

Corollaire 4.4. *Les fonctions coordonnées* $\mathbb{R}^n \ni X \mapsto x_i \in \mathbb{R}$ *pour* $1 \leq i \leq n$ *sont continues.*

Démonstration. les fonctions coordonnées sont les composantes de l' identité $X \mapsto X$ qui est évidemment continue. □

Le Lemme 4.3 permet donc de se ramener à des fonctions $f : D \to \mathbb{R}$. La proposition suivante se montre comme pour des fonctions d' une variable.

Proposition 4.5. *Soit* $D \subset \mathbb{R}^n$ *et* $f, g : D \to \mathbb{R}$ *deux fonctions continues sur* D.
(1) *les fonctions* $f + g, fg : D \to \mathbb{R}$ *sont continues sur* D.
(2) *si* $g(X) \neq 0$ *pour tout* $X \in D$ *la fonction* $fg^{-1} : D \to \mathbb{R}$ *est continue.*
(3) *soit* $I \subset \mathbb{R}$ *un intervalle tel que* $f(D) \subset I$ *et* $\varphi : I \to \mathbb{R}$ *une fonction continue. Alors* $\varphi \circ f : D \to \mathbb{R}$ *est continue sur* D.

On déduit donc de la Proposition 4.5 et du corollaire 4.4 que les fonctions polynômes :

$$P : X \mapsto \sum_{\alpha_1+\cdots+\alpha_n \leq D} a_{\alpha_1\dots\alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

sont continues sur $\mathbb{R}^n$.

4.2. **Comment montrer qu' un ensemble est ouvert ou fermé.** Très souvent une partie de $\mathbb{R}^n$ est définie par un certain nombre d' inégalités strictes ou larges. On peut alors utiliser la Proposition 4.2 *2)* pour montrer que cette partie est ouverte ou fermée. Le lemme suivant est laissé en exercice.

Lemme 4.6. *Soit* $I_1, \dots, \dots I_n$ *des intervalles ouverts resp. fermés. Alors* $I_1 \times \cdots \times I_n$ *est ouvert resp. fermé dans* $\mathbb{R}^n$.

L' ensemble

$$D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + 2x_1x_2x_3 < 0,\ \sin(x_1x_2) < x_3\}$$

est un ouvert de $\mathbb{R}^3$, car c'est l' image réciproque de l' ouvert $U =]-\infty, 0[\times]-\infty, 0[$ de $\mathbb{R}^2$ par l' application continue

$$F(x_1, x_2, x_3) = (x_1^2 + 2x_1x_2x_3, \sin(x_1x_2) - x_3).$$

De même l' ensemble

$$D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : 2x_1^2 + 3x_2^2 + 5x_3^2 \leq 1\}$$

est fermé comme image réciproque du fermé $]-\infty, 1]$ de $\mathbb{R}$ par l' application continue

$$F(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + 5x_3^2.$$

D' autres méthodes pour montrer qu' un ensemble est (ou n'est pas) ouvert ou fermé seront vues en travaux dirigés.

4.3. **Lien avec la compacité.**

Théorème 4.7. *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application continue et $K \subset \mathbb{R}^n$ un compact. Alors $f(K) \subset \mathbb{R}^p$ est compact.*

Démonstration. Soit $(Y_n) = (f(X_n))$ une suite dans $f(K)$. La suite (X_n) est à valeurs dans le compact K donc admet une sous-suite $X_{\varphi(n)}$ convergente vers $X \in K$. Comme f est continue la suite $(Y_{\varphi(n)}) = f(X_{\varphi(n)})$ converge vers $f(X) \in f(K)$. $\square$

Corollaire 4.8. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue et $K \subset \mathbb{R}^n$ un compact. Alors $f_{|K}$ est bornée et atteint ses bornes.*

Démonstration. $f(K)$ est compact dans $\mathbb{R}$ donc borné. Soit $a = \sup_{X \in K} f(X)$. Par définition il existe une suite (X_n) à valeurs dans K telle que $f(X_n) \to a$. (X_n) admet une sous-suite $(X_{\varphi(n)})$ qui converge vers $X \in K$. Comme f est continue $a = \lim f(X_{\varphi(n)}) = f(X)$ donc $\sup_K f$ est atteint. Le même argument montre que $\inf_K f$ est atteint. $\square$

4.4. **Continuité partielle.** Soit $D \subset \mathbb{R}^n$ un ouvert et $f : D \to \mathbb{R}$ une fonction. Comme D est ouvert, pour tout $A = (a_1, \ldots, a_n) \in D$, il existe des intervalles I_i de centre a_i tels que $I_1 \times \cdots \times I_n \subset D$. En introduisant les coordonnées cartésiennes, on peut donc définir les *fonctions partielles* au point A, qui sont des fonctions d' une seule variable

$$f_i(t) := f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n).$$

obtenues en gelant les coordonnées x_j pour $j \neq i$. En d' autres termes la fonction f_i est la restriction de f au segment $\{a_1\} \times \cdots \times \{a_{i-1}\} \times I_i \times \{a_{i+1}\} \times \cdots \times \{a_n\}$.

Définition 4.9. *La fonction $f : D \to \mathbb{R}$ est* partiellement continue *au point $A \in D$ si toutes les fonctions partielles f_i sont continues au point a_i. Elle est dite partiellement continue sur D si elle est partiellement continue en chaque point de D.*

Il est facile de voir que la continuité implique la continuité partielle. La réciproque est fausse, comme le montre l' exemple suivant. Soit

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2}, \text{ si } (x_1, x_2) \neq (0,0), \\ 0 \text{ si } (x_1, x_2) = (0,0). \end{cases}$$

On vérifie que f est continue (et donc séparément continue) en tout point différent de $(0,0)$. On a $f(0, x_2) = 0$ et $f(x_1, 0) = 0$ pour tout x_1, x_2 et donc f est partiellement continue en $(0,0)$. Si f était continue au point $(0,0)$, on aurait en particulier $\lim_{t \to 0} f(t,t) = f(0,0) = 0$, ce qui est impossible car $f(t,t) = \frac{1}{2}$ pour $t \neq 0$.

Remarque 4.10. *La bonne notion de continuité est celle donnée dans la Définition 4.1. La continuité partielle est une notion presque inutile, dont nous n' avons parlé que pour nous en débarrasser. La raison de son inutilité est son lien avec un système particulier de coordonnées.*

5. Dérivation des fonctions de plusieurs variables

On note par $\vec{e}_1, \dots, \vec{e}_n$, la base canonique de $\mathbb{R}^n$.

Définition 5.1. *Soit $D \subset \mathbb{R}^n$ un ouvert et $f : D \to \mathbb{R}$ une fonction. Soit $X_0 \in D$ et $\vec{u} \in \mathbb{R}^n$. On dit que f est* dérivable au point X_0 dans la direction $\vec{u}$ *si la fonction*

$$I \ni t \mapsto f(X_0 + t\vec{u}),$$

est dérivable en $t = 0$. Ici $I \subset \mathbb{R}$ est un intervalle ouvert contenant 0 tel que $X_0 + t\vec{u} \in D$ si $t \in I$, qui existe car D est ouvert.

Si f est dérivable en X_0 dans les directions $\vec{e}_1, \dots, \vec{e}_n$, on dit que f admet des dérivées partielles *en X_0 et on pose*

$$\frac{\partial f}{\partial x_i}(X_0) := \frac{d}{dt} f(X_0 + t\vec{e}_i)_{|t=0}.$$

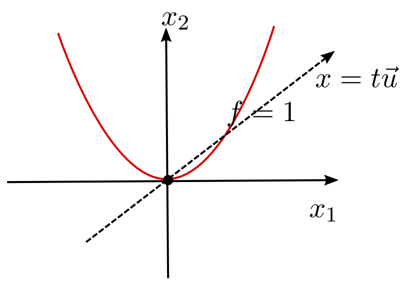
Notons que comme D est ouvert, la fonction $t \mapsto f(X_0 + t\vec{u})$, (qui correspond à restreindre f à un petit segment de la droite passant par X_0 et de vecteur directeur $\vec{u}$), est bien définie pour t dans un voisinage de 0.

De plus si $X_0 = (a_1, \dots, a_n)$, la dérivée partielle $\frac{\partial f}{\partial x_i}(X_0)$ est égal à la dérivée en a_i de la fonction partielle $f_i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n)$.

Remarque 5.2. *Une fonction f peut être dérivable dans toutes les directions en un point sans y être continue. En effet soit*

$$f(x_1, x_2) := \begin{cases} 1 \textit{ si } x_2 = x_1^2, \ (x_1, x_2) \neq (0,0), \\ 0 \textit{ sinon.} \end{cases}$$

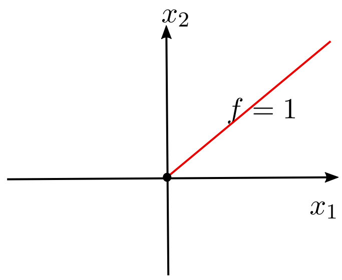
Cette fonction vaut 1 sur la parabole d' équation $x_2 = x_1^2$ privée de $(0,0)$ et 0 partout ailleurs. On remarque que $f(t\vec{u}) = 0$ pour $t \neq 0$ assez petit, et donc $\frac{d}{dt} f(t\vec{u})_{|t=0} = 0$, et f est dérivable dans toutes les directions en $(0,0)$. Par contre si $X_n = (n^{-1}, n^{-2})$ on a $\lim X_n = (0,0)$ mais $f(X_n) = 1$ donc $\lim f(X_n) \neq f(0,0)$ et donc f n'est pas continue en $(0,0)$.



Remarque 5.3. *Une fonction peut avoir des dérivées partielles en un point sans être dérivable dans toutes les directions. En effet soit*

$$f(x_1, x_2) := \begin{cases} 1 \ \textit{si} \ x_1 = x_2 > 0 \\ 0 \ \textit{sinon.} \end{cases}$$

Comme les fonctions partielles de f en $(0,0)$ sont identiquement nulles, on a $\frac{\partial}{\partial x_1} f(0,0) = \frac{\partial}{\partial x_2} f(0,0) = 0$. Par contre comme $f(t,t) = 1$ si $t > 0$ et $f(0,0) = 0$, f n'est pas dérivable dans la direction $(1,1)$.



Les deux exemples ci dessus montrent clairement que, comme la continuité partielle, les notions introduites dans la Définition 5.1 ne sont pas les bonnes.

Définition 5.4. *Soit $D \subset \mathbb{R}^n$ un ouvert et $f : D \to \mathbb{R}$ une fonction. f est* différentiable *au point $X_0 \in D$ s' il existe une application linéaire notée $Df(X_0) = \mathbb{R}^n \to \mathbb{R}$ telle que*

$$f(X_0 + X) = f(X_0) + Df(X_0)X + \|X\|\epsilon(X), \tag{5.1}$$

où $\epsilon(X)$ désigne une fonction définie sur une boule $B(0, \delta)$ telle que $\lim_{X \to 0} \epsilon(X) = 0$.

On sait que les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$ (appelées des *formes linéaires* sur $\mathbb{R}^n$) sont de la forme

$$TX = \vec{v} \cdot X, \text{ pour un vecteur } \vec{v} \in \mathbb{R}^n.$$

On a donc $Df(X_0)X := \vec{\nabla} f(X_0) \cdot X$, où le vecteur $\vec{\nabla} f(X_0)$ s' appelle le *gradient de f* au point X_0.

Lemme 5.5. *Soit $D \subset \mathbb{R}^n$ un ouvert et $f : D \to \mathbb{R}$ une fonction différentiable au point $X_0 \in D$. Alors f est continue en X_0. De plus f est dérivable dans toutes les directions au point X_0 et on a*

$$\vec{\nabla} f(X_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(X_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(X_0) \end{pmatrix} \tag{5.2}$$

Démonstration. L' identité (5.1) montre que $\lim_{X\to X_0} f(X) = f(X_0)$, donc f est continue en X_0. On applique ensuite (5.1) à $X = t\vec{u}$, on obtient $f(X_0+t\vec{u}) = f(X_0) + t\overrightarrow{\nabla} f(X_0)\cdot\vec{u} + t\epsilon(t)$, ce qui montre que $t \mapsto f(X_0+t\vec{u})$ est dérivable en $t = 0$ de dérivée $\overrightarrow{\nabla} f(X_0)\cdot\vec{u}$. En prenant $\vec{u} = \vec{e}_i$ on obtient (5.2). $\square$

Définition 5.6. *Soit $D \subset \mathbb{R}^n$ un ouvert et $f : D \to \mathbb{R}$ une fonction. On dit que f est* de classe *C^1 sur D et on écrit $f \in C^1(D)$ si f est différentiable en tout point X de D et si $D \ni X \mapsto \overrightarrow{\nabla} f(X) \in \mathbb{R}^n$ est continue.*

Théorème 5.7. *Soit $D \subset \mathbb{R}^n$ un ouvert. Une fonction $f : D \to \mathbb{R}$ est de classe C^1 sur D si et seulement si f admet des dérivées partielles $\frac{\partial f}{\partial x_i}(X)$, $1 \le i \le n$ en tout point X de D et les fonctions*

$$D \ni X \mapsto \frac{\partial f}{\partial x_i}(X) \in \mathbb{R}$$

sont continues.

Démonstration. On donne la démonstration pour $n = 2$, l' extension à n quelconque étant facile. On note $X_0 = (x_0, y_0) \in D$ et on applique le théorème des accroissements finis au point $y = y_0$ à la fonction d' une variable $y \mapsto f(x_0, y)$. On obtient

$$f(x_0,y) = f(x_0,y_0) + (y-y_0)\frac{\partial f}{\partial y}(x_0,\eta(y)), \tag{5.3}$$

où $\eta(y)$ est compris entre y_0 et y. Pour y *fixé* assez proche de y_0, on applique les accroissements finis à la fonction $x \mapsto f(x,y)$ au point $x = x_0$. On obtient

$$f(x,y) = f(x_0,y) + (x-x_0)\frac{\partial f}{\partial x}(\xi(x,y),y), \tag{5.4}$$

où $\xi(x,y)$ est compris entre x_0 et x. En utilisant aussi (5.3) on obtient :

$$f(x,y) = f(x_0,y_0) + (x-x_0)\frac{\partial f}{\partial x}(\xi(x,y),y) + (y-y_0)\frac{\partial f}{\partial y}(x_0,\eta(y)), \tag{5.5}$$

Comme $\eta(y)$ est compris entre y et y_0 on sait que $\lim_{y\to y_0}\epsilon(y) = y_0$ et comme $s \mapsto \frac{\partial f}{\partial y}(x_0,s)$ est continue, on a

$$\frac{\partial f}{\partial y}(x_0,\eta(y)) = \frac{\partial f}{\partial y}(x_0,y_0) + \epsilon_1(y),\ \lim_{y\to y_0}\epsilon_1(y) = 0.$$

De même comme $\xi(x,y)$ est compris entre x et x_0 on sait que $\lim_{x\to x_0}\xi(x,y) = x_0$ et comme $(x,y) \mapsto \frac{\partial f}{\partial x}(x,y)$ est continue (comme fonction de deux variables), on a :

$$\frac{\partial f}{\partial x}(\xi(x,y),y) = \frac{\partial f}{\partial x}(x_0,y_0) + \epsilon_2(x,y) \quad \lim_{(x,y)\to(x_0,y_0)}\epsilon_2(x,y) = 0.$$

On obtient donc

$$\begin{aligned} f(x,y) = \ & f(x_0,y_0) + (x-x_0)\tfrac{\partial f}{\partial x}(x_0,y_0) + (y-y_0)\tfrac{\partial f}{\partial y}(x_0,y_0) \\ & +(x-x_0)\epsilon_2(x,y) + (y-y_0)\epsilon_1(y). \end{aligned}$$

On pose ensuite

$$\epsilon(x,y) := \frac{(x-x_0)}{\|X-X_0\|}\epsilon_2(x,y) + \frac{(y-y_0)}{\|X-X_0\|}\epsilon_1(y),$$

et en utilisant que

$$\frac{|x-x_0|}{\|X-X_0\|} \leq 1, \ \frac{|y-y_0|}{\|X-X_0\|} \leq 1,$$

on obtient que $\lim_{X\to X_0} \epsilon(X) = 0$. Ceci termine la démonstration. □

5.1. **Extrema et points critiques.**

Définition 5.8. *Soit $f : D \to \mathbb{R}$ une fonction. On dit que $X_0 \in D$ est un*

(1) maximum *resp.* minimum *de f sur D si*

$$f(X) \leq f(X_0) \text{ resp. } f(X) \geq f(X_0) \ \forall X \in D,$$

(2) maximum *resp.* minimum local *de f sur D si il existe $\delta > 0$ tel que*

$$f(X) \leq f(X_0) \text{ resp. } f(X) \geq f(X_0) \ \forall X \in D \text{ avec } \|X - X_0\| < \delta.$$

Théorème 5.9. *Soit $D \subset \mathbb{R}^n$ un ouvert, $f \in C^1(D)$ et $X_0 \in D$ et $X_0 \in D$ un extremum local de f. Alors*

$$\overrightarrow{\nabla} f(X_0) = \overrightarrow{0},$$

et on dit que X_0 est un point critique *de f.*

Démonstration. Supposons que $\overrightarrow{\nabla} f(X_0) \neq 0$. On applique (5.1) à $X = t\overrightarrow{\nabla} f(X_0)$ pour $|t| < \delta$, avec δ assez petit et on obtient

$$f(X_0 + t\overrightarrow{\nabla} f(X_0)) - f(X_0) = t(\|\overrightarrow{\nabla} f(X_0)\|^2 + \epsilon(t)), \ \epsilon(t) \to 0 \text{ quand } t \to 0.$$

On a donc $\pm(f(X_0 + t\overrightarrow{\nabla} f(X_0)) - f(X_0)) > 0$ pour $0 < \pm t < \delta$, ce qui contredit l' hypothèse que X_0 est un extremum local. □

Remarque 5.10. *Comme pour les fonctions d' une variable, les extrema locaux sont des points critiques, mais la réciproque est fausse. Par exemple la fonction*

$$(x,y) \mapsto x^2 - y^2$$

possède le point critique $(0,0)$, qui n'est pas un extremum local.

5.2. **Dérivées partielles d' ordre supérieur.** Soit $D \subset \mathbb{R}^n$ un ouvert. On définit par récurrence l' ensemble $C^k(D)$ des fonctions de classe C^k sur D pour $k \geq 2$ en disant que $f \in C^k(D)$ si $f \in C^1(D)$ et les fonctions $X \mapsto \frac{\partial f}{\partial x_i}(X)$, $1 \leq i \leq n$ appartiennent à $C^{k-1}(D)$.

On utilise souvent la notation des *multiindices* pour les dérivées d' ordre supérieur, en posant pour $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$

$$\partial_x^\alpha f = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f.$$

Un résultat fondamental sur les dérivées partielles d' ordre supérieur est le suivant.

Théorème 5.11. *[Lemme de Schwarz] Soit $D \subset \mathbb{R}^n$ un ouvert et $f \in C^2(D)$. Alors*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(X) = \frac{\partial^2 f}{\partial x_j \partial x_i}(X),\ \forall 1 \leq i,j \leq n,\ X \in D.$$

La démonstration est donnée dans l' appendice A.2.

5.2.1. Il existe des fonctions ayant des dérivées partielles d' ordre 2 mais telles que $\frac{\partial^2 f}{\partial x_i \partial x_j} \neq \frac{\partial^2 f}{\partial x_j \partial x_i}$. En effet soit

$$f(x_1,x_2) = \begin{cases} x_1 x_2 \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2},\ \text{si } (x_1,x_2) \neq (0,0), \\ 0 \text{ si } (x_1,x_2) = (0,0). \end{cases}$$

Calculons $\frac{\partial f}{\partial x_2}(x_1,x_2)$. Si $x_1 = 0$ on a $f(0,x_2) = 0$ donc $\frac{\partial f}{\partial x_2}(0,x_2) = 0$. Si $x_1 \neq 0$ un calcul routinier donne que

$$\frac{\partial f}{\partial x_2}(x_1,x_2) = x_1 \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2},$$

et donc en particulier

$$\frac{\partial f}{\partial x_2}(x_1,0) = x_1, \forall x_1 \in \mathbb{R},$$

ce qui implique que

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = 1.$$

Comme $f(x_1,x_2) = -f(x_2,x_1)$, on obtient automatiquement que

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(0,0) = -1.$$

5.3. **Formule de Taylor à l' ordre 2.**

5.3.1. *Matrice hessienne.*

Définition 5.12. *Soit $D \subset \mathbb{R}^n$ un ouvert, $f \in C^2(D)$ et $X_0 \in D$. La* matrice hessienne *de f au point X_0 est la matrice $n \times n$ donnée par*

$$Hf(X_0) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X_0)\right]_{1 \leq i,j \leq n}.$$

A cause du lemme de Schwarz $Hf(X)$ est une matrice symétrique.

5.3.2. *Formule de Taylor à l' ordre* 2.

Théorème 5.13. *Soit $D \subset \mathbb{R}^n$ un ouvert , $f \in C^2(D)$ et $X_0 \in D$. Alors*

$$f(X_0 + X) = f(X_0) + \overrightarrow{\nabla} f(X_0) \cdot X + \frac{1}{2} X \cdot Hf(X_0) X + \|X\|^2 \epsilon(X), \tag{5.6}$$

où $\epsilon(X)$ désigne une fonction définie sur une boule $B(0,\delta)$ telle que $\lim_{X \to 0} \epsilon(X) = 0$.

Donnons une version explicite de (5.6) si $n=2$, à l' aide des coordonnées cartésiennes :

$$
\begin{array}{lll}
& & f(a_1+x_1,a_2+x_2)=f(a_1,a_2)+\frac{\partial f}{\partial x_1}(a_1,a_2)x_1+\frac{\partial f}{\partial x_2}(a_1,a_2)x_2 \\
(5.7) & + & \frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1^2}(a_1,a_2)x_1^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}(a_1,a_2)x_1x_2+\frac{\partial^2 f}{\partial x_2^2}(a_1,a_2)x_2^2\right) \\
& + & (x_1^2+x_2^2)\epsilon(x_1,x_2).
\end{array}
$$

Il est clair que l' écriture (5.6) est bien préférable à (5.7).

5.4. **Nature des points critiques.**

Théorème 5.14. *Soit $D\subset\mathbb{R}^n$ un ouvert , $f\in C^2(D)$ et $X_0\in D$ un point critique de f.*

(1) *si toutes les valeurs propres de $Hf(X_0)$ sont* strictement positives *resp.* strictement négatives, *alors X_0 est un* minimum local, *resp. un* maximum local.
(2) *si toutes les valeurs propres de $Hf(X_0)$ sont* non nulles, mais pas toutes de même signe, *X_0 n'est* pas un extremum local. *On dit que X_0 est un* point selle *ou encore un* point col.
(3) *si $Hf(X_0)$ a la valeur propre* 0, *X_0 est un* point critique dégénéré *et on ne peut rien dire sur la nature de X_0.*

La nouveauté par rapport au cas $n=1$ est l' apparition des points cols.
Démonstration. D' après (5.6) on a

$$
f(X_0+X)-f(X_0)=\frac{1}{2}X\cdot Hf(X_0)X+\|X\|^2\epsilon(X). \tag{5.8}
$$

Commençons par étudier la fonction $X\mapsto X\cdot Hf(X_0)X$. La matrice $Hf(X_0)$ est symétrique, donc possède une base orthonormée de vecteurs propres $(\vec{u}_1,\dots,\vec{u}_n)$, associés aux valeurs propres $\lambda_1,\dots,\lambda_n$, c'est à dire

$$
Hf(X_0)\vec{u}_i=\lambda_i\vec{u}_i,\ \vec{u}_i\cdot\vec{u}_j=\delta_{ij}, \tag{5.9}
$$

où $\delta_{ij}=1$ si $i=j$ et $\delta_{ij}=0$ sinon est le delta de Kronecker.

Soit $(x_1,\dots,x_n)$ les coordonnées de X dans la base $(\vec{u}_1,\dots,\vec{u}_n)$, c'est à dire $X=\sum_{i=1}^n x_i\vec{u}_i$. Grâce à (5.9) on a

$$
X\cdot Hf(X_0)X=\sum_{i=1}^n\lambda_i x_i^2.
$$

Si $\lambda_i>0$ pour tout i, on a

$$
X\cdot Hf(X_0)X\geq c\|X\|^2,\ c=\min\lambda_i,
$$

et donc en utilisant (5.8)

$$
f(X_0+X)-f(X_0)\geq\|X\|^2(c+\epsilon(X)),
$$

et donc $f(X_0+X)>f(X_0)$ pour $0<\|X\|\ll 1$, X_0 est un minimum local. Le même argument montre que si $\lambda_i<0$ pour tout i X_0 est un maximum local.

Supposons maintenant que $Hf(X_0)$ a une valeur propre $\lambda_1>0$ et une autre $\lambda_2<0$. En utilisant (5.8) on obtient

$$
f(X_0+t\vec{u}_1)-f(X_0)=\frac{1}{2}\lambda_1t^2+t^2\epsilon(t),
$$

donc $f(X_0 + t\vec{u}_1) > f(X_0)$ pour t assez petit, et de même :

$$f(X_0 + t\vec{u}_2) - f(X_0) = \frac{1}{2}\lambda_2 t^2 + t^2\epsilon(t),$$

donc $f(X_0+t\vec{u}_2) < f(X_0)$ pour t assez petit. X_0 n'est donc pas un extremum local. □

5.4.1. *Le cas de la dimension 2.* Pour étudier la nature d' un point critique, il est donc a priori nécessaire de calculer les valeurs propres de la matrice hessienne. Ce n'est pas nécessaire si $n = 2$. En effet soit $M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ une matrice symétrique. Son polynôme caractéristique est

$$P(\lambda) = \lambda^2 - 2(a + c)\lambda + ac - b^2,$$

et on reconnait que $ac - b^2 = \det A$, $a + c = \mathrm{Tr}A$, la *trace de A*, c'est à dire la somme des termes sur la diagonale.

D' autre part si un polynôme de degré 2 s' écrit comme $\lambda^2 - S\lambda + P$, alors S est la somme de ses racines et P leur produit, c'est à dire que $\det A$ est le produit des deux valeurs propres de A et $\mathrm{Tr}A$ leur somme. On obtient donc le résultat suivant, pour $A = Hf(X_0)$:

(1) si $\det A > 0$ et $\mathrm{Tr}A > 0$ X_0 est un minimum local.
(2) si $\det A > 0$ et $\mathrm{Tr}A < 0$ X_0 est un maximum local.
(3) si $\det A < 0$ X_0 est un point selle.
(4) si $\det A = 0$, X_0 est un point critique dégénéré.

6. Espaces vectoriel normés

6.1. **Rappels sur les espaces vectoriels.** Dans cette section E désigne un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, dont les éléments seront notés x, y etc. En cas d' attaque de panique, prendre $\mathbb{K} = \mathbb{R}$.

Le vecteur nul dans E sera noté 0_E, ou parfois simplement 0. Les scalaires seront notés λ, μ etc. Les applications linéaires entre deux espaces vectoriels seront notées $A : E \to F$, leur action sur $x \in E$ par Ax.

Nous aurons besoin de parler de la *dimension* d' un espace vectoriel. On rappelle que E est de dimension $n \in \mathbb{N}$ si E admet une base formée de n vecteurs. Toutes les autres bases de E ont alors aussi n éléments.

Si E n' admet pas de base formée d' un nombre fini de vecteurs, on dit que E est de *dimension infinie.* C'est le cas de la plupart des espaces vectoriels de fonctions qu' on considère en analyse. Pour montrer qu' un espace E est de dimension infinie, il suffit de trouver des familles libres avec un nombre arbitraire de vecteurs.

Rappelons encore que la dimension d' un espace vectoriel dépend du choix du corps des scalaires. Par exemple $\mathbb{C}$ est un espace vectoriel de dimension 1 sur $\mathbb{C}$ (une base étant celle formée du seul vecteur 1) mais de dimension 2 sur $\mathbb{R}$ (une base étant celle formée des deux vecteurs 1 et i).

6.2. **Norme sur un espace vectoriel.**

Définition 6.1. *Soit E un espace vectoriel sur $\mathbb{K}$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Une* norme *sur E est une application $N : E \to \mathbb{R}^+$ telle que pour tous $x, y \in E$ et $\lambda \in \mathbb{K}$ on a :*

(1) $N(\lambda x) = |\lambda| N(x)$.
(2) $N(x+y) \leq N(x) + N(y)$ *(inégalité triangulaire).*
(3) $N(x) = 0 \Rightarrow x = 0_E$.
Un espace vectoriel E muni d' une norme N se note (E, N) et s' appelle un espace vectoriel normé.

— il est habituel de noter une norme par $\|x\|$, plutôt que par $N(x)$. En rencontre parfois dans la littérature le symbole $|||\ A\ |||$ quand A est une application linéaire ou une matrice, qui est à éviter.

— une fonction $N : E \to \mathbb{R}^+$ vérifiant seulement les propriétés (1) et (2) est appelée une *semi-norme.*

— comme pour la fonction valeur absolue, on déduit facilement de (2) que

$$|N(x) - N(y)| \leq N(x-y), \ x, y \in E. \tag{6.1}$$

6.2.1. *Norme induite.* Si $(E, \|\cdot\|)$ est un evn et $F \subset E$ est un *sous-espace vectoriel*, la restriction de $\|\cdot\|$ à F fait de $(F, \|\cdot\|)$ un espace vectoriel normé.

6.2.2. *Exemples.* — Soit $E = \mathbb{K}^n$ dont les éléments sont notés $x = (x_1, \ldots, x_n)$, $x_i \in \mathbb{K}$. E est un espace vectoriel de dimension n sur $\mathbb{K}$. On pose

$$\begin{aligned} &\|x\|_1 = \textstyle\sum_{i=1}^n |x_i|, \\ &\|x\|_2 = (\textstyle\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}, \\ &\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \end{aligned}$$

et plus généralement

$$\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}, 1 < p < \infty.$$

Les fonctions $\|\cdot\|_p$ pour $1 \leq p \leq \infty$ sont des normes sur $\mathbb{K}^n$. C'est facile pour $p = 1, \infty$, déjà vu pour $p = 2$. Pour p quelconque, l 'inégalité triangulaire s' appelle l' *inégalité de Minkowski.*

– Soit X un ensemble et soit $B(X, \mathbb{K})$ l' ensemble des fonctions $f : X \to \mathbb{K}$ *bornées*, c'est à dire telles que $\|f\|_\infty := \sup_{x \in X} |f(x)| < \infty$. $(B(X, \mathbb{K}), \|\cdot\|_\infty)$ est un espace vectoriel normé.

$B(X, \mathbb{K})$ est de dimension finie n si et seulement si X possède exactement n éléments (exercice).

– Soit $R([a, b]; \mathbb{K})$ l' ensemble des fonctions à valeurs dans $\mathbb{K}$ intégrables au sens de Riemann sur $[a, b]$. $R([a, b]; \mathbb{K})$ est de dimension infinie. En effet si on note par $\mathbb{1}_{\{c\}}(x)$ la fonction qui vaut 1 si $x = a$ et 0 sinon, alors pour toute famille finie $\{c_1, \ldots, c_N\}$ avec $c_i \in [a, b]$ les vecteurs $\mathbb{1}_{\{c_1\}}, \ldots, \mathbb{1}_{\{c_N\}}$ forment une famille libre de $R([a, b]; \mathbb{K})$ (exercice).

La fonction

$$\|f\|_p = \left(\int_a^b |f(x|^p dx) \right)^{1/p}, \ 1 \leq p < \infty$$

est une semi-norme sur $R([a, b]; \mathbb{K})$. A nouveau l 'inégalité triangulaire pour $\|\cdot\|_p$ s' appelle l' inégalité de Minkowski (pour les intégrales). Ce n'est pas une norme car la fonction $f(x) = \mathbb{1}_{\{c\}}(x)$ pour $c \in [a, b]$ est non nulle mais $\|f\|_p = 0$.

– Considérons maintenant l 'espace $C([a,b];\mathbb{K})$ des fonctions *continues* sur $[a,b]$. $C([a,b],\mathbb{K})$ est de dimension infinie car les familles $(1,x,x^2,\ldots,x^n)$ sont libres pour tout $n\in\mathbb{N}$ (exercice).

$\|\cdot\|_p$ est une norme sur $C([a,b];\mathbb{K})$. En effet si $\int_a^b|u(x)|^p dx = 0$ et u continue, alors $u(x)=0$ sur $[a,b]$.

— Soit $\mathbb{K}^{\mathbb{N}} = \{a=(a_n)_{n\in\mathbb{N}}\}$ l' espace vectoriel des suites à valeurs dans $\mathbb{K}$.

Pour $1\leq p<\infty$ on pose

$$l^p(\mathbb{N};\mathbb{K}) = \{(a_n) : \sum_{n\in\mathbb{N}}|a_n|^p \text{ converge}\}.$$

On vérifie que $l^p(\mathbb{N};\mathbb{K})$ est un espace vectoriel et que

$$\|a\|_p = \left(\sum_{n=0}^{\infty}|a_n|^p\right)^{1/p}$$

est une norme sur $l^p(\mathbb{N};\mathbb{K})$.

De même

$$l^\infty(\mathbb{N};\mathbb{K}) = \{(a_n) : \sup_{n\geq 0}|a_n|<\infty\},$$

est un espace vectoriel et

$$\|a\|_\infty = \sup_{n\geq 0}|a_n|$$

est une norme sur $l^\infty(\mathbb{N};\mathbb{K})$. $\mathbb{K}^{\mathbb{N}}$ ainsi que $l^p(\mathbb{N};\mathbb{K})$ sont de dimension infinie. En effet en notant par δ_p la suite définie par $\delta_{p,n}=1$ si $n=p$ et 0 sinon, $(\delta_1,\ldots,\delta_N)$ est une famille libre pour tout $N\in\mathbb{N}$.

6.3. **Topologie des espaces vectoriels normés.**

6.3.1. *Distance induite par une norme.*

Définition 6.2. *Soit* $(E,\|\cdot\|)$ *un evn. La fonction* $E\times E\ni(x,y)\mapsto d(x,y)=\|x-y\|$ *est une distance sur* E*, appelée* distance induite *par la norme* $\|\cdot\|$*.*

On peut alors définir les boules de centre $a\in E$ de rayon r comme

$$B(a,r)=\{x\in E:\|x-a\|<r\},\ B_{\mathrm{f}}(a,r)=\{x\in E:\|x-a\|\leq r\},$$

les ensembles ouverts, fermés, compacts etc de E.

Définition 6.3. *Un espace vectoriel normé complet s' appelle un* espace de Banach.

Les espaces vectoriels normés de dimension *finie* sont complets (quelque soit la norme dont ils sont équipés).

Par contre un espace vectoriel normé de dimension *infinie* n'est pas nécessairement complet.

Les espaces $(B(X;\mathbb{K}),\|\cdot\|_\infty)$, $(C([a,b];\mathbb{K}),\|\cdot\|_\infty)$, $(l^p(\mathbb{N};\mathbb{K}),\|\cdot\|_p)$ pour $1\leq p\leq\infty$ sont complets.

L' espace $C([a,b];\|\cdot\|_p)$ pour $1\leq p<\infty$ n'est pas complet, voir 6.3.2.

Dans un espace de Banach les ensembles fermés et bornés ne sont pas toujours compacts, voir la Proposition 6.5. En fait on peut montrer que la boule unité fermée $B_{\mathrm{f}}(x_0,r)$ est compacte dans $(E,\|\cdot\|)$ si et seulement si E est de dimension finie.

6.3.2. *Exemples.* – Donnons un exemple d' espace vectoriel normé de dimension infinie qui n'est pas complet.

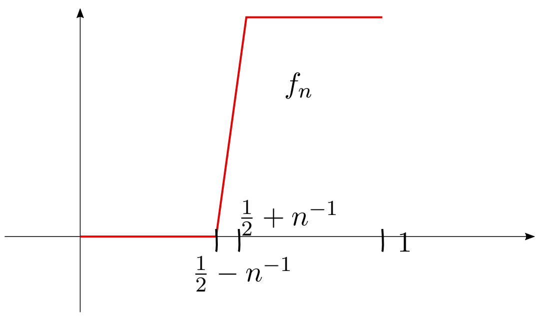
Proposition 6.4. *L' espace $E = C([0,1];\mathbb{R})$ muni de la norme $\|f\|_1 = \int_0^1 |f(x)|dx$ n'est pas complet.*

Démonstration. soit

$$f_n(x) = \begin{cases} 0 \text{ si } 0 \leq x \leq \frac{1}{2} - n^{-1}, \\ n/2(x - \frac{1}{2} - n^{-1}) \text{ si } \frac{1}{2} - n^{-1} \leq x \leq \frac{1}{2} + n^{-1}, \\ 1 \text{ si } \frac{1}{2} + n^{-1} \leq x \leq 1, \end{cases}$$

et

$$f(x) = \begin{cases} 0 \text{ si } 0 \leq x \leq \frac{1}{2}, \\ 1 \text{ si } \frac{1}{2} < x \leq 1. \end{cases}$$



On vérifie directement que $\|f - f_n\|_1 \to 0$ et donc $\forall \epsilon > 0\ \exists N \in \mathbb{N}$ tel que $\|f_n - f\|_1 \leq \epsilon$ si $n \geq N$. Si $n, n \geq N$ on a donc $\|f_n - f_p\|_1 \leq \|f_n - f\|_1 + \|f_p - f\|_1 \leq 2\epsilon$, donc (f_n) est de Cauchy dans $(C([0,1];\mathbb{R}), \|\cdot\|_1)$. Supposons qu' il existe $g \in C([0,1];\mathbb{R})$ tel que $\|f_n - g\|_1 \to 0$. On a donc $\|f - g\|_1 = 0$, donc

$$\int_0^{\frac{1}{2}} |f - g|(x)dx = \int_{\frac{1}{2}}^1 |f - g|(x)dx = 0.$$

Comme f et g sont continues sur $[0, \frac{1}{2}[$ et sur $]\frac{1}{2}, 1]$, on a $f = g$ sur $[0, \frac{1}{2}[$ et sur $]\frac{1}{2}, 1]$, ce qui contredit la continuité de g au point $\frac{1}{2}$. La suite de Cauchy (f_n) ne converge donc pas dans $(C([0,1];\mathbb{R}), \|\cdot\|_1)$. □

– Voici maintenant un exemple d' espace de Banach de dimension infinie dans lequel les ensembles fermés bornés ne sont pas nécessairement compacts.

Proposition 6.5. *Dans l' espace $E = (l^1(\mathbb{N};\mathbb{R}), \|\cdot\|_1)$ la boule unité fermée $B_{\rm f}(0,1)$ n'est pas compacte.*

Démonstration. pour éviter des complications de notation, on note un élément de E (c' est à dire une suite réelle) par u, et le n-ième terme de u par $u(n)$. On note que pour $n \in \mathbb{N}$ fixé on a

$$|u(n)| \leq \|u\|_1, \ u \in E. \tag{6.2}$$

Soit $(u_p)_{p\in\mathbb{N}}$ la suite d' éléments de E définie par $u_p(n) = \delta_{np}$. On a $\|u_p\|_1 = 1$, donc $(u_p)_{p\in\mathbb{N}}$ est une suite dans $B_{\mathrm{f}}(0,1)$ qui est un ensemble fermé et borné dans E. Supposons qu' il existe une sous-suite $(u_{\varphi(p)})$ qui converge vers $u \in E$. Par l' inégalité triangulaire (6.1) on a

$$|\|u\|_1 - \|u_{\varphi(p)}\|_1| \leq \|u - u_{\varphi(p)}\|_1$$

donc en faisant $p \to \infty$ on a $\|u\|_1 = 1$. D'après (6.2) appliqué à $u_{\varphi(p)} - u$ on a

$$|u_{\varphi(p)}(n) - u(n)| \leq \|u_{\varphi(p)} - u\|_1,$$

donc pour n *fixé* $\lim_{p\to\infty} u_{\varphi(p)}(n) - u(n) = 0$. Comme $u_{\varphi(p)}(n) = 0$ dès que $p > n$ on en déduit que $u(n) = 0$. Comme n est arbitraire, u est la suite nulle ce qui contredit le fait que $\|u\|_1 = 1$. La boule unité fermée dans E est donc fermée et bornée mais non compacte. □

6.4. **Normes équivalentes.** Un point important à garder à l' esprit est que la notion d' ensembles ouverts, d' applications continues, d' ensembles compacts etc dans un espace vectoriel E *dépend* a priori du choix d' une norme sur E. Deux normes différentes sur E conduisent en général à deux topologies différentes.

Définition 6.6. *Soit N_1, N_2 deux normes sur un espace vectoriel E.*

(1) *on dit que N_1 et N_2 sont* topologiquement équivalentes *si (E, N_1) et (E, N_2) ont les mêmes ensembles ouverts.*

(2) *on dit que N_1 et N_2 sont* equivalentes *(on écrit parfois $N_1 \sim N_2$) si il existe $C_1, C_2 > 0$ telles que*

$$N_1(X) \leq C_1 N_2(X),\ N_2(X) \leq C_2 N_1(X),\ \forall X \in E.$$

On peut réécrire (2) de manière plus élégante comme : il existe $C > 0$ tel que

$$C^{-1} N_2(X) \leq N_1(X) \leq C N_2(X),\ \forall X \in E,$$

(prendre $C = \max(C_1, C_2)$. La taille des constantes C_1, C_2 ou C n'a que peu d' importance.

Il est facile de voir que si $N_1 \sim N_2$ et $N_2 \sim N_3$ alors $N_1 \sim N_3$ (la relation $\sim$, comme la relation d' équivalence topologique sont des relations d' équivalence).

Théorème 6.7. *Deux normes sur E sont topologiquement équivalentes si et seulement si elles sont équivalentes.*

La démonstration sera faite dans 6.6.3.

Remarque 6.8. *Il est facile de voir que les normes $\|\cdot\|_p$ sur $\mathbb{K}^n$ sont équivalentes entre elles. En effet on vérifie (exercice) que*

$$\|x\|_p \leq n^{1/p} \|x\|_\infty,\ \|x\|_\infty \leq \|x\|_p,\ x \in \mathbb{K}^n.$$

On verra dans la suite un résultat a priori surprenant qui affirme que toutes *les normes sur $\mathbb{K}^n$ sont équivalentes.*

Par contre les normes $\|\cdot\|_p$ sur $C([a,b]; \mathbb{K})$ ne sont pas *équivalentes entre elles. Par exemple si $[a,b] = [-1,1]$ et*

$$f_n(x) = \begin{cases} 1 - n|x| \text{ si } |x| \leq n^{-1}, \\ 0 \text{ sinon} \end{cases}$$

(tracer son graphe), on a $\|f\|_\infty = 1$ *et* $\|f\|_1 = n^{-1}$*. Ceci entraine qu' il n' existe pas de constante* C *telle que* $\|f\|_\infty \leq C\|f\|_1$ *pour tout* $f \in C([-1,1];\mathbb{K})$ *(prendre* $f = f_n$ *et faire* $n \to \infty$*).*

6.4.1. *Equivalence des normes en dimension finie.*

Théorème 6.9. *Soit* E *un espace vectoriel sur* $\mathbb{K}$ *de dimension finie. Alors* E *possède au moins une norme et toutes les normes sur* E *sont équivalentes.*

Démonstration. On peut supposer que $\mathbb{K} = \mathbb{R}$ (en identifiant $\mathbb{C}$ à $\mathbb{R}^2$). On fixe alors une base $\mathcal{B} = (e_1, \dots, e_n)$ de E et on identifie E à $\mathbb{R}^n$ à l' aide de $\mathcal{B}$. On peut donc supposer que $E = \mathbb{R}^n$ et que $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$. La norme $\|\cdot\|_\infty$ est une norme sur $\mathbb{R}^n$, équivalente à la norme usuelle, et définissant donc les mêmes ouverts.

Soit $S_\infty(0,1) = \{x \in \mathbb{R}^n : \|x\|_\infty = 1\}$ la sphère de rayon 1 pour $\|\cdot\|_\infty$. $S_\infty(0,1)$ est fermée, car l' application $E \ni x \mapsto \|x\| \in \mathbb{R}$ est continue sur tout evn $(E, \|\cdot\|)$ (exercice). $S_\infty(0,1)$ est évidemment bornée, donc compacte par le Théorème de Borel-Lebesgue.

Soit maintenant $N : \mathbb{R}^n \to \mathbb{R}^+$ une norme. On a

$$N(x) = N(\sum_{i=1}^n x_i e_i) \leq \sum_{i=1}^n |x_i| N(e_i) \leq C\|x\|_\infty, \ C = \sum_{i=1}^n N(e_i).$$

Grâce à (6.1), on obtient

$$|N(x) - N(y)| \leq N(x-y) \leq C\|x-y\|_\infty,$$

et donc la fonction $N : \mathbb{R}^n \to \mathbb{R}$ est continue sur $(\mathbb{R}^n, \|\cdot\|_\infty)$. Par le Corollaire (4.8) N est bornée et atteint ses bornes sur $S_\infty(0,1)$. Comme $N(x) > 0$ pour tout $x \in S_\infty(0,1)$ il existe donc des constantes $0 < a < b$ telles que

$$a \leq N(x) \leq b \ \forall x \in S_\infty(0,1).$$

En appliquant cette inégalité à $\|x\|_\infty^{-1} x$ on en déduit que

$$a\|x\|_\infty \leq \|x\|_\infty N(\|x\|_\infty^{-1} x) = N(x) \leq b\|x\|_\infty, \ \forall x \neq 0,$$

et donc N et $\|\cdot\|_\infty$ sont équivalentes. Toutes les normes sur $\mathbb{R}^n$ sont équivalentes à $\|\cdot\|_\infty$ et donc équivalentes entre elles. □

6.5. **Compléments sur les espaces vectoriels normés.**

6.5.1. *Norme* L^∞ *et convergence uniforme.* Soit X un ensemble (par exemple un intervalle de $\mathbb{R}$) et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions $f_n : X \to \mathbb{R}$.

On rappelle que la suite de fonctions (f_n) *converge simplement* vers une fonction $f : X \to \mathbb{R}$ si pour tout $x \in X$ la suite réelle $(f_n(x))$ converge vers $f(x)$. La convergence simple est la notion la plus faible de convergence pour des suites de fonctions. Exprimée à l' aide de quantificateurs, elle s' écrit :

$$\forall x \in X, \forall \epsilon > 0, \ \exists N \in \mathbb{N} \text{ tel que } |f_n(x) - f(x)| \leq \epsilon, \ \forall n \geq N, \tag{6.3}$$

l' entier N dépendant a priori de ϵ *et* de x.

La suite de fonctions (f_n) *converge uniformément* vers une fonction $f : X \to \mathbb{R}$ si

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N} \text{ tel que } |f_n(x) - f(x)| \leq \epsilon, \ \forall n \geq N, \forall x \in X. \tag{6.4}$$

L' entier N ne dépend maintenant que de ϵ. On peut réécrire (6.4) comme

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N} \text{ tel que } \sup_{x \in X} |f_n - f| < \epsilon. \tag{6.5}$$

On voit donc que sur l' espace vectoriel $B(X, \mathbb{R})$ des fonctions *bornées* sur X, une suite (f_n) converge uniformément vers f si et seulement si (f_n) converge vers f pour la norme $\| \cdot \|_\infty$.

6.5.2. *Limites uniformes de fonctions continues.* Prenons $X = [a, b]$. L' espace $C([a, b]; \mathbb{R})$ des fonctions continues sur $[a, b]$ est un sous espace vectoriel de l' espace $B([a, b]; \mathbb{R})$ des fonctions bornées. On sait que la limite uniforme d' une suite de fonctions continues sur $[a, b]$ est une fonction continue sur $[a, b]$.

Avec le language que nous avons appris, ceci signifie que $C([a, b]; \mathbb{R})$ est un sous-espace vectoriel *fermé* de $B([a, b]; \mathbb{R})$ muni de la norme $\| \cdot \|_\infty$.

6.5.3. *Séries à valeurs dans un espace vectoriel normé.* Soit $(E, \| \cdot \|)$ un espace vectoriel normé et $(u_n)_{n \in \mathbb{N}}$ une suite dans E. On peut former la suite

$$S_N = \sum_{n=0}^{N} u_n,$$

appelée suite des *sommes partielles de la série* $\sum_{n \in \mathbb{N}} u_n$. La définition suivante est exactement la même que pour les séries numériques.

Définition 6.10. *La série $\sum_{n \in \mathbb{N}} u_n$ converge dans $(E, \| \cdot \|)$ si la suite des sommes partielles $(S_N)_{N \in \mathbb{N}}$ converge dans $(E, \|\cdot\|)$. L' élément* $\lim_{N \to +\infty} \sum_{n=1}^{N} u_n$ *est noté $\sum_{n=0}^{\infty} u_n$ et appelé* somme de la série $\sum_{n \in \mathbb{N}} u_n$.

Les propriétés suivantes sont immédiates.

Proposition 6.11. *Soit $\sum_{n \in \mathbb{N}} u_n$, $\sum_{n \in \mathbb{N}} v_n$ deux séries convergentes à valeurs dans E et $\lambda \in \mathbb{K}$. Alors $\sum_{n \in \mathbb{N}} u_n + v_n$ et $\sum_{n \in \mathbb{N}} \lambda u_n$ sont convergentes et on a :*

$$\sum_{n=0}^{\infty} u_n + v_n = \sum_{n=0}^{\infty} u_n + \sum_{n=0}^{\infty} v_n,\ \sum_{n=0}^{\infty} \lambda u_n = \lambda \sum_{n=0}^{\infty} u_n.$$

6.5.4. *Convergence normale.*

Définition 6.12. *Soit $(E, \| \cdot \|)$ un espace vectoriel normé et $(u_n)_{n \in \mathbb{N}}$ une suite dans E. La série $\sum_{n \in \mathbb{N}} u_n$* converge normalement *si la série numérique $\sum_{n \in \mathbb{N}} \|u_n\|$ converge.*

La convergence normale est un outil important pour montrer la convergence d' une série à valeurs dans un espace de Banach. On peut retenir sa définition en pensant à la 'convergence de la série des normes'.

Théorème 6.13. *Soit $(E, \| \cdot \|)$ un espace de Banach. Alors toute série $\sum_{n \in \mathbb{N}} u_n$ normalement convergente est convergente et on a*

$$\| \sum_{n=0}^{\infty} u_n \| \leq \sum_{n=0}^{\infty} \|u_n\|.$$

Si $E = \mathbb{R}$ ou $\mathbb{C}$, il s'agit de séries numériques, et la convergence normale correspond à la convergence absolue. Il existe donc des séries convergentes mais non normalement convergentes, comme la série $\sum_{n\geq 1}(-1)^n n^{-1}$.

Démonstration. Soit $S_N = \sum_{n=0}^{N} u_n$. Si $M > N$ on a

$$S_M - S_N = \sum_{n=N+1}^{M} u_n,$$

et donc par l 'inégalité triangulaire :

$$\|S_M - S_N\| \leq \sum_{n=N+1}^{M} \|u_n\| = T_M - T_N, \tag{6.6}$$

pour $T_N = \sum_{n=0}^{N} \|u_n\|$. Comme la série $\sum_{n\in\mathbb{N}} \|u_n\|$ est convergente, la suite $(T_N)_{N\in\mathbb{N}}$ converge donc est de Cauchy. L ' inégalité (6.6) entraine que la suite $(S_N)_{N\in\mathbb{N}}$ est de Cauchy, donc convergente comme E est complet. La série $\sum_{n\in\mathbb{N}} u_n$ est donc convergente.

D' autre part comme l' application $u \mapsto \|u\|$ est continue on a

$$\|\sum_{n=0}^{\infty} u_n\| = \|\lim_{N\to\infty} S_N\| = \lim_{N\to\infty} \|S_N\| \leq \lim_{N\to\infty} T_N = \sum_{n=0}^{\infty} \|u_n\|. \ \square$$

6.6. **Applications linéaires continues.** Soit E, F deux espaces vectoriels normés, dont on note les normes respectives par $\|\cdot\|_E$ et $\|\cdot\|_F$. On note par $L(E, F)$ l' espace des applications linéaires de E dans F et on note $L(E, E)$ simplement par $L(E)$ On rappelle que $L(E, F)$ est un espace vectoriel si on pose

$$\begin{aligned}&(\lambda A)x := \lambda Ax,\\ &(A + B)x = Ax + Bx,\ x \in E,\end{aligned}$$

et on note encore par $0 \in L(E, F)$ l' *application nulle* telle que $0x = 0_F$ pour tout $x \in E$.

L' espace $L(E)$ a une structure supplémentaire, comme on peut composer entre elles des applications de $L(E)$. On note traditionellement par AB l' application $A \circ B : x \mapsto A(B(x))$. On a les propriétés suivantes :

$$\begin{aligned}&(\lambda A)B = A(\lambda B) = \lambda AB,\\ &A(B + C) = AB + AC,\ (B + C)A = BA + CA,\\ &0A = A0 = 0,\end{aligned}$$

mais évidemment $AB \neq BA$ en général (comme pour le produit de matrices).

Théorème 6.14. *Soit $A \in L(E, F)$. Les propriétés suivantes sont equivalentes :*

(1) *$A : E \to F$ est continue.*

(2) *$A : E \to F$ est continue au point 0_E.*

(3) *il existe $C \geq 0$ tel que $\|AX\|_F \leq C\|X\|_E$ pour tout $X \in E$.*

Démonstration. (1) $\Rightarrow$ (2) est évident.

(2) $\Rightarrow$ (3) : supposons A continue au point 0_E. En prenant $\epsilon = 1$ dans la définition de la continuité, il existe donc $\delta > 0$ tel que $\|x\|_E \leq \delta \ \Rightarrow \|Ax\|_F \leq 1$. Pour $x \neq 0_E \in E$ on a $\|\alpha\|x\|_E^{-1}x\|_E = \alpha$ et donc

$$\|A\alpha\|x\|_E^{-1}x\|_F = \alpha\|x\|_E^{-1}\|Ax\|_F \leq 1,$$

donc

$$\|Ax\|_F \leq \alpha^{-1}\|x\|_E, \ \forall x \in E.$$

(3) $\Rightarrow$ (1) : montrons que A est continue en $x_0 \in E$. On a par (3)

$$\|Ax - Ax_0\|_F = \|A(x - x_0)\|_F \leq C\|x - x_0\|_E.$$

Pour $\epsilon > 0$, si $\|x - x_0\|_E \leq \delta = C^{-1}\epsilon$, on a donc $\|Ax - Ax_0\|_F \leq \epsilon$. A est donc continue au point x_0. $\square$

Définition 6.15. *L' espace des applications linéaires continues de E dans F se note $B(E, F)$. Si $F = E$ $B(E, F)$ se note simplement $B(E)$. Une application linéaire continue entre deux espaces vectoriels normés est dite* bornée.

L' adjectif borné vient du fait que la condition (3) dans le Théorème 6.14 est équivalente au fait que A soit *bornée* sur la boule unité $B_{\rm f}(0, 1)$ de E.

Lemme 6.16. *Soit E, F deux espaces vectoriels normés et E de dimension finie. Alors $L(E, F) = B(E, F)$.*

Démonstration. en fixant une base de E on peut supposer que $E = \mathbb{R}^n$, et soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Comme toutes les normes sur $\mathbb{R}^n$ sont équivalentes, on peut munir $\mathbb{R}^n$ de la norme $\|\cdot\|_1$ définie dans 6.2.2.

On a pour $x = \sum_{i=1}^n x_i e_i$:

$$\begin{aligned} \|Ax\|_F = \|\textstyle\sum_{i=1}^n x_i e_i\|_F &\leq \textstyle\sum_{i=1}^n |x_i|\|Ae_i\|_F \\ &\leq C\textstyle\sum_{i=1}^n |x_i| = C\|x\|_1, \end{aligned}$$

où $C = \max \|Ae_i\|$. $\square$

6.6.1. *Norme sur $B(E, F)$.*

Proposition 6.17. *Soit $A \in B(E, F)$. On pose*

$$\|A\| := \sup_{x \in E, \|x\|_E \leq 1} \|Ax\|_F.$$

(1) *$\|\cdot\|$ est une norme sur $B(E, F)$, appelée* norme uniforme.
(2) *on a $\|Ax\|_F \leq \|A\|\|x\|_E$ pour tout $x \in E$.*
(3) *$\|A\|$ est la plus petite constante C telle que*

$$\|Ax\|_F \leq C\|x\|_E, \ \forall x \in E.$$

Démonstration. $\|A\|$ est bien définie, par le Théorème 6.14. On a pour $A \in B(E, F)$ $\lambda \in \mathbb{K}$ et $\|x\|_E \leq 1$:

$$\|(\lambda A)x\|_F = \|\lambda Ax\|_F = |\lambda|\|Ax\|_F$$

donc

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

De même pour $A, B \in B(E, F)$ et $\|x\|_E \leq 1$:

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

donc $\|A+B\| \leq \|A\| + \|B\|$. On a $\|Ax\|_F = \|x\|_E \|Ay\|_F$ pour $y = \|x\|_E^{-1} x$. Comme $\|y\|_E = 1$ on a donc $\|Ax\|_F \leq \|A\| \|x\|_E$, ce qui montre (2). (2) implique aussi que si $\|A\| = 0$ alors $Ax = 0_F$ pour tout $x \in E$ c'est à dire que $A = 0$. On a donc aussi montré (1).

Montrons (3). Soit $I \subset \mathbb{R}^+$ l' ensemble des $C \in \mathbb{R}^+$ tels que $\|Ax\| \leq C\|x\|$ pour tout $x \in E$. Par (2) on sait que $\|A\| \in I$. Il faut montrer que $\|A\|$ est le plus petit élément de I. Soit donc $C \in I$, c'est à dire que $\|Ax\| \leq C\|x\|$ pour tout $x \in E$. Pour $\|x\| \leq 1$ on a donc $\|Ax\| \leq C$, c'est à dire que $\|A\| = \sup_{\|x\| \leq 1} \|Ax\| \leq C$. $\|A\|$ est donc bien le plus petit élément de I. □

On peut montrer assez facilement le résultat suivant.

Théorème 6.18. *Soit E, F deux espaces de Banach. Alors $B(E, F)$ muni de la norme uniforme est un espace de Banach.*

6.6.2. *L' espace $B(E)$.*

Proposition 6.19. *Si $A, B \in B(E)$ alors $AB \in B(E)$ et :*

$$\|AB\| \leq \|A\|\|B\|, \; A, B \in B(E).$$

Démonstration. on a $\|ABx\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|$, pour tout $x \in E$. □

6.6.3. *Preuve du Théorème 6.7.* En utilisant la Proposition 4.2 (2), qui est valable pour (E, N), on voit que N_1 et N_2 sont topologiquement équivalentes si et seulement si

$$Id : (E, N_1) \to (E, N_2) \text{ et } Id : (E, N_2) \to (E, N_1)$$

sont continues. Comme ce sont des applications linéaires, elle sont continues si et seulement si elles sont bornées, c'est à dire si il existe des constantes $C_1, C_2 > 0$ telles que

$$\mathcal{N}_2(x) \leq C_1 N_1(x), \; N_1(x) \leq C_2 N_2(x), \; \forall x \in E,$$

ce qui signifie exactement que N_1 et N_2 sont équivalentes. □

6.7. **Le cas des matrices.** On identifie une matrice $A \in M_n(\mathbb{C})$ avec une application linéaire, encore notée $A \in L(\mathbb{C}^n)$ de la manière usuelle. Comme espace vectoriel $M_n(\mathbb{C})$ s' identifie à $\mathbb{C}^{n \times n}$ (une matrice $n \times n$ possède $n \times n$ entrées).

Toutes les normes sont donc équivalentes sur $M_n(\mathbb{C})$. Néanmoins la norme la plus utile est la norme uniforme obtenue à partir de la norme euclidienne sur $\mathbb{C}^n$ définie dans la sous Section 2.4.

6.7.1. *La norme uniforme sur $M_n(\mathbb{C})$.* Cette norme possède un certain nombre de propriétés importantes. Rappelons d' abord une définition.

Définition 6.20. *Soit $A = [a_{ij}]_{1\leq i,j\leq n} \in M_n(\mathbb{C})$. La matrice B telle que $b_{ij} = \overline{a}_{ji}$ est appelée* matrice adjointe *de A et notée A^*. On a*

$$(y|Ax) = (A^*y|x),\ \ x, y \in \mathbb{C}^n.$$

A est autoadjointe *si $A = A^*$.*

Lemme 6.21. *On a*

$$\|A^*\| = \|A\| = \|A^*A\|^{\frac{1}{2}}.$$

Démonstration. notons d' abord que par le Lemme 2.5 on a

$$\|A\| = \sup_{\|x\|\leq 1} \|Ax\| = \sup_{\|x\|,\|y\|\leq 1} |(y|Ax)|. \tag{6.7}$$

La première inégalité suit alors de (6.7), comme $(y|Ax) = (A^*y|x)$. On a donc aussi $\|A^*A\| \leq \|A^*\|\|A\| = \|A\|^2$. D 'autre part

$$\|A\|^2 = \sup_{\|x\|\leq 1} \|Ax\|^2 = \sup_{\|x\|\leq 1} (Ax|Ax) = \sup_{\|x\|\leq 1} (x|A^*Ax) \leq \|A^*A\|,$$

par (6.7), ce qui montre la deuxième inégalité. □

6.7.2. *Comment calculer la norme d' une matrice.* On suit la convention habituelle qui est de répéter les valeurs propres d' une matrice symétrique avec leur multiplicité, c'est à dire la dimension du sous espace propre associé.

Théorème 6.22. *Soit $A \in M_n(\mathbb{C})$. On a*

$$\|A\| = \max_{1\leq i\leq n} \mu_i,$$

où $\mu_i = \lambda_i^{\frac{1}{2}}$, et $\lambda_1, \dots, \lambda_n$ sont les valeurs propres de la matrice A^*A.

Démonstration. Soit d' abord B une matrice autoadjointe. On sait que B possède une base orthonormale de vecteurs propres $(e_1, \dots, e_n)$ avec $Be_i = \lambda_i e_i$. Comme $|\lambda_i| = \|Be_i\| \leq \|B\|\|e_i\| = \|B\|$, on a $\max_i |\lambda_i| \leq \|B\|$.

D' autre part si $x = \sum_i x_i e_i$, on a $Bx = \sum_i \lambda_i x_i e_i$ et donc

$$\|x\|^2 = \sum_i |x_i|^2,\ \ \|Bx\|^2 = \sum_i |\lambda_i|^2 |x_i|^2,$$

ce qui entraine que

$$\|Bx\|^2 \leq \max_i |\lambda_i|^2 \sum_i |x_i|^2 = (\max_i |\lambda_i|^2)\|x\|^2$$

et donc $\|B\|^2 \leq \max_i |\lambda_i|^2$.

Soit maintenant $A \in M_n(\mathbb{C})$ arbitraire. Comme $\|A\|^2 = \|A^*A\|$, il suffit d' appliquer ce qui précède à la matrice autoadjointe $B = A^*A$. □

6.7.3. *Comment majorer la norme d' une matrice.* En pratique il peut être impossible de calculer exactement les valeurs propres de la matrice A^*A pour appliquer le Théorème 6.22. On peut dans ce cas utiliser une majoration explicite de $\|A\|$, donnée dans la proposition suivante.

Proposition 6.23. *on a*

$$\|A\| \leq \|A\|_{\text{HS}} \tag{6.8}$$

où

$$\|A\|_{\text{HS}} := (\sum_{1\leq i,j\leq n} |a_{ij}|^2)^{\frac{1}{2}}.$$

Il est évident que $\|\cdot\|_{HS}$, appelée *norme de Hilbert-Schmidt* est une norme sur $M_n(\mathbb{C})$, comme c'est la norme euclidienne canonique sur $\mathbb{C}^{n\times n} \simeq M_n(\mathbb{C})$.
Démonstration. pour $x, y \in \mathbb{C}^n$ on a

$$(y|Ax) = \sum_{i=1}^{n} y_i(Ax)_i = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{y}_i a_{ij} x_j.$$

On applique l' inégalité de Cauchy-Schwarz pour le produit scalaire dans $\mathbb{C}^{n\times n}$ à la somme double et on obtient :

$$\begin{aligned} |(y|Ax)| \leq \ & (\textstyle\sum_{1\leq i,j\leq n} |\overline{y}_i x_j|^2)^{\frac{1}{2}} (\sum_{1\leq i,j\leq n} |a_{ij}|^2)^{\frac{1}{2}} \\ = \ & \|A\|_{\text{HS}} (\textstyle\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}} (\sum_{j=1}^{n} |y_j|^2)^{\frac{1}{2}} = \|A\|_{\text{HS}} \|x\| \|y\|, \end{aligned}$$

ce qui entraine (6.8), grâce à (6.7). □

7. Systèmes d' équations différentielles à coefficients constants

7.1. **Introduction.** Pour $I \subset \mathbb{R}$ un intervalle, on note par $I \ni t \mapsto x(t) \in \mathbb{R}^n$ une fonction de I dans $\mathbb{R}^n$, et on écrira parfois $x(t) = (x_1(t), \dots, x_n(t))$ où $I \ni t \mapsto x_i(t) \in \mathbb{R}$ sont des fonctions à valeurs réelles. On s' intéresse à des systèmes d' équations différentielles *à coefficients constants* de la forme :

$$\left\{ \begin{array}{ccccccc} x_1'(t) & = & a_{11}x_1(t) & + & \cdots & + & a_{1n}x_n(t) + f_1(t) \\ \vdots & = & \vdots & + & \vdots & + & \vdots \\ x_i'(t) & = & a_{i1}x_i(t) & + & \cdots & + & a_{in}x_n(t) + f_i(t) \\ \vdots & = & \vdots & + & \vdots & + & \vdots \\ x_n'(t) & = & a_{i1}x_n(t) & + & \cdots & + & a_{nn}x_n(t) + f_n(t), \end{array} \right.$$

où les coefficients a_{ij} sont des *constantes* réelles et les fonctions $I \mapsto f_i(t) \in \mathbb{R}$ sont fixées. En introduisant

$$A = [a_{ij}]_{1\leq i,j\leq n} \in M_n(\mathbb{R}), \ f(t) = (f_1(t), \dots, f_n(t)) : I \to \mathbb{R}^n$$

on peut réécrire ce système sous la forme condensée

$$(E) \ x'(t) = Ax(t) + f(t).$$

La fonction $f : I \to \mathbb{R}^n$ s' appelle le *second membre* de l' équation (E). L' équation

$$(H) \ x'(t) = Ax(t)$$

est appelée *l' équation homogène* associée à (E).

Souvent on est conduit à résoudre un *problème de Cauchy*, qui consiste à rajouter à (E) des *conditions initiales.* Si $x_0 \in \mathbb{R}^n$ et $0 \in I$ on considère le problème

$$(C) \left\{ \begin{array}{l} x'(t) = Ax(t) + f(t), \\ x(0) = x_0. \end{array} \right.$$

Définition 7.1. *Soit $I \subset \mathbb{R}$ un intervalle (avec $0 \in I$ dans le cas (C)) et $f : I \to \mathbb{R}^n$ une fonction continue. On appelle* solution *de (E) (ou de (H), (C)) sur I toute fonction $I \ni t \mapsto x(t) \in \mathbb{R}^n$ de classe C^1 sur I qui vérifie (E) (ou (H), (C)) en tout point $t \in I$.*

7.2. **Forme normale de Jordan d' une matrice.**

7.2.1. *Rappels.* – tout polynôme P à coefficients dans $\mathbb{C}$ ou $\mathbb{R}$ est *scindé* sur $\mathbb{C}$, c'est à dire que P peut se factoriser comme :

$$P(z) = \prod_{i=1}^{p} (z - \lambda_i)^{m_i},$$

avec $\lambda_1, \ldots, \lambda_p \in \mathbb{C}$ les racines de P et $m_i = m(\lambda_i)$ est la *multiplicité algébrique* de la racine λ_i.

– $\lambda \in \mathbb{K}$ est *valeur propre* de $A \in M_n(\mathbb{K})$ si $E_\lambda = \mathrm{Ker}(A - \lambda \mathbb{1}) = \{x \in \mathbb{K}^n : Ax = \lambda x\}$ est différent de $\{0_n\}$.

– on note par $\sigma_{\mathbb{C}}(A)$ le *spectre complexe* de A, c'est à dire l' ensemble des valeurs propres de A dans $\mathbb{C}$ (*même si* $A \in M_n(\mathbb{R})$, ie A est réelle).

– les valeurs propres de A sont les racines (dans $\mathbb{C}$) du *polynôme caractéristique* de A, $P_A(z) = \det(A - z\mathbb{1})$. Si $\lambda \in \sigma_{\mathbb{C}}(A)$ la multiplicité de λ comme racine de $P_A(z)$ est appelée *multiplicité algébrique* de λ, et notée $m(\lambda)$.

– si A est réelle on note par $\sigma_{\mathbb{R}}(A) = \sigma_{\mathbb{C}}(A) \cap \mathbb{R}$ le *spectre réel* de A, c'est à dire l' ensemble des valeurs propres réelles de A.

– si $\lambda \in \sigma_{\mathbb{C}}(A)$ on pose $\mu(\lambda) = \dim_{\mathbb{K}} E_\lambda$, la dimension de l' espace propre associé à λ, appelé *multiplicité géométrique* de A. On a $0 \leq \mu(\lambda) \leq m(\lambda)$.
– si A est une matrice réelle, et $\lambda \in \sigma_{\mathbb{R}}(A)$, $dim_{\mathbb{C}} E_\lambda = \dim_{\mathbb{R}} E_\lambda$. De plus si $\lambda \in \sigma_{\mathbb{C}}(A) \setminus \mathbb{R}$ alors $\overline{\lambda} \in \sigma_{\mathbb{C}}(A) \setminus \mathbb{R}$ et $\mu(\lambda) = \mu(\overline{\lambda})$, $m(\lambda) = m(\overline{\lambda})$.

– deux matrices $A, B \in M_n(\mathbb{K})$ sont *semblables* si et seulement si il existe $P \in GL_n(\mathbb{K})$ telle que $A = PBP^{-1}$.

Définition 7.2. *Une matrice $N \in M_n(\mathbb{C})$ est* nilpotente *d'ordre p si*

$$N^p = 0, \ N^k \neq 0 \textit{ pour } 0 \leq k \leq p - 1.$$

Pour $p \in \mathbb{N}$ on note par $N_p \in M_p(\mathbb{R})$ la matrice nilpotente d' ordre p canonique *:*

$$N_p = [n_{ij}]_{1 \leq i,j \leq p}, \ n_{ij} = \left\{ \begin{array}{l} 1 \textit{ si } j = i + 1 \\ 0 \textit{ sinon.} \end{array} \right.$$

Par exemple on a

$$N_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \ N_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ etc.}$$

Théorème 7.3. *[Théorème de Jordan] Soit $A \in M_n(\mathbb{C})$. Alors il existe $P \in GL_n(\mathbb{C})$ telle que*

$$A = P(D+N)P^{-1},$$

où D est diagonale et N nilpotente, et plus précisément :

$$D + N = \begin{pmatrix} T_1 & 0 & \dots & 0 \\ 0 & T_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & T_d \end{pmatrix}, \tag{7.1}$$

où $T_j = \lambda_j \mathbb{1} + N_j \in M_j(\mathbb{C})$ avec $\lambda_j \in \sigma_{\mathbb{C}}(A)$ et N_j la matrice nilpotente canonique d' ordre j.

7.3. **L' exponentielle d' une matrice.**

7.3.1. *Rappels.* – On munit $\mathbb{C}^n$ de sa norme euclidienne canonique. On identifie une matrice $A \in M_n(\mathbb{C})$ avec son application linéaire associée, notée encore $A \in L(\mathbb{C}^n)$. On a vu dans la Sous section 6.7 quelques propriétés importantes de la norme uniforme $\|A\|$, associée à la norme euclidienne.

– $M_n(\mathbb{C})$ muni de la norme uniforme est complet (comme tout evn de dimension finie), et donc toute série normalement convergente de matrices est convergente, d' après le Théorème 6.13.

– Le sous ensemble des matrices *inversibles* dans $M_n(\mathbb{C})$ est noté $GL_n(\mathbb{C})$ (pour 'groupe linéaire').

– On rappelle que la définition de l' exponentielle d' un nombre réel ou complexe est

$$e^x := \sum_{i=0}^{\infty} \frac{x^n}{n!}, \; x \in \mathbb{C},$$

qui converge absolument, par exemple en utilisant le critère de d' Alembert pour les séries numériques.

– Dans presque toute la suite, l' écriture concrète d' une matrice comme un tableau de nombres sera inutile. Seules comptent les propriétés de la norme matricielle par rapport à l' addition et au produit.

Théorème 7.4. *Soit $A \in M_n(\mathbb{C})$.*

(1) *la série*

$$\sum_{n \in \mathbb{N}} \frac{A^n}{n!}$$

converge dans $M_n(\mathbb{C})$, sa somme

$$e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

s' appelle l' exponentielle de A.

(2) *on a*

$$\|e^A\| \leq e^{\|A\|}.$$

(3) *on a*

$$\|e^A - \sum_{n=0}^{N} \frac{A^n}{n!}\| \leq \frac{\|A\|^{N+1}}{N+1!} e^{\|A\|}, \; N \in \mathbb{N}.$$

(4) *on a*

$$e^A e^B = e^{A+B} \text{ si } AB = BA.$$

(5) *si $AB = BA$ alors $e^A B = B e^A$ et en particulier $A e^A = e^A A$.*

Démonstration. comme $\|AB\| \leq \|A\|\|B\|$, on a $\|A^n\| \leq \|A\|^n$ par récurrence sur n et donc

$$\|\frac{A^n}{n!}\| = \frac{\|A^n\|}{n!} \leq \frac{\|A\|^n}{n!}.$$

La série $\sum_{n\in\mathbb{N}} \frac{A^n}{n!}$ converge donc normalement dans $M_n(\mathbb{C})$ donc converge, ce qui montre (1). Par l' inégalité triangulaire on a

$$\|\sum_{n=0}^{N} \frac{A^n}{n!}\| \leq \sum_{n=0}^{N} \frac{\|A\|^n}{n!},$$

ce qui entraine (2) en faisant $N \to +\infty$.

Montrons (3) : on a

$$\begin{array}{ll} & \|e^A - \sum_{n=0}^{N} \frac{A^n}{n!}\| = \|\sum_{n=N+1}^{\infty} \frac{A^n}{n!}\| \\ \leq & \sum_{n=N+1}^{\infty} \frac{\|A\|^n}{n!} = \frac{\|A\|^{N+1}}{N+1!} \int_0^1 (1-t)^N e^{t\|A\|} dt \leq \frac{\|A\|^{N+1}}{N+1!} e^{\|A\|}, \end{array}$$

où on a appliqué la formule de Taylor-Lagrange avec reste intégral à la fonction e^x en $x = 0$ dans la deuxième ligne.

Montrons (4). Comme $AB = BA$ on a

$$(A+B)^n = \sum_{p=0}^{n} C_n^p A^p B^{n-p}.$$

On montre alors (4) exactement comme l' identité $e^a e^b = e^{a+b}$ pour $a, b \in \mathbb{C}$ en regroupant les termes. (5) est évident et laissé en exercice. □

Remarque 7.5. *Si A et B ne commutent pas alors $e^A e^B \neq e^{A+B}$. Prenons par exemple $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. On a $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = BA$.*

On a $A^2 = B^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ donc $e^A = \mathbb{1}_2 + A$, $e^B = \mathbb{1}_2 + B$ et

$$e^A e^B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

D' autre part $A + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ donc $(A+B)^2 = \mathbb{1}_2$. On écrit pour $C = A + B$:

$$e^C = \sum_{p=0}^{\infty} \frac{C^{2p}}{2p!} + \sum_{p=0}^{\infty} \frac{C^{2p+1}}{2p+1!},$$

en séparant les puissances paires et impaires. Comme $C^2 = \mathbb{1}_2$ *la première somme vaut* $\mathbb{1}_2 \times \sum_{p=0}^{\infty} \frac{1}{2p!} = \mathrm{ch}(1)\mathbb{1}_2$*. La deuxième vaut* $C \times \sum_{p=0}^{\infty} \frac{1}{2p+1!} = C\mathrm{sh}1$*. On a donc*

$$\mathrm{e}^{A+B} = \begin{pmatrix} \mathrm{ch}1 & \mathrm{sh}1 \\ \mathrm{sh}1 & \mathrm{ch}1 \end{pmatrix} \neq \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \mathrm{e}^{A}\mathrm{e}^{B}.$$

Proposition 7.6. *On a pour* $A \in M_n(\mathbb{C})$ *:*

$$\begin{aligned}
&(1) \quad \mathrm{e}^{tA}\mathrm{e}^{sA} = \mathrm{e}^{(s+t)A}, \ t, s \in \mathbb{R},\\
&(2) \quad (\mathrm{e}^{tA})^{-1} = \mathrm{e}^{-tA}, \ t \in \mathbb{R},\\
&(3) \quad \tfrac{\mathrm{d}}{dt}(\mathrm{e}^{tA}) = A\mathrm{e}^{tA}, \ t \in \mathbb{R}.
\end{aligned}$$

Démonstration. (1) suit du Théorème 7.4 (4) comme tA et sA commutent. On en déduit que $\mathrm{e}^{tA}\mathrm{e}^{-tA} = \mathrm{e}^{-tA}\mathrm{e}^{tA} = \mathrm{e}^{0A} = \mathbb{1}$ ce qui entraine (2). Soit $t \in \mathbb{R}$. On a

$$\epsilon^{-1}(\mathrm{e}^{(t+\epsilon)A} - \mathrm{e}^{tA}) = \epsilon^{-1}(\mathrm{e}^{\epsilon A} - \mathbb{1})\mathrm{e}^{tA},$$

donc (3) se déduit de

$$\lim_{\epsilon \to 0} \epsilon^{-1}(\mathrm{e}^{\epsilon A} - \mathbb{1}) = A, \tag{7.2}$$

où on considère la limite dans l' evn $(M_n(\mathbb{C}), \|\cdot\|)$. Montrons (7.2). On a

$$\begin{aligned}
e^{\epsilon A} - \mathbb{1} &= \textstyle\sum_{n=1}^{\infty} \frac{\epsilon^n A^n}{n!} = (\epsilon A) \sum_{n=0}^{\infty} \frac{(\epsilon A)^n}{n+1!}\\
&= (\epsilon A)(\mathbb{1} + \textstyle\sum_{n=1}^{\infty} \frac{(\epsilon A)^n}{n+1!}) =: \epsilon A(\mathbb{1} + R_\epsilon(A)).
\end{aligned}$$

On a

$$\begin{aligned}
\|R_\epsilon(A)\| &\leq \textstyle\sum_{n=1}^{\infty} \frac{(\epsilon\|A\|)^n}{n+1!} = |\epsilon| \sum_{n=0}^{\infty} \|A\| \frac{(\epsilon\|A\|)^n}{n+2!}\\
&\leq |\epsilon|\|A\| \textstyle\sum_{n=0}^{\infty} \frac{\|A\|^n}{n+2!}, \ \text{pour } |\epsilon| \leq 1.
\end{aligned}$$

On a donc

$$\epsilon^{-1}(\mathrm{e}^{\epsilon A} - \mathbb{1}) = A + AR_\epsilon(A),$$

et $\|AR_\epsilon(A)\| \leq C|\epsilon|$, ce qui entraine (7.2) et termine la démonstration. □

7.4. **Application aux systèmes d' équations différentielles.**

Considérons d'abord le cas homogène :

$$(H) \ x'(t) = Ax(t), \ A \in M_n(\mathbb{R}).$$

Proposition 7.7. *Toutes les solutions de* (H) *sont données par*

$$x(t) = \mathrm{e}^{tA}v, \ v \in \mathbb{R}^n.$$

Démonstration. si $x(t) = \mathrm{e}^{tA}v$, alors $x(t)$ est solution de (H), par la Proposition 7.6 (2). Inversement soit $x(t)$ une solution de (H) sur I et $\tilde{x}(t) = \mathrm{e}^{-tA}x(t)$. On a

$$\tilde{x}'(t) = (\mathrm{e}^{-tA})'x(t) + \mathrm{e}^{-tA}x'(t) = -A\mathrm{e}^{-tA}x(t) + \mathrm{e}^{-tA}Ax(t) = \mathrm{e}^{-tA}(A - A)x(t) = 0,$$

et donc $\tilde{x}(t) = v$ pour un $v \in \mathbb{R}^n$, donc $x(t) = \mathrm{e}^{tA}v$. □

Regardons maintenant le cas non homogène (E) et le problème de Cauchy (C) :

$$(E) \ x'(t) = Ax(t) + f(t),$$

$$(C)\quad \begin{cases} x'(t) = Ax(t) + f(t), \\ \qquad\qquad x(0) = x_0. \end{cases}$$

Théorème 7.8. *Soit $I \ni t \mapsto f(t) \in \mathbb{R}^n$ une fonction continue et bornée, et $x_0 \in \mathbb{R}^n$.*

(1) *la solution générale de (E) est*

$$x(t) = \mathrm{e}^{tA}v + \int_0^t \mathrm{e}^{(t-s)A} f(s)ds, \ v \in \mathbb{R}^n.$$

(2) *l' unique solution de (C) est*

$$x(t) = \mathrm{e}^{tA}x_0 + \int_0^t \mathrm{e}^{(t-s)A} f(s)ds.$$

Démonstration. on adapte bêtement la méthode de la variation de la constante. Cherchons la solution de (E) sous la forme $x(t) = \mathrm{e}^{tA}y(t)$, où $y : I \to \mathbb{R}^n$. On a $x'(t) = A\mathrm{e}^{tA}y(t) + \mathrm{e}^{tA}y'(t)$ et donc $x(t)$ est solution de (E) si et seulement si $y'(t) = \mathrm{e}^{-tA}f(t)$, c'est à dire

$$y(t) = v + \int_0^t \mathrm{e}^{-sA} f(s)ds, v \in \mathbb{R}^n.$$

Ceci donne bien

$$x(t) = \mathrm{e}^{tA}v + \int_0^t \mathrm{e}^{(t-s)A} f(s)ds.$$

L' unique solution de (E) telle que $x(0) = x_0$ est bien celle donnée dans (2). □

7.5. **Comment calculer e^{tA}.** En pratique il est relativement facile de calculer directement e^{tA} à l' aide du Théorème de Jordan, à l' aide des remarques suivantes.

7.5.1. *Action des similitudes.* Soit $P \in GL_n(\mathbb{R})$ telle que $A = PBP^{-1}$ comme dans le Théorème 7.3. Ceci entraine que $A^n = PB^nP^{-1}$ et donc que

$$\mathrm{e}^{tA} = P\mathrm{e}^{tB}P^{-1}.$$

Si on sait calculer explicitement P et e^{tB}, on sait calculer e^{tA}.

7.5.2. *Décomposition en blocs.* Supposons que la matrice A se décompose en blocs

$$A = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)$$

Alors

$$A^n = \left(\begin{array}{c|c} A_1^n & 0 \\ \hline 0 & A_2^n \end{array}\right)$$

et donc

$$\mathrm{e}^{tA} = \left(\begin{array}{c|c} \mathrm{e}^{tA_1} & 0 \\ \hline 0 & \mathrm{e}^{tA_2} \end{array}\right).$$

Par récurrence ce résultat s' étend à un nombre quelconque de blocs.

7.5.3. *Le cas des blocs de Jordan.* Supposons que

$$A = \lambda \mathbb{1}_p + N_p,\ \lambda \in \mathbb{C},$$

où N_p est la matrice nilpotente canonique d' ordre p. On a alors, comme $\lambda \mathbb{1}_p$ et N_p commutent :

$$\mathrm{e}^{t(\lambda \mathbb{1}_p + N_p)} = \mathrm{e}^{t\lambda}\mathrm{e}^{tN_p} = \mathrm{e}^{t\lambda} \sum_{k=0}^{p-1} \frac{t^k N_p^k}{k!},$$

car $N^k = 0$ pour $k \geq p$. De plus la forme de N_p^k et donc celle de e^{tN_p} sont très simples. Par exemple

$$N_3^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, N_4^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ N_4^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

et donc :

$$\mathrm{e}^{tN_2} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},\ \mathrm{e}^{tN_3} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}, e^{tN_4} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} \\ 0 & 1 & t & \frac{t^2}{2!} \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

7.6. **Calcul pratique en dimension 2 ou 3.** Nous allons maintenant détailler la forme des solutions de (H) dans le cas $n = 2$ ou 3.

7.6.1. *Cas* $n = 2$. Soit $A \in M_2(\mathbb{R})$, P son polynôme caractéristique, qui est un polynôme de degré 2 dont le discriminant est noté Δ. On distingue différents cas.

1) $\Delta > 0$: P a deux racines réelles simples $\lambda_1 \neq \lambda_2$, A est donc diagonalisable. On cherche $\vec{u}_1, \vec{u}_2$ vecteurs propres de A pour λ_1, λ_2. La solution générale de (H) est

$$x(t) = c_1 \mathrm{e}^{\lambda_1 t}\vec{u}_1 + c_2 \mathrm{e}^{\lambda_2 t}\vec{u}_2,\ c_1, c_2 \in \mathbb{R}.$$

2) $\Delta < 0$: P a deux racines complexes conjuguées, $z = \lambda + \mathrm{i}\omega$, $\overline{z} = \lambda - \mathrm{i}\omega$, $\omega > 0$. On cherche un vecteur propre *complexe* de A, $\vec{u} \in \mathbb{C}^2$ pour z, et $\overline{\vec{u}}$ est alors vecteur propre de A pour $\overline{z}$, car A est une matrice réelle. La solution générale *réelle* de (H) est :

$$x(t) = \mathrm{Re}(c\vec{u}\mathrm{e}^{zt}),\ c \in \mathbb{C},$$

ou de manière équivalente :

$$\begin{aligned} x(t) = \ & c_1 \mathrm{e}^{\lambda t}(\cos(\omega t)\mathrm{Re}\vec{u} - \sin(\omega t)\mathrm{Im}\vec{u}) \\ & + c_2 \mathrm{e}^{\lambda t}(\sin(\omega t)\mathrm{Re}\vec{u} + \cos(\omega t)\mathrm{Im}\vec{u}),\ c_1, c_2 \in \mathbb{R}. \end{aligned}$$

3) $\Delta = 0$. P a une racine réelle double λ. Si A est diagonalisable, alors $A = \lambda \mathbb{1}_2$ et la solution générale de (H) est

$$x(t) = \mathrm{e}^{\lambda t}\vec{v},\ \vec{v} \in \mathbb{R}^2.$$

Si A n'est pas diagonalisable (c'est à dire $A \neq \lambda \mathbb{1}_2$), on cherche $\vec{u} \in \mathbb{R}^2$ vecteur propre de A pour λ. On cherche ensuite $\vec{v} \in \mathbb{R}^2$ tel que

$$A\vec{v} - \lambda\vec{v} = \vec{u}.$$

La solution générale de (H) est

$$x(t) = c_1 e^{\lambda t}\vec{u} + c_2 e^{\lambda t}(t\vec{u} + \vec{v}),\ c_1, c_2 \in \mathbb{R}.$$

7.6.2. *Cas $n = 3$.* Soit $A \in M_3(\mathbb{R})$. Le polynôme caractéristique P est un polynôme réel de degré 3, donc possède toujours une racine réelle λ_1. On peut donc factoriser $P(z) = (z - \lambda_1)Q(z)$ et chercher les racines du trinôme Q, dont on note à nouveau par Δ le discriminant. A nouveau il faut distinguer différents cas :

1) P a trois racines réelles distinctes $\lambda_1, \lambda_2, \lambda_3$ (et donc $\Delta > 0$). On cherche $\vec{u}_1, \vec{u}_2, \vec{u}_3$ vecteurs propres de A pour $\lambda_1, \lambda_2, \lambda_3$. La solution générale de (H) est

$$x(t) = c_1 e^{\lambda_1 t}\vec{u}_1 + c_2 e^{\lambda_2 t}\vec{u}_2 + c_3 e^{\lambda_3 t}\vec{u}_3,\ c_1, c_2, c_3 \in \mathbb{R}.$$

2) P a une racine réelle λ_1 et deux valeurs propres complexes conjuguées $z = \lambda + \mathrm{i}\omega$, $\overline{z} = \lambda - \mathrm{i}\omega$, $\omega > 0$ (et donc $\Delta < 0$). On cherche un vecteur propre réel $\vec{u}_1 \in \mathbb{R}^3$ pour λ_1 et un vecteur propre complexe $\vec{u} \in \mathbb{C}^3$ pour z. La solution générale de (H) est :

$$x(t) = c_1 e^{\lambda_1 t}\vec{u}_1 + \mathrm{Re}(c\vec{u}e^{zt}),\ c_1 \in \mathbb{R},\ c \in \mathbb{C},$$

ou de manière équivalente :

$$\begin{aligned} x(t) = \ & c_1 e^{\lambda_1 t}\vec{u}_1 \\ & + c_2 e^{\lambda t}(\cos(\omega t)\mathrm{Re}\vec{u} - \sin(\omega t)\mathrm{Im}\vec{u}) \\ & + c_3 e^{\lambda t}(\sin(\omega t)\mathrm{Re}\vec{u} + \cos(\omega t)\mathrm{Im}\vec{u}),\ c_1, c_2, c_3 \in \mathbb{R}. \end{aligned}$$

3) P a une racine réelle λ_1 et une racine réelle double $\lambda_2 \neq \lambda_1$ (et donc $\Delta = 0$).

Si A est diagonalisable, on cherche un vecteur propre $\vec{u}_1 \in \mathbb{R}^3$, pour λ_1, et deux vecteurs propres linéairement indépendants $\vec{u}_2, \vec{v}_2 \in \mathbb{R}^3$ pour λ_2. La solution générale de (H) est :

$$x(t) = c_1 e^{\lambda_1 t}\vec{u}_1 + c_2 e^{\lambda_2 t}\vec{u}_2 + c_3 e^{\lambda_2 t}\vec{v}_2,\ c_1, c_2, c_3 \in \mathbb{R}.$$

Si A n'est pas diagonalisable, on cherche $\vec{u}_2 \in \mathbb{R}^3$ vecteur propre de A pour λ_2. On cherche ensuite $\vec{v}_2 \in \mathbb{R}^3$ tel que

$$A\vec{v}_2 - \lambda\vec{v}_2 = \vec{u}_2.$$

La solution générale de (H) est

$$x(t) = c_1 e^{\lambda_1 t}\vec{u}_1 + c_2 e^{\lambda_2 t}\vec{u}_2 + c_3 e^{\lambda_2 t}(t\vec{u}_2 + \vec{v}_2),\ c_1, c_2, c_3 \in \mathbb{R}.$$

Pour être complet il faudrait aussi considérer le cas où P a une racine réelle triple. Nous laisserons ce cas de coté par manque de temps.

Annexe A.

A.1. **Preuve du Théorème 3.34.** Commençons par quelques définitions.

Définition A.1. $F \subset \mathbb{R}^d$ *a la* propriété d' intersection finie *si pour toute famille $(F_n)_{n\in\mathbb{N}}$ de fermés telle que $F_{n+1} \subset F_n$ et $F_n \cap F \neq \emptyset$ pour tout $n \in \mathbb{N}$ on a*

$$\Big(\bigcap_{n\in\mathbb{N}} F_n\Big) \cap F \neq \emptyset.$$

Une famille $(F_n)_{n\in\mathbb{N}}$ d' ensembles telle que $F_{n+1} \subset F_n$ pour tout n est appelée une *famille décroissante* d' ensembles.

Proposition A.2. *$F \subset \mathbb{R}^d$ est séquentiellement compact si et seulement si F a la propriété d' intersection finie.*

Démonstration. Supposons F séquentiellement compact et soit $(F_n)_{n\in\mathbb{N}}$ une famille de fermés comme dans la Définition A.1. En choisissant un $X_n \in F_n \cap F$ on obtient une suite (X_n) dans F. Par compacité séquentielle, il existe une sous-suite $(Y_n) = (X_{\varphi(n)})$ qui converge vers $X \in F$. Comme $Y_n = X_{\varphi(n)} \in F_{\varphi(n)}$ on a $Y_n \in F_p$ pour $\varphi(n) \geq p$ et donc comme F_p est fermé, on a $X = \lim Y_n \in F_p$. On en déduit que $X \in \bigcap_{p\in\mathbb{N}} F_p \cap F$, c'est à dire que F a la propriété d' intersection finie.

Inversement supposons que F a la propriété d' intersection finie et soit (X_n) une suite dans F. Posons $F_n = Adh(T_n)$ pour $T_n = \{X_k : k \geq n\}$). F_n est fermé, $F_{n+1} \subset F_n$ et $F_n \cap F$ contient X_n et donc est non vide. Il existe donc un $X \in \bigcap_{n\in\mathbb{N}} F_n \cap F$. On va construire une sous-suite $X_{\varphi(n)}$ telle que $\lim X_{\varphi(n)} = X$ de la manière suivante : comme $X \in Adh(T_n)$, on peut choisir des points $X_{\varphi(n)} \in T_n$ avec $d(X_{\varphi(n)}, X) \leq n^{-1}$ et $\varphi(n+1) > \varphi(n)$ pour tout n. La sous-suite $(X_{\varphi(n)})$ converge bien vers X, et donc F est séquentiellement compact. □

Lemme A.3. *Soit $F \subset \mathbb{R}^d$ séquentiellement compact. et $(U_i)_{i\in I}$ un recouvrement ouvert de F. Alors il existe $\delta > 0$ tel que pour tout $X \in F$ il existe $i \in I$ tel que $B(X, \delta) \cap F \subset U_i \cap F$.*

Démonstration. On raisonne par contradiction : supposons que pour tout $\delta > 0$ il existe $X \in F$ tel que pour tout $i \in I$ $B(X, \delta) \cap F$ n'est pas inclus dans $U_i \cap F$. En prenant $\delta = n^{-1}$ on obtient une suite (X_n) dans F telle que $B(X_n, n^{-1}) \cap F$ n'est pas inclus dans $U_i \cap F$ pour tout $i \in I$.

Par compacité séquentielle de F, il existe une sous-suite $(X_{\varphi(n)})$ qui converge vers $X \in F$. Comme $(U_i)_{i\in I}$ recouvre F il existe au moins un $i \in I$ tel que $X \in U_i$. Comme U_i est ouvert, il existe $\epsilon > 0$ tel que $B(X, \epsilon) \subset U_i$. Comme $X_{\varphi(n)} \to X$ il existe $n \in \mathbb{N}$ assez grand pour que $d(X, X_{\varphi(n)}) < \epsilon/2$ et $\varphi(n)^{-1} < \epsilon/2$. Alors $B(X_{\varphi(n)}, \varphi(n)^{-1}) \subset B(X, \epsilon) \subset U_i$, donc $B(X_{\varphi(n)}, \varphi(n)^{-1}) \cap F \subset B(X, \epsilon) \cap F \subset U_i \cap F$. Mais ceci contredit le choix de la suite (X_n), qui demande que $B(X_n, n^{-1}) \cap F$ n'est pas inclus dans $U_i \cap F$ pour tout $i \in I$. □

A.1.1. *Preuve du Théorème 3.34.* Supposons d' abord que F est compact. Soit $(F_n)_{n\in\mathbb{N}}$ une famille de fermés comme dans la Définition A.1. Supposons que $\bigcap_{n\in\mathbb{N}} F_n \cap F = \emptyset$, c'est à dire que F n' a pas la propriété d' intersection finie. Soit $U_n = \mathbb{R}^d \setminus F_n$. Les U_n sont ouverts et comme $\bigcap_{n\in\mathbb{N}} F_n \cap F = \emptyset$ on a $F \subset \bigcup_n U_n$, c'est à dire que $(U_n)_{n\in\mathbb{N}}$ forme un recouvrement ouvert de F.

Comme F est compact il existe un sous recouvrement fini $\{U_{n_1}, \dots, U_{n_k}\}$ de F, avec $n_1 < \cdots < n_k$. On a $F \subset \bigcup_{i=1}^k U_{n_i}$ donc $\bigcap_{i=1}^k F_{n_i} \cap F = \emptyset$. Comme la famille (F_n) est décroissante pour l 'inclusion, on a $\bigcap_{i=1}^k F_{n_i} = F_{n_k}$ et donc $F_{n_k} \cap F = \emptyset$ ce qui contredit le fait que $F_n \cap F \neq \emptyset$ pour tout n.

On en déduit que F a la propriété d' intersection finie, et est donc séquentiellement compact, par la Proposition A.2.

Supposons maintenant que F est séquentiellement compact, et montrons d' abord que F est *totalement borné*, c'est à dire que pour tout $\epsilon > 0$, il existe $X_1, \dots, X_n \in F$ tels que $F \subset \bigcup_{i=1}^n B(X_i, \epsilon)$.

On raisonne par contradiction et on suppose qu 'il existe $\epsilon_0 > 0$ tel que F n 'est recouvert par aucune union d' un nombre fini de boules de rayon ϵ_0 et de centres appartenant à F. On construit une suite (X_n) dans F de la manière suivante : on choisit $X_1 \in F$. Comme $F \not\subset B(X_1, \epsilon_0)$, il existe $X_2 \in F$ avec $d(X_1, X_2) \geq \epsilon_0$. Comme $F \not\subset B(X_1, \epsilon_0) \cup B(X_2, \epsilon_0)$, il existe $X_3 \in F$ avec $d(X_1, X_3) \geq \epsilon_0$ et $d(X_2, X_3) \geq \epsilon_0$. Par récurrence on construit ainsi une suite (X_n) avec $d(X_{n+1}, X_k) \geq \epsilon_0$ pour tous $1 \leq k \leq n$.

Comme F est séquentiellement compact, (X_n) possède une sous-suite convergente $(X_{\varphi(n)})$. On a $d(X_{\varphi(n+1)}, X_{\varphi(n)}) \geq \epsilon_0$, ce qui contredit la convergence de la suite $(X_{\varphi(n)})$. On en déduit que F est totalement borné.

Prenons maintenant un recouvrement ouvert $(U_i)_{i\in I}$ de F et soit $\delta > 0$ comme dans le Lemme A.3. Comme F est totalement borné, il existe $X_1, \dots, X_n \in F$ tels que $F \subset \bigcup_{j=1}^n B(X_j, \delta)$, et par le Lemme A.3 un $U_{i(j)}$ tel que $B(X_j, \delta) \subset U_{i(j)}$. On a donc $F \subset \bigcup_{j=1}^n U_{i(j)}$. F est donc compact. $\square$

A.2. **Preuve du Théorème 5.11.** Il suffit de traiter le cas $n = 2$, en gelant les coordonnées x_k pour $k \neq i, j$. Fixons $X_0 = (a_1, a_2) \in D$. On va montrer que

$$|\frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, a_2) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2)| \leq 2\epsilon, \ \forall \epsilon > 0. \tag{A.1}$$

Fixons donc $\epsilon > 0$. Soit $\vec{u} \in B(0, r)$ et

$$\begin{aligned} \varphi_1(t) &= f(a_1 + t, a_2 + u_2) - f(a_1 + t, a_2) - t u_2 \tfrac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2), \\ \varphi_2(t) &= f(a_1 + u_1, a_2 + t) - f(a_1, a_2 + t) - t u_1 \tfrac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, a_2), \end{aligned}$$

qui sont bien définies si $|t|$ et $\|\vec{u}\|$ sont assez petits. On a

$$\begin{aligned} \varphi_1'(t) = &\ \tfrac{\partial f}{\partial x_1}(a_1 + t, a_2 + u_2) - \tfrac{\partial f}{\partial x_1}(a_1 + t, a_2) - u_2 \tfrac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2) \\ = &\ \theta_t(u_2) - \theta_t(0), \end{aligned}$$

pour

$$\theta_t(s) = \frac{\partial f}{\partial x_1}(a_1 + t, a_2 + s) - s \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2).$$

On a

$$\frac{d}{ds}\theta_t(s) = \frac{\partial^2 f}{\partial x_2 \partial x_1} f(a_1 + t, a_2 + s) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2),$$

et par la continuité de $X \mapsto \frac{\partial^2 f}{\partial x_2 \partial x_1}(X)$ il existe $\delta > 0$ tel que si $|t|, |s| \leq \delta$ alors $|\frac{d}{ds}\theta_t(s)| \leq \epsilon$. En appliquant les accroissements finis à la fonction $s \mapsto \theta_t(s)$ entre $s = 0$ et $s = u_2$, on obtient que si $|u_2|, |t| \leq \delta$ on a $|\varphi_1'(t)| \leq \epsilon |u_2|$.

En appliquant à nouveau par les accroissements finis à la fonction $t \mapsto \varphi_1(t)$ entre $t = 0$ et $t = u_1$, on obtient que si $|u_1|, |u_2| \leq \delta$ on a

$$|\varphi_1(u_1) - \varphi_1(0)| \leq \epsilon |u_1| u_2|. \tag{A.2}$$

D' autre part

(A.3)
$$\begin{aligned}\varphi_1(u_1)-\varphi_1(0) &= f(a_1+u_1,a_2+u_2)-f(a_1+u_1,a_2)\\ &-f(a_1,a_2+u_2)+f(a_1,a_2)-u_1u_2\tfrac{\partial^2 f}{\partial x_1\partial x_2}(a_1,a_2).\end{aligned}$$

Le même raisonnement appliqué à la fonction φ_2 donne

(A.4)
$$|\varphi_2(u_2)-\varphi_2(0)|\leq \epsilon|u_1||u_2|.$$

D' autre part

(A.5)
$$\begin{aligned}&\varphi_2(u_2)-\varphi_2(0)=f(a_1+u_1,a_2+u_2)\\ &-f(a_1+u_1,a_2)-f(a_1,a_2+u_2)+f(a_1,a_2)-u_1u_2\tfrac{\partial^2 f}{\partial x_2\partial x_1}(a_1,a_2).\end{aligned}$$

En faisant la différence de (A.3) et (A.5) on obtient

$$\begin{aligned}&|(\varphi_1(u_1)-\varphi_1(0))-(\varphi_2(u_2)-\varphi_2(0))|\\ =\ &|u_1u_2||\tfrac{\partial^2 f}{\partial x_1\partial x_2}(a_1,a_2)-\tfrac{\partial^2 f}{\partial x_2\partial x_1}(a_1,a_2)|\end{aligned}$$

et par (A.2), (A.4) on a

$$|(\varphi_1(u_1)-\varphi_1(0))-(\varphi_2(u_2)-\varphi_2(0))|\leq 2\epsilon|u_1u_2|.$$

On obtient donc

$$|u_1u_2||\frac{\partial^2 f}{\partial x_1\partial x_2}(a_1,a_2)-\frac{\partial^2 f}{\partial x_2\partial x_1}(a_1,a_2)|\leq 2\epsilon|u_1u_2|$$

si $|u_1|,|u_2|\leq\delta$. En divisant par $|u_1u_2|$ on obtient (A.1). $\square$