Cycle Préparatoire Polytechnique

2$^{\text{ème}}$ année

Espaces Vectoriels Normés et Topologie

Polycopié de cours

Rédigé par Yannick Privat

Bureau 321 - Institut Élie Cartan Nancy (Mathématiques) - Université Henri Poincaré Nancy 1
B.P. 239, F-54506 Vandoeuvre-lès-Nancy Cedex.
e-mail : Yannick.Privat@iecn.u-nancy.fr

Introduction

Ce cours présente les grands concepts à l'origine de la Topologie et de l'Analyse fonctionnelle. L'étymologie du mot « topologie » est éloquente. En effet, en Grec, *topos* signifie *lieu* tandis que *logos* signifie *étude*.

Ce domaine des Mathématiques s'intéresse donc à l'étude des lieux, appelés en général *espaces* et aux propriétés qui les caractérisent. L'Analyse Fonctionnelle est très liée à la Topologie. En effet, dans cete branche des Mathématiques, on s'intéresse plus précisément aux espaces de fonctions. Un espace fonctionnel que vous connaissez probablement très bien est $\mathcal{C}([0,1])$, l'espace des fonctions continues sur le segment $[0,1]$.

Pour vous donner un exemple assez concret, vous connaissez peut-être le résultat suivant : si f est continue sur $[0,1]$, alors il existe x_0 et x_1, deux éléments de $[0,1]$ qui, respectivement, maximise et minimise f sur ce segment. Nous verrons qu'il existe un résultat bien plus général permettant de démontrer l'existence de minima et maxima d'une fonction. On comprendra aisémnent l'intérêt que cela présente dans le domaine de l'Optimisation par exemple. En Physique notamment, il est courant que l'on cherche à maximiser ou minimiser une énergie.

Historiquement, c'est Leonhard Euler (1707-1783) qui a initié la Topologie. En 1736, il présenta le *problème des sept ponts de Königsberg*. Kaliningrad (Königsberg jusqu'en 1946) est une ville de Russie, située dans une enclave territoriale totalement isolée du territoire russe, (jusqu'en 1945 « Prusse orientale ») au bord de la mer Baltique, entre la Pologne et la Lituanie. L'histoire veut que Léonhard Euler, en visite dans cette ville, ait eu à résoudre le problème qui préoccupait fortement ces habitants :

« Est-il possible de trouver un circuit qui emprunte une fois et une seule chacun des sept ponts de la ville ? »

La réponse, négative, fut trouvée par Léonhard Euler. Son intérêt principal réside dans le fait que ce résultat ne dépend d'aucune mesure (aucune distance).

La Topologie a connu une avancée considérable à la fin du XIX$^{\text{ème}}$ siècle et tout au long du XX$^{\text{ème}}$ siècle. Quelques grands noms de la Topologie sont :

- Henri Poincaré (1854-1912) ; (homotopie, cohomologie)
- David Hilbert (1862-1943) ; (bases de Hilbert, espaces de Hilbert)
- Maurice Fréchet (1878-1973) ; (convergence uniforme, convergence compacte, d'équicontinuité)
- Stefan Banach (1892-1945) ; (fondateur de l'Analyse Fonctionnelle, espaces de Banach)

Table des matières

Chapitre 1

Espaces vectoriels normés

1.1 Quelques rappels d'Algèbre linéaire

Si A et B désignent deux ensembles, on définit de prime abord le produit cartésien de A et de B, noté $A \times B$. Cette notation sera utilisée très régulièrement.

Définition 1.1. *Produit cartésien.*
- *Le* ***produit cartésien*** *de deux ensembles E et F, noté $E \times F$ est l'ensemble des couples dont le premier élément appartient à E et le second à F.*
- *Si $E = F$, on note $E^2 = E \times E$.*
- *Cette définition se généralise aisément. Si E_1, ..., E_n désignent n ensembles. On note $E = E_1 \times ... \times E_n$ le produit cartésien défini par :*

$$E = \{(e_1, ..., e_n), \text{ tel que } e_1 \in E_1, ..., e_n \in E_n\}.$$

Exemple : on définit par exemple l'ensemble $\mathbb{N} \times \mathbb{R}_+$. L'élément $(2, \pi)$ appartient à $\mathbb{N} \times \mathbb{R}_+$.

Remarque : si on considère des ensembles finis (i.e. dont le nombre d'éléments de l'ensemble est fini), on appelle cardinal de l'ensemble, le nombre d'éléments de l'ensemble. Et, si E et F sont finis, on a :

$$\text{card } (E \times F) = \text{card } E \times \text{card } F.$$

1.1.1 Groupes

Définissons au préalable la notion de loi de composition interne.

Définition 1.2. *Une loi de composition interne $*$ sur un ensemble E est une application de $E \times E$ dans E.*
Si $(a, b) \in E^2$, l'image de (a, b) est notée $a * b$.

Exemples : $+$ et $\times$ sont des lois de composition interne dans $\mathbb{R}$.
Sur $\mathbb{N}$, la loi $*$ définie pour $(a, b) \in \mathbb{N}^2$ par $a * b = a^b$ est une loi de composition interne.

Remarque : on note généralement lci pour désigner une loi de composition interne.

En Algèbre linéaire, on parle fréquemment de lois commutatives ou associatives.

Définition 1.3.
(*i*) *Une loi* $*$ *sur un ensemble* E *est dite commutative si, et seulement si :* $\forall(a, b) \in E^2$, $a*b = b*a$. *Si deux éléments* a *et* b *deux* E *sont tels que* $a * b = b * a$*, on dit qu'ils commutent.*
(*ii*) *Une loi* $*$ *sur* E *est dite associative si, et seulement si :* $\forall(a, b, c) \in E^3$, $(a*b)*c = a*(b*c)$. *Dans ce cas, on peut noter* $a * b * c$.
(*iii*) *Une loi* T *est dite distributive par rapport à une autre loi* $*$ *sur* E *si, et seulement si :*

$$\forall(a, b, c) \in E^3,\ (aTb) * c = (a * c)T(b * c) \text{ et } a * (bTc) = (a * b)T(a * c).$$

Exemple : sur $\mathbb{R}$, il est bien évident que les lois $+$ et $\times$ sont commutatives et associatives, et $\times$ est distributive par rapport à $+$.

Définition 1.4. *Élément neutre.*
Si $*$ *est une loi sur* E*, on dit que* $e \in E$ *est neutre si, et seulement si :* $\forall x \in E$, $e * x = x * e = x$.
Si e *est un élément neutre, alors* e *est nécessairement unique.*

Ces définitions étant établies, nous pouvons introduire la notion de groupe.

Définition 1.5. *On dit que* $(G, *)$ *est un groupe si* $*$ *est une loi de composition interne associative sur* G *pour laquelle il existe un élément neutre* e*, et tel que tout élément* $x \in G$ *est inversible, c'est à dire qu'il existe un élément* $x' \in G$ *tel que* $x * x' = x' * x = e$.

Remarque : si $x \in G$ est inversible, on note en général son inverse x^{-1}, mais attention ! Il ne s'agit là que d'une notation, sans rapport a priori avec un quotient dans $\mathbb{R}$.

Remarque 2 : si de plus, $*$ est commutative, on dit que le groupe est commutatif ou abélien.

Exemples : $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Z}, +)$, $(\mathbb{C}, +)$ sont des groupes. En revanche, $(\mathbb{Z}, \times)$ n'est pas un groupe. À votre avis, pourquoi ?

On définit très simplement la notion de sous-groupe :

Définition 1.6. *Soit* $(G, *)$*, un groupe. On dit que* H *est un sous-groupe de* G *pour* $*$ *si* $H \neq \emptyset$, $H \subset G$*, et si* $*$ *est une loi de composition interne sur* H *qui le munit d'une structure de groupe.*

1.1.2 Structure d'espace vectoriel

<u>Attention</u> ! Dans le paragraphe qui va suivre, les notations $+$ et $\times$ désigneront respectivement des lois commutative et associative. Définissons au préalable les notions d'anneau et de corps.

Définition 1.7. *Anneau.*
Soit A*, un ensemble,* $+$ *et* $\times$*, deux lois de composition interne sur* A*. Supposons que* $\times$ *est associative et distributive par rapport à* $+$.
On dit que A *est un anneau si* $(A, +)$ *est un groupe abélien, et s'il existe un élément neutre pour la loi* $\times$ *noté* 1_A.
On parle d'anneau commutatif lorsque $\times$ *est une loi commutative sur* A.

Remarque : le fait que $(A, +)$ soit un groupe impose l'existence d'un neutre pour la loi $+$, noté traditionnellement 0_A.

Exemples : je n'en donne que très peu car la notion de corps m'intéresse davantage que la notion d'anneau.
- $(\mathbb{N}, +, \times)$ est un anneau.
- $\{a + b\sqrt{2}, (a, b) \in \mathbb{Z}^2\}$ est un anneau.
- Si A est un anneau, $A[X]$, l'ensemble des polynômes à coefficients dans A est encore un anneau.
- L'ensemble des matrices carrées de type $n \times n$, avec $n \in \mathbb{N}$ fixé est un anneau non commutatif.

J'énonce à présent deux propriétés caractéristiques des anneaux que vous avez certainement déjà rencontrées.

Propriété 1.1. *Soit $(A, +, \times)$, un anneau* ***commutatif****. Soient a et b, deux éléments de A.*

1. ***Formule du binôme de Newton.***
$$(a+b)^n = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} a^k b^{n-k}, \text{ avec } \begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!(n-k)!}$$

2. ***Formule des anneaux.***
$$a^n - b^n = (a-b) \sum_{k=0}^{n-1} a^k b^{n-1-k}.$$

Remarque : la notation $(a+b)^n$ par exemple doit bien-sûr être comprise dans le sens suivant :
$$(a+b)^n = \underbrace{(a+b) \times ... \times (a+b)}_{n \text{ fois}}.$$

Passons à présent à la notion de corps.

Définition 1.8. *Corps.*
On dit que $(\mathbb{K}, +, \times)$ est un corps si $(\mathbb{K}, +)$ est un groupe abélien, $\times$ est une loi de composition interne associative, commutative et distributive par rapport à $+$, pour laquelle il y a un neutre $1_{\mathbb{K}}$ (neutre pour $\times$) distinct de $0_{\mathbb{K}}$ (neutre pour $+$), et si tout élément non nul de $\mathbb{K}$ est inversible pour $\times$.

Remarque 1 : un corps est en particulier un anneau.

Remarque 2 : dans la littérature, vous pourrez peut-être trouver une définition des corps un peu différente de celle que je donne ici : en effet, on parle parfois de corps, même lorsque la loi $\times$ n'est pas commutative. Mais ça n'est là qu'une convention et il s'agit de la fixer dès le départ.

Exemples et contre-exemple :
- $(\mathbb{Q}, +, \times)$ est un corps.
- $(\mathbb{C}, +, \times)$ est un corps.
- En revanche, $(\mathbb{Z}, +, \times)$ n'en est pas un. Dans $\mathbb{Z}$, 2 est non inversible pour $\times$.
- Un corps célèbre, souvent noté $\mathbb{F}_p$, où p désigne un nombre premier, est :
$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \text{ (autre notation) } = \left\{\overline{0}, \overline{1}, ..., \overline{p}\right\}, \text{ et } \forall k \in \{1, ..., p\},\ x \in \overline{k} \Longleftrightarrow x \equiv k[p].$$
$\overline{k}$ est une classe d'équivalence. Mais cet exemple s'éloigne déjà un peu du thème que je souhaite traiter ici, et si vous ne le trouvez pas parlant, laissez-le de côté, car il n'est en rien essentiel pour comprendre la suite.

Définissons à présent la notion d'espace vectoriel.

Définition 1.9. *Soit $\mathbb{K}$, un corps. On dit que $(E, +, .)$ est un $\mathbb{K}$-espace vectoriel si $(E, +)$ est un groupe abélien et si . est une loi externe sur E ayant $\mathbb{K}$ pour domaine d'opérateur, vérifiant les quatre points suivants :* $\forall(\lambda, \mu) \in \mathbb{K}^2, \forall(x, y) \in E^2$,
(i) $1_{\mathbb{K}}.x = x$;
(ii) $\lambda.(x + y) = \lambda x + \lambda y$;
(iii) $(\lambda + \mu)x = \lambda x + \mu x$;
(iv) $(\lambda\mu)x = \lambda(\mu x)$.

Exemples : Appelons $\mathbb{R}[X]$, l'ensemble des polynômes à coefficients réels. $\mathbb{R}[X]$ est un $\mathbb{R}$-espace vectoriel.
$(\mathbb{R}^2, +, .)$ est un $\mathbb{R}$-espace vectoriel.

On définit de façon assez classique la notion de sous espace vectoriel.

Définition 1.10. *Soit E un $\mathbb{K}$-espace vectoriel et $X \subset E$, un sous-ensemble de E. On dit que X est un sous-espace vectoriel de E s'il satisfait aux conditions de stabilité linéaire, i.e. :* $\forall(x, y) \in E^2$ *et* $\forall\lambda \in \mathbb{K}$, $x + y \in X$ *et* $\lambda x \in X$.

Exemple : Considérons le $\mathbb{R}$-espace vectoriel $\mathbb{R}[X]$, des polynômes à coefficients réels à une indéterminée. On appelle $\mathbb{R}_n[X]$ l'ensemble des polynômes à coefficients réels de degré au plus n. Alors, il est immédiat que $\mathbb{R}_n[X]$ est un sous-espace vectoriel de $\mathbb{R}[X]$.

Enfin, si E et F désignent deux espaces vectoriels, je rappelle à toute fin utile ce que l'on entend lorsque l'on écrit $E + F$ ou $E \oplus F$.

Définition 1.11. *Soient F et F', deux sous espaces vectoriels d'un espace vectoriel E. On définit l'espace :*

$$F + F' := \left\{f + f', f \in F \text{ et } f \in F'\right\}.$$

Remarque : $F + F'$ est donc l'espace des éléments s'écrivant sous la forme $f + f'$, avec $f \in F$ et $f \in F'$.

Propriété 1.2. *Soient F et F', deux sous espaces vectoriels d'un espace vectoriel E. l'espace $F + F'$ est un sous espace vectoriel de E.*
Notation : *si F et F' sont tels que $F \cap F' = \{0_E\}$, on dit que $F + F'$ est une **somme directe** et on note $F \oplus F'$.*

1.2 Quelques généralités sur les espaces vectoriels normés, introduction à l'Analyse fonctionnelle

1.2.1 Quelques éléments sur les normes

Nous allons définir successivement deux notions fondamentales de l'Analyse fonctionnelle. Il s'agit des notions de **distance** et **norme**. Nous donnerons des exemples dans chaque cas.

On se place dorénavant dans E, un $\mathbb{K}$-espace vectoriel.

Définition 1.12. *Notion de norme.*
Une application $N : E \longrightarrow \mathbb{R}$ est appelée ***norme*** *si, et seulement si les trois propriétés suivantes sont vérifiées :*
(i) $N(x) = 0 \Longrightarrow x = 0$, pour $x \in E$.
(ii) Soit $\lambda \in \mathbb{R}$. $N(\lambda x) = |\lambda|.N(x)$.
(iii) $\forall (x, y) \in E^2$, $N(x + y) \leq N(x) + N(y)$. (Inégalité triangulaire)

Remarque 1 : la propriété suivante découle directement de la notion de norme : « si N désigne une norme, on a pour tout élément x de l'espace vectoriel E, $N(x) \geq 0$ ». Par conséquent, on peut écrire : $N : E \longrightarrow \mathbb{R}_+$.

Remarque 2 : la norme la plus connue est la norme euclidienne définie sur $\mathbb{R}^2$ par :

$$\begin{array}{cccc} \|.\|_2 : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}_+ \\ & (x, y) & \longmapsto & (x^2 + y^2)^{\frac{1}{2}}. \end{array}$$

On va vérifier que $\|.\|_2$ est une norme. Mais auparavant, définissons rapidement la notion de produit scalaire. Rappelons au préalable qu'une forme linéaire sur un $\mathbb{K}$-espace vectoriel E est une application linéaire dont l'ensemble de définition est E et à valeurs dans $\mathbb{K}$. En général, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.
Dans tout ce qui suit, E désignera un $\mathbb{R}$-espace vectoriel. On étudiera le cas $\mathbb{K} = \mathbb{C}$ dans le chapitre 3 consacré notamment à l'étude des formes hermitiennes dans des $\mathbb{C}$ espaces vectoriels.

Définition 1.13. *On appelle* ***produit scalaire*** *sur E toute forme bilinéaire symétrique définie positive autrement dit, toute application φ de $E \times E$ dans $\mathbb{R}$ vérifiant :*
(i) $\forall x \in E$, $\varphi_x : y \longmapsto \varphi(x, y)$ est linéaire ;
(ii) $\forall (x, y) \in E^2$, $\varphi(x, y) = \varphi(y, x)$;
(iii) $\forall x \in E \backslash \{0\}$, $\varphi(x, x) > 0$.

Vocabulaire et remarques :
- La propriété (ii) s'appelle **symétrie** du produit scalaire, la propriété (i) associée à la propriété (ii) (symétrie du produit scalaire) permettent de montrer qu'un produit scalaire est linéaire par rapport à chacune des deux variables. Cette propriété s'appelle **bilinéarité** du produit scalaire. Enfin,
- Un produit scalaire peut se noter $\varphi(x, y)$ ou $< x, y >$.
- La donnée d'un produit scalaire permet de définir une norme N par la relation :

$$N^2(x) =< x, x > \Longleftrightarrow N(x) = \sqrt{< x, x >}.$$

- Si E, est un espace de dimension finie muni d'un produit scalaire, on dit que E est un espace **euclidien**.
- Si E, est un espace de dimension finie ou infinie muni d'un produit scalaire, on dit que E est un espace **préhilbertien réel**.

Le résultat qui suit est très important. Il énonce une inégalité au moins aussi récurrente dans les exercices que l'inégalité triangulaire. Il s'agit de l'inégalité de Cauchy-Schwarz.

Propriété 1.3. *Inégalité de Cauchy-Schwarz.*
Soit E, un $\mathbb{R}$-espace vectoriel et $< ., . >$, un produit scalaire sur E.
Pour tous $(x, y) \in E^2$, on a :

$$| < x, y > | \leq \sqrt{< x, x >}\sqrt{< y, y >}.$$

Démonstration. Soient x et y, deux éléments quelconques de E. Soit $\lambda \in \mathbb{R}$. On appelle ϕ, la fonction définie pour tout $x \in E$ par : $\varphi(x) =< x, x >$. Considérons le polynôme : $P(\lambda) = \phi(x+\lambda y)$. (le fait que P est un polynôme découle de la bilinéarité du produit scalaire, on est d'ailleurs en mesure de préciser le degré de P : $d^\circ P = 2$) Le produit scalaire étant une forme définie positive, on en déduit que $\phi(x + \lambda y) \geq 0$. Or, $P(\lambda) =< x + \lambda y, x + \lambda y >= \lambda^2\phi(y) + 2\lambda < x, y > +\phi(x)$ et P puisque P est un polynôme du second degré, positif, quel que soit $\lambda \in \mathbb{R}$, on en déduit que son discriminant Δ est négatif. Ainsi :

$$\Delta = 4\left((< x, y >)^2 - \phi(x)\phi(y)\right) \leq 0.$$

Cette inégalité s'écrit encore :

$$| < x, y > | \leq \sqrt{< x, x >}\sqrt{< y, y >}.$$

□

Remarque : on peut s'intéresser au cas d'égalité dans l'inégalité de Cauchy-Schwarz. Soient donc x et y, deux éléments de E tels que :

$$| < x, y > | = \sqrt{< x, x >}\sqrt{< y, y >}.$$

Dans ce cas, on a, avec les notations de la démonstration ci-dessus : $\Delta = 0$ et par conséquent, il existe $\lambda_0 \in \mathbb{R}$ tel que : $P(\lambda_0) = 0$, autrement dit $\phi(x + \lambda_0 y) = 0$, et puisque le produit scalaire est défini positif, on en déduit que $x = -\lambda_0 y$, autrement dit, x et y sont liés.
Réciproquement, on vérifie aisément que si x et y sont liés, alors, on a un cas d'égalité dans l'inégalité de Cauchy-Schwarz.

Nous avons à présent tous les outils nécessaires pour démontrer que $\|.\|_2$ est une norme. Le produit scalaire associé à $\|.\|_2$ est, en notant $x = (x_1, x_2)$ et $y = (y_1, y_2)$: $< x, y >_2= x_1y_1+x_2y_2$.

Preuve que $\|.\|_2$ est une norme :
(i) $\|x\|_2 = 0 \Longrightarrow \|x\|_2^2 = 0 \Longrightarrow x_1^2 + x_2^2 = 0 \Longrightarrow x_1 = 0$ et $x_2 = 0$.
(ii) On vérifie immédiatement que $\|\lambda x\|_2 = |\lambda|\|x\|_2$, pour tout $\lambda \in \mathbb{R}$.
(iii) $\|x+y\|_2^2 = (x_1+y_1)^2+(x_2+y_2)^2 = (x_1^2+x_2^2)+(y_1^2+y_2^2)+2(x_1y_1+x_2y_2) = \|x\|_2^2+\|y\|_2^2+2 < x, y >_2\leq \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2\|y\|_2 = (\|x\|_2 + \|y\|_2)^2$, d'après l'inégalité de Cauchy-Schwarz, puis on obtient l'inégalité triangulaire par passage à la racine.

D'autres exemples de normes : on appelle $\|.\|_1$, la norme définie par :

$$\begin{array}{rccc} \|.\|_1 : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}_+ \\ & (x_1, x_2) & \longmapsto & |x_1| + |x_2|. \end{array}$$

On va démontrer que $\|.\|_1$ est encore une norme.

Preuve que $\|.\|_1$ est une norme :
(i) $\|x\|_1 = 0 \Longrightarrow |x_1 + |x_2| = 0 \Longrightarrow x_1 = 0$ et $x_2 = 0$.
(ii) Pour tout $\lambda \in \mathbb{R}$, $\|\lambda x\|_1 = |\lambda x_1| + |\lambda x_2| = |\lambda|(|x_1| + |x_2|)$.
(iii) $\|x + y\|_1 = \|(x_1 + y_1, x_2 + y_2)\|_1 = |x_1 + y_1| + |x_2 + y_2| \leq |x_1| + |y_1| + |x_2| + |y_2|$, en utilisant l'inégalité triangulaire sur les valeurs absolues, et on a : $\|x+y\|_1 \leq \|x\|_1 + \|y\|_1$, en regroupant convenablement les termes.

On appelle $\|.\|_\infty$, la norme définie par :

$$\begin{array}{rrcl} \|.\|_\infty : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}_+ \\ & (x_1, x_2) & \longmapsto & \max(|x_1|, |x_2|). \end{array}$$

On va démontrer que $\|.\|_\infty$ est encore une norme.

Preuve que $\|.\|_\infty$ est une norme :

(i) $\|x\|_\infty = 0 \Longrightarrow \max(|x_1|, |x_2|) = 0 \Longrightarrow |x_1| = |x_| = 0$, car $|x_i| \geq 0$, pour $i \in \{1, 2\}$ et donc, $x_1 = x_2 = 0$.

(ii) Pour tout $\lambda \in \mathbb{R}$, $\|\lambda x\|_\infty = \max(|\lambda x_1|, |\lambda x_2|) = \max(|\lambda| \times |x_1|, |\lambda| \times |x_2|) = |\lambda| \max(|x_1|, |x_2|)$.

(iii) $\|x + y\|_\infty = \|(x_1 + y_1, x_2 + y_2)\|_\infty = \max(|x_1 + y_1|, |x_2 + y_2|)$. Supposons par exemple que l'on ait : $\max(|x_1 + y_1|, |x_2 + y_2|) = |x_1 + y_1| \leq |x_1| + |y_1|$, et $|x_1| \leq \max(|x_1|, |x_2|) = \|x\|_\infty$, et de même, $|y_1| \leq \|y\|_\infty$, et ce raisonnement fonctionne encore dans le cas où le max vaut $|x_2 + y_2|$.
Ceci achève la preuve. On a : $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$.

1.2.2 Normes dans $\mathbb{R}^n$

Il s'agit juste d'une généralisation du cas précédent. Si on se place dans $\mathbb{R}^n$, on choisit deux éléments $x = (x_1, x_2, ..., x_n)$ et $y = (y_1, y_2, ..., y_n)$. Le produit scalaire dans $\mathbb{R}^n$ est défini par :

$$< x, y >_{\mathbb{R}^n} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + ... + x_n y_n.$$

Par conséquent, la norme euclidienne dans $\mathbb{R}^n$ est définie par :

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + ... + x_n^2} = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{< x, x >_{\mathbb{R}^n}}.$$

Les autres normes se généralisent aussi au cas de la dimension n. On a :

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + ... + |x_n| \text{ et } \|x\|_\infty = \max_{1 \leq i \leq n} (|x_i|) = \max(|x_1|, |x_2|, ..., |x_n|).$$

La preuve que ces fonctions désignent effectivement des normes constitue une généralisation des cas précédents. Il est possible de raisonner par récurrence si l'on souhaite généraliser les preuves précédentes. Les détails d'une telle preuve sont laissés aux soins du lecteur.

On peut encore munir $\mathbb{R}^n$ de la norme $\|.\|_p$, pour p entier naturel non nul, définie par :

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}.$$

La preuve que $\|.\|_p$ consitue une norme repose sur les inégalités de Hölder et Minkowski. Cette preuve fait l'objet d'un des exercices du chapitre.

Définition 1.14. *Notion de distance.*
Soit X, un ensemble. Une ***distance*** *sur X est une application d de $X \times X = X^2$ dans $\mathbb{R}_+$ vérifiant les trois propriétés suivantes :*
(i) $d(x,y) = 0 \Longleftrightarrow x = y$.
(ii) $\forall (x,y) \in X^2$, $d(x,y) = d(y,x)$.
(iii) $\forall (x,y,z) \in X^3$, $d(x,z) \leq d(x,y) + d(y,z)$ *(inégalité triangulaire).*

Exemples :

- **Dans $\mathbb{R}$:** on définit la distance entre deux éléments a et b de $\mathbb{R}$ par : $d(a,b) = |a-b|$. Par exemple, la distance entre 2 et -3 est $d(2,-3) = |2-(-3)| = 5$.
- **Dans $\mathbb{R}^n$:** on suppose que $\mathbb{R}^n$ est muni d'une norme N. Si x et y sont deux éléments de $\mathbb{R}^n$, on définit la distance entre x et y par :

  $$d(x,y) = N(x-y).$$

  Le fait que d définit bien une distance découle directement des propriétés de la norme (inégalité triangulaire, etc.) Le lecteur s'en convaincra facilement à l'aide d'un petit calcul.

Définition 1.15. *Espace métrique.*
Un espace métrique est la donnée d'un couple (X,d), où X est un espace et d une distance sur cet espace.

1.2.3 Notions sur les ouverts et les fermés

Dans ce cours, nous évoquerons fréquemment les notions d'ouvert et de fermés. De quoi s'agit-il ? Supposons que nous nous soyons placés dans un espace vectoriel normé E, c'est à dire muni d'une norme N. En général, nous utiliserons la norme euclidienne (la norme $\|.\|_2$) et l'espace vectoriel sera souvent $\mathbb{R}^n$, avec $n \in \mathbb{N}^*$.

Définition 1.16. *Boule ouverte.*
La boule ouverte de centre ω et de rayon r désigne l'ensemble des éléments x de l'espace vectoriel tels que $N(x-\omega) < r$. On la note indifféremment $\mathcal{B}(\omega,r)$ ou $B(\omega,r)$.

Exemple 1 : dans $\mathbb{R}^2$ muni de la norme $\|.\|_2$, la boule ouverte de centre $O(0,0)$ et de rayon 1 est l'intérieur du cercle unité privé de sa frontière.
Si on munit $\mathbb{R}^2$ de la norme $\|.\|_\infty$, la boule ouverte unité est l'intérieur du carré de bord délimité par les points de coordonnées $(-1,-1)$, $(-1,1)$, $(1,1)$ et $(1,-1)$. À titre de comparaison, on a tracé dans un même repère les frontières des trois boules unités correspondant respectivement aux normes $\|.\|_1$ (petit carré), $\|.\|_2$ (cercle) et $\|.\|_\infty$ (grand carré).
Remarque 1 : on peut facilement comprendre l'origine de ces représentations.

- Pour la norme $\|.\|_2$, tout élément $X = (x,y)$ de la boule unité, notée $\mathcal{B}_2$ vérifie : $\|X - O\|_2 < 1 \Longleftrightarrow x^2 + y^2 < 1$, en passant au carré dans l'inégalité obtenue. En se souvenant que l'équation cartésienne $x^2 + y^2 = 1$ est celle d'un cercle centré en O et de rayon 1, on comprend que la boule unité $\mathcal{B}_2$ est, dans $\mathbb{R}^2$, le disque unité privé de son bord.
- Pour la norme $\|.\|_1$, tout élément $X = (x,y)$ de la boule unité, notée $\mathcal{B}_1$ vérifie : $\|X - O\|_1 < 1 \Longleftrightarrow |x| + |y| < 1 \Longleftrightarrow -1 - |x| < y < 1 - |x|$, par définition de la valeur absolue. En distinguant les cas $x \geq 0$ et $x < 0$, on peut exprimer aisément les équations des courbes

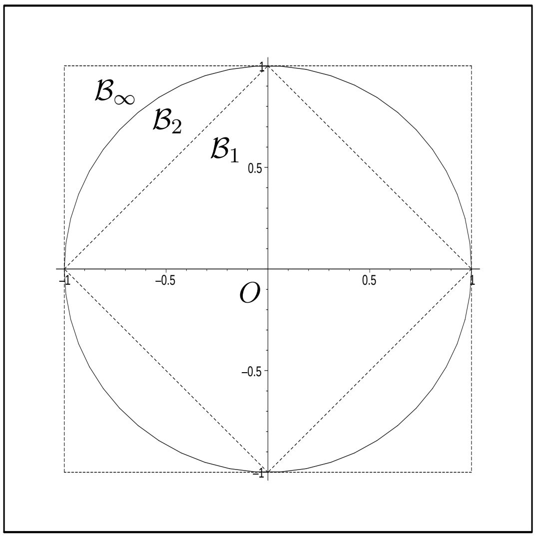

FIG. 1.1 – Représentation des boules unités associées aux normes $\|.\|_1$, $\|.\|_2$ et $\|.\|_\infty$

d'équations respectives $y = -1 - |x|$ et $y = -1 + |x|$, sans le symbole « valeur absolue », ce qui permet de montrer que la boule unité $\mathcal{B}_1$ est le petit carré retourné (formé de quatre segments) de la figure ci-dessus.
- Pour la norme $\|.\|_\infty$, le principe est exactement le même que pour $\|.\|_1$. On appelle $\mathcal{B}_\infty$, la boule unité associée à cette norme.

Remarque 2 : attention aux « faux-amis » ! ! On parle de boule ouverte associée à une norme, bien qu'en général, une boule n'ait rien de rond. En effet, si l'on considère par exemple la norme $\|.\|_\infty$, la boule unité ouverte est un carré...

Exemple 2 : considérons à présent la norme $\|.\|_2$ (norme euclidienne) et caractérisons, pour $n \in \{1, 2, 3\}$, les différentes boules unités dans $\mathbb{R}^n$ muni de cette norme.
- Si $x \in \mathbb{R}$, alors $\|x\|_2 = |x|$ et l'équation cartésienne de la boule unité est donc $|x| < 1$, autrement dit, il s'agit du segment $]-1, 1[$.
- Si $x = (x_1, x_2) \in \mathbb{R}^2$, alors $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ et l'équation cartésienne de la boule unité est donc $x_1^2 + x_2^2 < 1$. On reconnaît donc l'équation du disque de centre O, l'origine du repère et de rayon 1, privé de sa frontière.
- Si $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, alors $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2}$ et l'équation cartésienne de la boule unité est donc $x_1^2 + x_2^2 + x_3^2 < 1$. On reconnaît donc l'équation de la boule de centre O, l'origine du repère et de rayon 1, privé de son bord. Le bord de cette boule est bien-sûr la sphère de centre O et de rayon 1.

Définition 1.17. *Notion d'ouvert.*
Un ensemble $\mathcal{O}$ est appelé ***ouvert*** *si, pour tout point $x \in \mathcal{O}$, on peut trouver une boule ouverte de centre x, incluse dans l'ensemble $\mathcal{O}$.*

Remarque : une boule ouverte est un ouvert. Cela se conçoit assez bien géométriquement.

Exemple 1 : si l'on s'est placé dans un espace vectoriel normé (E, N), $\emptyset$ et E sont ouverts dans E.

Exemple 2 : si a et b sont deux nombres réels tels que $a < b$, alors , on démontre aisément que $]a, b[$ est un ouvert.

Preuve : en effet, si $x_0 \in]a, b[$, on considère la boule ouverte de centre x_0 et de rayon $r = \dfrac{\min(b - x_0, x_0 - a)}{2}$, par exemple. On a alors : $\forall x \in]a, b[, \ |x - x_0| < r \Longleftrightarrow x_0 - \dfrac{\min(b - x_0, x_0 - a)}{2} < x < x_0 + \dfrac{\min(b - x_0, x_0 - a)}{2}$. Or, $x_0 - \dfrac{\min(b - x_0, x_0 - a)}{2} = \dfrac{\min(x_0 - b, x_0 + a)}{2} = \dfrac{x_0 + a}{2} > a$, car $x_0 < b$, et de même, $x_0 + \dfrac{\min(b - x_0, x_0 - a)}{2} = \dfrac{\min(3x_0 - a, x_0 + b)}{2} < b$. Donc $B(x_0, r) \subset]a, b[$, et $]a, b[$ est donc un ouvert.

Définition 1.18. *Notion de fermé.*
Un fermé se définit comme le complémentaire d'un ouvert dans l'ensemble considéré.

Exemple 1 : si l'on s'est placé dans un espace vectoriel normé (E, N), $\emptyset$ et E désignent toujours à la fois des ouverts et de fermés de E.

Exemple : si a et b sont deux nombres réels tels que $a < b$, alors , on démontre aisément que $[a, b]$ est un fermé, en démontrant tout simplement que son complémentaire $\mathbb{R} \backslash [a, b] =]-\infty, a[\cup]b, +\infty[$ est un ouvert. On utilise exactement la même méthode que celle utilisée pour démontrer que $]a, b[$ est un ouvert.

Définition 1.19. *Boule fermée.*
La boule fermée de centre ω et de rayon r désigne l'ensemble des éléments x de l'espace vectoriel tels que $N(x - \omega) \leq r$. On la note indifféremment $\mathcal{B}_f(\omega, r)$ ou $B_f(\omega, r)$.

Remarque : comme on s'y attend, on peut facilement démontrer que :
- Une boule ouverte dans un espace vectoriel normé E est un ouvert.
- Une boule fermée dans un espace vectoriel normé E est un fermé.

Nous allons à présent définir la notion de topologie sur un espace vectoriel normé $(E, \|.\|)$.

Définition 1.20.
- *La topologie de E est l'ensemble des ouverts de E. On la note $\mathcal{T}$.*
- *Deux normes N_1 et N_2 sont dites équivalentes s'il existe deux nombres réels $\alpha > 0$ et $\beta > 0$ tels que :*
$$\forall x \in E, \ \alpha N_1(x) \leq N_2(x) \leq \beta N_1(x).$$

L'intérêt de cette notion réside dans le théorème suivant : dans le cas d'équivalence de deux normes N_1 et N_2, les ouverts de l'espace vectoriel normé sont les mêmes.

Théorème 1.1. *Deux normes équivalentes sur un espace vectoriel normé E donnent la même topologie.*

Démonstration. Soient N_1 et N_2, deux normes équivalentes. on appelle $\mathcal{T}_1$ et $\mathcal{T}_2$ les topologies associées. On veut montrer que $\mathcal{T}_1 = \mathcal{T}_2$.
Soit $A \in \mathcal{T}_1$, autrement dit, un ouvert pour N_1. Si $A \neq \emptyset$, soit $x \in A$. Il existe donc $r_x > 0$ tel

que $B_1(x,r_x) \subset A$, où $B_1(x,r_x)$ désigne bien-sûr la boule ouverte de centre x et de rayon r_x, pour la norme N_1. Or, puisque N_1 et N_2 sont supposées équivalentes :

$$\exists(\alpha,\beta) \in (\mathbb{R}_+^*)^2 \ : \ \forall x \in E,\ \alpha N_1(x) \leq N_2(x) \leq \beta N_1(x).$$

On va démontrer à présent que A est un ouvert pour la norme N_2. Soit $y \in B_2(x,\alpha r_x)$. Alors, $N_2(y-x) < \alpha r_x \Longrightarrow N_1(y-x) < r_x$, d'après la relation ci-dessus. On en déduit que $B_2(x,\alpha r_x) \subset B_1(x,r_x) \subset A$. Posons $r'_x = \alpha r_x$. Quel que soit $x \in A$, il existe donc $r'_x > 0$ tel que $B_2(x,r'_x) \subset A$, autrement dit, A est un ouvert de N_2, puis $\mathcal{T}_1 \subset \mathcal{T}_2$.
Exactement de la même façon, on démontre que $\mathcal{T}_1 \subset \mathcal{T}_2$, et le résultat souhaité s'ensuit. □

Exemple : les normes $\|.\|_1$, $\|.\|_2$ et $\|.\|_\infty$ sont équivalentes dans $\mathbb{R}^n$, $n \geq 1$.
Souvenons-nous que si $x = (x_1, ..., x_n) \in \mathbb{R}^n$, on a :

- $\|x\|_1 = \sum_{i=1}^{n} |x_i|$.
- $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$.
- $\|x\|_\infty = \max_{i\in\{1,...,n\}} |x_i|$.

Alors, quel que soit $x \in \mathbb{R}^n$, si l'on suppose que $\max_{i\in\{1,...,n\}} |x_i| = |x_{i_0}|$ on a :

$$|x_{i_0}| \leq \sum_{i=1}^{n} |x_i| \leq n|x_{i_0}|.$$

On a majoré chaque terme de la somme par leur max. On en déduit que les normes $\|.\|_1$ et $\|.\|_\infty$ sont équivalentes. De la même façon, on a :

$$|x_{i_0}| \leq \sqrt{\sum_{i=1}^{n} x_i^2} \leq \sqrt{n|x_{i_0}|^2} = \sqrt{n}|x_{i_0}|.$$

Conclusion : les trois normes ci-dessus sont bien équivalentes si l'on se place sur un espace vectoriel de dimension finie. Attention, d'après les inégalités ci-dessus, cela ne sera plus forcément vrai en dimension infinie, comme nous allons le voir dans le contre exemple qui suit.

Contre-exemple : le cas de $\mathbb{K}[X]$.
On rappelle que, si $\mathbb{K}$ désigne un corps, $\mathbb{K}[X]$ désigne l'ensemble des polynômes à coefficients dans $\mathbb{K}$. X désigne bien-sûr l'indéterminée. On introduit à présent les normes N_1, N_2 et N_∞ sur l'espace $\mathbb{K}[X]$ définies pour $P = a_0 + a_1 X + ... + a_n X^n \in \mathbb{K}[X]$ par :

- $N_1(P) = \sum_{n\in\mathbb{N}} |a_n|$.
- $N_2(P) = \sqrt{\sum_{n\in\mathbb{N}} |a_n|^2}$.
- $N_\infty(P) = \sup_{n\in\mathbb{N}} |a_n|$.

Considérons à présent la famille de polynômes $(P_q)_{q\in\mathbb{N}^*}$ définie par la relation :

$$P_q = 1 + X + ... + X^{q-1}.$$

On a : $N_1(P_q) = q$, $N_2(P_q) = \sqrt{q}$ et $N_\infty(P_q) = 1$. Remarquons de plus que :

$$\frac{N_1(P_q)}{N_2(P_q)} = \sqrt{q} \text{ et } \frac{N_2(P_q)}{N_\infty(P_q)} = \sqrt{q}.$$

Il suffit de faire tendre q vers $+\infty$ pour comprendre que les quotients ci-dessus ne sont pas bornés, ce qui contredit l'équivalence des trois normes.

Propriété 1.4. *En dimension finie, toutes les normes sont équivalentes.*

Nous venons de le démontrer pour les trois cas particuliers ci-dessus qui, en pratique, sont très utilisés. La démonstration générale de cette propriété sera présentée dans le chapitre 2.

On donne à présent des propriétés caractéristiques des ouverts et des fermés, dont une application importante sera révêlée dans la recherche d'extrema d'une fonction de plusieurs variables.

Propriété 1.5. *Soit $(E, \|.\|)$, un espace vectoriel normé.*
(i) $\emptyset$ et E sont à la fois des ouverts et des fermés de E.
(ii) Si A_1 et A_2 sont ouverts, alors $A_1 \cap A_2$ est ouvert.
(iii) Si $(A_i)_{i\in I}$ est une famille d'ouverts (I désigne un ensemble d'indices), alors $\bigcup_{i\in I} A_i$ est ouvert.
(iv) Si B_1 et B_2 sont deux fermés de E, alors $B_1 \cup B_2$ est encore fermé.
(v) Si $(B_i)_{i\in I}$ désigne une famille de fermés et I, un ensemble d'indices, alors $\bigcap_{i\in I} B_i$ est un fermé.

Remarque : un raisonnement par récurrence et l'application de la propriété précédente permettent de montrer de façon instantanée que si A_1, ..., A_n sont des parties ouvertes de l'espace E, alors leur intersection $\bigcap_{i=1}^{n} A_i$ est encore une partie ouverte de E. Exactement de la même façon, on montre que si B_1, ..., B_n sont des parties fermées de E, alors $\bigcup_{i=1}^{n} B_i$ est encore une partie fermée de E.

Démonstration. Quelques idées pour démontrer cette proposition :

- Soit $x \in \bigcup_{i\in I} A_i$. Il existe donc un indice i_0 tel que A_{i_0} contienne x. Or, puisque i_0 est ouvert, il existe $\varepsilon > 0$ tel que $\mathcal{B}(x,\varepsilon) \subset A_{i_0}$, et donc $\mathcal{B}(x,\varepsilon) \subset \bigcup_{i\in I} A_i$.
- La propriété relative aux sous-espaces fermés se démontre immédiatement, à partir de la démonstration ci-dessus, par passage au complémentaire...

□

Contre-exemple : considérons la famille dénombrable d'ouverts $(A_n)_{n\in\mathbb{N}^*}$ définie par : $A_n = \left]-\frac{1}{n}, \frac{1}{n}\right[$.

Notes : une famille dénombrable est une famille qu'il est possible d'indicer par $\mathbb{N}$, l'ensemble des entiers naturels.

Il est ici tout à fait évident que $(A_n)_{n\in\mathbb{N}^*}$ est une famille d'ouvert. En revanche, on a :

$$\bigcap_{n\in\mathbb{N}^*} A_n = \bigcap_{n\in\mathbb{N}^*} \left]-\frac{1}{n}, \frac{1}{n}\right[= \{0\}.$$

$\{0\}$ est bien-sûr fermé. Cela montre que l'intersection dénombrable d'une famille d'ouverts n'est pas nécessairement un ouvert.

1.2.4 Intérieur et Adhérence d'un ensemble

La notion de voisinage est essentielle. Dans le chapitre qui suit, nous l'utiliserons abondamment lorsque nous nous intéresserons, par exemple, aux calculs de développement limité.

Définition 1.21. *Voisinage.*
Soit $(E, \|.\|)$, un espace vectoriel normé, et $a \in E$. On dit que $\mathcal{V}$ est un voisinage de a s'il existe $r > 0$ tel que $\mathcal{B}(a, r) \subset \mathcal{V}$.

Exemple : $[1, 2]$ est un voisinage de $\frac{3}{2}$. En revanche, $[1, 2]$ n'est pas un voisinage de 1.

Propriété 1.6. *Soit $(E, \|.\|)$, un espace vectoriel normé, et $a \in E$.*
(i) Si A est un voisinage de a et si B est un autre voisinage de a, alors $A \cap B$ est encore un voisinage de a.
(ii) Si A est un voisinage de a et si $A \subset B$, avec $B \subset E$, alors B est un voisinage de a.
(iii) ***Caractérisation des ouverts :*** *A est ouvert si, et seulement si A est voisinage de chacun de ses points.*

Nous allons à présent définir les notions d'intérieur et d'adhérence.

Définition 1.22. *Intérieur d'un ensemble.*
Soit $(E, \|.\|)$, un espace vectoriel normé, et $a \in E$. On dit que a est intérieur à $A \subset E$ si, et seulement si A est voisinage de a, ou encore, a est intérieur à A si, et seulement s'il existe $r > 0$ tel que $\mathcal{B}(a, r) \subset A$.
L'intérieur de A, noté $\mathring{A}$ est l'ensemble des points intérieurs.

Les propriétés suivantes caractérisent l'intérieur de A.

Propriété 1.7. *Soit A, un sous-ensemble d'un espace vectoriel normé E.*
(i) $\mathring{A}$ est un ouvert.
(ii) $\mathring{A}$ est le plus grand ouvert inclus dans A.
(iii) A est ouvert si, et seulement si $A = \mathring{A}$.

Démonstration.
(i) Si $x \in \mathring{A}$, il existe $r > 0$ tel que $\mathcal{B}(x, r) \subset A$. Mais $\mathcal{B}(x, r)$ est un ouvert. Donc $\mathcal{B}(x, r)$ est voisinage de chacun de ses points, donc A est un voisinage de chacun des points de $\mathcal{B}(x, r)$, d'où $\mathcal{B}(x, r) \subset \mathring{A}$, puis $\mathring{A}$ est ouvert.
(ii) $\mathring{A} \subset A$. Si B est un ouvert inclus dans A, soit $x \in B$. B est donc voisinage de x donc A aussi, car $B \subset A$. Par conséquent, $x \in \mathring{A}$, puis $B \subset \mathring{A}$.

(iii) C'est une conséquence directe des points (i) et (ii).

$\square$

Les propriétés que je cite à présent sont aisées à établir. Je laisse au lecteur le soin de s'en convaincre :

Propriété 1.8. *Soient A et B, deux sous ensembles d'un espace vectoriel normé E. Alors :*

$$\overset{\circ}{\overline{A \cap B}} = \mathring{A} \cap \mathring{B} \text{ et } A \subset B \Longrightarrow \mathring{A} \subset \mathring{B}.$$

Passons à présent à l'adhérence d'un ensemble dans un espace vectoriel normé.

Définition 1.23. *Soit $(E, \|.\|)$, un espace vectoriel normé, A, un sous-ensemble de E, $a \in A$. Il existe deux caractérisations équivalentes de la notion d'adhérence :*
*(i) a est dit **adhérent** à A si, et seulement s'il existe un voisinage $\mathcal{V}(a)$ de a tel que, quel que soit $v \in \mathcal{V}(a)$, on a : $\mathcal{V}(a) \cap A \neq \emptyset$.*
*(ii) a est dit **adhérent** à A si, et seulement si quel que soit $r > 0$ tel que $\mathcal{B}(a, r) \cap A \neq \emptyset$.*

On définit donc assez naturellement la notion d'adhérence d'un sous-ensemble d'un espace vectoriel normé, exactement de la même façon que nous avions défini l'intérieur d'un sous-ensemble d'un espace vectoriel normé.

Définition 1.24. ***L'adhérence** de A, notée $\overline{A}$, est l'ensemble des points adhérents de A.*

De même que pour la notion d'intérieur, l'adhérence possède des propriétés caractéristiques :

Propriété 1.9. *Soit $(E, \|.\|)$, un espace vectoriel normé, A, un sous-ensemble de E, $a \in A$.*
(i) $\overline{A}$ est un fermé qui contient A.
(ii) $\overline{A}$ est le plus petit fermé contenant A.

Démonstration. Soit $\mathcal{F} := \{F,\ F \text{ fermé tel que } A \subset F\}$, et $\mathcal{I} := \bigcap_{F \in \mathcal{F}} F$. $\mathcal{I}$ est un fermé, car il se définit comme intersection de fermés. De plus, $\mathcal{I}$ contient A. $\mathcal{I}$ est donc le plus petit fermé contenant A. Reste à montrer que $\mathcal{I} = \overline{A}$. Souvenons-nous de la caractérisation de tout élément $a \in \overline{A}$:

$$a \notin \overline{A} \Longleftrightarrow \exists r > 0 : \mathcal{B}(a, r) \cap A = \emptyset.$$

Cette dernière assertion est encore équivalente à l'assertion suivante :

$$\exists F := E \backslash \mathcal{B}(a, r) \text{ (fermé, car complémentaire d'un ouvert...), tel que : } A \subset F.$$

Or, d'après la définition de $\mathcal{I}$ comme intersection d'ensembles, puisque nous sommes parvenus à exhiber un fermé qui contient A, on peut encore en déduire que l'assertion ci-dessus en encore équivalente à l'assertion : $a \notin \mathcal{I}$.

Conclusion : $\overline{A} = \mathcal{I}$. Toutes les propriétés que l'on souhaitait démontrer s'en déduisent donc très simplement. $\square$

Les propriétés qui suivent constituent un bon entraînement pour le lecteur désireux de se tester sur cette notion. Elles sont assez simples à démontrer mais nécessitent une bonne compréhension de la notion d'adhérence.

Propriété 1.10. *Soit $(E, \|.\|)$, un espace vectoriel normé, A et B, deux sous-ensembles de E.*
(i) $A \subset B \Longrightarrow \overline{A} \subset \overline{B}$.
(ii) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.
(iii) A *fermé* $\Longleftrightarrow$ $A = \overline{A}$.

Enfin, la propriété qui suit fait le lien entre les boules ouvertes et les boules fermées.

Propriété 1.11. *Soit $(E, \|.\|)$, un espace vectoriel normé, et $a \in E$. Soit r, un nombre réel strictement positif. Alors :*
(i) $\overline{\mathcal{B}(a,r)} = \mathcal{B}_f(a,r)$.
(ii) $\overset{\circ}{\overline{\mathcal{B}_f(a,r)}} = \mathcal{B}(a,r)$.

Démonstration.
(i) On sait déjà que $\mathcal{B}(a,r) \subset \mathcal{B}_f(a,r)$. Par conséquent, $\overline{\mathcal{B}(a,r)} \subset \overline{\mathcal{B}_f(a,r)} = \mathcal{B}_f(a,r)$. Il reste donc à montrer que $\mathcal{B}_f(a,r) \subset \overline{\mathcal{B}(a,r)}$.
- **Premier cas :** si $x \in \mathcal{B}(a,r)$, alors $x \in \overline{\mathcal{B}(a,r)}$.

  Avant de poursuivre, définissons la notion de **sphère de centre a et rayon $r > 0$.**

  Définition 1.25. *Soit $(E, \|.\|)$, un espace vectoriel normé, et $a \in E$. On appelle* ***sphère*** *de centre a et rayon r l'ensemble :*

$$\mathcal{S}(a,r) := \{x \in E : \|x - a\| = r\}.$$

- **Deuxième cas :** si $x \in \mathcal{S}(a,r)$.
  Soit $\rho > 0$. On veut démontrer que $x \in \mathcal{B}_f(a,r) \Longrightarrow x \in \overline{\mathcal{B}(a,r)}$, autrement dit que $\mathcal{B}(x,\rho) \cap \mathcal{B}(a,r) \neq \emptyset$. Soit $\lambda \in]0,1[$ et $y_\lambda := \lambda a + (1-\lambda)x$. On va vérifier que y_λ est candidat pour appartenir à $\mathcal{B}(x,\rho) \cap \mathcal{B}(a,r)$.
  On a d'une part : $\|y_\lambda - a\| = (1-\lambda)\|x - a\| = (1-\lambda)r < r$, et donc $y_\lambda \in \mathcal{B}(a,r)$. D'autre part : $\|y_\lambda - x\| = \lambda\|a - x\| = \lambda r$. Or si l'on choisit λ suffisamment petit (i.e. positif, proche de 0), tel que $\lambda r < \rho$, ce qui est toujours possible, alors il est clair que $y_\lambda \in \mathcal{B}(x,\rho)$. Cela clôt la démonstration.

(ii) La seconde assertion se démontre exactement sur le même principe que la première. Elle est laissée aux soins du lecteur.

□

Définition 1.26. *Frontière d'un ensemble.*
Soit $(E, \|.\|)$, un espace vectoriel normé et X, un sous espace de E. On appelle ***frontière*** *de X et on note $Fr(X)$, l'ensemble constitué des éléments appartenant à l'adhérence de X privée de l'intérieur de X, c'est à dire :*

$$Fr(X) = \overline{X} \backslash \overset{\circ}{X}.$$

Propriété 1.12. *$Fr(X)$ est une partie fermée de E.*

Remarque : cette propriété sera démontrée en exercice, ainsi que d'autres propriété caractéristiques de la frontière d'un ensemble.

Cette notion se visualise aisément, géométriquement.

Chapitre 2

Suites et continuité dans un espace vectoriel normé

2.1 Convergence et continuité dans un espace vectoriel normé

2.1.1 Suites et convergence dans un espace vectoriel normé

À la fin de cette sous-partie, je fournirai, à titre d'exemple, une application des résultats énoncés ci-dessous aux suites à valeurs dans $\mathbb{R}$.

Définition 2.1. *Convergence dans un espace vectoriel normé.*
Soit $(E, \|.\|)$, un espace vectoriel normé, et $u = (u_n)_{n\in\mathbb{N}}$, une suite de $(E, \|.\|)$.
On dit que u a pour limite $\ell \in E$ dans $(E, \|.\|)$ si, et seulement si $\lim\limits_{n\to+\infty} \|u_n - \ell\| = 0$. cela s'écrit encore :

$$\forall\varepsilon > 0,\ \exists N \in \mathbb{N} : n > N \Longrightarrow \|u_n - \ell\| \leq \varepsilon.$$

Remarque : quelques conséquences immédiates découlent de cette définition.

- Si une suite u converge dans $(E, \|.\|)$, alors sa limite est unique.
- Toute suite convergente de $(E, \|.\|)$ est bornée. On en fera la preuve en exercice.
- Soit φ, une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. Alors, si u est une suite convergente dans $(E, \|.\|)$, la suite v dont le terme général est donné pour tout $n \in \mathbb{N}$ par $v_n = u_{\varphi(n)}$ est convergente de même limite que u. Ce résultat s'énonce encore de la façon suivante : toute suite extraite d'une suite convergente dans un espace vectoriel normé est convergente dans cet espace de même limite.
- La notion de convergence dans un espace vectoriel normé est donc étroitement liée au choix de la norme sur cet espace. La propriété suivante montre qu'il est possible de généraliser la notion de convergence de suite dans un espace vectoriel normé, dans le cas où l'on choisit deux normes équivalentes sur cet espace.

Propriété 2.1. *Soit E, un espace vectoriel, et N_1 et N_2, deux normes sur E.*
N_1 et N_2 sont équivalentes si et seulement si toute suite convergeant vers 0_E pour une norme converge vers 0_E pour l'autre norme.

Démonstration. • **Sens $\Longrightarrow$:** c'est trivial, en revenant à la notion de normes équivalentes. De plus ce résultat peut être généralisé : si N_1 et N_2 sont équivalentes, toute suite u convergeant dans E au sens de N_1 converge également dans E (vers la même limite) au sens de N_2.

- **Sens $\Longleftarrow$:** on raisonne par l'absurde. Supposons donc que $\frac{N_2}{N_1}$ soit non-bornée et que toute suite convergeant vers 0_E pour une norme converge vers 0_E pour l'autre norme. Dire que $\frac{N_2}{N_1}$ est non bornée s'obtient en niant l'assertion « $\frac{N_1}{N_2}$ est bornée ». $\frac{N_2}{N_1}$ est bornée s'écrirait :

$$\exists M > 0 : \forall x \in E \backslash \{0_E\},\ \frac{N_2(x)}{N_1(x)} \leq M.$$

La négation de cette proposition s'écrit donc :

$$\forall M > 0, \exists x \in E \backslash \{0_E\} :\ \frac{N_2(x)}{N_1(x)} \geq M.$$

En choisissant successivement $M = 1$, $M = 2$, ..., etc., il est donc possible de construire une suite d'éléments de E, notée $(x_n)_{n\in\mathbb{N}}$ telle que :

$$\forall n \in \mathbb{N},\ \frac{N_2(x_n)}{N_1(x_n)} > n.$$

Définissons alors la suite $(y_n)_{n\in\mathbb{N}}$ par la relation : $y_n = \frac{x_n}{\sqrt{n}N_1(x_n)}$. Il est clair que : $N_1(y_n) = \frac{1}{\sqrt{n}} \xrightarrow[n\to+\infty]{} 0$ et $N_2(y_n) > nN_1(y_n) = \sqrt{n} \xrightarrow[n\to+\infty]{} +\infty$, ce qui est absurde. La proposition s'ensuit.

□

Dans le cas d'un espace vectoriel normé s'écrivant sous la forme d'un produit d'espaces vectoriels normés, la notion de continuité se traite de façon similaire.

Soient donc $p \in \mathbb{N}$ et E_1, ..., E_p, des espaces vectoriels normés, respectivement numis de normes notées N_1, ..., N_p. Posons $E := E_1 \times ... \times E_p = \prod_{k=1}^{p} E_k$. Posons, pour tout $x = (x_1, ..., x_p) \in E$:

$$\|x\| = \sup_{i\in\{1,...,p\}} (N_i(x_i)).$$

On a la caractérisation suivante :

Théorème 2.1. *Sous les hypothèses rappelées précédemment, et si on appelle $(x^n)_{n\in\mathbb{N}}$, une suite de E notée $(x_1^n, ..., x_p^n)$, où l'on a bien-sûr $x_i^n \in E_i$. Alors :*

$$x^n \xrightarrow[n\to+\infty]{\text{dans } E} L := (L_1, ..., L_p) \in E \Longleftrightarrow \forall i \in \{1, ..., p\},\ x_i^n \xrightarrow[n\to+\infty]{\text{dans } E_i} L_i.$$

Remarque : dans ce cas, on peut également écrire :

$$\lim_{n\to+\infty}(x^n) = \left(\lim_{n\to+\infty}(x_1^n), ..., \lim_{n\to+\infty}(x_p^n)\right).$$

Exemple : dans la seconde partie de cet ouvrage, nous détaillerons en particulier le cas où $\forall i \in \{1, ..., p\}$, on a : $E_i = \mathbb{R}$, muni de la norme $|.|$. La caractérisation précédente s'écrit alors :

$$x^n := (x_1^n, ..., x_p^n) \xrightarrow[n \to +\infty]{\text{dans } \mathbb{R}^p} L := (L_1, ..., L_p) \in \mathbb{R}^p \Longleftrightarrow \forall i \in \{1, ..., p\},\ |x_i^n - L_i| \xrightarrow[n \to +\infty]{\text{dans } \mathbb{R}} 0.$$

Rappelons à présent proprement ce que l'on appelle suite extraite ou sous-suite d'une suite :

Définition 2.2. *Suite extraite.*
Soit u, une suite dans un espace vectoriel normé $(E, \|.\|)$. Soit $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$, une application strictement croissante. Appelons v, la suite définie pour tout $n \in \mathbb{N}$ par :

$$\forall n \in \mathbb{N},\ v_n = u_{\varphi(n)}.$$

*On dit que v est une **suite extraite** ou **sous-suite** de u.*
*De plus, si v converge dans E vers ℓ, alors on dit que ℓ est une **valeur d'adhérence** de u. (voir ci-après)*

Remarque : souvenons-nous que dans $\mathbb{R}$ les suites extraites possèdent des propriétés bien particulières. Par exemple :

- **Propriété de Bolzano-Weierstrass :** de toute suite bornée de réels, on peut extraire une sous-suite convergente.
- **Caractérisation des suites convergentes :** une suite u est dite convergente si, et seulement si toutes les suites extraites de u convergent et ont même limite.
- etc.

Dans la seconde partie de ce chapitre, nous présenterons la notion de compacité qui permettra de généraliser à un espace vectoriel normé quelconque les propriétés ci-dessus.

Comme pour le cas des suites réels, il est possible de caractériser précisément les valeurs d'adhérence d'une suite définie dans un espace vectoriel normé E.

Théorème 2.2. *Soit $(E, \|.\|)$, un espace vectoriel normé. Les trois propositions suivantes sont équivalentes :*
(i) $\exists a \in E : \forall \varepsilon > 0, \forall N \in \mathbb{N},\ \exists n > N : \|x_n - a\| < \varepsilon$.
(ii) $\forall \varepsilon > 0$, *$\mathcal{B}(a, \varepsilon)$ contient x_n pour une infinité d'indices.*
(iii) a est limite d'une suite extraite de $(x_n)_{n \in \mathbb{N}}$.

Remarque : la notion de suite extraite est fortement utilisée pour démontrer qu'une suite n'est pas convergente. En effet, lorsque l'on souhaite montrer qu'une suite n'est pas convergente, une méthode fréquemment utilisée consiste à extraire deux sous-suites de la suite initiale, dont les limites sont différentes. C'est l'exemple de la suite $(u_n)_{n \in \mathbb{N}}$ de réel, définie pour tout entier naturel n par : $u_n = (-1)^n$. Il suffit de considérer la sous-suite de u extraite en ne choisissant que les éléments de rang pair, puis la suite extraite construire en ne considérant que les éléments de rang impair.

Définition 2.3. *Valeur d'adhérence.*
Soit $(u_n)_{n \in \mathbb{N}}$, une suite d'éléments d'un espace vectoriel normé (E, N). On dit que $a \in E$ est une valeur d'adhérence de la suite $(u_n)_{n \in \mathbb{N}}$ si a s'écrit comme la limite d'une suite extraite de $(u_n)_{n \in \mathbb{N}}$, autrement dit s'il existe une application strictement croissante $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ telle que $u_{\varphi(n)} \xrightarrow[n \to +\infty]{\text{dans } E} a$.

Remarque : le théorème 2.2 présente différentes caractérisations de la notion de valeur d'adhérence.

Enfin, les suites dans les espaces vectoriels normés peuvent être utilisées pour caractériser un fermé. La propriété qui va suivre, appelée « caractérisation séquentielle des fermés » s'avérera donc particulièrement intéressante. Je présente au préalable une propriété caractérisant l'adhérence d'un ensemble dans un espace vectoriel normé.

Propriété 2.2. *Caractérisation séquentielle de l'adhérence.*
Soit $(E, \|.\|)$, un espace vectoriel normé, et A, un sous-ensemble de E. Alors, les propositions suivantes sont équivalentes :
(i) $a \in \overline{A}$.
(ii) *a est limite d'une suite d'éléments de A.*

Démonstration. (i) **Sens $\Longrightarrow$:** $a \in \overline{A}$ revient exactement à affirer que si on se fixe $\varepsilon > 0$, alors $\mathcal{B}(a, \varepsilon) \cap A \neq \emptyset$. Choisissons pour tous les n entier naturels $\varepsilon = \frac{1}{n}$. On construit donc une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de A tels que $\forall n \in \mathbb{N}$, $x_n \in \mathcal{B}(a, \frac{1}{n}) \cap A$ (qui est non nul d'après ce que l'on vient de voir). $(x_n)_{n\in\mathbb{N}}$ est donc une suite d'éléments de A telle que : $\forall n \in \mathbb{N}$, $\|x_n - a\| < \dfrac{1}{n}$, autrement dit, $a = \lim\limits_{n\to+\infty}(x_n)$, au sens de la norme de E.

(ii) **Sens $\Longleftarrow$:** supposons que $a \in E$ soit limite d'une suite d'éléments de A, i.e. $a = \lim\limits_{n\to+\infty}(x_n)$, au sens de la norme de E, avec $\forall n \in \mathbb{N}$, $x_n \in A$. Soit $\varepsilon > 0$ fixé. Alors il est clair qu'il existe un certain rang $N \in \mathbb{N}$ à partir duquel $\mathcal{B}(a, \varepsilon)$ contient tous les termes de la suite, c'est à dire les $(x_n)_{n\geq N}$, autrement dit : $\|x_n - a\| < \varepsilon$, autrement dit $\mathcal{B}(a, \varepsilon) \cap A \neq \emptyset$, pour tout $\varepsilon > 0$ donné. Cela signifie exactement que $a \in \overline{A}$.

□

Propriété 2.3. *Caractérisation séquentielle des fermés.*
Soit $(E, \|.\|)$, un espace vectoriel normé, et A, un sous-ensemble de E. Alors, les propositions suivantes sont équivalentes :
(i) *A est fermé dans E.*
(ii) *Toute suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers $x \in E$ implique que $x \in A$.*

Démonstration. (i) **Sens $\Longrightarrow$:** supposons que A est fermé. Soit $(a_n)_{n\in\mathbb{N}}$ une suite d'éléments de A qui converge vers $x \in E$. La caractérisation séquentielle de l'adhérence nous montre que l'on peut en déduire que $x \in \overline{A}$. or, $\overline{A} = A$. Par conséquent, $x \in A$.

(ii) **Sens $\Longleftarrow$:** supposons que toute suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers $x \in E$ implique que $x \in A$. On va démontrer que A est fermé en prouvant que $\overline{A} \subset A$. Soit $x \in \overline{A}$. Alors, d'après la propriété précédente, x est limite d'une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A. Par hypothèse, on en déduit que $x \in A$. C'est ce que l'on voulait démontrer.

□

2.1.2 Notion de densité dans un espace vectoriel normé

Le concept de densité dans un espace vectoriel normé, bien qu'un peu théorique, conduit à des résultats des plus intéressants. Nous insisterons particulièrement sur le théorème de Stone Weierstrass.

Définition 2.4. *Sous-espace dense.*
Soit A, un espace muni d'une norme N, et X, une partie de A. On dit que X est **dense** *dans A si, et seulement si $\overline{X} = A$.*

Il existe d'autres caractérisations de la densité. Par exemple :

Propriété 2.4. *Soit (E, N), un espace vectoriel normé, et A, une partie de E. On dit que X est dense dans A si, et seulement si tout élément x de A est limite d'une suite d'éléments de X, autrement dit, en notant $X^{\mathbb{N}}$, l'ensemble des suites à valeurs dans X :*

$$\exists (x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}} : x_n \xrightarrow[n\to+\infty]{\textit{dans } E} x.$$

Cette propriété découle directement de la caractérisation séquentielle de l'adhérence, dans la propriété 2.2. Elle permet notamment de démontrer que :

Propriété 2.5.
(i) $\mathbb{Q}$ est dense dans $\mathbb{R}$.
(ii) $\mathbb{R}\backslash\mathbb{Q}$ est dense dans $\mathbb{R}$.

Démonstration.
(i) Soit $x \in \mathbb{R}$. Appelons $(u_n)_{n\in\mathbb{N}}$, la suite définie par :

$$u_n = \frac{1}{10^n}.E\left(10^n x\right).$$

La fonction E désigne ici la fonction partie entière. Il est tout à fait clair que $u_n \in \mathbb{Q}$. De plus, en utilisant la propriété $\forall x \in \mathbb{R}$, $x - 1 < E(x) \leq x$, on montre que quel que soit $n \in \mathbb{N}$, on a :

$$\frac{1}{10^n}.(10^n x - 1) < \frac{1}{10^n}.E\left(10^n x\right) \leq \frac{1}{10^n}.10^n x.$$

Par théorème d'encadrement, on montre immédiatement que :

$$\lim_{n\to+\infty} \frac{1}{10^n}.(10^n x - 1) = \lim_{n\to+\infty} \frac{1}{10^n}.10^n x = x$$

, donc tout élément $x \in \mathbb{R}$ est limite d'une suite d'éléments de $\mathbb{Q}$.
On en déduit que $\mathbb{Q}$ est dense dans $\mathbb{R}$.

(ii) Le principe de la démonstration pour $\mathbb{R}\backslash\mathbb{Q}$ est exactement le même que précédemment. En effet, on considère ici la suite $(v_n)_{n\in\mathbb{N}}$ définie pour tout entier naturel n non nul par :

$$v_n = u_n\left(1 + \frac{\sqrt{2}}{n}\right).$$

Le théorème relatif à la limite d'un produit nous permet d'affirmer que $\lim\limits_{n\to+\infty}(v_n) = x$. De plus $\forall n \in \mathbb{N}^*,\ v_n \in \mathbb{R}\backslash\mathbb{Q}$. En effet, si l'on supposait le contraire, on obtiendrait l'existence de deux suites d'entiers non nuls $(p_n)_{n\in\mathbb{N}^*}$ et $(q_n)_{n\in\mathbb{N}^*}$ tels que $\dfrac{p_n}{q_n} = v_n$. On en déduirait donc immédiatement que $\sqrt{2} \in \mathbb{Q}$, ce qui est absurde. Par conséquent, $\forall n \in \mathbb{N}^*,\ v_n \in \mathbb{R}\backslash\mathbb{Q}$.

Remarque : la notion de norme n'apparaît pas clairement dans les démonstrations ci-dessus.

Pourtant, la norme sur $\mathbb{R}$ est la valeur absolue qui intervient implicitement lorsque l'on écrit ci-dessus des limites. Par exemple, dire que $\lim\limits_{n\to+\infty} \frac{1}{10^n}.10^n x = x$ signifie en fait que :

$$\lim_{n\to+\infty} \left| \frac{1}{10^n}.10^n x - x \right| = 0.$$

□

Prenons un exemple concret pour nous en convaincre.

Exemple : la constante de Neper, notée e est définie par la relation :

$$e = \sum_{k=0}^{+\infty} \frac{1}{k!}.$$

La série associée a cette somme est convergente en vertu du théorème de d'Alembert, par exemple. Posons $\forall n \in \mathbb{N},\ u_n = \sum\limits_{k=0}^{n} \frac{1}{k!}$. Puisque $(\mathbb{Q}, +)$ est un groupe, chaque terme u_n est dans $\mathbb{Q}$.
De plus, $e \notin \mathbb{Q}$. La preuve est rapide quoique un peu technique. Je la fais figurer pour information. On raisonne par l'absurde en supposant l'existence de deux nombres entiers strictement positifs p et q tels que : $e = \frac{p}{q}$. On définit ensuite le réel X par la relation :

$$X := q! \left(e - \sum_{k=0}^{q} \frac{1}{k!} \right) = q! \sum_{k=q+1}^{+\infty} \frac{1}{k!}.$$

De plus, $X = p(q-1)! - \sum\limits_{k=0}^{q} \frac{q!}{k!}$, et $k!$ divise $q!$, car $k \in \{0, ..., q\}$. Il s'ensuit que $X \in \mathbb{Z}$. Plus précisément, $X \in \mathbb{N}$, d'après l'expression de X :

$$X = q! \sum_{k=q+1}^{+\infty} \frac{1}{k!}.$$

Toujours d'après cette expression, on peut encore écrire :

$$X = \sum_{n=1}^{+\infty} \frac{1}{(q+1)...(q+n)} \leq \sum_{n=1}^{+\infty} \frac{1}{(q+1)^n} = \frac{1}{q} < 1.$$

On a reconnu la somme d'une série géométrique.

Récapitulons : $X \in \mathbb{N}^*$ et $X \in]0, 1[$. C'est absurde. Par conséquent, e est irrationnel et s'écrit effectivement comme la limite d'une suite d'éléments de $\mathbb{Q}$.

Énonçons à présent une autre propriété qui permet de caractériser un sous espace dense :

Propriété 2.6. *Soit A, un sous-espace d'un espace vectoriel normé (E, N). Une partie X de A est dense dans A si, et seulement si l'intersection avec tout ouvert non vide de A est non vide.*

Démonstration. Il s'agit là d'une conséquence quasi-directe de la définition de la notion de densité. En effet, X est dense dans A si, et seulement si $\overline{X} = A$, autrement dit : $\overset{\circ}{\overline{A \backslash X}} = \emptyset$. Donc, aucun ouvert non vide de A n'est inclus dans $A \backslash X$ ou encore tout ouvert non vide de A possède au moins un élément qui est dans X. C'est ce que l'on souhaitait démontrer. □

Application : on sait que dans $\mathbb{R}$, muni de la norme $|.|$, les ouverts sont les réunions d'intervalles ouverts du type $]a, b[$. Puisque $\mathbb{Q}$ est dense dans $\mathbb{R}$, d'après la propriété que l'on vient d'énoncer, on est sûr que quels que soient les réels a et b tels que $a < b$, il existe toujours $q \in \mathbb{Q}$ tel que $a < q < b$ et il s'agit là d'une autre caractérisation de la densité de $\mathbb{Q}$ dans $\mathbb{R}$.
Cela peut donner lieu à une nouvelle démonstration de la densité de $\mathbb{Q}$ dans $\mathbb{R}$.

Nous retrouverons ultérieurement la notion de densité à travers le théorème de Stone-Weierstrass et ses diverses applications. Cela sera traité dans la partie consacrée à l'étude de l'espace $\mathcal{C}(X, \mathbb{R})$, des fonctions continues sur un espace métrique X que l'on supposera **compact** (sections suivantes) et à valeurs réelles.

2.1.3 Limite et continuité dans un espace vectoriel normé

Définition 2.5. *Soient (E, N) et $(F, \|.\|)$, deux espaces vectoriels normés et $f : (E, N) \longrightarrow (F, \|.\|)$. Soit A, un sous-ensemble de E et $x_0 \in \overline{A}$. On dit que f a pour limite L quand x tend vers x_0, selon A, si :*

$$\forall \varepsilon > 0,\ \exists \alpha > 0 : N(x - x_0) < \alpha \text{ et } x \in A \Longrightarrow \|f(x) - L\| < \varepsilon.$$

On écrira alors : $\lim\limits_{\substack{x\to x_0\\ x\in A}} f(x) = L$ *ou* $f(x) \xrightarrow[\substack{x\to x_0\\ x\in A}]{} L$.

Cette notion de limite reste inchangée si l'on utilise des normes équivalentes.

Exemple : considérons la fonction de plusieurs variables :

$$\begin{array}{rccl} f: & (\mathbb{R}^2, \|.\|_2) & \longrightarrow & (\mathbb{R}, |.|) \\ & (x, y) & \longmapsto & \begin{cases} xy & \text{si } (x, y) \neq (1, 1) \\ 0 & \text{si } (x, y) = (1, 1). \end{cases} \end{array}$$

$\|.\|_2$ désigne la norme euclidienne de $\mathbb{R}^2$. On note $A = \mathbb{R}^2 \backslash \{(1, 1)\}$. On a d'une part $\lim\limits_{\substack{(x,y)\to(1,1)\\ (x,y)\in A}} f(x) = 1$ et d'autre part $\lim\limits_{\substack{(x,y)\to(1,1)\\ (x,y)\in\mathbb{R}^2}} f(x)$ n'existe pas. On dira en effet que f est discontinue en $(1, 1)$.
Nous allons y revenir.

Définition 2.6. *Continuité d'une application dans un e.v.n.*
Soient (E, N) et $(F, \|.\|)$, deux espaces vectoriels normés et $f : (E, N) \longrightarrow (F, \|.\|)$.
(i) f est continue en $x_0 \in E$ si, et seulement si :

$$\forall \varepsilon > 0,\ \exists \alpha > 0 : N(x - x_0) < \alpha \Longrightarrow \|f(x) - f(x_0)\| < \varepsilon.$$

(ii) On dit que f est continue sur $X \subset E$ si, et seulement si f est continue en tout point de X.

Remarque : f est continue en x_0 peut s'écrire de façon plus synthétique : $\lim\limits_{x\to x_0} f(x) = f(x_0)$.

Exemple : continuité d'une fonction de plusieurs variables.
On définit la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ par :

$$f(x, y) := \begin{cases} \dfrac{\sin(x + y)}{x + y} & \text{si } x \neq -y \\ 1 & \text{si } x = -y \end{cases}$$

f est continue sur $\mathbb{R}^2$. En effet, si $x \neq y$, cela ne pose aucun problème, puisque f s'écrit comme un quotient de fonctions continues dont le dénominateur ne s'annule pas. De plus, si on étudie la continuité de f au voisinage des points de coordonnées $(x_0, -x_0)$, avec x_0 réel, alors, en posant $u = x + y$ et en remarquant que $\lim\limits_{(x,y)\to(0,0)} u = 0$, puisque $\lim\limits_{u\to 0} \dfrac{\sin u}{u} = 1$, on a : $\lim\limits_{(x,y)\to(x_0,-x_0)} f(x,y) = 1 = f(x_0, -x_0)$, et f est continue.

La propriété qui suit s'intitule « caractérisation séquentielle de la continuité ». En pratique, cette propriété nous servira surtout à prouver qu'une fonction est discontinue.

Propriété 2.7. *Une fonction $f : (E, N) \longrightarrow (F, \|.\|)$ est continue en un point a élément d'un ensemble $X \subset E$ si, et seulement si, pour toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de l'espace X convergeant vers $a \in X$, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$.*

Démonstration. On va démontrer les deux sens de l'équivalence :

- **Sens $\Longrightarrow$:** soit $(x_n)_{n\in\mathbb{N}}$, une suite d'éléments de l'espace X convergeant vers $a \in X$. Puisque f est continue en a, alors, si l'on se fixe $\varepsilon > 0$, on a :

$$\exists \eta > 0 : N(x - a) < \eta \Longrightarrow \|f(x) - f(a)\| < \varepsilon.$$

Pour le nombre $\eta > 0$ donné, on a encore, puisque $(x_n)_{n\in\mathbb{N}}$ est convergente :

$$\exists N \in \mathbb{N} : n \geq N \Longrightarrow \|x_n - a\| < \eta.$$

Il est alors parfaitement clair, en combinant les deux inégalités précédentes que :

$$\exists N \in \mathbb{N} : n \geq N \Longrightarrow \|f(x_n) - f(a)\| < \varepsilon.$$

Puisque ε a été choisi de façon arbitraire, on en déduit que la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$.

- **Sens $\Longleftarrow$:** pour démontrer cette implication, on va raisonner par contra-posée. Supposons donc que f n'est pas continue en $a \in X$. Alors, il existe un certain $\varepsilon > 0$ tel que l'on puisse construire une suite d'éléments de X, notée $(x_n)_{n\in\mathbb{N}}$ vérifiant :

$$\forall n \in \mathbb{N}, N(x_n - a) < \frac{1}{n} \text{ et } \|f(x) - f(a)\| > \varepsilon.$$

On a donc montré l'existence d'une suite $(f(x_n))_{n\in\mathbb{N}}$ ne convergeant pas vers $f(a)$, alors que $(x_n)_{n\in\mathbb{N}}$ converge vers $a \in X$.

□

Poursuivons en donnant à présent une autre caractérisation de la continuité, à l'aide des ouverts et des fermés, autrement dit, de la Topologie d'un espace. L'intérêt de cette caractérisation réside dans le fait qu'elle peut être généralisée à des espaces sur lesquels aucune métrique n'a été définie, c'est à dire des espaces non munis d'une distance ou d'une norme, mais ça n'est pas l'objet de ce cours, puisqu'il est consacré aux espaces vectoriels munis d'une norme.

Rappelons au préalable la notion d'image réciproque d'un ensemble par une application.

Définition 2.7. *Image réciproque.*
Soient X et Y, deux ensembles, et $f : X \longrightarrow Y$, une application. Soit $A \subset Y$. On définit ***l'image réciproque de** A **par*** f *et on note $f^{-1}(A)$, l'ensemble :*

$$f^{-1}(A) = \{x \in X : f(x) \in A\} \subset X.$$

Remarque 1 : attention ! Il n'est absolument pas nécessaire que f soit bijective pour définir la notion d'image réciproque d'un ensemble par f. La notation est trompeuse. L'ensemble $f^{-1}(A)$ désigne ici le sous-ensemble de X constitué des éléments dont l'image par f appartient à A.

Remarque 2 : si f est bijective, on définit l'application réciproque de f que l'on note f^{-1}, par la relation : $f^{-1} \circ f = Id_X$ et $f \circ f^{-1} = Id_Y$. Dans ce cas, $f^{-1}(A)$ désigne à la fois l'image directe de A par l'application f^{-1} et l'image réciproque de A par l'application f.

Propriété 2.8. *Soient X et Y, deux espaces inclus dans des espaces vectoriels (E, N) et $(F, \|.\|)$, respectivement munis des normes N et $\|.\|$. Soit $f : X \longrightarrow Y$, une application. Alors, les propositions suivantes sont équivalentes :*
(i) f est continue sur X.
(ii) Pour tout ouvert $U \subset Y$, l'ensemble $f^{-1}(U)$ est un ouvert de X.
(iii) Pour tout fermé $V \subset Y$, l'ensemble $f^{-1}(V)$ est un fermé de X.

Démonstration. Cela se fait en deux étapes :

- **Démontrons que $(i) \Longleftrightarrow (ii)$:**
  - **Démontrons de prime abord que $(i) \Longrightarrow (ii)$:** supposons donc f continue et soit U un ouvert de Y. Soit $a \in f^{-1}(U)$. On a bien-sûr $f(a) \in U$, et puisque U est ouvert, il existe $\varepsilon > 0$ tel que $\mathcal{B}(f(a), \varepsilon) \subset U$. Traduisons à présent le fait que f est une fonction continue en a. Pour le ε trouvé précédemment :

  $$\exists \eta > 0 : N(x-a) < \eta \Longrightarrow \|f(x) - f(a)\| < \varepsilon.$$

  Par conséquent, l'ensemble $\mathcal{B}(a, \eta) \cap X$ est un ouvert (intersection de deux ouverts) inclus dans $f^{-1}(U)$, puisque l'image de tout point de cet ensemble est incluse dans $\mathcal{B}(f(a), \varepsilon) \subset U$. Il s'ensuit, par définition d'un ouvert, et puisque $a \in \mathcal{B}(a, \eta) \cap X$, qu'il existe une boule ouverte de centre a incluse dans $f^{-1}(U)$. C'est exactement ce que l'on souhaitait démontrer.
  - **Démontrons à présent que $(ii) \Longrightarrow (i)$:** supposons donc que l'image réciproque de tout ouvert par la fonction f est encore un ouvert. Soit alors $a \in X$. Alors, $f(a) \in Y$ et si ε est une constante positive fixée, on sait que $\mathcal{B}(f(a), \varepsilon)$ est un ouvert de Y donc $f^{-1}(\mathcal{B}(f(a), \varepsilon))$ est encore un ouvert de X, par hypothèse, et cet ouvert contient a. Donc, il existe $\eta > 0$ tel que $\mathcal{B}(a, \eta) \subset f^{-1}(\mathcal{B}(f(a), \varepsilon))$, par définition d'un ouvert. Par conséquent, on a démontré l'existence d'une constante $\eta > 0$ telle que :

  $$N(x-a) < \eta \Longrightarrow \|f(x) - f(a)\| < \varepsilon.$$

- Démontrons que $(ii) \Longleftrightarrow (iii)$: cette équivalence est immédiate, puisque, par définition, les fermés se définissent comme les complémentaires des ouverts. De plus, il est aisé de démontrer que, si A^c désigne le complémentaire d'un ensemble A donné dans un espace normé, alors : $f^{-1}(A^c) = \left(f^{-1}(A)\right)^c$. En effet, $f^{-1}(A^c) = \{x \in X : f(x) \in A^c\}$. Donc en particulier, si $a \in f^{-1}(A^c)$, alors $f(a) \notin A$ et donc $a \in \left(f^{-1}(A)\right)^c$. La réciproque se démontre exactement sur le même principe. Je laisse au lecteur le soin de s'en convaincre.

□

Remarque 1 : attention à ne pas confondre ! La caractérisation de la continuité sous cette forme n'a été établie qu'à l'aide des images réciproques d'ouverts ou de fermés par une application donnée. En revanche, une application peut très bien être continue, telle que l'image d'un ouvert ne soit pas un ouvert ou l'image d'un fermé ne soit pas un fermé. Par exemple, considérons la fonction f définie par :

$$\begin{array}{rcl} \mathbb{R} & \longrightarrow & [0,1[\\ x & \longmapsto & f(x) = \dfrac{|x|}{1+|x|}. \end{array}$$

Cette fonction est bien sûr continue comme produit de fonctions continues. En revanche, une étude de fonction très rapide (et le fait que f est paire) montrent que $f(\mathbb{R}) = [0,1[$, et $[0,1[$ n'est ni ouvert ni fermé.

Remarque 2 : on peut reformuler la caractérisation de la continuité à l'aide des ouverts de la façon suivante : f est continue si, et seulement si l'image réciproque d'une boule ouverte de centre $f(a)$ contient toujours une boule ouverte de centre a.

La proposition qui suit est un corollaire assez immédiat de la proposition que nous venons de voir. Elle se révêlera fort utile en pratique.

Propriété 2.9. *Soit f, une application continue sur un sous-espace X d'un espace vectoriel E, à valeurs réelles. Soit $\alpha \in \mathbb{R}$. Alors :*
(i) Les ensembles $\{x \in E : f(x) < \alpha\}$ et $\{x \in E : f(x) > \alpha\}$ sont ouverts dans E.
(ii) Les ensembles $\{x \in E : f(x) \leq \alpha\}$, $\{x \in E : f(x) \geq \alpha\}$ et $\{x \in E : f(x) = \alpha\}$ sont fermés dans E.

Preuve rapide : il suffit de remarquer par exemple que l'intervalle $]-\infty, \alpha[$ est un ouvert et de voir l'ensemble $\{x \in E : f(x) < \alpha\}$ comme l'image réciproque par la fonction f de $]-\infty, \alpha[$. En appliquant la caractérisation précédente de la continuité, c'est évident.

Enfin, pour conclure sur les notions de continuité, présentons rapidement le concept d'**homéomorphisme**.

Définition 2.8. *Homéomorphisme.*
Soient (E, N) et $(F, \|.\|)$, deux espaces vectoriels normés, X, une partie de E, Y, une partie de F, et f, une application de E dans F. On dit que f est un ***homéomorphisme*** *de X sur Y si, et seulement si, les trois propriétés suivantes sont vérifiées :*
(i) f est continue.
(ii) f réalise une bijection de X sur Y.
(iii) L'application réciproque de f, f^{-1} est continue.

Remarque : on rappelle que si f est une fonction bijective de X sur Y, alors, on définit sa bijection réciproque f^{-1} par la formule suivante :

$$\forall (x,y) \in X \times Y,\ y = f(x) \Longleftrightarrow x = f^{-1}(y).$$

Exemple : si f est une fonction continue sur un intervalle $I \subset \mathbb{R}$, à valeurs réelles et strictement monotone sur cet intervalle, alors f réalise une bijection de I sur $f(I)$, et de ce fait, f est un homéomorphisme de I sur $f(I)$.
Pour donner un exemple concret, on peut dire que la fonction logarithme népérien de base e est un homéomorphisme de $]0, +\infty[$ sur $\mathbb{R}$ tout entier.

Conséquences : quelques conséquences découlent de la définition d'un homéomorphisme. Par exemple, f possède toutes les propriétés d'une fonction continue, mais f^{-1} également. Par conséquent, l'image réciproque d'un ouvert par un homéomorphisme est un ouvert (caractérisation de la continuité), mais on peut également dire que l'image de tout ouvert par un homéomorphisme est encore un ouvert, puisque f^{-1} est continue. De telles considérations tiennent également avec les fermés.

2.1.4 Applications Lipschitziennes et uniforme continuité

Définition 2.9. *Soient (E, N) et $(F, \|.\|)$, deux espaces vectoriels normés, X, une partie de E, Y, une partie de F, et f, une application de E dans F. On dit que l'application f est Lipschitzienne de rapport $k > 0$ si, et seulement si :*

$$\forall (x, y) \in E^2, \|f(x) - f(y)\| \leq kN(x - y).$$

Conséquence immédiate : une application Lipschitzienne est continue. En effet, si $\varepsilon > 0$ est donné, on peut trouver, en réutilisant les notations de la définition de la continuité, $\eta = \frac{\varepsilon}{k}$, tel que l'inégalité traduisant la continuité de f soit vérifiée.

Exemple : soit (E, N), un espace vectoriel normé. Alors, N est une application 1-Lipschitzienne. En effet, on a toujours :

$$\forall (x, y) \in E^2, \ |N(x) - N(y)| \leq N(x - y).$$

En effet, il s'agit ici d'une conséquence de l'inégalité triangulaire : $\forall (x, y) \in E^2$, $N(x) = N(x - y + y) \leq N(x - y) + N(y)$...

Enfin, dans le cas particulier ou f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ dérivable, il existe une caractérisation très simple du caractère Lipschitzien d'une fonction.

Propriété 2.10. *Soit $f : I \longrightarrow \mathbb{R}$, dérivable sur $\mathring{I}$, où I désigne un intervalle de $\mathbb{R}$. Alors, f est Lipschitzienne si, et seulement si f' est bornée sur $\mathring{I}$.*

Preuve rapide : cette propriété provient directement de l'inégalité des accroissements finis. Je la redonne pour information.

Propriété 2.11. *Inégalité des accroissements finis.*
Soit I, un intervalle de $\mathbb{R}$ et $f : I \longrightarrow \mathbb{R}$, une fonction dérivable sur $\mathring{I}$. Si $\forall x \in \mathring{I}$, $\exists k > 0$: $|f'(x)| < k$, alors :

$$\forall (x, y) \in \mathring{I}^2, \ |f(x) - f(y)| < k|x - y|.$$

Cette propriété traduit donc exactement le fait que f est une fonction Lipschitzienne si sa dérivée est bornée.
L'équivalence s'obtient en étudiant la réciproque : si quels que soient x et y éléments de $\mathring{I}^2$, tels que $x \neq y$, le taux d'accroissement $\dfrac{f(x) - f(y)}{x - y}$ est borné, alors la dérivée de f le sera également. En effet, cela s'obtient grâce à un passage à la limite, en écrivant $y = x + h$ et en faisant $h \longrightarrow 0$ pour écrire le nombre dérivé de f en $x \in \mathring{I}$.

Exemple : considérons la fonction $\phi : \begin{array}{ccc} [0, 1] & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sqrt{x} \end{array}$. ϕ n'est pas Lipschitzienne sur $]0, 1]$ car sa dérivée est donnée par la formule $\phi'(x) = \dfrac{1}{2\sqrt{x}}$, et ϕ' n'est pas bornée sur cet intervalle.

Abordons à présent la notion de **continuité uniforme** :

Définition 2.10. *Continuité uniforme.*
Soient (E, N) et $(F, \|.\|)$, deux espaces vectoriels normés, X, une partie de E, Y, une partie de

F, et f, une application de E dans F. On dit que l'application f est ***uniformément continue*** *ur X si, et seulement si :*

$$\forall \varepsilon > 0,\ \exists \eta > 0 : \forall (x, y) \in X^2,\ N(x - y) < \eta \Longrightarrow \|f(x) - f(y)\| < \varepsilon.$$

Remarque 1 : il est clair qu'une application uniformément continue est nécessairement continue. En effet, si l'on souhaite démontrer la continuité de f en $x_0 \in X$, il suffit de poser $y = x_0$ pour retrouver la définition de la continuité à partir de la définition précédente de l'uniforme continuité.

Remarque 2 : si f est une application k-Lipschitzienne, alors f est uniformément continue. Il suffit de choisir $\eta = \frac{\varepsilon}{k}$ pour retrouver la définition de l'uniforme continuité.

Exemple : nous avons démontré précédemment que la fonction ϕ n'est pas Lipschitzienne. En revanche, ϕ est uniformément continue. En effet, soient x et y, deux éléments de $]0, 1]$ tels que $y > x$ (pour fixer les idées). Alors, si $y - x < \eta$, on peut écrire :

$$(\sqrt{y} - \sqrt{x})^2 = x + y - 2\sqrt{xy} < x + y - 2x = x - y.$$

En effet, cela provient du fait que $x < \sqrt{xy} < y$. Il suffit alors de choisir $\eta = \varepsilon^2$ pour prouver l'uniforme continuité de ϕ.
Cet exemple permet également de montrer qu'une fonction uniformément continue n'est pas nécessairement Lipschitzienne.

2.2 Notion de complétude dans un espace vectoriel normé

Cette notion est réellement **fondamentale** en Analyse. On attend de vous dans ce cours que vous sachiez démontrer la complétude d'espaces vectoriels normés afin d'en déduire des propriétés de convergence. Définissons au préalable la notion de suite de Cauchy.

2.2.1 Suites de Cauchy dans un E.V.N.

Définition 2.11. *Suite de Cauchy.*
Soit $(E$, un espace vectoriel muni d'une norme N. Une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de E est dite ***« de Cauchy »*** *si, et seulement si :*

$$\forall \varepsilon > 0,\ \exists N_0 \in \mathbb{N} : \forall (n, p) \in \mathbb{N}^2,\ n \geq N_0 \Longrightarrow N(u_{n+p} - u_n) < \varepsilon.$$

On peut encore donner la caractérisation suivante :

$$\forall \varepsilon > 0,\ \exists N_0 \in \mathbb{N} : \forall (n, q) \in \mathbb{N}^2,\ q > n \geq N_0 \Longrightarrow N(u_q - u_n) < \varepsilon.$$

Abordons à présent quelques conséquences de cette définition.

Propriété 2.12. *Soit (E, N), un espace vectoriel normé. Alors :*
(i) Toute suite convergente est de Cauchy.
(ii) Toute suite de Cauchy est bornée.
(iii) L'image d'une suite de Cauchy par une application f uniformément continue est de Cauchy.

(iv) Si deux normes N_1 et N_2 sont équivalentes sur E, alors, toute suite de Cauchy pour N_1 est également une suite de Cauchy pour N_2.

Démontrons à présent ces propositions :

Démonstration. Dans toute la preuve, $(u_n)_{n\in\mathbb{N}}$ désignera une suite de Cauchy de l'espace vectoriel normé E.

(*i*) Supposons que $(u_n)_{n\in\mathbb{N}}$ est une suite convergente. Fixons $\varepsilon > 0$. Alors, par définition, il existe $\ell \in E$ telle que :

$$\exists N \in \mathbb{N} : n \geq N \Longrightarrow N(u_n - \ell) < \frac{\varepsilon}{2}.$$

Soient n et p, deux entiers supérieurs à N. Alors $N(u_n - u_p) \leq N(u_n - \ell) + N(u_p - \ell) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, et $(u_n)_{n\in\mathbb{N}}$ est donc une suite de Cauchy dans E.

(*ii*) Puisque $(u_n)_{n\in\mathbb{N}}$ est de Cauchy, on peut écrire, par inégalité triangulaire, et si on s'est fixé $\varepsilon > 0$:

$$\exists N \in \mathbb{N} : \forall (n,p) \in \mathbb{N}^2,\ n \geq N \Longrightarrow N(u_{n+p}) - N(u_n) < \varepsilon.$$

Posons $p = N + 1$. Alors : $\forall n > N$, $N(u_n) < N(u_{N+1}) + \varepsilon$. On en déduit que :

$$\forall n \in \mathbb{N},\ N(u_n) \leq \max\{N(u_1), ..., N(u_N), N(u_{N+1}) + \varepsilon\}.$$

(*iii*) Soit $f : (E, N) \longrightarrow (F, \|.\|)$, une application uniformément continue sur X, une partie de E, et supposons que $(u_n)_{n\in\mathbb{N}}$ prend toutes ses valeurs dans X. Rappelons l'inégalité vérifiée par f. Fixons $\varepsilon > 0$. Alors :

$$\exists \eta > 0 : N(x - y) < \eta \Longrightarrow \|f(x) - f(y)\| < \varepsilon.$$

Appelons $v_n := f(u_n)$. On souhaite donc démontrer que $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy. Puisque $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy, alors :

$$\exists N \in \mathbb{N} : \forall (n,p) \in \mathbb{N}^2 : n > p \geq N \Longrightarrow N(u_n - u_p) < \eta.$$

Par conséquent, puisque f est uniformément continue, on a : $\|f(u_n) - f(u_p)\| < \varepsilon$. Autrement dit, $(v_n)_{n\in\mathbb{N}}$ est une suite de Cauchy.

(*iv*) La démonstration est immédiate. En effet, il existe deux constantes α et β strictement positives telles que $\forall x \in E,\ \alpha N_1(x) < N_2(x) < \beta N_1(x)$. Du coup, si $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy pour N_1, alors, pour $\varepsilon > 0$ fixé, il existe $N \in \mathbb{N}$ tel que $n > p \geq N \Longrightarrow N_1(u_n - u_p) < \dfrac{\varepsilon}{\beta}$.

Et donc :

$$N_2(u_n - u_p) \leq \beta N_1(u_n - u_p) < \varepsilon.$$

Par conséquent, $(u_n)_{n\in\mathbb{N}}$ est bien une suite de Cauchy pour N_2.

□

Remarque : on a montré qu'une suite convergente est de Cauchy ; attention, en revanche, une suite de Cauchy n'est pas nécessairement convergente. Utilisons un exemple que nous avons déjà introduit au cours de ce chapitre. On appelle $(u_n)_{n\in\mathbb{N}}$, la suite définie par :

$$u_n = \sum_{k=0}^{n} \frac{1}{k!}.$$

Plaçons-nous dans l'espace vectoriel normé $(\mathbb{Q}, |.|)$ (muni de la norme euclidienne usuelle). $(u_n)_{n\in\mathbb{N}}$ est bien-sûr de Cauchy. En effet, soient n et p, deux entiers tels que $n > p$. Alors :

$$u_n - u_p = \sum_{k=p+1}^{n} \frac{1}{k!} \leq \sum_{k=p+1}^{n} \frac{1}{2^{k-1}} = \frac{1}{2^{p-1}}\left(1 - \left(\frac{1}{2}\right)^{n-p-2}\right) < \frac{1}{2^{p-1}} \xrightarrow[p\to+\infty]{} 0.$$

Donc $(u_n)_{n\in\mathbb{N}}$ est bien une suite de Cauchy dans $\mathbb{Q}$ muni de la distance euclidienne et nous avons déjà montré que $\lim\limits_{n\to+\infty}(u_n) = e$, et que e est un nombre irrationnel. Nous venons donc d'exhiber une suite de Cauchy dans un espace vectoriel normé qui ne converge pas dans cet espace. Nous allons donc faire appel à la notion de **complétude** d'un espace vectoriel normé pour préciser ces cas un peu litigieux.

Propriété 2.13. *Soit (E, N), un espace vectoriel normé et $(u_n)_{n\in\mathbb{N}}$, une suite de Cauchy dans cet espace.*
Si $(u_n)_{n\in\mathbb{N}}$ possède une valeur d'adhérence, alors $(u_n)_{n\in\mathbb{N}}$ converge dans E au sens de N, bien-sûr.

Démonstration. Appelons a, la valeur d'adhérence de $(u_n)_{n\in\mathbb{N}}$. Il existe donc une application strictement croissante $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ telle que $u_n \xrightarrow[n\to+\infty]{\text{dans } E} a$. Pour éviter un formalisme un peu pénible à écrire, je donne l'argument principal de cette démonstration. Il suffit d'écrire que pour $n > N$ (autrement dit, n suffisamment grand ou encore n choisit plus grand qu'un certain entier N), on a :

$$N(u_n - a) \leq N(u_n - u_{\varphi(n)}) + N(u_{\varphi(n)} - a).$$

Il suffit de conclure en remarquant que la quantité $N(u_{\varphi(n)} - a)$ peut être rendue aussi petite que l'on veut, puisque la suite extraite $(u_{\varphi(n)})_{n\in\mathbb{N}}$ est convergente, et d'autre part, la suite $(u_n)_{n\in\mathbb{N}}$ étant de Cauchy, on sait que : $N(u_n - u_{\varphi(n)}) \xrightarrow[n\to+\infty]{} 0$ donc, on peut rendre cette quantité aussi petite que souhaité pourvu que l'on choisisse N suffisamment grand. □

Exercice : rédiger proprement la démonstration précédente.

2.2.2 Espaces complets et exemples

Passons à présent à la notion d'espace vectoriel normé **complet**. Nous allons tenter de donner des exemples de preuve de complétude d'un espace.

Définition 2.12. *Espace Vectoriel Normé complet.*
*Soit (E, N), un espace vectoriel normé. On dit que E est **complet** si, et seulement si toute suite de Cauchy de E converge dans E.*

Nous continuons tout de suite par un exemple d'espace vectoriel normé.

Propriété 2.14. *L'espace vectoriel $\mathbb{R}$ muni de la norme euclidienne est un espace vectoriel normé complet.*

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$, une suite de Cauchy de $\mathbb{R}$. Alors nous avons vu que $(u_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathbb{R}$. D'après la propriété de Bolzano-Weierstrass, on peut extraire de $(u_n)_{n\in\mathbb{N}}$ une sous-suite convergente dans $\mathbb{R}$. Par conséquent, $(u_n)_{n\in\mathbb{N}}$ possède une valeur d'adhérence dans $\mathbb{R}$. Or, toute suite de Cauchy possédant une valeur d'adhérence converge. On en déduit que $(u_n)_{n\in\mathbb{N}}$ converge, et donc, $\mathbb{R}$ est un espace vectoriel normé complet. $\square$

On va démontrer à présent que tous les esaces $\mathbb{R}^n$ munis de la distance euclidienne sont complets. Pour cela, il nous suffit de démontrer que le produit de deux espaces complets est encore complet et nous en déduirons la complétude de $\mathbb{R}^n$ à l'aide d'un raisonnement par récurrence quasi immédiat.

Propriété 2.15. *Soient (E_1, N_1) et (E_2, N_2), deux espaces vectoriels normés complets. Alors, l'espace produit $E_1 \times E_2$ est également complet.*

Démonstration. Soit $x \in E_1 \times E_2$. Alors $x = (x_1, x_2)$, où $x_1 \in E_1$ et $x_2 \in E_2$. On munit $E_1 \times E_2$ de la norme N définie pour tout x de cet espace par : $N(x) = N_1(x_1) + N_2(x_2)$. N est une norme (nous l'avons déjà démontré en exercice). Soit $(x_n)_{n\in\mathbb{N}}$, une suite de Cauchy de E. On peut donc écrire $\forall n \in \mathbb{N},\ x_n = (x_n^1, x_n^2)$, où $(x_n^1)_{n\in\mathbb{N}}$ est une suite de E_1 et $(x_n^2)_{n\in\mathbb{N}}$ est une suite de E_2. Reste à montrer que $(x_n)_{n\in\mathbb{N}}$ converge dans E. Soit $\varepsilon > 0$ fixé. $(x_n)_{n\in\mathbb{N}}$ étant une suite de Cauchy, on en déduit que :

$$\exists N \in \mathbb{N},\ p > q > q \Longrightarrow N(x_p - x_q) < \varepsilon.$$

On va montrer que chacune des suites $(x_n^1)_{n\in\mathbb{N}}$ et $(x_n^2)_{n\in\mathbb{N}}$ est de Cauchy. C'est assez immédiat. En effet, il suffit de remarquer que $N_1(x_p - x_q) < N(x_p - x_q)$ et cette remarque s'applique également à la norme N. Par conséquent $(x_n^1)_{n\in\mathbb{N}}$ et $(x_n^2)_{n\in\mathbb{N}}$ sont des suites de Cauchy. Les espaces E_1 et E_2 étant complets on en déduit l'existence de deux élements $x_\infty^1 \in E_1$ et $x_\infty^1 \in E_2$ qui sont les limites respectives de $(x_n^1)_{n\in\mathbb{N}}$ et $(x_n^2)_{n\in\mathbb{N}}$. Je laisse au lecteur le soin de terminer la preuve. C'est élémentaire : $x_\infty = (x_\infty^1, x_\infty^2)$ est la limite de la suite $(x_n)_{n\in\mathbb{N}}$. $\square$

Conséquence : par récurrence immédiate, on démontre la propriété suivante :

Propriété 2.16. *L'espace $\mathbb{R}^n$ est complet pour tout $n \in \mathbb{N}$.*

Remarque : $\mathbb{R}^n$ est un espace de dimension finie dans lequel toutes le normes sont équivalentes. Donc, inutile, dans la propriété ci-dessus de préciser la norme dont on munit $\mathbb{R}^n$.

Les deux propriétés qui suivent s'avèreront fort utiles pour démontrer qu'un espace vectoriel est complet, en se ramenant à la démonstration qu'un espace vectoriel est fermé.

Propriété 2.17. *Soit (E, N), un espace vectoriel normé complet. Soit X, une partie de E. Alors, X est complète si, et seulement si X est fermée.*

Démonstration. On va démontrer les deux implications :

- **Sens $\Longrightarrow$:** supposons X complète dans E complet. Soit $(u_n)_{n\in\mathbb{N}}$, une suite de X convergeant dans E. On note L sa limite dans E. $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy (car convergente) dans E, donc en particulier dans X qui est complet. On en déduit l'existence de $u \in X$ tel que $u_n \xrightarrow[n\to+\infty]{\text{dans } X} u$. Or, $(u_n)_{n\in\mathbb{N}}$ étant convergente dans E, par unicité de la limite, $L = u \in X$. On retrouve la caractérisation séquentielle des fermés. Ainsi, X est une partie fermée.

- **Sens $\Longleftarrow$:** supposons X fermée. Soit $(u_n)_{n\in\mathbb{N}}$, une suite de Cauchy d'éléments de X. En particulier, $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de E, donc est convergente vers $u \in E$. Mais puisque X est fermée, toujours d'après la caractérisation séquentielle des fermés, il s'ensuit que $u \in X$. Donc $(u_n)_{n\in\mathbb{N}}$ converge dans X.

$\square$

Quelques applications de cette propriété.

1. $[a,b]$, où $a < b$ est complet.
2. $[a,+\infty[$, où $a \in \mathbb{R}$ est complet.
3. $\mathbb{R}^*$ n'est pas complet.
4. $\{1,...,n\}$ est complet.

Le théorème qui suit fournit encore une caractérisation de la complétude à l'aide d'une application uniformément continue. Nous la démontrerons en exercice. Rappelons juste que si $f : X \longrightarrow Y$ est une application, alors on définit l'ensemble $f(X)$ par : $f(X) = \{f(x), x \in X\}$.

Théorème 2.3. *Soient (E,N) et $(F,\|.\|)$, deux espaces vectoriels normés, et X une partie de E. On suppose que :*
(i) $f : X \longrightarrow f(X)$ est une bijection ;
(ii) f est continue sur X ;
(iii) f^{-1} est uniformément continue sur $f(X)$.
Si X est une partie complète de E, alors $f(X)$ est une partie complète de F.

Passons à présent à des exemples importants d'espaces complets. Un exemple d'espace non complet est donné par le contre-exemple étudié dans le paragraphe 2.2.1, où l'on montre qu'il existe une suite de Cauchy de $(\mathbb{Q},|.|)$ non convergente dans cet espace.

Comment démontrer qu'un espace est complet ?

Le cas de $\ell^\infty(\mathbb{R})$: $\ell^\infty(\mathbb{R})$ est un espace complet.
Définissons au préalable cet espace : $\ell^\infty(\mathbb{R})$ est l'ensemble des suites réelles bornées. Si $u = (u_n)_{n\in\mathbb{N}}$ désigne une suite réelle bornée, on munit $\ell^\infty(\mathbb{R})$ de la norme : $\|u\| = \sup\limits_{n\in\mathbb{N}}(|u_n|)$. Je laisse au lecteur le soin de vérifier que $\ell^\infty(\mathbb{R})$ est un espace vectoriel.
Pour démontrer qu'un espace est complet, il faut revenir à la définition. On se donne une suite de Cauchy d'éléments de cet espace et on démontre qu'elle est convergente.

Soit $(u^p)_{p\in\mathbb{N}}$, une suite d'éléments de $\ell^\infty(\mathbb{R})$. $(u^p)_{p\in\mathbb{N}}$ est donc une suite de suites. On peut encore noter pour p fixé :
$$u^p = (u^p_n)_{n\in\mathbb{N}} = (u^p_0, u^p_1, ..., u^p_n, ...).$$
Supposons à présent que $(u^p)_{p\in\mathbb{N}}$ est une suite de Cauchy. Redonnons la définition :
$$\forall\varepsilon > 0, \exists N \in \mathbb{N} : r > s > N \Longrightarrow \|u^r - u^s\| < \varepsilon.$$
Par conséquent, on peut encore écrire que pour tout n, entier naturel, on a : $r > s > N \Longrightarrow |u^r_n - u^s_n| < \varepsilon$. Il s'ensuit donc que , $(u^p_n)_{p\in\mathbb{N}}$ est une suite de Cauchy à valeurs réelles, et puisque $\mathbb{R}$ est complet, $(u^p_n)_{p\in\mathbb{N}}$ converge dans $\mathbb{R}$ vers un réel ℓ_n. On peut écrire : $u^p_n \xrightarrow[p\to+\infty]{} \ell_n$. On dispose donc d'une suite $(\ell_n)_{n\in\mathbb{N}}$. On va montrer que cette suite est la limite de la suite de Cauchy

initiale. Reprenons l'inégalité ci-dessus. On avait montré que : $r > s > N \Longrightarrow |u_n^r - u_n^s| < \varepsilon$. Faisons tendre r vers $+\infty$. On obtient alors pour tout entier naturel n : $s > N \Longrightarrow |\ell_n - u_n^s| < \varepsilon$. Un passage au sup fournit immédiatement : $s > N \Longrightarrow \|\ell - u^s\| < \varepsilon$.
À ce stade, nous avons presque achevé notre démonstration. Reste à prouver que la suite $(\ell_n)_{n\in\mathbb{N}}$ est encore dans l'espace vectoriel ℓ^∞, autrement dit, il faut montrer que la suite $(\ell_n)_{n\in\mathbb{N}}$ est bornée. Or, en considérant $s > N$, d'après l'inégalité $\forall n \in \mathbb{N},\ |\ell_n - u_n^s| < \varepsilon$, on sait que la suite $(\ell - u^s)$ est dans $\ell^\infty(\mathbb{R})$, et puisque $\ell^\infty(\mathbb{R})$ est un espace vectoriel, alors $\ell \in \ell^\infty(\mathbb{R})$, car $u^s \in \ell^\infty(\mathbb{R})$.

Conclusion : on a démontré que :

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} : s > N \Longrightarrow \|\ell - u^s\| < \varepsilon.$$

Autrement dit $(u^s)_{s\in\mathbb{N}}$ converge dans $\ell^\infty(\mathbb{R})$, ce qui prouve que $\ell^\infty(\mathbb{R})$ est complet.

Remarques :
- En utilisant le même modèle, on démontre aisément que $\ell^\infty(\mathbb{C})$ est également complet.
- La méthode pour démontrer que ℓ^∞ est complet est une méthode assez générale. L'idée est de considérer une suite de Cauchy de ℓ^∞ et d'essayer de se ramener à une suite de Cauchy de $\mathbb{R}$ (ou un autre espace complet). Ensuite, on utilise le fait que $\mathbb{R}$ est un espace complet, ce qui nous fournit une première notion de convergence. Il reste ensuite à déduire de cette information (lorsque cela est possible) que la suite de Cauchy de ℓ^∞ converge au sens de la norme dont ℓ^∞ est muni et que sa limite est bien dans ℓ^∞.

Passons à présent à un deuxième exemple qui sera un peu développé ultérieurement.

Cas des fonctions continues sur I : on appelle $\mathcal{C}(I, \mathbb{R})$, l'espace des fonctions continues sur un intervalle $I = [a, b]$ et à valeurs réelles. On munit $\mathcal{C}(I, \mathbb{R})$ de la norme suivante : $\|.\|_\infty$: $\begin{array}{rcl} \mathcal{C}(I, \mathbb{R}) & \longrightarrow & \mathbb{R}_+ \\ f & \longmapsto & \|f\|_\infty = \sup\limits_{x\in I} |f(x)| \end{array}$. Alors, $\mathcal{C}(I, \mathbb{R})$ **est un espace vectoriel normé complet pour la norme** $\|.\|_\infty$. Je laisse au lecteur le soin de démontrer que $\mathcal{C}(I, \mathbb{R})$ est un espace vectoriel. Montrons qu'il est complet.
Une première remarque essentielle est de vérifier que le choix de la norme $\|.\|_\infty$ est judicieux pour les éléments de l'epsace $\mathcal{C}(I, \mathbb{R})$, autrement dit que la norme infinie de tout élément de $\mathcal{C}(I, \mathbb{R})$ est finie. C'est le cas, car une fonction continue sur un intervalle réel est nécessairement majorée. Ce résultat peut encore se voir d'une autre façon en admettant un théorème que nous énoncerons dans la section qui suivra : le théorème de Heine. Ce théorème permet de montrer que f est uniformément continue sur l'intervalle I, car I est un compact de $\mathbb{R}$ et que f est continue sur I. Ainsi, f est nécessairement bornée sur I ce qui justifie la définition de la norme.

Attaquons-nous à la preuve qui nous intéresse : soit $(f_n)_{n\in\mathbb{N}}$, une suite de Cauchy d'éléments de $\mathcal{C}(I, \mathbb{R})$. Alors :

$$\forall \varepsilon > 0 : \exists N \in \mathbb{N} : p > q > N \Longrightarrow \|f_p - f_q\|_\infty < \varepsilon.$$

Par conséquent, il est immédiat que :

$$\exists N \in \mathbb{N} : p > q > N \Longrightarrow \forall x \in I,\ |f_p(x) - f_q(x)| < \varepsilon.$$

Cette inégalité traduit le fait qu'à x fixé, élément de I, la suite $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy de réels. Par conséquent, et puisque $\mathbb{R}$ est complet, la suite $(f_n(x))_{n\in\mathbb{N}}$ est convergente dans $\mathbb{R}$. On note $f(x)$ sa limite, ce qui définit une application f dans $\mathcal{C}(I, \mathbb{R})$.
Reste à démontrer que f est bien limite uniforme de la suite $(f_n)_{n\in\mathbb{N}}$ et surtout que f est encore

un élément de $\mathcal{C}(I,\mathbb{R})$. La première assertion est presque évidente, puisque $(f_n)_{n\in\mathbb{N}}$ est une suite de Cauchy. Par conséquent, on, peut faire tendre p vers $+\infty$ dans l'inégalité ci-dessus, ce qui garantit que :

$$\forall q \in \mathbb{N} : q > N,\ \|f - f_q\|_\infty < \varepsilon.$$

Pour démontrer que f est continue, on se fixe $x_0 \in I$ et $\varepsilon > 0$. On sait, en particulier que si x est fixé, on a : $f_n(x) \xrightarrow[n\to+\infty]{} f(x)$, donc, il existe un rang $N_0 \in \mathbb{N}$ tel que : $n \geq N_0 \Longrightarrow |f_n(x) - f(x)| \leq \frac{\varepsilon}{3}$ et $n \geq N_0 \Longrightarrow |f_n(x_0) - f(x_0)| \leq \frac{\varepsilon}{3}$. de plus, les applications éléments de $(f_n)_{n\in\mathbb{N}}$ étant continues, il existe $\eta > 0$ tel que : $|x-x_0| < \eta \Longrightarrow |f(x) - f(x_0)| \leq \frac{\varepsilon}{3}$. Fixons-nous alors $n \geq N_0$. On peut alors écrire par inégalité triangulaire :

$$\forall x \in I : |x - x_0| < \eta,\ |f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \leq 3.\frac{\varepsilon}{3}.$$

Cela démontre que f est continue sur I. I étant compact, f est bornée sur I. On le aussi le voir en remarquant, toujours par inégalité triangulaire et en choisissant $n > N$ que :

$$\|f\|_\infty \leq \|f - f_n\|_\infty + \|f_n\|_\infty \leq \varepsilon + \|f_n\|_\infty.$$

Cela démontre que $\mathcal{C}(I,\mathbb{R})$ muni de le norme de la convergence uniforme (norme infinie) est complet.

2.2.3 Le théorème du point fixe et ses applications

Définissons au préalable la notion d'application contractante :

Définition 2.13. *Soient (E,N) et $(F,\|.\|)$, deux espaces vectoriels normés et $f : X \subset E \longrightarrow F$. f est dite* ***contractante*** *si, et seulement s'il existe $k \in]0,1[$ tel que :*

$$\forall (x,y) \in X^2, \|f(x) - f(y)\| \leq kN(x-y).$$

Je présente maintenant le théorème du point fixe dont je donnerai deux applications essentielles : l'une en Analyse Numérique, pour déterminer les solutions d'une équation, et l'autre pour démontrer l'existence de solutions d'une équation différentielle dans un cas très simple.

Théorème 2.4. *Soit (E,N), un espace vectoriel normé* ***complet****. Soit X, une partie* ***fermée*** *de E et $f : X \longrightarrow X$, une application* ***contractante****.*
Alors, l'équation $f(x) = x$ possède dans X une solution unique x_0.
De plus, x_0 est limite de toutes les suites $(u_n)_{n\in\mathbb{N}}$ définies par :

$$\begin{cases} u_0 = \lambda \in X \\ u_{n+1} = f(u_n). \end{cases}$$

Démonstration. La démonstration s'articule en deux étapes :
- **1$^{\text{ère}}$ partie : preuve de l'unicité.** L'existence étant (provisoirement) admise, prouvons que le point fixe est nécessairement unique. Bien évidemment, on utilise un raisonnement par l'absurde. Supposons qu'il existe deux éléments x_0 et x_1, non égaux, tels que $f(x_0) = x_0$ et $f(x_1) = x_1$. f étant contractante sur X, on peut écrire que : $N(x_0 - x_1) = N(f(x_0) - f(x_1)) \leq kN(x_0 - x_1)$, ce qui est absurde, puisque $k < 1$. L'unicité est donc démontrée.

- **$2^{\text{ème}}$ partie : preuve de l'existence.** Soit $\lambda \in X$ et considérons la suite $(u_n)_{n\in\mathbb{N}}$ définie par :

$$\begin{cases} u_0 = \lambda \in X \\ u_{n+1} = f(u_n). \end{cases}$$

L'idée est de démontrer que la suite $(u_n)_{n\in\mathbb{N}}$ est de Cauchy. On peut successivement écrire pour n et p entiers tels que $n > p$:

$$\begin{array}{rcl} N(u_{n+2} - u_{n+1}) & \leq & kN(u_{n+1} - u_n) \\ N(u_{n+3} - u_{n+2}) & \leq & kN(u_{n+2} - u_{n+1}) \\ & \vdots & \\ N(u_{n+p} - u_{n+p-1}) & \leq & kN(u_{n+p-1} - u_{n+p-2}) \end{array}$$

Multiplions alors membre à membre ces inégalités. Des simplifications s'opèrent en cascade et on obtient alors :

$$N(u_{n+p} - u_{n+p-1}) \leq k^{p-1}N(u_{n+1} - u_n) \leq k^{n+p-1}N(u_1 - u_0).$$

Pour prouver que $(u_n)_{n\in\mathbb{N}}$ est de Cauchy, utilisons une inégalité triangulaire. Si $n > p$, on a :

$$\begin{array}{rcl} N(u_{n+p} - u_n) & \leq & N(u_{n+p} - u_{n+p-1}) + \ldots + N(u_{n+1} - u_n) \\ & \leq & (k^{n+p-1} + k^{n+p-2} + \ldots + k^n)N(u_1 - u_0) = k^n \dfrac{1-k^p}{1-k}.N(u_1 - u_0). \end{array}$$

Or, puisque $|k| < 1$, la suite $n \longmapsto k^n$ est convergente de limite nulle. Un passage à la limite dans l'inégalité ci-dessus ($p, n \longrightarrow +\infty$) prouve, par théorème d'encadrement, que la suite $(u_n)_{n\in\mathbb{N}}$ est de Cauchy. Puisque X est fermé dans un espace vectoriel complet, X est complet lui-même, et $(u_n)_{n\in\mathbb{N}}$ converge donc vers un élément de X noté x_0. Un passage à la limite ($n \longrightarrow +\infty$) dans l'égalité $u_{n+1} = f(u_n)$ prouve que x_0 vérifie l'égalité $x_0 = f(x_0)$, autrement dit que x_0 est un point fixe pour f. f est contractante donc continue. L'existence est donc démontrée, en utilisant la caractérisation séquentielle de la continuité.

$\square$

Application n°1 : convergence de la suite de Fibonacci :

On définit la suite de Fibonacci par la relation de récurrence :

$$\begin{cases} u_{n+2} = u_{n+1} + u_n \quad \forall n \geq 0 \\ u_0 = 0 \\ u_1 = 1 \end{cases}$$

Les premies termes de cette suite se calculent aisément : $u_2 = 1$, $u_3 = 2$, $u_4 = 3$, $u_4 = 5$, $u_5 = 8$, $u_6 = 13$, $u_7 = 21$, et $u_8 = 34$.
Des raisonnements par récurrence très simples permettent de démontrer que :
- $\forall n \geq 1\ u_n > 0$.
- $\forall n \geq 4,\ u_n \geq n$.
- $(u_n)_{n\in\mathbb{N}}$ est strictement croissante.

Cela prouve en particulier que $\lim\limits_{n\to+\infty}(u_n) = +\infty$. Intéressons-nous à présent à la suite $(v_n)_{n\in\mathbb{N}}$ définie par la relation explicite :

$$\forall n \in \mathbb{N}^*,\ v_n = \frac{u_{n+1}}{u_n}.$$

Remarquons, en utilisant la relation de récurrence définissant la suite de FIBONACCI que $(v_n)_{n\in\mathbb{N}^*}$ est également définie par la relation de récurrence :

$$\begin{cases} v_{n+1} = \dfrac{u_{n+2}}{u_{n+1}} = 1 + \dfrac{1}{v_n} & \forall n \geq 1 \\ v_1 = 1 \end{cases}$$

On peut encore écrire que $\forall n \in \mathbb{N}^*$, $v_{n+1} = f(v_n)$, où f désigne la fonction définie pour $x \neq 0$ par :

$$f(x) = 1 + \frac{1}{x}.$$

On peut, comme précédemment, calculer les premiers termes de la suite $(v_n)_{n\geq 1}$. On a : $v_1 = 1$, $v_2 = 2$ $v_3 = \frac{3}{2}$, $v_4 = \frac{5}{3}$, $v_5 = \frac{8}{5}$.
L'idée est d'utiliser dans cet exemple le théorème du point fixe afin de démontrer la convergence de la suite $(v_n)_{n\in\mathbb{N}^*}$. La statégie à adopter est la suivante :

- f doit être restreinte à un intervalle I sur lequel elle est contractante pour pouvoir appliquer le théorème du point fixe.
- Nous devrons donc démontrer qu'à partir d'un certain rang n_0, tous les éléments de la suite $(v_n)_{n\geq n_0}$ appartiennent à I.

Tous les éléments de $(v_n)_{n\geq 1}$ étant positifs, on peut étudier f sur l'intervalle $]0; +\infty[$. Pour tout élément x de cet intervalle : $f'(x) = -\dfrac{1}{x^2}$. On en déduit que f est strictement décroissante sur $\mathbb{R}^*$. Pour rechercher l'intervalle I mentionné précédemment, il est nécessaire de rechercher un intervalle stable par la fonction f, c'est à dire tel que : $f(I) \subset I$. Nous allons comprendre pourquoi dans un instant.

D'après l'inégalité des accroissements finis et l'expression de la dérivée de f, f est contractante sur $[1 + \varepsilon; +\infty[$, avec $\varepsilon > 0$, aussi faible que souhaité. En observant les premiers termes de la suite $(v_n)_{n\geq 1}$, il semble que si $n \geq 2$, $v_n \in [\frac{3}{2}; 2]$. Vérifions-le à l'aide d'un petit raisonnement par récurrence. Je vous laisse le soin de le rédiger proprement. L'hérédité résulte du fait que, en appelant I l'intervalle $[\frac{3}{2}; 2]$, I est stable par f. En effet, puisque f est décroissante, $f(I) \subset [f(2); f(\frac{3}{2})] = [\frac{3}{2}, \frac{5}{3}] \subset I$. Ainsi, si $v_n \in I$, alors $f(v_n) \in I$, autrement dit $v_{n+1} \in I$.

Conclusion : $(v_n)n \geq 2$ est définie par la relation de récurrence $v_{n+1} = f(v_n)$, où $f : I \longrightarrow I$ est contractante de constante de Lipschitz $\frac{4}{9}$. (Cela résulte directement de l'expression de la dérivée de f). D'après le théorème du pont fixe, la suite $(v_n)_{n\in\mathbb{N}}$ est convergente et l'on note ℓ sa limite.

La limite de $(v_n)_{n\in\mathbb{N}}$ se trouve aisément. En effet, f ést contractante, donc continue, et $(v_n)_{n\in\mathbb{N}}$ converge vers ℓ. D'après la caractérisation séquentielle de la continuité, un passage à la limite dans la relation de récurrence définissant $(v_n)_{n\in\mathbb{N}}$ prouve que $\ell = f(\ell)$. (i.e. ℓ est un point fixe de f) Enfin, l'équation $x = f(x)$ est équivalente à une équation du second degré très simple. On trouve deux solutions dont une négative que l'on élimine : puisque tous les termes de $(v_n)_{n\in\mathbb{N}}$ sont positifs, un passage à la limite prouve que $\ell \geq 0$. Finalement, on en déduit que :

$$\ell = \frac{1 + \sqrt{5}}{2}.$$

ℓ est le célèbre **nombre d'or**.

Application n°2 : la méthode de Newton-Raphson en Analyse Numérique.

Cette méthode d'Analyse numérique, dont je vais donner les grandes lignes et un exemple d'application est utilisée pour calculer les zéros d'une fonction f donnée. Bien que cette méthode puisse être généralisée aux fonctions de plsuieurs variables, considérons une fonction f d'une variable réelle définie sur un intervalle de $\mathbb{R}$ noté $I = [a, b]$. On cherche donc à résoudre l'équation :

$$(E)\ f(x) = 0,\ \text{avec } x \in I.$$

Supposons de plus que f soit deux fois dérivable et strictement monotone, sans changement de convexité, telle que $f(a)f(b) < 0$, ce qui garantit, grâce au théorème des valeurs intermédiaires, l'existence d'une solution unique à l'équation (E). On introduit la fonction g définie sur I par :

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Remarquons que g est bien définie sur I tout entier puisque f est strictement monotone (sa dérivée ne s'annule donc pas). f étant supposée deux fois dérivable sur I à dérivées continues, les théorèmes classiques de dérivabilité garantissent que g est dérivable sur I et sa dérivée est donnée par :

$$\forall x \in I,\ g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}.$$

De plus, remarquons que l'équation (E) est équivalente à l'équation :

$$(E')\ g(x) = x,\ x \in I.$$

L'idée générale est de choisir un sous-intervalle de I, noté J, contenant le point fixe α de g, c'est à dire la solution recherchée, et tel que $g(J) \subset J$, autrement dit, J est stable par g. En remarquant de plus que g' est continue et que $\lim\limits_{x \to \alpha} g'(x) = g'(\alpha) = 0$, on peut choisir l'intervalle J pour que : $\exists k < 1 : \forall x \in J,\ |g'(x)| < k$, autrement dit, g est contractante, grâce à l'inégalité des accroissements finis. On s'est ainsi placé dans les conditions d'application du théorème du point fixe qui garantit que tout suite du type :

$$\begin{cases} u_0 = \lambda \in J \\ u_{n+1} = g(u_n). \end{cases}$$

sera convergente. En appliquant l'inégalité des accroissements finis à la fonction g, et en utilisant le fait que J est stable par g, on démontre très facilement que tous les u_n sont dans J et l'inégalité suivante s'ensuit :

$$\forall n \in \mathbb{N},\ |u_{n+1} - \alpha| \leq k|u_n - \alpha|.$$

k désigne la constante de Lipschitz dans l'inégalité ci-dessus. Un raisonnement par récurrence permet alors de prouver que :

$$\forall n \in \mathbb{N},\ |u_n - \alpha| \leq k^n|u_0 - \alpha|.$$

Ce résultat traduit notamment la convergence de la suite $(u_n)_{n\in\mathbb{N}}$ vers le réel α. En réalité il existe même un résultat plus fort. On pourrait montrer que, dans le cas d'une méthode de Newton, pour une fonction g donnée, il existe une constante strictement positive C donnée telle que :

$$\forall n \in \mathbb{N},\ |u_{n+1} - \alpha| \leq C|u_n - \alpha|^2.$$

Cette inégalité traduit le fait que la suite $(u_n)_{n\in\mathbb{N}}$ aura tendance à se rapprocher rapidement de sa limite. On parle alors de convergence quadratique ou d'ordre 2.

À titre d'exemple, on peut choisir, si $a > 0$, $f(x) = x^2 - a$. Alors g est définie par : $g(x) = x - \dfrac{x^2 - a}{2x} = \dfrac{1}{2}\left(x + \dfrac{a}{x}\right)$. Alors, en choisissant convenablement l'intervalle de définition de f et g noté J, on obtient un algorithme pour calculer rapidement une bonne approximation de $\sqrt{a}$. On utilise donc la suite :

$$\begin{cases} u_0 = \lambda \in J \\ u_{n+1} = \frac{1}{2}\left(u_n + \frac{a}{u_n}\right). \end{cases}$$

Remarque : cet algorithme de calcul d'une approximation de racine est historiquement dû à Héron d'Alexandrie (1$^{\text{er}}$ siècle après JC), bien que déjà connu des babylonnien quelques siècles auparavant.

Le théorème de Cauchy-linéaire pour les équations différentielles.
Je vais redonner dans cette partie un résultat classique sur les systèmes différentiels qui sera bien davantage exploité dans le chapitre qui suit. Je me contente donc, dans un premier temps, de démontrer un résultat théorique d'existence et d'unicité qui utilise le théorème du point fixe. Par la suite, nous nous intéresserons plus particulièrement à la résolution de systèmes différentiels. Voici l'énoncé du théorème :

Théorème 2.5. *Théorème de Cauchy-Linéaire.*
Soit $I = [a, b]$, où $a < b$, un intervalle de $\mathbb{R}$. Soit $(\mathcal{S})$, un système différentiel d'ordre n, c'est à dire une équation différentielle de la forme :

$$(\mathcal{S}) \begin{cases} Y' = AY + B \\ Y(t_0) = Y_0 \in \mathbb{R}^n, \text{ avec } t_0 \in I. \end{cases}$$

où A et B sont des fonctions définies sur I telles que :
- *$A : I \longrightarrow \mathcal{M}_n(\mathbb{R})$ est continue. (A est donc une matrice carrée de taille n.)*
- *$B : I \longrightarrow \mathbb{R}^n$ est continue. (B est donc un vecteur colonne.)*

Alors, $(\mathcal{S})$ possède une solution unique $Y \in \mathbb{R}^n$, de classe $\mathcal{C}^1$ sur I.

Remarque : si les notions d'Algèbre Linéaire dont on fait mention dans ce théorème ne sont pas limpides, allez urgemment les réviser. Des rappels sont fournis dans la deuxième partie de cet ouvrage. Même remarque pour les applications de classe $\mathcal{C}^1$ sur un intervalle.

Démonstration. Remarquons immédiatement que si le système ci-dessus possède une solution, alors cette solution est nécessairement de classe $\mathcal{C}^1$ sur I, puisque A et B sont continues sur I (ce qui implique directement que Y' est encore continue sur I). Une intégration membre à membre de chaque ligne du système $(\mathcal{S})$ ci-dessus fournit immédiatement :

$$\forall t \in I,\ Y(t) = Y_0 + \int_{t_0}^{t} [A(s)Y(s) + B(s)]ds.$$

Nous avons démontré précédemment que l'espace $\mathcal{C}(I, \mathbb{R})$ est complet. Les mêmes arguments que ceux que nous avions alors utilisés servent à démontrer que $\mathcal{C}(I, \mathbb{R}^n)$ est complet, pour n entier naturel quelconque. On admet de plus que l'on peut normer cet espace vectoriel en considérant l'application : $\|.\|_\infty : \begin{array}{ccc} \mathcal{C}(I, \mathbb{R}^n) & \longrightarrow & \mathbb{R}_+ \\ Y = (Y_i)_{1 \le i \le n} & \longmapsto & \max\limits_{1 \le i \le n} \|Y_i\|_\infty \end{array}$. On admet encore, mais ce résultat est très classique et simple à démontrer que l'espace $\mathcal{C}(I, \mathcal{M}_n(\mathbb{R})$ des matrices carrées dont les coefficients sont des applications continues est un espace vectoriel normé complet (c'est encore la même démonstration que précédemment), normable par l'application N : $\begin{array}{rcl} \mathcal{C}(I, \mathcal{M}_n(\mathbb{R})) & \longrightarrow & \mathbb{R}_+ \\ A = (a_{i,j})_{i,j} & \longmapsto & \max\limits_{i,j} \|a_{i,j}\|_\infty \end{array}$ (ce que nous étudierons en fin de chapitre). Considérons à présent l'application ψ définie par :

$$\psi : \begin{array}{rcl} \mathcal{C}(I, \mathbb{R}^n) & \longrightarrow & \mathcal{C}(I, \mathbb{R}^n) \\ Y & \longmapsto & \psi(Y) \end{array}, \text{ où } \psi(Y) : \begin{array}{rcl} I & \longrightarrow & \mathbb{R}^n \\ t & \longmapsto & Y_0 + \displaystyle\int_{t_0}^t [A(s)Y(s) + B(s)]ds. \end{array}$$

Soient Y et Z, deux éléments de $\mathcal{C}(I, \mathbb{R}^n)$. On a, en utilisant le fait que N est une norme et des propriétés ultra classiques sur les intégrales :

$$\begin{aligned} \| (\psi(Y) - \psi(Z)) \|_\infty &= \left\| \left(\int_{t_0}^t A(s)[Y(s) - Z(s)]ds \right) \right\|_\infty \\ &\leq N(A) \|Y - Z\|_\infty \left| \int_{t_0}^t ds \right| = N(A) \|Y - Z\|_\infty |t - t_0| \end{aligned}$$

Or, si Y et Z sont des éléments de $\mathcal{C}(I, \mathbb{R}^n)$, alors $\psi(Y)$ et $\psi(Z)$ le sont encore ($\mathcal{C}(I, \mathbb{R}^n)$ est stable par ψ). Par conséquent, une récurrence immédiate permet de démontrer l'inégalité :

$$\forall (Y, Z) \in \mathcal{C}(I, \mathbb{R}^n)^2, \forall n \in \mathbb{N} \ \| (\psi^n(Y) - \psi^n(Z)) \|_\infty \leq \frac{N^n(A) |t - t_0|^n}{n!} \|Y - Z\|_\infty.$$

Dans l'inégalité ci-dessus, on a noté : $\psi^n = \psi \circ \psi \circ ... \circ \psi$ (n fois).
Or, la suite $\left(\dfrac{N^n(A) |t - t_0|^n}{n!} \|Y - Z\|_\infty \right)_{n \in \mathbb{N}}$ étant convergente de limite nulle, on en déduit l'existence de $N \in \mathbb{N}$ tel que $N > N$ implique que la fonction ψ^n est contractante. L'application du théorème du point fixe prouve donc que la fonction ψ^n admet un unique point fixe Y, c'est à dire tel que : $\psi^n(Y) = Y$. Or, si nous composons cette égalité par la fonction ψ, on obtient $\psi^{n+1}(Y) = \psi(\psi^n(Y)) = \psi(Y) = \psi^n(\psi(Y))$, et par unicité du point fixe de ψ^n, on en déduit que $\psi(Y) = Y$.
Le théorème de Cauchy-Linéaire est donc démontré. □

Voici un exemple concret. Si x, y, z et t sont des fonctions de classe $\mathcal{C}^1(\mathbb{R})$, on considère le système différentiel :

$$\left\{ \begin{array}{rcccccc} x' & = & x & + & y & & \\ y' & = & & & y & + & z \\ z' & = & & & & & z + t \\ t' & = & & & & & t \end{array} \right.$$

D'après le théorème de Cauchy-linéaire, le système ci-dessus admet une solution unique.

Exercice : trouver la solution de ce système, à l'aide de calculs élémentaires et de résolutions d'équations différentielles.

2.3 Compacité dans un espace vectoriel normé

2.3.1 Généralités

Définition 2.14. *Soit (E, N), un espace vectoriel normé.*
*Une partie X de E est dite **compacte** si de toute suite d'éléments de X, on peut extraire une sous-suite convergente dans X.*

Exemple : considérons l'espace normé $\mathbb{R}$, muni de la norme usuelle $|.|$. Soit- $[a, b]$, où $a < b$, un intervalle de $\mathbb{R}$. **Alors $[a, b]$ est un compact de $\mathbb{R}$.** En effet, soit $(u_n)_{n\in\mathbb{N}}$, une suite d'éléments de $[a, b]$. Alors, en vertu de la propriété de Bolzano-Weierstrass, $(u_n)_{n\in\mathbb{N}}$ étant bornée, on peut extraire de $(u_n)_{n\in\mathbb{N}}$ une sous-suite convergente dans $[a, b]$.

Contre-exemple : $(\mathbb{R}, |.|)$ n'est pas compact. En effet, considérons la suite $(u_n)_{n\in\mathbb{N}}$ définie par $u_n = n$. u_n est croissante non majorée, donc on ne peut pas extraire de $(u_n)_{n\in\mathbb{N}}$ une sous-suite convergente dans $\mathbb{R}$. En effet, toute suite extraite de $(u_n)_{n\in\mathbb{N}}$ serait non bornée donc non convergente.

La propriété qui suit est fondamentale. Elle donne des propriétés générales des compacts. Elle sera précisée dans le cas où l'on considère un espace vectoriel normé de dimension finie, ce qui simplifiera grandement notre étude.

Propriété 2.18. *Soit (E, N), un espace vectoriel normé, et X, un compact de E.*
*Alors, X est une partie **fermée** et **bornée** de E*

Démonstration. On va démontrer séparément ces deux propriétés.

(*i*) X **est bornée :** raisonnons par l'absurde et supposons donc X non bornée. Traduisons mathématiquement cette assertion. Si X était bornée, on aurait : $\exists N > 0 : \forall x \in X,\ N(x) \leq M$. Donc, dans notre cas, on a : $\forall N > 0, \exists x \in X : N(x) \geq M$. Alors, en choisissant succesivement pour valeurs de M $1,\ 2,\ 3, ...$, on construit donc une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que : $\forall n \in \mathbb{N},\ N(x_n) \geq n$. Mais, puisque X est compacte, il existe une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$, notée φ, telle que $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converge dans X vers un certain $\ell \in X$. Mais ceci est absurde, car $\lim\limits_{n\to+\infty} N(x_n) = +\infty$, et par inégalité triangulaire, on a : $|N(x_n) - N(\ell)| \leq N(x_n - \ell) \xrightarrow[n\to+\infty]{} 0$. On en déduit que X est nécessairement bornée.

(*ii*) X **est fermée :** on va utiliser la caractérisation séquentielle des fermés. Si $(x_n)_{n\in\mathbb{N}}$ est une suite de points de X convergeant vers $\ell \in E$, alors, par compacité de X, on peut extraire de $(x_n)_{n\in\mathbb{N}}$ une sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ convergente dans X. Par unicité de la limite de $(x_n)_{n\in\mathbb{N}}$, la limite de $(x_{\varphi(n)})_{n\in\mathbb{N}}$ est nécessairement ℓ. On en déduit que $\ell \in X$, autrement dit, X est fermée.

□

Remarque 1 : attention ! La réciproque de cette propriété n'est pas toujours vraie. À savoir, un compact de E est toujours fermé et borné, mais un fermé borné n'est pas nécessairement compact. En dimension finie, on a l'équivalence entre ces deux notions.

Remarque 2 : un espace vectoriel normé (E, N) n'est jamais compact. En effet, E n'est pas borné...

Propriété 2.19. *Soit (E, N), un espace vectoriel normé, et X, un compact de E.*
Si Y est une partie fermée de X, alors Y est compacte.

Démonstration. C'est immédiat. En effet, soit $(y_n)_{n\in\mathbb{N}}$, une suite d'éléments de Y. Alors, puisque $Y \subset X$ et que X est compact, on peut extraire de Y une sous-suite convergente dans X. Mais, puisque Y est fermé, toute suite convergente d'éléments de Y converge dans Y. On a donc démontré que, de toute suite d'éléments de Y, on peut extraire une sous-suite convergente. □

Passons à présent à des caractérisations concrètes de compacts. Il existe des cas dans lesquels on sait très bien montrer qu'un ensemble est compact. Par exemple lorsqu'il peut s'écrire comme produit cartésien de compacts, où lorsque (on l'a déjà évoqué) on s'est placé dans un espace vectoriel normé de dimension finie.

Théorème 2.6. *Soient X et Y, deux parties compactes respectivement incluses dans les espaces vectoriels normés (E, N_1) et (F, N_2).*
Alors, $X \times Y$ est compact.

Remarque : dans les conditions rappelées dans le théorème ci-dessus, on dit que $X \times Y$ est compact dans l'espace vectoriel $E \times F$, muni de la norme induite N définie pour $x \in X$ et $y \in Y$ par : $N(x, y) = N_1(x) + N_2(y)$.

Démonstration. Soit $(u_n)_{n\in\mathbb{N}}$, une suite de $X \times Y$. Alors, $\forall n \in \mathbb{N},\ u_n = (x_n, y_n)$, où $x_n \in X$ et $y_n \in Y$. Puisque X est compact, il existe une application $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$, strictement croissante, telle que $(x_{\varphi(n)})_{n\in\mathbb{N}}$ est convergente, de limite dans X, x. De même, Y étant compact, on en déduit l'existence d'une application strictement croissante $\psi : \mathbb{N} \longrightarrow \mathbb{N}$ telle que $(y_{\psi(n)})_{n\in\mathbb{N}}$ converge vers $y \in Y$. Toute suite extraite d'une suite convergente étant convergente, on en déduit que $u_n = (x_n, y_n) \xrightarrow[n\to+\infty]{} (x, y) \in X \times Y$, autrement dit, $X \times Y$ est compact. □

Remarque 1 : une récurrence immédiate associée au théorème que l'on vient de démontrer permettent de prouver que si X_1, ..., X_n désignent n compacts de E, alors le produit cartésien $\displaystyle\prod_{k=1}^{n} X_i$ est encore compact.

Remarque 2 : intérêt de la remarque précédente... Si l'on souhaite démontrer par exemple que la partie $\displaystyle\prod_{k=1}^{n} [a_k, b_k]$ de $\mathbb{R}^n$ muni de la norme euclidienne usuelle, constitue une partie compacte de $\mathbb{R}^n$ $(a_k < b_k,\ \forall k \in \{1, ..., n\}$, il nous suffit de démontrer que $[a, b] \subset \mathbb{R}$ est compact et le tour est joué, en vertu du théorème précédent. C'est ce que l'on va faire à présent.

Propriété 2.20. *Compacts en dimension finie.*
(i) Les parties compactes de $\mathbb{R}$ sont les parties fermées bornées non vide.
(ii) Les parties compactes de $\mathbb{R}^n$, où n est un entier sont les parties fermées bornées non vides.
(iii) Les parties compactes de $\mathbb{C}^n$, où n est un entier sont les parties fermées bornées de $\mathbb{C}^n$.

Démonstration. On va démontrer ces assertions. Compte tenu de ce que l'on a fait jusqu'à présent, c'est assez élémentaire. En effet, on a déjà montré que si X est un compact, alors X est fermé borné. Supposons donc, réciproquement, que X est une partie fermée bornée dans $\mathbb{R}$. X étant bornée, il est aisé de voir que $X \subset [a, b]$, où a et b sont deux réels tels que $a < b$. Or, $[a, b]$ est compact, on l'a déjà vu et X étant fermée dans un compact est une partie compacte de $\mathbb{R}$.
Pour la deuxième assertion, on utilise exactement la même idée. En effet, on sait déjà que si X est un compact de $\mathbb{R}^n$, alors X est nécessairement fermé et borné. Et si l'on suppose réciproquement que X est fermé borné dans $\mathbb{R}^n$, alors il existe $(a_k)_{k\in\{1,...,n\}}$ et $(b_k)_{k\in\{1,...,n\}}$, deux suites

telles que $\forall k \in \{1, ..., n\}$, $a_k < b_k$ et $X \subset \prod_{k=1}^{n} [a_k, b_k]$. Or, on a déjà vu que $\prod_{k=1}^{n} [a_k, b_k]$ est un produit de compacts et est donc compact dans $\mathbb{R}^n$, et puisque X est fermé dans un compact, X est lui-même compact.
Enfin, on démontre très facilement la troisième assertion puisque $\mathbb{C}^n$ est isomorphe à $\mathbb{R}^{2n}$, en appliquant la deuxième assertion. □

Remarque : le résultat énoncé dans la remarque précédente peut encore être généralisé : si I désigne une famille d'indices **quelconques** et $(X_i)_{i \in I}$, une famille de compacts, alors le produit $\prod_{i \in I} X_i$ est encore compact. Il s'agit du **théorème de Tychonov**.

La propriété qui suit fait le lien entre les espaces complets, que nous avons déjà étudiés et les espaces compacts.

Propriété 2.21. *Soit (E, N), un espace vectoriel normé et X, une partie compacte de E. Alors, X est nécessairement une partie complète de E.*

Démonstration. Soit $(u_n)_{n \in \mathbb{N}}$, une suite de Cauchy de X. X étant compact, $(u_n)_{n \in \mathbb{N}}$ admet une valeur d'adhérence $u_\infty \in X$. Il existe donc une application strictement croissante $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ telle que $u_{\varphi(n)} \xrightarrow[n \to +\infty]{\text{dans } X} u_\infty$. r, $\forall n \in \mathbb{N}$, $\varphi(n) \geq n$. Ainsi, puisque $(u_n)_{n \in \mathbb{N}}$ est de Cauchy, fixons $\varepsilon > 0$, et on a alors : $\exists N \in \mathbb{N} : N(u_n - u_{\varphi(n)}) < \dfrac{\varepsilon}{2}$. De plus, puisque $(u_{\varphi(n)})_{n \in \mathbb{N}}$ est convergente, il existe un rang $N_1 \in \mathbb{N}$ tel que : $n \geq N_1 \Longrightarrow N(u_{\varphi(n)} - u_\infty) < \frac{\varepsilon}{2}$. On en déduit que $n \geq \max(N_1, N) \Longrightarrow N(u_n - u_\infty) < N(u_n - u_{\varphi(n)}) + N(u_{\varphi(n)} - u_\infty) < 2.\frac{\varepsilon}{2} = \varepsilon$, ce qui prouve que $(u_n)_{n \in \mathbb{N}}$ converge et donc X est complet. □

2.3.2 Lien entre applications continues et uniformément continues

Dans cette partie, je commence par donner quelques résultats sur les images de compacts par des applications continues et sur le lien entre des applications continues sur un compact et des applications uniformément continues sur un compact (théorème de Heine).

Théorème 2.7. *Soient (E, N) et $(F, \|.\|)$, deux espaces vectoriels normés, et $f : E \longrightarrow F$, une application continue sur E. Soit $X \subset E$, une partie compacte de E. Alors $f(X)$ est une partie compacte de F.*

Démonstration. Soit $(v_n)_{n \in \mathbb{N}}$, une suite de $f(X)$. Il existe donc $(u_n)_{n \in \mathbb{N}}$ telle que $\forall n \in \mathbb{N}$, $v_n = f(u_n)$. $(u_n)_{n \in \mathbb{N}}$ est une suite d'éléments d'un compact, donc $(u_n)_{n \in \mathbb{N}}$ admet une valeur d'adhérence u_∞ (i.e. il existe $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$, strictement croissante, telle que $u_{\varphi(n)} \xrightarrow[n \to +\infty]{\text{dans } X} u_\infty$). Puis, f étant continue, on en déduit que $f(u_{\varphi(n)}) \xrightarrow[n \to +\infty]{\text{dans } F} f(u_\infty) \in f(X)$. Ceci montre que $(v_n)_{n \in \mathbb{N}}$ admet une valeur d'adhérence, donc $f(X)$ est compact. □

Corollaire 2.1. *Soit X, un compact non vide de E et f, une application continue sur X et à valeurs réelles. Alors, f est bornée et atteint ses bornes.*

Démonstration. On sait que $f(X)$ est compact d'après le théorème précédent. Il est donc borné, ce qui justifie que f est bornée.

De plus f atteint ses bornes car $f(X)$ est fermé. En effet, cela a pour conséquence que $\inf_{x\in X} f(x) \in f(X)$ et $\sup_{x\in X} f(x) \in f(X)$. $\square$

Utilisation de ce théorème en Optimisation : dans ce domaine, on souhaite en général minimiser des fonctions ou fonctionnelles. On s'intéresse souvent aux fonctions dites *coercives*, c'est-à-dire aux fonctions $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ telles que $\lim_{\|x\|_{\mathbb{R}^n}\to+\infty} f(x) = +\infty$, pour la raison suivante : on peut démontrer aisément que, si l'on souhaite trouver le minimum de f sur une partie fermée $F \subset \mathbb{R}^n$, alors il est aisé de démontrer que f est minorée et qu'elle atteint sa borne inférieure. En effet, si $a \in F$, étant donné que f est coercive, il existe $R > 0$ tel que $x > R \Longrightarrow f(x) > f(a)$. Ainsi, il est donc parfaitement clair que : $\inf_{x\in F} f(x) = \inf_{x\in B_f(0,R)\cap F} f(x)$. Or, puisque F est fermé et que $B_f(0,R)$ est compact (car fermé borné dans un espace de dimension finie), on en déduit que $B_f(0,R)\cap F$ est encore compact et on est ramené à démonter l'existence d'un minimum de f sur un compact. D'après le théorème précédent, c'est immédiat. Nous étudierons en exercice l'exemple de la minimisation d'une fonctionnelle f définie par une relation du type :

$$f(x) = \frac{1}{2} < Ax, x > - < b, x >,\ \forall x \in \mathbb{R}^n.$$

A désigne une matrice symétrique définie positive et b, un vecteur de $\mathbb{R}^n$.

Le théorème qui suit est le célèbre théorème de Heine. Une démonstration complète de ce théorème sera proposée en TD. J'en donnerai ici une autre, un peu différente.

Théorème 2.8. *Théorème de Heine.*
Soient (E,N) et $(F,\|.\|)$, deux espaces vectoriels normés, et $f : E \longrightarrow F$, une application continue sur $X \subset E$. Supposons de plus que X est compact.
Alors, f est uniformément continue sur X.

Démonstration. Raisonnons par l'absurde et supposons f non uniformément continue sur X. Alors, $\exists\varepsilon > 0 : \forall\eta > 0, \exists(x,y) \in X^2 : N(x-y) < \eta$ et $\|f(x) - f(y)\| > \varepsilon$. On peut donc construire, en faisant successivement prendre à η les valeurs 1, $\frac{1}{2}$, $\frac{1}{3}$, ... $\frac{1}{n}$, une suite $(x_n, y_n)_{n\in\mathbb{N}}$, d'éléments de X^2 telle que $N(x_n - y_n) < \frac{1}{n}$ et $\|f(x_n) - f(y_n)\| \geq \varepsilon$. Or, un produit de compacts étant compact, on en déduit que X^2 est encore compact. Ainsi, la suite $(x_n, y_n)_{n\in\mathbb{N}}$ possède une valeur d'adhérence. Il existe donc $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$, strictement croissante, telle que $(x_{\varphi(n)}, y_{\varphi(n)})_{n\in\mathbb{N}}$ converge vers un élément de X^2 noté (x_∞, y_∞), autrement dit $x_{\varphi(n)} \xrightarrow[n\to+\infty]{\text{dans } X} x_\infty$ et $y_{\varphi(n)} \xrightarrow[n\to+\infty]{\text{dans } X} y_\infty$. Enfin, puisque $N\left(x_{\varphi(n)} - y_{\varphi(n)}\right) < \frac{1}{\varphi(n)} \xrightarrow[n\to+\infty]{} 0$, on en déduit que $x_\infty = y_\infty$. f étant continue, on a bien entendu, par caractérisation séquentielle de la continuité, $f(x_{\varphi(n)}) \xrightarrow[n\to+\infty]{\text{dans } F} f(x_\infty)$ et $f(y_{\varphi(n)}) \xrightarrow[n\to+\infty]{\text{dans } F} f(y_\infty) = f(x_\infty)$. Un passage à la limite dans l'inégalité vérifiée pour tout $n \in \mathbb{N}$: $\|f(x_{\varphi(n)}) - f(y_{\varphi(n)})\| \geq \varepsilon$ conduit à l'inégalité $0 \geq \varepsilon$, ce qui est absurde.
Donc, f est uniformément continue sur X. $\square$

Application du théorème de Heine : les sommes de Riemann.
Le théorème suivant est une conséquence directe du théorème de Heine. Il s'agit ici d'un résultat de limites qui permet, notamment d'approcher des intégrales de fonctions continues à l'aide de sommes.

Théorème 2.9. *Sommes de Riemann.*
Soient a et b, deux réels tels que $a < b$. On appelle ***subdivision*** *de ce segment toute suite finie $a_0 = a < a_1 < ... < a_{n-1} < a_n = b$. On appelle* ***pas de cette subdivision*** *et on note δ, le nombre : $\delta := \max\limits_{1 \leq i \leq n}(a_{i+1} - a_i)$. On définit de plus les sommes S_m et S_M, appelées* ***Sommes de Darboux*** *par :*

$$S_m := \sum_{i=1}^{n}(x_i - x_{i-1}) \inf_{\xi \in [x_{i-1}, x_i]} f(\xi) \text{ et } S_M := \sum_{i=1}^{n}(x_i - x_{i-1}) \sup_{\xi \in [x_{i-1}, x_i]} f(\xi).$$

Alors, on a :

$$S_m \leq \int_a^b f(t)dt \leq S_M \text{ et } \lim_{\delta \to 0}(S_m) = \lim_{\delta \to 0}(S_M) = \int_a^b f(t)dt.$$

En particulier, soit $n \in \mathbb{N}^*$. On peut démontrer que si : $\forall i \in \{1, ..., n\}$, $x_i = a + i\dfrac{b-a}{n}$ (on dit alors que la subdivision est régulière car $\forall i \in \{1, ..., n-1\}$, $x_{i+1} - x_i = \frac{b-a}{n}$), on a :

$$\lim_{n \to +\infty} \frac{b-a}{n} \sum_{i=1}^{n} f\left(a + k\frac{b-a}{n}\right) = \int_a^b f(t)dt.$$

La démonstration de ce résultat est proposée en exercice. Des exemples seront alors donnés.

2.3.3 Notion de densité et approximations uniformes

Pour conclure cette partie, nous allons nous intéresser de plus près à l'espace fonctionnel $\mathcal{C}(X, Y)$, où X est une partie de E et Y, une partie de F. Nous allons voir que sous certaines conditions, il est possible de donner des propriétés fort intéressantes des éléments de cet espace. Commençons par le résultat caractérisatique suivant.

Propriété 2.22. *Soit X une partie* ***compacte*** *d'un espace vectoriel normé (E, N) et Y, une partie d'un espace vectoriel normé $(F, \|.\|)$. Considérons l'application $\|.\|_\infty$ définie par :*

$$\begin{array}{rccl} \|.\|_\infty : & \mathcal{C}(X, Y) & \longrightarrow & \mathbb{R} \\ & f & \longmapsto & \sup\limits_{x \in X} \|f(x)\| \end{array}.$$

Alors, $(\mathcal{C}(X, Y), \|.\|_\infty)$ est un espace vectoriel normé et $\|.\|_\infty$ s'appelle ***la norme de la convergence uniforme****.*
De plus, si Y est un espace ***complet****, alors $(\mathcal{C}(X, Y), \|.\|_\infty)$ est un espace complet.*

Le résultat qui suit est un résultat de densité. Pour mieux le comprendre, commençons par étudier un petit exemple de convergence uniforme qui s'incrit comme une application du théorème que nous allons énoncer.

Un exemple de convergence uniforme : la fonction $x \longmapsto \sin x$ est un élément de $\mathcal{C}([0, 1])$ par exemple. Puisque $[0, 1]$ est un compact de $\mathbb{R}$, les résultats classiques sur les séries enières nous permettent d'écrire que :

$$\forall x \in [0, 1], \ \sin x = \sum_{k=0}^{+\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Plus précisément, appelons f_n, la fonction définie sur $[0,1]$ par : $f_n(x) = \displaystyle\sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{(2k+1)!}$.
La suite de fonctions $(f_n)_{n\geq 0}$ **converge uniformément** vers la fonction sinus sur l'intervalle $[0,1]$. Il s'agit là d'un résultat bien plus puissant que celui traduit par la limite écrite ci-dessus. Autrement dit, on a, en appelant f la fonction sinus définie sur $[0,1]$:

$$\|f_n - f\|_\infty := \sup_{x\in[0,1]} |f_n(x) - f(x)| \xrightarrow[n\to+\infty]{} 0.$$

Il est clair que cela implique en particulier la convergence dite **simple** de la suite de fonctions $(f_n)_{n\in\mathbb{N}}$, autrement dit, la première limite écrite ci-dessus, que nous pouvons encore traduire par : $\forall x \in [0,1]$, $\lim\limits_{n\to\infty}(f_n(x)) = f(x)$.

Vers une généralisation : le théorème de Bernstein va nous permettre d'introduire avec précision le théorème de Weierstrass, qui constitue l'objectif de ce paragraphe. En effet, il permet d'exprimer toute fonction continue sur $[0,1]$ comme la limite uniforme d'une suite de polynômes (appelés d'ailleurs polynômes de Bernstein).

Théorème 2.10. *Théorème de Bernstein.*
Soit f, une fonction continue sur le segment $[0,1]$, et $(B_n)_{n\geq 0}$, la suite de fonctions polynômes définis pour tout entier naturel n par :

$$B_n(x) := \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} f\left(\frac{k}{n}\right) x^k(1-x)^{n-k}.$$

Alors, la suite B_n converge uniformément vers f, c'est-à-dire que : $\|B_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0$.

Remarque 1 : ce théorème est fort intéressant à différents points de vue. En effet, il ne se contente pas d'énoncer un résultat théorique d'approximation uniforme d'une fonction continue à l'aide de polynômes, mais il construit cette suite de polynômes.

Remarque 2 : dans l'énoncé de ce théorème, nous avons choisi $x \in [0,1]$, mais si nous considérons à présent une fonction f continue sur un intervalle $[a,b]$ quelconque, avec $a < b$, il est tout à fait évident que le résultat énoncé ci-dessus se généralise à de telles fonctions. En effet, il suffit de considérer la bijection affine : $x \in [0,1] \longmapsto a + t(b-a) \in [a,b]$. C'est la raison pour laquelle la démonstration ci-dessous sera faite en considérant des fonctions continues sur $[0,1]$.

Démonstration. f est continue sur un segment, donc $|f|$ est bornée par une constante que l'on note $M > 0$. De plus, d'après le théorème de Heine, f est uniformément continue sur $[0,1]$. Soit donc $\varepsilon > 0$. Il existe donc $\eta > 0$ tel que $\forall k \in \{0,...,n\}$, $\left|x - \frac{k}{n}\right| < \eta \Longrightarrow \left|f(x) - f\left(\frac{k}{n}\right)\right| < \frac{\varepsilon}{2}$.
Utilisons alors l'identité : $\forall x \in [0,1]$, $1 = \displaystyle\sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k}$, qui résulte du binôme de Newton. On peut alors écrire pour $x \in [0,1]$ et n entier naturel non nul que :

$$|f(x)-B_n(x)| \leq \sum_{\left\{k:\left|x-\frac{k}{n}\right|<\varepsilon\right\}} \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k} + \sum_{\left\{k:\left|x-\frac{k}{n}\right|\geq\varepsilon\right\}} \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k}.$$

Appelons alors :

- $S_1 := \displaystyle\sum_{\left\{k:\left|x-\frac{k}{n}\right|<\varepsilon\right\}} \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k}$.

• $S_2 := \displaystyle\sum_{\{k : |x-\frac{k}{n}|\geq\varepsilon\}} \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k}.$

On règle assez rapidement le cas de S_1, en utilisant l'uniforme continuité de f associée à une formule du binôme :

$$S_1 \leq \frac{\varepsilon}{2}\sum_{k=0}^{n} x^k(1-x)^{n-k} = \frac{\varepsilon}{2}.$$

À présent, en utilisant la majoration de f, on obtient pour la seconde somme :

$$S_2 \leq 2M \sum_{\{k : |x-\frac{k}{n}|\geq\varepsilon\}} x^k(1-x)^{n-k} \leq \frac{2M}{\eta^2}\sum_{\{k : |x-\frac{k}{n}|\geq\varepsilon\}} \left(x-\frac{k}{n}\right)^2 x^k(1-x)^{n-k}.$$

En développant le carré ci-dessus et en réarrangeant les termes à l'aide par exemple du binôme de Newton, on prouve aisément que :

$$S_2 \leq \frac{2M}{\eta^2}\left[\frac{x(1-x)}{n} - x^2\right] \leq \frac{2M}{4n\eta^2} \xrightarrow[n\to+\infty]{} 0.$$

Il suffit alors de choisir n suffisamment grand, ou plus présicément $n \geq N_0 := E(\frac{M}{\delta^2\varepsilon}) + 1$ et il s'ensuit alors que :

$$\exists n \in \mathbb{N} : n \geq N_0 \Longrightarrow \forall x \in [0,1],\ |f(x) - B_n(x)| \leq 2.\frac{\varepsilon}{2} = \varepsilon.$$

Cela démontre la convergence uniforme de la suite de polynômes $(B_n)_{n\in\mathbb{N}}$ vers la fonction continue f. □

Terminons cette section en éonçant deux théorèmes d'approximation uniforme : les théorèmes de Weierstrass et Weierstrass-Trigonométrique.

Théorème 2.11. *Théorème de Weierstrass.*
Soit X, une partie compacte de $\mathbb{R}$.
L'ensemble des fonctions polnômiales est dense dans $\mathcal{C}(X,\mathbb{R})$.

Il s'agit là plus ou moins d'une généralisation du théorème de Bernstein que nous venons d'énoncer.

Théorème 2.12. *Théorème de Weierstrass-Trigonométrique.*

- *Un* ***polynôme trigonométrique*** *a pour expression :*

$$P_n = \sum_{k=0}^{n}[a_k\cos(kx) + b_k\sin(kx)].$$

Dans l'expression ci-dessus, $(a_k)_{k\geq 0}$ et $(b_k)_{k\geq 0}$ sont deux sites numériques.

- *Toute fonction continue sur $\mathbb{R}$, 2π-périodique, est limite* ***uniforme*** *d'une suite de polynômes trigonométriques.*

Remarque : la démonstration du théorème de Weierstrass Trigonométrique est proposée en exercice.

On fera le lien entre le théorème de Weierstrass trigonométrique et la théorie des séries de Fourier.

Je redonne ici quelques éléments assez rudimentaires pour bien comprendre la portée du théorème de Weierstrass. Des explications plus complètes ur cette notion seront fournies dans le chapitre 3.

Cadre et notations : considérons une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$, 2π-périodique et localement intégrable. On définit alors les *coefficients de Fourier* de f par :

$$a_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos(nt)dt \text{ et } b_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(nt)dt.$$

On appelle alors **série de Fourier** associée à f, la série trigonométrique :

$$S(f)(x) = \frac{a_0(f)}{2} + \sum_{n=1}^{+\infty}[a_n(f)\cos(nx) + b_n(f)\sin(nx)].$$

On a (notamment) le résultat d'approximation uniforme suivant :

Théorème 2.13. *Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$, une fonction continue, 2π-périodique et $\mathcal{C}^1$ par morceaux. Alors, la série de Fourier de f converge uniformément vers f sur $\mathbb{R}$.*

Exemple : on note, comme à notre habitude $\|.\|_\infty$, la norme de la convergence uniforme sur $\mathbb{R}$. En développant $f_\lambda : x \longmapsto \cos(\lambda x)$ en série de Fourier, on démontre que l'on a $\|f_\lambda - f_{N,\lambda}\| \xrightarrow[N\to+\infty]{} 0$, où l'on a noté $f_{N,\lambda}$, la fonction définie pour $N \in \mathbb{N}$ par : $f_{N,\lambda}(x) = \dfrac{\sin(\lambda x)}{\lambda\pi} + \displaystyle\sum_{n=1}^{N}(-1)^n \frac{2\lambda\sin(\lambda\pi)}{\pi(\lambda^2 - n^2)}\cos(nx)$. Il s'agit bien-sûr d'un exemple de convergence uniforme. En particulier, la convergence simple est vérifiée, autrement dit :

$$\forall x \in \mathbb{R},\ \cos(\lambda x) = \frac{\sin(\lambda x)}{\lambda\pi} + \sum_{n=1}^{+\infty}(-1)^n\frac{2\lambda\sin(\lambda\pi)}{\pi(\lambda^2 - n^2)}\cos(nx).$$

2.3.4 Propriété de Borel-Lebesgue et recouvrements

J'introduis dans ce paragraphe un autre point de vue sur la compacité. Il rejoint bien-sûr le précédent, mais son originalité réside dans le fait qu'il peut être étendu pour déboucher sur une théorie abstraite plus générale que celle que j'ai présentée jusqu'à maintenant. En effet, avec cette nouvelle définition, il n'est plus nécessaire que l'espace soit muni d'une distance et encore moins d'une norme.
Commençons par définir la notion de recouvrement d'une partie.

Définition 2.15. *Recouvrement.*
Soit X une partie (éventuellement de dimension infinie), d'un espace vectoriel normé, et soit $(U_i)_{i\in I}$, une famille de parties de cet ensemble.

- *On dit que $(U_i)_{i\in I}$ est un* ***recouvrement*** *de X lorsque :*

$$\bigcup_{i\in I} U_i = X.$$

- *De plus, si I est une partie d'indices finie (i.e. card$(J) < +\infty$), on dit que $(U_i)_{i\in I}$ est un* ***recouvrement fini****.*
- *Enfin, si $J \subset I$ et si $\bigcup_{i\in J} U_i = X$, on dira que $(U_i)_{i\in J}$ est un* ***sous recouvrement*** *de X.*

On a alors la caractérisation suivante de la notion de compacité :

Théorème 2.14. *Propriété de Borel-Lebesgue.*
Pour qu'un espace X muni d'une norme soit compact, il est nécessaire et il suffit qu'il vérifie la propriété : « Pour tout recouvrement de X par des ouverts $(O_i)_{i\in I}$, il existe $J \subset I$ tel que J est fini et $(O_i)_{i\in J}$ recouvre X. ».

Un cas particulier, la précompacité :

Définition 2.16. *Un espace métrique X est dit* ***précompact*** *si, et seulement si pour tout $r > 0$, il existe une partie finie $Y \subset X$ telle que :*

$$X = \bigcup_{y\in Y} B(y, r).$$

Enfin, on va démontrer le théorème suivant, qui caractérise la compacité au sein d'un espace métrique.

Théorème 2.15. *Un espace normé est précompact si, et seulement s'il est précompact et complet.*

Pour vous entraîner, je propose l'exercice ultra détaillé suivant, qui démontre le théorème qui vient d'être énoncé. Je n'en ferai donc pas la correction.

Exercice : compacité et précompacité

Soit E, un espace vectoriel normé muni d'une norme N.

1. Soit X, un espace compact. On a déjà vu que X est complet. On va démontrer dans cette question que X est précompact. Soit $\varepsilon > 0$ et soit x, un élément de X. On dit que x est un point isolé s'il existe un voisinage $\mathcal{V}$ de x dans X tel que $\mathcal{V} \cap X = \{x\}$. On admet que toute partie de X qui a tous ses points isolés est une partie finie.
   (a) Construire, par récurrence une suite $(x_n)_{n\geq 0}$ d'éléments de X tels que : $x_n \in X \backslash \bigcup_{k=1}^{n-1} B(x_k, \varepsilon)$.
   On partira du principe que, s'il existe $n_0 \in \mathbb{N}$ tel que $X \subset \bigcup_{k=1}^{n_0} B(x_k, \varepsilon)$, on arrête la construction à x_{n_0} ce qui garantit le résultat que l'on souhaite démontrer.
   (b) Démontrer très simplement que tous les points de l'ensemble $\{x_k, k \in \mathbb{N}^*\}$ sont isolés.
   (c) Conclure.
2. Supposons ici que X est précompact et complet. Soit $(u_n)_{n\in\mathbb{N}}$, une suite d'éléments de X prenant un nombre infini de valeurs (autrement, cette suite admettrait clairement une valeur d'adhérence et le résultat serait démontré).
   (a) Démontrer qu'il existe $N \in \mathbb{N}$ et une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X tels que $X = \bigcup_{n=1}^{N} B(x_n, 1)$.
   (b) Démontrer qu'il existe $m \in \{1, ..., n\}$ et φ_0, une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ tels que $\forall n \in \mathbb{N},\ u_{\varphi_0(n)} \in B(x_m, 1)$. On pose alors $x_m = a_0$.
   (c) En procédant par récurrence, démontrer que pour tout entier naturel p, on peut construire une suite $(u_{\varphi_0\circ\varphi_1\circ\ldots\circ\varphi_p(n)})_{n\in\mathbb{N}}$ et une boule $B(a_p, \frac{1}{2^p})$ telles que :

$$\forall n \in \mathbb{N},\ u_{\varphi_0\circ\varphi_1\circ\ldots\circ\varphi_p(n)} \in \bigcap_{k=0}^{p} B(a_k, \frac{1}{2^k}).$$

(d) En déduire qu'il existe $\psi : \mathbb{N} \longrightarrow \mathbb{N}$, strictement croissante, telle que : $\forall (p,q) \in \mathbb{N}$: $p \leq q,\ N(u_{\psi(p)} - u_{\psi(q)}) < \frac{1}{2^{p-1}}$.

(e) Conclure.

2.4 Connexité dans les espaces vectoriels normés

2.4.1 Connexité par arcs

Définition 2.17. *Notion de chemin.*
Soit (E, N), un espace vectoriel normé. On appelle ***chemin*** *dans E toute application continue $\varphi : [0,1] \longrightarrow E$. On dit alors que le chemin « joint » $\varphi(0)$ à $\varphi(1)$.*

Exemple : considérons l'espace $E = \mathbb{R}^2$ muni de la norme euclidienne. L'application $\gamma : t \longmapsto (\cos t, \sin t)$ est un chemin qui joint $(1,0)$ à $(-1,0)$.

Définition 2.18. *Connexité par arcs.*
Soit E, un espace vectoriel normé. $X \subset E$ est dit ***connexe par arcs*** *si tute couple de points de X est relié par un chemin qui reste dans X, c'est-à-dire que $t \in [0,1] \Longrightarrow \varphi(t) \in X$.*

Exemple : notion de convexe. Cette notion sera présentée plus en détail dans le chapitre qui suit.

Définition 2.19. *Ensemble convexe.*
On dit que $C \subset E$ est convexe si :

$$\forall t \in [0,1],\ \forall (a,b) \in C^2,\ ta + (1-t)b \in C.$$

Un convexe est connexe par arcs.
En particulier, une boule ouverte est fermée est convexe donc connexe par arcs. Démontrons-le dans le cas d'une boule ouverte. Considérons donc E, un espace vectoriel muni d'une norme N. Soient $x \in E$ et $r > 0$.

$$\begin{aligned} \forall (y,z) \in B(x,r),\ \forall t \in [0,1],\ N(ty + (1-t)z - x) &= N\left(t(y-x) + (1-t)(z-x)\right) \\ &< tN(y-x) + (1-t)N(z-x) < t + (1-t) = 1. \end{aligned}$$

Exemple 2 : $\mathbb{R}^2 \backslash \{(0,0)\}$ est connexe par arcs.

Théorème 2.16. *Soient E_1, ..., E_p, p espaces vectoriels munis de normes N_1, ..., N_p. Soit $E = E_1 \times ... \times E_p$.*
Si pour tout $i \in \{1, ..., p?\}$, A_i est connexe par arcs dans (E_i, N_i), alors $A = A_1 \times ... \times A_p$ est connexe par arcs dans E.

Démonstration. Soit $a = (a_1, ..., a_p) \in A$ et $b = (b_1, ..., b_p) \in A$. Pour tout $i \in \{1, ..., p\}$, il existe $\varphi_i : [0,1] \longrightarrow A_i$, continue, telle que $\varphi_i([0,1]) = A_i$, telle que $\varphi_i(a_i) = b_i$.
Considérons l'application :

$$\begin{array}{rccl} \Psi : & [0,1] & \longrightarrow & A \\ & t & \longmapsto & (\varphi_1(t), ..., \varphi_p(t)) . \end{array}$$

Il est alors évident que :

- Ψ est continue car ses composantes le sont.
- Ψ est un chemin. $\Psi([0,1]) \subset A$.
- On a $\Psi(0) = (a_1, ..., a_p) = a$ et $\Psi(1) = (b_1, ..., b_p) = b$, donc Ψ joint a à b en restant dans A.

□

Théorème 2.17. *Soient E et F, deux espaces vectoriels normés. Soit $f : E \longrightarrow F$, une application continue, et soit A, connexe par arcs dans E.*
Alors, $f(A)$ est connexe par arcs dans F.

Démonstration. Soit $a' \in f(A)$ et $b' \in f(A)$. Alors, il existe a et b, deux éléments de A tels que $a' = f(a)$ et $b' = f(b)$. Il existe φ, une application continue de $[0,1]$ dans A, telle que $\varphi(0) = a$ et $\varphi(1) = b$. Posons alors $\psi = f \circ \varphi$. On a : $\psi : [0,1] \longrightarrow f(A)$, et ψ est tel que $\psi(0) = a'$, $\psi(1) = b'$, et ψ est continue comme composée d'applications continues. Donc ψ est un chemin joignant a' et b' dans $f(A)$. □

Intéressons-nous à présent aux parties connexes par arcs de $\mathbb{R}$. On a la propriété :

Propriété 2.23. *X est une partie de $\mathbb{R}$ connexe par arcs si, et seulement si X est un intervalle, ou encore si, et seulement si X est convexe.*

Démonstration. Il est trivial de remarquer que X est un intervalle si, et seulement si X est convexe (propriété des convexes), ce qui implique que X est connexe par arcs. Reste à démontrer la réciproque. Supposons X connexe par arcs. Soit $(a,b) \in X^2$. Il existe $\varphi : [0,1] \longrightarrow A$, un chemin joignant a à b dans X, tel que $\varphi(0) = a$ et $\varphi(b) = 1$. Or, $\varphi([0,1]) = [m,M]$, où m et M sont deux réels tels que $m < M$, d'après une propriété bien connue des fonctions continues sur un intervalle de $\mathbb{R}$. On a donc $[m,M] \subset A$. On a $a \in [m,M]$ et $b \in [m,M]$, donc $[a,b] \subset [m,M] \subset X$. On en déduit que nécessairement, X est un intervalle de $\mathbb{R}$ (propriété des intervalles de $\mathbb{R}$). □

2.4.2 Introduction aux espaces connexes

Dans cette partie, nous étendons la notion de *connexité par arcs*, introduite dans le paragraphe précédent. Ce paragraphe n'a pas l'ambition d'être exhaustif. Il ne s'agit là que d'une introduction à la notion de connexité et des nombreuses propriétés des espaces connexes ne seront pas citées ici.

Cadre de notre étude : comme à notre habitude, plaçons-nous dans un espace vectoriel normé (E,N). Soit $X \subset E$. Souvenons-nous que par définition, un ouvert de X se définit comme l'intersection de X et d'un ouvert de E. Il en est de même pour les fermés. Dans ce cadre, on parle parfois d'ouverts et de fermés relatifs.

Définition 2.20. *Notion de connexité dans un E.V.N.*
Une partie X de E est dite **connexe** *s'il n'existe pas de partition de X en deux ouverts (relatifs) non vides.*

Il existe deux autres définitions équivalentes de cette notion. Nous avons déjà démontré que E et $\emptyset$ sont à la fois ouverts et fermés. En termes d'ouverts et de fermés relatifs (dans X), X et $\emptyset$ sont à la fois ouverts et fermés également. On a la propriété :

Propriété 2.24. *$X \subset E$ est connexe si, et seulement si l'une ou l'autre des propositions suivantes est vérifiée :*

(i) Il n'existe pas de partition de X en deux fermés (relatifs) non vides.
*(ii) Les **seuls** ensembles à la fois ouverts et fermés de X sont X lui-même et l'ensemble vide.*

Remarque : intuitivement, on peut interpréter les connexes de la façon suivante : un ensemble est connexe s'il est en seul « morceau ». Par exemple, $\mathbb{R}^*$ n'est pas connexe (il est en deux morceaux : $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$). Un cas particulier de la connexité et plus précisément de la connexité par arcs est le théorème des valeurs intermédiaires que je redonne ici pour information :

Théorème 2.18. *Théorème des valeurs intermédiaires.*
Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$, continue sur un intervalle I.
Alors pour tous réels a et b de I, pour tout réel k compris entre $f(a)$ et $f(b)$, il existe au moins un réel c compris entre a et b tel que $f(c) = k$.

La propriété qui suit va nous permettre de faire le lien entre les ensembles connexes et connexes par arcs.

Propriété 2.25. *Soit $X \subset E$. Si X est connexe par arcs, alors X est connexe.*

Démonstration. Raisonnons par l'absurde et supposons X connexe par arcs mais non connexe. Il existerait donc deux ouverts O_1 et O_2 tels que $X = O_1 \cup O_2$, avec $O_1 \cap O_2 = \emptyset$ et pour tout $i \in \{1, 2\}$, $O_i \neq \emptyset$ et $O_i \neq X$. L'idée ici est d'utiliser la caractérisation des ensembles connexes par arcs à l'aide de la continuité. Appelons f, la fonction caractéristique de l'ouvert O_1, c'est à dire la fonction définie par :

$$f(x) = \begin{cases} 1 & \text{si } x \in O_1 \\ 0 & \text{si } x \in O_2 \end{cases} .$$

Soit à présent U, un ouvert de $\mathbb{R}$. Montrons que f est continue. En effet, distinguons plusieurs cas :

- Si $\{0\} \in U$ et si $\{1\} \notin U$, $f^{-1}(U) = O_2$, qui est ouvert.
- Si $\{1\} \in U$ et si $\{0\} \notin U$, $f^{-1}(U) = O_1$, qui est ouvert.
- Si $\{0, 1\} \in U$, $f^{-1}(U) = X$, qui est ouvert (relatif).
- Si $\{0, 1\} \notin U$, $f^{-1}(U) = \emptyset$, qui est ouvert (relatif).

Finalement, on retrouve ici la caractérisation de la continuité à l'aide des images réciproques. On en déduit que f est une application continue. Or, d'après la définition de la connexité par arcs étudiée précédemment, puisque X est connexe par arcs, alors $f(X)$ est un connexe par arcs de $\mathbb{R}$, autrement dit un intervalle. Or, $f(X) = \{0, 1\}$, ce qui est absurde. La propriété est donc démontrée. $\square$

Remarque : rappelons que la notation $f^{-1}(\{0\})$ ne signifie en rien que f est bijective. Nous l'avons déjà introduite dans la section 2.1.3.

Dans le cas de $\mathbb{R}$, on a l'équivalence suivante :

Théorème 2.19. *Dans $\mathbb{R}$, les ensembles connexes par arcs sont exactement les ensembles connexes, autrement dit, ce sont exactement les intervalles de $\mathbb{R}$.*

Nous démontrerons ce théorème en exercice.

2.5 Applications linéaires et continuité

2.5.1 Cas des applications linéaires

Dans cette partie, nous nous intéressons à l'espace des fonctions linéaires continues. Nous allons chercher à normer cet espace, puis à le caractériser. Dans toute cette partie, (E, N) et $(F, \|.\|)$ désigneront deux espaces vectoriels normés sur un corps $\mathbb{K}$. On désigne par :
- $\mathcal{L}(E, F)$, l'ensemble des applications linéaires de E dans F.
- $L(E, F)$, l'ensemble des applications linéaires **continues** de E dans F.

Objectif de cette partie : nous allons normer l'espace des applications linéaires continues, donner des propriétés de cet espace puistenter de caractériser cet espace en dimension finie. Nous montrerons en particulier qu'en dimension finie, on a : $\mathcal{L}(E, F) = L(E, F)$.

On rappelle la définition suivante :

Définition 2.21. *Soient E et F, deux espaces vectoriels.*
Une application f de E dans F est dite linéaire si :
(i) $\forall (x, y) \in E^2$, $f(x + y) = f(x) + f(y)$;
(ii) $\forall a \in \mathbb{R}$, $\forall x \in E$, $f(ax) = af(x)$.

Le théorème qui suit va fournir une réponse partielle aux questions que nous nous posions.

Théorème 2.20. *Soit $u \in \mathcal{L}(E, F)$. Alors, les propositions suivantes sont équivalentes :*
(i) u est continue sur E.
(ii) u est continue en 0_E.
(iii) u est bornée sur $\mathcal{B} := \mathcal{B}_f(0_E, 1)$.
(iv) u est bornée sur $\mathcal{S} := \{x \in E : N(x) = 1\}$ (sphère unité).
(v) Il existe $k \in \mathbb{R}_+$ tel que $\forall x \in E$, $\|u(x)\| \leq kN(x)$.
(vi) u est uniformément continue sur E.

Démonstration. Démontrons ces propriétés dans l'ordre :
- $(i) \Longrightarrow (ii)$ est trivial.
- $(ii) \Longrightarrow (iii)$: Par définition de la continuité de u en 0_E (en choisissant $\varepsilon = 1$ dans la définition de la continuité), on obtient l'existence de $\eta > 0$ tel que $N(x - 0_E) < \eta \Longrightarrow \|u(x) - u(0_E)\| < 1$. Soit à présent $y \in \mathcal{B}$. On a donc $N(y) \leq 1$, et par conséquent, $N\left(\frac{\eta}{2}y\right) \leq \frac{\eta}{2} < \eta$ et on en déduit que $\left\|u\left(\frac{\eta}{2}y\right)\right\| < 1$, autrement dit : $\forall y \in \mathcal{B}$, $\|u(y)\| < \frac{2}{\eta}$.
- $(iii) \Longrightarrow (iv)$ est trivial, car $\mathcal{S} \subset \mathcal{B}$.
- $(iv) \Longrightarrow (v)$: u est bornée sur $\mathcal{S}$, c'est-à-dire qu'il existe $k > 0$ tel que $\forall x \in \mathcal{S}$, $\|u(x)\| \leq k$. Si $x \neq 0_E$, on a $\frac{x}{N(x)} \in \mathcal{S}$. On en déduit que $\left\|u\left(\frac{x}{N(x)}\right)\right\| \leq k$, et puisque u est linéaire, on en déduit que $\forall x \in E$, $\|u(x)\| \leq kN(x)$, ce qu'on voulait démontrer.
- $(v) \Longrightarrow (vi)$: c'est clair. En effet, si x et y sont deux éléments de E, alors $\|u(x - y)\| \leq kN(x - y)$, donc u est Lipschitzienne, donc uniformément continue.
- $(vi) \Longrightarrow (i)$ est trivial. C'est donc fini.

$\square$

Le théorème suivant va nous permettre d'introduire une définition de la norme d'une application linéaire continue (que nous appellerons triple norme). De plus, nous démontrerons prochainement que, en dimension finie, toutes les applications linéaires sont nécessairement continues. Mais tant que cela n'est pas établi, je continuerai à utiliser l'ensemble des applications linéaires continues plutôt que l'ensemble des applications linéaires.

Théorème 2.21. *Soit $u \in \mathcal{L}(E,F)$. Supposons de plus que u est continue. Posons :*
- $N_1 := \sup\limits_{x \in E\setminus\{0_E\}} \left(\dfrac{\|u(x)\|}{N(x)}\right)$.
- $N_2 := \sup\limits_{x\in\mathcal{S}}(\|u(x)\|)$.
- $N_3 := \sup\limits_{x\in\mathcal{B}}(\|u(x)\|)$.
- $N_4 := \inf\{k \in \mathbb{R}_+^* : \forall x \in E,\ \|u(x)\| \leq kN(x)\}$.

Alors, on a : $N_1 = N_2 = N_3 = N_4$.

Démonstration. On procéde en établissant des inégalités :
- $N_2 \leq N_1$: par définition du suprémum, pour tout $x \neq 0_E$, on a : $\dfrac{\|u(x)\|}{N(x)} \leq N_1$. Donc, si $x \in \mathcal{S}$, on en déduit immédiatement que $\|u(x)\| \leq N_1$ donc $N_2 \leq N_1$.
- $N_1 \leq N_2$: en effet, $\dfrac{x}{N(x)} \in \mathcal{S}$ donc $\left\{\dfrac{\|u(x)\|}{N(x)}, x \in E\backslash 0_E\right\} \subset \{\|u(x)\|, x \in \mathcal{S}\}$. Un passage au sup fournit l'inégalité souhaitée.
- $N_2 \leq N_3$: en effet, cela vient du fait que $\mathcal{S} \subset \mathcal{B}$.
- $N_3 \leq N_4$: Si $x \in \mathcal{S}$, alors $\|u(x)\| \leq N_4$ d'où l'inégalité.
- $N_2 \leq N_4$: on utilise le même raisonnement que ci-dessus.
- $N_4 \leq N_2$: $\forall x \in E\setminus\{0_E\}$, on a : $\left\|u\left(\dfrac{x}{N(x)}\right)\right\| \leq N_2$. On en déduit que $\forall x \in E,\ \|u(x)\| \leq N_2N(x)$. On en déduit que $N_4 \leq N_2$.

□

Ce théorème étant établi, il est possible de norme l'espace $\mathcal{L}(E,F)$. C'est l'objet de la définition qui suit.

Définition 2.22. *Si u est une application linéaire continue de E dans F, on pose :*

$$|||u||| := N_1 = N_2 = N_3 = N_4.$$

*$|||.|||$ s'appelle la **triple norme** ou norme subordonnée à la norme de E et F. On a en particulier :*

$$\forall x \in E,\ \|u(x)\| \leq |||u|||\ .\|x\|.$$

La conséquence directe de cette définition est énoncée dans le théorème qui suit :

Théorème 2.22.
(i) $L(E,F)$ est un sous-espace vectoriel de $\mathcal{L}(E,F)$.
(ii) L'application $\begin{array}{ccc} L(E,F) & \longrightarrow & \mathbb{R} \\ u & \longmapsto & |||u||| \end{array}$ *est une norme sur $L(E,F)$.*

Démonstration. (i) et (ii) vont être démontrés simultanément. $L(E,F) \neq \emptyset$ car l'application identiquement nulle est continue. De plus, il est bien évident que si λ est un scalaire et u un élément de $L(E,F)$, alors $\lambda.u$ est encore un élément de $L(E,F)$. Il nous reste donc à démontrer que si $u \in L(E,F)$ et $v \in L(E,F)$, alors $u+v \in L(E,F)$. Soit $x \in \mathcal{B}$, la boule unité fermée de E. Alors, $\|u(x)+v(x)\| \leq \|u(x)\| + \|v(x)\|$. Or, $\|u(x)\| \leq |||u|||$ et $\|v(x)\| \leq |||v|||$. On en déduit donc que $\|u(x)+v(x)\| \leq |||u||| + |||v|||$, ce qui montre que $v+v \in L(E,F)$ (propriété énoncée précédemment). Un passage à la borne supérieure, pour $x \in \mathcal{B}$ démontre donc que :

$$|||u+v||| \leq |||u||| + |||v||| .$$

Les autres propriétés sur cette norme étant aisées à vérifier, je laisse ce soin au lecteur. □

Théorème 2.23. *Soient E, F et G, trois espaces vectoriels normés, $u \in L(E,F)$ et $v \in L(F,G)$. Alors, $v \circ u \in L(E,G)$ et $|||v \circ u||| \leq |||v|||\,.\,|||u|||$.*

Démonstration. Il est clair que $v \circ u$ est linéaire (si cela n'est pas totalement clair pour vous, n'hésitez pas à le redémontrer rapidement...). De plus, soit $x \in E$. Alors, $\|v \circ u(x)\| = \|v[u(x)]\| \leq |||v|||\,\|u(x)\| \leq |||v|||\,.\,|||u|||\,.\|x\|$. Une division par $\|x\|$, puis un passage au suprémum prouvent que $v \circ u$ est bien continue et que $|||v \circ u||| \leq |||v|||\,.\,|||u|||$. □

Quelques exemples :

- **Exemple 1 :** on considère $\mathcal{C}([0,\pi])$, l'espace des fonctions f continues sur $[0,\pi]$ à valeurs réelles . Si $f \in \mathcal{C}([0,\pi])$, il est tout à fait clair que les intégrales $\int_0^\pi |f(x)|dx$ et $\int_0^\pi [f(x)]^2 dx$ existent et sont finies. On peut normer $\mathcal{C}([0,\pi])$ par la norme $\|.\|_1$ définie pour $f \in \mathcal{C}([0,\pi])$ par : $\|f\|_1 = \int_0^\pi |f(x)|dx < +\infty$, ou encore par la norme $\|.\|_2$ définie pour $f \in \mathcal{C}([0,\pi])$ par : $\|f\|_2 = \left(\int_0^\pi [f(x)]^2 dx\right)^2$. Je vous laisse le soin de que ces applications définissent des normes.
  Soit a_0, une fonction continue sur $[0,\pi]$. Considérons alors l'application φ définie par :

$$\begin{array}{rccc} \varphi : & (\mathcal{C}([0,\pi]), \|.\|_2) & \longrightarrow & (\mathcal{C}([0,\pi]), \|.\|_1) \\ & f & \longmapsto & a_0 f \end{array}$$

  Alors, φ est linéaire et continue. Il est évident que φ est une appllication linéaire. En effet, si λ est un réel et f et g, deux fonctions de $\mathcal{C}([0,\pi])$, on vérifie très facilement que l'on a : $\varphi(f+\lambda.g) = \varphi(f) + \lambda.\varphi(g)$. Reste à prouver que φ est continue. On va le démontrer, puis tenter de calculer $|||\varphi|||$. Mais au préalable, il faut montrer que l'application φ est bien définie, autrement dit, que si l'on choisit $f \in \mathcal{C}([0,\pi])$, alors, $\varphi(f) \in \mathcal{C}([0,\pi])$. C'est immédiat, car un produit de fonctions continues est continu.
  Prouvons à présent que l'application φ est continue. Soit $f \in \mathcal{C}([0,\pi])$. Posons $g := \varphi(f)$. On a, en utilisant l'inégalité de Cauchy-Schwarz :

$$\|\varphi(f)\|_1 := \int_0^\pi |g(x)|dx = \int_0^\pi |a_0(x)f(x)|dx \leq \left(\int_0^\pi a_0^2(x)dx\right)^{\frac{1}{2}} \left(\int_0^\pi f^2(x)dx\right)^{\frac{1}{2}} = \|a_0\|_2.\|f\|_2.$$

  Cela prouve donc que φ est continue en $0_{\mathcal{C}([0,\pi])}$ donc sur $\mathcal{C}([0,\pi])$ tout entier d'après le théorème 2.20. De plus, on obtient également l'information suivante : $|||\varphi||| \leq \|a_0\|_2$. On va démontrer qu'en fait, cette inégalité est une égalité. En effet, il suffit de supposer que f est la fonction a_0 (cas d'égalité dans l'inégalité de Cauchy-Schwarz). Alors, $\|\varphi(f)\|_1 = \|a_0\|_2^2$ et on en déduit (par définition de la borne supérieure) que $|||\varphi||| \geq \|a_0\|_2$. Finalement, on a donc : $|||\varphi||| = \|a_0\|_2$.

- **Exemple 2 :** on appelle $\mathbb{R}[X]$ l'espace des polynômes à coefficients réels. On munit cet espace de la norme $\|.\|$ définie pour $P \in \mathbb{R}[X]$, s'écrivant sous la forme $P = \sum_{n=0}^{\deg P} a_n X^n$, par : $\|P\| := \sup_{n\in\{0,\ldots,\deg P\}} |a_n|$. Soit $x_0 \in \mathbb{R}$, tel que $|x_0| < 1$. On appelle u l'application (linéaire, mais à vous de le vérifier) définie par :
$$u = \begin{array}{rcl} (\mathbb{R}[X], \|.\|) & \longrightarrow & (\mathbb{R}, |.|) \\ P & \longmapsto & P(x_0) \end{array}.$$
On va démontrer que u est continue et calculer sa norme subordonnée. Soit $P \in \mathbb{R}[X]$, s'écrivant sous la forme $P = \sum_{n=0}^{\deg P} a_n X^n$. On sait que $u(P) = \sum_{n=0}^{\deg P} a_n x_0^n$. On peut donc écrire que :
$$|P(x_0)| \leq \|P\|. \sum_{n=0}^{\deg P} |x_0|^n = \|P\|.\frac{1 - |x_0|^{\deg P+1}}{1 - |x_0|} \leq \frac{\|P\|}{1 - |x_0|}.$$
Cette inégalité nous prouve en particulier que u est continue. Nous allons à présent construire une suite d'éléments de $\mathbb{R}[X]$ de norme 1, que nous allons tenter de faire converger vers $|||u|||$. (caractérisation de la borne supérieure à l'aide de suites) Choisissons pour tout entier n, $Q_n := \sum_{k=0}^{n} \varepsilon_k X^k$, où ε_k désigne le signe de x_0^k. Il est clair que pour tout entier n, $\|Q_n\| = 1$. On a :
$$u(Q_n) = \sum_{k=0}^{n} |x_0|^k = \frac{1 - |x_0|^{n+1}}{1 - |x_0|}.$$
On en déduit que pour tout entier naturel n, on a : $\frac{|u(Q_n)|}{\|Q_n\|} \leq |||u||| \leq \frac{1}{1 - |x_0|}$. Un passage à la limite démontre donc que : $|||u||| = \frac{1}{1 - |x_0|}$.

- **Exemple 3 :** réutilisons les hypothèses précédentes, mais imposons cette fois que $|x_0| = 1$. Alors, pour le même choix de polynômes $(Q_n)_{n\in\mathbb{N}}$ que précédemment, on a : $u(Q_n) = 1 + 1 + \ldots + 1 = n + 1$. Donc, le rapport $\frac{|u(Q_n)|}{\|Q_n\|}$ est non borné, et par conséquent, u n'est pas continue.

À présent que l'espace $L(E, F)$ est normé, on est en droit de se demander si cet espace est complet. Le théorème qui suit répond à cette question.

Théorème 2.24. *Si F est un $\mathbb{K}$-espace vectoriel normé complet, alors $L(E, F)$ est également un espace vectoriel normé complet.*

Démonstration. Soit $(f_n)_{n\in\mathbb{N}}$, une suite de Cauchy d'éléments de $L(E, F)$. Soit x, un élément quelconque fixé de E. On a donc :
$$\forall (p, q) \in \mathbb{N}^2,\ \|f_p(x) - f_q(x)\| \leq |||f_p - f_q|||\ .N(x).$$
On déduit de cette inégalité que la suite $(f_n(x))_{n\in\mathbb{N}}$ est de Cauchy dans F qui est complet ; par conséquent, cette suite est convergente dans F. Notons $f(x)$, sa limite. Il nous faut vérifier que la suite $(f_n)_{n\in\mathbb{N}}$ converge bien vers f au sens de $|||.|||$ et que f est linéaire et continue. (autrement dit que f est bien un élément de $L(E, F)$).

- On sait que $(f_n)_{n\in\mathbb{N}}$ est une suite de Cauchy. Soit $\varepsilon > 0$. Il existe donc $N \in \mathbb{N}$ tel que : $q > p \geq N \Longrightarrow |||f_p - f_q||| \leq \varepsilon$. En particulier, pour $x \in E$, cela implique que : $\|f_p(x) - f_q(x)\| \leq \varepsilon N(x)$. Faisons alors tendre q vers $+\infty$. On obtient que $\|f_p(x) - f(x)\| \leq \varepsilon N(x)$, autrement dit que $|||f_p - f||| \leq \varepsilon$, donc $(f_n)_{n\in\mathbb{N}}$ converge vers f dans $L(E, F)$.

- La linéarité découle simplement d'un passage à la limite, en écrivant que : $\forall n \in \mathbb{N}, \forall (\lambda, \mu) \in \mathbb{K}^2, \forall (x,y) \in E^2,\ f_n(\lambda.x + \mu.y) = \lambda f_n(x) + \mu.f_n(y)$.
- Pour la continuité, on peut procéder de plein de façons. Puisque $(f_n)_{n\in\mathbb{N}}$ est une suite de Cauchy, elle est bornée dans $L(E,F)$. Un passage à la limite prouve alors que f est également bornée dans $L(E,F)$, autrement dit que f est continue.

Conclusion : $L(E,F)$ est complet. □

2.5.2 Applications linéaires en dimension finie

Dans ce paragraphe, nous allons appliquer les résultats précédents à la dimension finie. Mais au préalable, complétons les informations que nous possédons déjà dans ce domaine. Souvenons-nous qu'en dimension finie, toutes les normes sont équivalentes.

Théorème 2.25. *Si E est un $\mathbb{K}$-espace vectoriel de dimension finie, alors : $X \subset E$ est compacte, si et seulement si X est fermée et bornée.*

Démonstration. On suppose que E est de dimension n. Puisque toutes les normes sont équivalentes, on munit E de la norme N définie pour tout $x = (x_1, ..., x_n) \in E$ par $\|x\| = \sup_{i\in\{1,...,n\}} |x_i|$.
On a déjà démontré précédemment que si X est compacte, alors X est nécessairement fermée et bornée.
Supposons donc que X désigne une partie fermée et bornée de E. Nous allons démontrer que X est compacte. Appelons φ l'application définie par : $\varphi : \begin{array}{ccc} \mathbb{K}^n & \longrightarrow & E \\ (x_1, ..., x_n) & \longmapsto & \sum_{i=1}^n x_i e_i \end{array}$, où $(e_i)_{1\le i\le n}$ désigne une base de E. On munit $\mathbb{K}^n$ de la norme $\|.\|_\infty$ usuelle, définie dans le chapitre précédent. φ est donc linéaire et **isométrique**, c'est-à-dire que $\forall x \in E,\ N(\varphi(x)) = \|x\|_\infty$. φ étant isométrique est donc continue et bijective. $\varphi(X)$ est fermée car φ est continue. De plus $\varphi^{-1}(X)$ est bornée car φ est isométrique. On en déduit que $\varphi^{-1}(X)$ est un compact de $\mathbb{K}^n$. Puisque φ est continue, on en déduit que $X = \varphi[\varphi^{-1}(X)]$ est un compact de X. □

Théorème 2.26. *Soit F, un espace vectoriel de dimension finie. Alors, F est complet.*
Plus généralement, tout sous-espace vectoriel F, de dimension finie dans un espace vectoriel E de dimension quelconque est complet.

Démonstration. Pour le démontrer, on réutilise l'application φ introduite précédemment. Il ets aisé de démontrer que φ est continue, que φ^{-1} l'est aussi. Et puisque $\mathbb{K}^n$ est complet et que $E = \varphi(\mathbb{K}^n)$, on en déduit que E est complet, d'après le théorème 2.3. □

Venons-en au théorème principal de cette section.

Théorème 2.27. *Soient E et F, deux espaces vectoriels normés, avec E de dimension finie. Alors $\mathcal{L}(E,F) = L(E,F)$, autrement dit, toute application linéaire dans E de dimension finie est continue.*

Démonstration. Soit $\|.\|$, une norme de F et N, une norme de E. Soit $\mathcal{B} = (e_i)_{1\le i\le n}$, une base de E. Appelons u, une application linéaire de E dans F. On va montrer que u est continue.

Tout élément x de E s'écrit donc : $x = \sum_{i=1}^{n} x_i e_i$, et $\|u(x)\| = \left\| \left[\sum_{i=1}^{n} x_i u(e_i) \right] \right\| \leq \sum_{i=1}^{n} |x_i| \|u(e_i)\|$.
Prenons alors $N(x) = \sup_{1 \leq i \leq n} |x_i|$, dans E. On a :

$$\|u(x)\| \leq N(x) \sum_{i=1}^{n} \|u(e_i)\|. \leq k.N(x), \text{ où } k \text{ est une constante.}$$

On retrouve ici la condition de continuité : u est continue. □

Remarque : la norme subordonnée définit également une norme matricielle, puisque toute application linéaire peut être repérée par sa matrice en dimension finie. Si $u \in \mathcal{L}(E, F)$, avec E de dimension n et F de dimension finie p, alors, en appelant M, la matrice de u dans une base $\mathcal{B}$ donnée, on a :

$$|||M||| = |||u||| = \sup_{x \in E \setminus \{0_E\}} \left(\frac{\|u(x)\|}{N(x)} \right) = \sup_{X \in \mathbb{R}^n \setminus \{0_{\mathbb{R}^n}\}} \left(\frac{\|MX\|_{\mathbb{R}^p}}{\|X\|_{\mathbb{R}^n}} \right).$$

Exercice : Démontrer que si $f : (\mathbb{R}^n, \|.\|_\infty) \longrightarrow (\mathbb{R}^n, \|.\|_\infty)$ est une application linéaire dont la matrice dans une base $\mathcal{B}$ donnée est $M = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, alors :

$$|||M||| = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

2.5.3 Cas des applications multilinéaires

Procédons au préalable à quelques rappels. On introduit ici la notion de forme multilinéaire. Dans tout ce paragraphe, E désignera un $\mathbb{K}$-espace vectoriel de dimension finie.

Définition 2.23. *Forme multilinéaire ou forme p-linéaire.*
Soit $f : E^p \longrightarrow \mathbb{K}$. On dit que f est une forme p-linéaire si, et seulement si pour tous $(x_1, ..., x_p) \in E^p$, l'application $x \longmapsto f(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_p)$ est linéaire de E dans $\mathbb{K}$.
On note alors $\mathcal{L}_p(E, \mathbb{K})$ l'ensemble des formes p-linéaires de E^p dans $\mathbb{K}$.

Exemple : dans $\mathbb{R}^n$, considérons l'application :

$$\begin{array}{rrcl} \varphi : & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x = (x_1, ..., x_n) & \longmapsto & x_1 x_2 ... x_n = \prod_{k=1}^{n} x_i. \end{array}$$

φ est une application n-linéaire.

Remarque 1 : il faut veiller à ne pas confondre $\mathcal{L}_p(E, \mathbb{K})$ avec $\mathcal{L}(E^p, \mathbb{K})$, l'ensemble des applications linéaires de E^p à valeurs dans $\mathbb{K}$.

Remarque 2 : l'espace $\mathcal{L}_p(E, \mathbb{K})$ consitue un $\mathbb{K}$-espace vectoriel.

Cadre de notre étude : considérons $E = E_1 \times E_2 \times ... \times E_p$, avec $p \geq 2$, un espace vectoriel s'écrivant comme le produit cartésien de n espaces vectoriels. On munit chaque espace vectoriel E_i d'une norme $\|.\|_{E_i}$ et E de la norme $\|.\|_E$ définie pour $x = (x_1, ..., x_p) \in E$ par :

$$\|x\|_E = \sup_{i \in \{1,...,p\}} \|x_i\|_{E_i}.$$

F désignera un autre espace vectoriel muni d'une norme $\|.\|_F$.

Notation : appelons $\mathcal{B}$, la boule unité de E. Il est aisé de vérifier que $\mathcal{B} = \mathcal{B}_1 \times \mathcal{B}_2 \times ... \times \mathcal{B}_n$, où B_i désigne a boule unité de E_i, avec $i \in \{1, ..., p\}$. On appelle $\mathcal{S}$, la sphère unité de E.

Nous allons à présent énoncer un théorème qui ressemble fort au résultat que nous avons énoncé précédemment, relatif aux applications linéaires continues. Il s'agit d'ailleurs d'une généralisation du théorème précédent.

Théorème 2.28. *Soit u, une application p-linéaire de E dans F. Alors, les conditions suivantes sont équivalentes :*
(i) u est continue sur E.
(ii) u est continue en 0_E.
(iii) u est bornée sur $\mathcal{B}$.
(iv) $\exists k > 0 : \forall x \in E,\ \|u(x)\|_F \leq k\|x_1\|_{E_1} \times ... \times \|x_p\|_{E_p}$, avec $x = (x_1, ..., x_p)$.

Démonstration. Les techniques utilisées précédemment sont à peu près les mêmes :

- $(i) \Longrightarrow (ii)$ est trivial.
- $(ii) \Longrightarrow (iii)$: On sait que $u(0_E) = 0_F$. Fixons-nous $\varepsilon = 1$ dans la définition de la continuité de u en 0_E. On obtient donc qu'il existe $\eta > 0$ tel que $\|x\|_E < \eta \Longrightarrow \|u(x)\|_F < 1$.Si $y \neq 0_E$, tel que $y \in \mathcal{B}$, alors il est clair que $\left\| \frac{\eta}{2}.y \right\|_E < \eta$. On en déduit immédiatement que $\|u(y)\|_F < \frac{\eta}{2}$, autrement dit (iii) est vérifiée.
- $(iii) \Longrightarrow (iv)$, car $\mathcal{S} \subset \mathcal{B}$.
- $(iv) \Longrightarrow (v)$: supposons l'existence de $k > 0$ tel que : $\forall x \in \mathcal{S},\ \|u(x)\|_F \leq k$. Soit $x = (x_1, ..., x_p) \in E$. Si au moins une coordonnée x_i de x vaut 0_{E_i}, alors $u(x) = 0$. Sinon, posons $y = \left(\frac{x_1}{\|x_1\|_{E_1}}, ..., \frac{x_1}{\|x_p\|_{E_p}} \right)$. Il est clair que $y \in \mathcal{S}$. traduisons donc le fait que $\|u(y)\|_F \leq k$. Cela s'écrit encore :
$$\left\| u\left(\frac{x_1}{\|x_1\|_{E_1}}, ..., \frac{x_1}{\|x_p\|_{E_p}} \right) \right\| \leq k.$$
Il suffit juste d'utiliser la multilinéarité de u pour obtenir qu'alors : $\|u(x)\|_F \leq k\|x_1\|_{E_1} \times ... \times \|x_p\|_{E_p}$, et c'est ce que l'on souhaitait prouver.
- $(v) \Longrightarrow (i)$: la preuve étant un peu technique, je fais la preuve dans le cas où $p = 2$, ce qui me permettra d'alléger les notations. Soit $(a_1, a_2) \in E_1 \times E_2 = E$. Écrivons que $\forall (x, y) \in E_1 \times E_2,\ u(x, y) - u(a_1, a_2) = u(x - a_1, y) + u(a_1, y - a_2)$. Par conséquent, $\|u(x, y) - u(a_1, a_2)\|_F \leq k\|x - a_1\|_{E_1}\|y\|_{E_2} + k\|a_1\|_{E_1}\|y - a_2\|_{E_2}$. On se place alors dans la boule de centre (a_1, a_2) et de rayon 1 (on s'intéresse ici à démontrer localement la continuité de u). Par conséquent, $\|y - a_2\| \leq 1 \Longrightarrow \|y\| \leq \|a_2\| + 1$, par inégalité triangulaire. Cela signifie donc que :
$$\|u(x, y) - u(a_1, a_2)\|_F \leq k(\|a_2\|_{E_2} + 1)\|x - a_1\|_{E_1} + k\|a_1\|_{E_1}.\|y - a_2\|_{E_2}.$$
Or, d'après la norme choisie sur E, on en déduit que :
$$\|u(x, y) - u(a_1, a_2)\|_F \leq \lambda\|(x, y) - (a_1, a_2)\|_E,\ \text{avec } \lambda = \max(\|a_1\|_{E_1}, \|a_2\|_{E_2} + 1).$$
Ceci prouve que u est localement lipschitzienne donc continue.

$\square$

Remarque : les applications multilinéaires continues ne sont pas *a priori* uniformément continues ; cependant, si $(u_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de E, alors $(f(u_n))_{n\in\mathbb{N}}$ est une suite de Cauchy de F.

Propriété 2.26. *Notons $\mathcal{L}(E_1, ..., E_p, F)$, l'espace des applications multilinéaires de $E_1 \times ... \times E_p$ dans F. Cet espace est un espace vectoriel normé dont la norme est donnée par :*

$$\forall u \in \mathcal{L}(E_1, ..., E_p, F),\ |||u||| := \sup_{x_i \in E_i} \frac{\|u(x_1, ..., x_p)\|_F}{\|x_1\|_{E_1} \times ... \times \|x_p\|_{E_p}}.$$

La démonstration ressemble beaucoup à celle du cas des applications linéaires continues. Je la laisse au lecteur.

Exemples :

- Si E est un $\mathbb{K}$-espace vectoriel, alors l'application $\begin{array}{ccc} \mathbb{K} \times E & \longrightarrow & E \\ (\lambda, x) & \longmapsto & \lambda.x \end{array}$. En effet, on utilise la condition (v), car $\|\lambda.x\| = |\lambda|.\|x\|$.
- De la même façon, soit u, un endomorphisme continu de E, c'est-à dire $u \in L(E, E) =: L(E)$. L'application $\begin{array}{ccc} E \times L(E) & \longrightarrow & E \\ (x, u) & \longmapsto & u(x) \end{array}$ est bien-sûr bilinéaire et continue, car : $\forall (x, u) \in E \times L(E)$, $\|u(x)\|_E \leq |||u\ |||\ .\|x\|_E$. On utilise encore ici la condition (v).
- Soit un espace euclidien E, c'est-à-dire un espace muni d'un produit scalaire noté $< .,. >$. Le produit scalaire est une forme bilinéaire continue en vertu de l'inégalité de Cauchy-Schwarz : $\forall (x, y) \in E^2,\ | < x, y > | \leq \|x\|_E.\|y\|_E$.

Plus généralement, et de façon analogue au cas des applications linéaires continues, on a le théorème :

Théorème 2.29. *Soit $u : E := E_1 \times ... \times E_p \longrightarrow F$, une application p-linéaire.*
Supposons que E est un espace de dimension finie. Alors, u est une application continue.

C'est très simple à démontrer car l'idée de preuve s'inspire largement des démonstrations précédentes. Il est préférable de la faire dans le cas où $p = 2$, toujours dans le but d'alléger au maximum les notations, ce qui permet d'étendre ensuite cette démonstration au cas p quelconque.
Supposons que F est normé par la norme $\|.\|$. Supposons que $\dim E_1 = r$ et $\dim E_2 = s$, et que E_1 et E_2 sont normés par les application N_1 et N_2 respectivement définies par : $N_1(x_1, ..., x_r) = \sup_{1\leq i\leq r} |x_i|$ et $N_2(y_1, ..., y_s) = \sup_{1\leq j\leq s} |y_j|$. On appelle alors $(e_1, ..., e_r)$, une base de E_1 et $(f_1, ..., f_s)$, une base de E_2.
Soit $x \in E$. Alors x s'écrit sous la forme $x = \left(\sum_{i=1}^{r} \lambda_i.e_i, \sum_{i=1}^{s} \mu_j.f_j\right)$. Il s'ensuit que $u(x) = \sum_{i=1}^{r}\sum_{j=1}^{s} \lambda_i\mu_j u(e_i, f_j)$. On a alors :

$$\|u(x)\| \leq \sum_{i=1}^{r}\sum_{j=1}^{s} |\lambda_i\mu_j|\|u(e_i, f_j)\| \leq N_1(x_1)N_2(x_2)\sum_{i=1}^{r}\sum_{j=1}^{s} \|u(e_i, f_j)\|.$$

En posant $\kappa = \sum_{i=1}^{r}\sum_{j=1}^{s} \|u(e_i, f_j)\|$, on a :

$$\forall x \in E,\ \|u(x)\| \leq \kappa N_1(x_1)N_2(x_2).$$

On retrouve la caractérisation de la continuité d'une application linéaire. u est donc continue.

Chapitre 3

Introduction à l'Analyse Fonctionnelle

3.1 Espaces préhilbertiens réels et complexes

Remarque sur les notations : dans toute cette partie, on munira les espaces condidérés de produits. Ces produits seront indifféremment notés $< .,. >$ ou $\varphi(.,.)$.

3.1.1 Espaces euclidiens et préhilbertiens réels

Définition 3.1. *Espaces euclidien et préhilbertien réel.*
- *Soit E, un $\mathbb{R}$-espace vectoriel* ***de dimension finie****. Si E est muni d'un produit scalaire $< .,. >$, on dit que E est un espace* ***euclidien****.*
- *Dans le cas où E est un $\mathbb{R}$-espace vectoriel de dimension quelconque, muni d'un produit scalaire $< .,. >$, on dit que E est un espace* ***préhilbertien réel****.*

Remarque 1 : un espace euclidien est en particulier préhilbertien réel.

Remarque 2 : dans un espace préhilbertien réel, l'inégalité de Cauchy-Schwarz est réalisée. Si E est muni d'une norme N,

$$\forall (x,y) \in E^2,\ | < x,y > | \leq N(x)N(y).$$

Rappel important : tout espace vectoriel normé E muni d'un produit scalaire peut être normé par la norme dite **induite**, notée $\|.\|$ et définie par : $\forall x \in E,\ \|x\| = \sqrt{< x,x >}$.

Je ne redonne pas ici les résultats relatifs aux espaces euclidiens, que le lecteur pourra aisément trouver dans des ouvrages d'Algèbre linéaire.

Remarque 3 : si E est un espace préhilbertien réel, complet, on dit que E est un espace de Hilbert. Nous l'étudierons ultérieurement.

Je donne à présent quelques exemples d'espaces préhilbertiens réels :
- Soit $[a,b]$, un intervalle de $\mathbb{R}$. On appelle $\mathcal{C}([a,b],\mathbb{R})$, l'espace des fonctions continues sur le segment $[a,b]$. On peut munir $\mathcal{C}([a,b],\mathbb{R})$ du produit scalaire $< .,. >$ défini par :

$$\forall (f,g) \in \mathcal{C}([a,b],\mathbb{R}),\ < f,g >= \int_a^b f(t)g(t)dt.$$

$\mathcal{C}([a,b],\mathbb{R})$ est donc un espace préhilbertien réel.

- On désigne par $\ell^2(\mathbb{R})$ l'espace défini par : $\ell^2(\mathbb{R}) := \left\{(u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \sum_{n=0}^{+\infty} u_n^2 < +\infty\right\}$. On peut munir $\ell^2(\mathbb{R})$ du produit scalaire $< .,. >$ défini par :
$$\forall (u,v) \in \left(\ell^2(\mathbb{R})\right)^2,\ < u,v >= \sum_{n=0}^{+\infty} u_n v_n.$$
$\ell^2(\mathbb{R})$ est donc un espace préhilbertien réel. On remarque, en utilisant l'inégalité $\forall n \in \mathbb{N},\ |u_n v_n| \leq \frac{1}{2}(u_n^2 + v_n^2)$ que le produit scalaire est toujours défini.
- On désigne par $LC^2(I,\mathbb{R})$, l'espace des fonctions continues de carré intégrable, c'est-à-dire $\{f \in LC^2(I,\mathbb{R}), \int_I f^2(t)dt < +\infty\}$. Cet espace est muni du produit scalaire $< .,. >$ défini par :
$$\forall (f,g) \in \left(LC^2(I,\mathbb{R})\right)^2,\ < f,g >= \int_I f(t)g(t)dt.$$

3.1.2 Espaces préhilbertiens complexes

Dans ce paragraphe, E désignera un $\mathbb{C}$-espace vectoriel.

Définition 3.2.

(i) ***Forme sesquilinéaire :*** *une forme sesquilinéaire sur E est une application de E^2 dans $\mathbb{C}$, notée $< .,. >$, semi-linéaire à gauche et linéaire à droite, c'est-à-dire :*
$$\begin{aligned}\forall (x,y,z) \in E^3,\ \forall (\lambda,\mu) \in \mathbb{C}^2,\ < \lambda x + \mu y, z > &= \overline{\lambda}. < x,z > + \overline{\mu}. < y,z > \\ \text{et } < x, \lambda.y + \mu.z > &= \lambda < x,y > + \mu < x,z > .\end{aligned}$$

(ii) ***Forme hermitienne :*** *une forme sesquilinéaire $< .,. >$ sur un $\mathbb{C}$-espace vectoriel est dite* ***hermitienne*** *ou* ***à symétrie hermitienne*** *lorsque :*
$$\forall (x,y) \in E^2,\ < x,y >= \overline{< y,x >}.$$

(iii) ***Espace préhilbertien complexe :*** *un espace préhilbertien complexe est un espace muni d'un produit hermitien.*

(iv) ***Espace hermitien :*** *on appelle espace hermitien un espace préhilbertien complexe de dimension finie.*

(v) ***Norme induite :*** *la norme induite par le produit hermitien, notée $\|.\|$ est définie par :*
$$\forall x \in E,\ \|x\| = \sqrt{< x,x >}.$$

Remarque : un espace préhilbertien réel est en particulier un espace préhilbertien complexe.

Écrivons et démontrons l'inégalité de Cauchy-Schwarz dans un espace préhilbertien complexe (ou hermitien).

Propriété 3.1. *Inégalité de Cauchy-Schwarz.*
Soit E, un espace préhilbertien complexe muni d'un produit hermitien $< .,>$. On note $\|.\|$, la norme induite associée. Alors,
$$\forall (x,y) \in E^2,\ | < x,y > | \leq \|x\|.\|y\|.$$
Cette inégalité est une égalité si, et seulement si x et y sont liés, c'est-à-dire s'il existe $\lambda \in \mathbb{C}$ tel que $x = \lambda.y$.

Démonstration. Soient x et y, deux éléments de E. On appelle P, le polynôme défini pour $\lambda \in \mathbb{C}$ par : $P(\lambda) = \|x + \lambda y\|^2$. P et un polynôme de degré 2. De plus, $\forall \lambda \in \mathbb{C}$, $P(\lambda \geq 0$. Or,

$$P(\lambda) = \|x\|^2 + \lambda < y, x > +\lambda < x, y > +|\lambda|^2\|y\|^2.$$

Désignons par θ, un argument du nombre complexe $< x, y >$, avec la convention $\theta \in [0, 2\pi[$. Choisissons alors $\lambda = \mu.e^{-i\theta}$, où μ décrit $\mathbb{R}$ tout entier. L'égalité précédente devient donc :

$$P(\lambda) = \|x\|^2 + 2| < x, x > |\mu + \mu^2\|y\|^2.$$

P est donc un polynôme en μ. La fin de la démonstration ressemble fortement au cas du produit scalaire. En effet, P étant positif de degré 2, son discriminant Δ est nécessairement positif. Cela s'écrit exactement :

$$\forall (x, y) \in E^2, \ | < x, y > | \leq \|x\|.\|y\|.$$

Le cas d'égalité est obtenu en traitant le cas $\Delta = 0$. Supposons $y \neq 0$ (autrement, c'est trivial). Le trinôme du second degré P possède donc une racine double μ^*, et donc $x = -\mu^* e^{-i\theta} y$, ce qui démontre la proposition. □

Quelques exemples d'espaces préhilbertiens complexes :

- Soit $[a, b]$, un intervalle de $\mathbb{R}$. On appelle $\mathcal{C}([a, b], \mathbb{C})$, l'espace des fonctions continues sur le segment $[a, b]$. On peut munir $\mathcal{C}([a, b], \mathbb{C})$ du produit hermitien $< ., . >$ défini par :

$$\forall (f, g) \in \mathcal{C}([a, b], \mathbb{C}), \ < f, g >= \int_a^b \overline{f(t)} g(t) dt.$$

  $\mathcal{C}([a, b], \mathbb{R})$ est donc un espace préhilbertien complexe.
- On désigne par $\ell^2(\mathbb{C})$ l'espace défini par : $\ell^2(\mathbb{C}) := \left\{ (u_n)_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}} : \sum_{n=0}^{+\infty} |u_n|^2 < +\infty \right\}$. On peut munir $\ell^2(\mathbb{R})$ du produit hermitien $< ., . >$ défini par :

$$\forall (u, v) \in \left(\ell^2(\mathbb{C})\right)^2, \ < u, v >= \sum_{n=0}^{+\infty} \overline{u_n} v_n.$$

  $\ell^2(\mathbb{R})$ est donc un espace préhilbertien complexe.

Enfin, on énonce à présent une égalité remarquable dans les espaces préhilbertiens. Cette égalité s'appelle *l'égalité du parallélogramme*.

Propriété 3.2. *Égalité du parallélogramme.*
Soit E, un espace préhilbertien complexe, φ, une forme hermitienne, et $\|.\|$, la norme associée dans cet espace. Alors,

$$\forall (x, y) \in E^2, \ \|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Démonstration. Il suffit de remarquer que :

$$\forall (x, y) \in E^2, \ \|x+y\|^2 = \|x\|^2+\|y\|^2+2\Re e\left(\varphi(x, y)\right) \text{ tandis que } \|x-y\|^2 = \|x\|^2+\|y\|^2-2\Re e\left(\varphi(x, y)\right).$$

□

De cette égalité, on déduit immédiatement le :

Théorème 3.1. *Théorème de la médiane. Soit E, un espace préhilbertien complexe, ϕ, une forme sesquilinéaire, et Φ, la forme hermitienne associée dans cet espace. Alors,*

$$\forall (x,y) \in E^2,\ \|x\|^2 + \|y\|^2 = 2\left\|\frac{x+y}{2}\right\|^2 + \frac{1}{2}\|x-y\|^2.$$

Démonstration. Il suffit d'appliquer l'égalité du parallélogramme à $x+y$ et $x-y$. □

Mentionnons enfin ce théorème célèbre en Géométrie : le théorème de Pythagore, qui s'applique dans un espace préhilbertien.

Théorème 3.2. *Théorème de Pythagore.*
Soit E, un espace préhilbertien complexe, ϕ, une forme sesquilinéaire, et Φ, la forme hermitienne associée dans cet espace. Soient x_1, ..., x_n, n éléments de E orthogonaux deux à deux, c'est à dire tels que : $\forall (i,j) \in \{1,...,n\} : i \neq j,\ \varphi(x_i, x_j) = 0$. Alors,

$$\Phi\left(\sum_{k=1}^{n} x_k\right) = \sum_{k=1}^{n} \Phi(x_k).$$

3.1.3 Comment rendre des bases orthonormées ?

3.1.3.1 Le procédé d'orthonormalisation de Gram-Schmidt

Le théorème que nous énonçons à présent est valable dans les espaces préhilbertiens d'une façon générale. Il est important de bien en cerner la preuve. En effet, on y trouve un résultat d'existence d'une base orthonormée, dont la preuve nous fournit la construction.

Théorème 3.3. *Procédé d'orthonormalisation de Gram-Schmidt.*
Soit E, un espace préhilbertien réel ou complexe muni d'un produit noté ici pour des raisons de commodité d'écriture $< ., . >$. Soit $(x_i)_{i\in\{1,...,n\}}$, une famille libre de vecteurs de E. Alors, il existe une famille orthonormée $(y_i)_{i\in\{1,...,n\}}$ telle que :

$$\forall k \in \{1,...,n\},\ \text{vect}(x_1,...,x_k) = \text{vect}(y_1,...,y_k).$$

Démonstration. Comme je l'avais annoncé précédemment, la démonstration s'appuie sur la construction de la base $(y_i)_{i\in\{1,...,n\}}$; on dit que c'est une démonstration algorithmique. Nous allons procéder par récurrence. Nous allons passer par une base intermédiaire que nous noterons $(z_i)_{i\in\{1,...,n\}}$.
- On initialise : posons $z_1 = x_1$.
- Posons alors $z_2 = x_2 + \lambda.z_1$. On veut que $< z_2, z_1 >= 0$, ce qui équivaut encore à $< x_2, x_1 > +\lambda.\|x_1\|^2$. Choisissons donc $\lambda = -\dfrac{< x_2, x_1 >}{\|x_1\|^2}$, ce qui assure que $< z_2, z_1 >= 0$. De plus, on a construit z_2 tel que $\text{vect}(z_1, z_2) = \text{vect}(x_1, x_2)$, et il est clair que $z_2 \neq 0_E$.
- Caractère héréditaire : supposons construite $(z_1, ..., z_k)$, deux à deux orthogonaux, tels que $\text{vect}(z_1, ..., z_i) = \text{vect}(x_1, ..., x_i)$, pour tout $i \in \{1, ..., k\}$. Posons alors $z_{k+1} = x_{k+1} + \sum_{i=1}^{k} \mu_i z_i$. Soit $i \in \{1, ..., k\}$. On souhaite que la relation $< z_{k+1}, z_i >= 0$ soit vérifiée. Cela s'écrit encore : $0 =< x_{k+1}, z_i > +\mu_i\|z_i\|^2$, et donc $\mu_i = -\dfrac{< x_{k+1}, z_i >}{\|z_i\|^2}$. Les vecteurs $(z_1, ..., z_{k+1})$ sont donc deux à deux orthogonaux par construction. De plus, si $z_{k+1} = 0$, alors $x_{k+1} = -\displaystyle\sum_{i=1}^{k} \mu_i z_i \in$ $\text{vect}(z_1, ..., z_k) = \text{vect}(x_1, ..., x_k)$, ce qui est contraire aux hypothèses, donc $z_{k+1} \neq 0$. De plus,

on a $z_{k+1} = x_{k+1} + \sum_{i=1}^{k} \mu_i z_i$.. On en déduit que $\mathrm{vect}(z_1, ..., z_{k+1}) = \mathrm{vect}(x_1, ..., x_{k+1})$. Normons à présent la famille $(z_i)_{i\in\{1,...,n\}}$, en posant :

$$\forall i \in \{1, ..., n\},\ y_i = \frac{z_i}{\|z_i\|}.$$

La famille $(y_i)_{i\in\{1,...,n\}}$ satisfait alors aux conditions de l'énoncé.

□

Présentation algorithmique du procédé de Gram-Schmidt : attention, je donne ici une présentation algorithmique des résultats précédents, mais uniquement pour obtenir une base **orthogonale**, et non orthonormée. Autrement dit, dans l'algorithme, la normalisation finale des vecteurs obtenus dans la dernière étape de la preuve n'apparaitra pas.

- On considère $(x_i)_{i\in\{1,...,n\}}$, une famille libre de vecteurs de E.
- On définit la suite $(z_i)_{i\in\{1,...,n\}}$ de la façon suivante :

$$\begin{array}{rcl} z_1 & = & x_1 \\ z_2 & = & x_2 - \dfrac{< z_1, x_2 >}{< z_1, z_1 >}.z_1 \\ z_3 & = & x_3 - \dfrac{< z_1, x_3 >}{< z_1, z_1 >}.z_1 - \dfrac{< z_2, x_3 >}{< z_2, z_2 >}.z_2 \\ \vdots & & \\ z_n & = & x_n - \dfrac{< z_1, x_n >}{< z_1, z_1 >}.z_1 - \dfrac{< z_2, x_n >}{< z_2, z_2 >}.z_2 - \dfrac{< z_{n-1}, x_n >}{< z_{n-1}, z_{n-1} >}.z_{n-1} \end{array}$$

De plus, si on s'arrange pour normer la famille $(z_i)_{i\in\{1,...,n\}}$, on obtient : $\forall i \in \{1, ..., n\},\ < z_i, z_i >= \|z_i\|^2 = 1$.

Exemple : considèrons l'espace vectoriel $\mathbb{R}^3$ muni de la norme euclidienne. Plutôt que de choisir la base canonique, on considère la famille $\mathcal{F} := (e_1, e_2, e_3)$ telle que, dans la base canonique, on ait :

$$e_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ e_2 = \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix} \text{ et } e_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Il est assez facile de vérifier que $\mathcal{F}$ est une base de $\mathbb{R}^3$. En effet, on a :

$$\begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 2 & 3 & 1 \end{vmatrix} = 1 \text{ et } \mathrm{card}(\mathcal{F}) = 3.$$

On peut chercher à construire la base $\mathcal{B}$, associée à $\mathcal{F}$ à l'aide du procédé d'orthonormalisation de Schmidt. Notons $\mathcal{B} = (e'_1, e'_2, e'_3)$.

- Choisissons pour commencer $e'_1 = e_1$. e_1 est déjà normé, ce qui nous facilite la tâche.
- Cherchons alors (sans se préoccuper de la norme dans un premier temps), le vecteur e'_2 sous la forme $\alpha.e'_1 + e_2$ (à une constante multiplicative près). On note $u = \alpha.e'_1 + e_2$. Il nous faut déterminer α. α doit être choisi pour que, en appelant $< ., . >$ le produit scalaire usuel de $\mathbb{R}^3$,

on ait : $< u, e'_1 >= 0$, autrement dit, $3+\alpha = 0$, puis $\alpha = -3$. On en déduit que $u = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

On en déduit que $e'_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

- Il nous reste à trouver e'_3. Toujours en s'inspirant de la démonstration algorithmique précédente, on se ramène à trouver un vecteur v, sous la forme $\beta.e'_1 + \gamma.e'_2 + e_3$, où β et γ sont des constantes réelles choisies pour que $< v, e'_1 >=< v, e'_2 >= 0$. La première égalité s'écrit encore $-\frac{1}{\sqrt{2}} + \beta = 0$, tandis que la seconde s'écrit $\gamma + 2 = 0$. On en déduit que $v = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, et ainsi, $e'_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Les composantes de la famille $\mathcal{B} = (e'_1, e'_2, e'_3)$ se décomposent dans la base canonique sous la forme :

$$e'_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ e'_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \text{ et } e'_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Remarque : une application très importante du procédé d'orthonoàrmalisation de Schmidt consiste à déterminer une base d'espaces fonctionnels très célèbres que nous introduirons dans la section suivante : les espaces L^p, pour $p \in [1, +\infty]$.

3.1.3.2 Factorisation QR d'une matrice inversible

Afin de nous rafraîchir la mémoire et parce que cela n'a pas été fait jusqu'à maintenant, je propose un algorithme mettant en œuvre le procédé d'orthonormalisation de Gram-Schmidt.

Algorithme de Gram-Schmidt

- On se donne une famille $(x_i)_{i\in\{1,...,n\}}$ d'élements, de E, un espace préhilbertien réel muni d'un produit scalaire $<.,.>$.
- On pose $z_1 = x_1$, $r_{11} = \|z_1$, puis $q_1 = \dfrac{z_1}{r_{11}}$.

$\longrightarrow$ Pour $j = 1$ à n, faire :
$\longrightarrow$ Pour $i = 1$ à $j-1$, faire :

$$r_{ij} =< q_i, x_j > .$$

Cela termine la boucle sur i.

- On pose alors :

$$\begin{aligned} z_j &= x_j - \sum_{i=1}^{j-1} r_{ij} q_i \ ; \\ r_{jj} &= \|z_j\| \ ; \\ q_j &= \frac{z_j}{r_{jj}}. \end{aligned}$$

Cela termine la boucle sur j.

- La famille $(q_j)_{j\in\{1,..,n\}}$ est alors une famille orthonormée.

On suppose que tous les vecteurs $(q_j)_{j\in\{1,...,n\}}$ sont des vecteurs colonnes. On appelle alors R et Q, les matrices définies par :

$$R = (r_{ij})_{\substack{i\in\{1,...,n\}\\ j\in\{1,...,n\}}} \text{ et } Q = (q_1, ..., q_n).$$

La démonstration même du procédé d'orthonormalisation de Gram-Schmidt prouve, en supposant que tous les vecteurs $(x_i)_{i\in\{1,...,n\}}$ sont des vecteurs colonnes, la factorisation de X suivante :

$$X = QR.$$

Remarque : il est bien évident que R, telle qu'elle a été définie ci-dessus est une matrice triangulaire supérieure, tandis que les colonnes de Q forment, par construction, une famille orthogonale. Q est donc une matrice orthogonale.
Pour ceux qui auraient oublié la définition d'une matrice orthogonale, j'en redonne la définition, sans trop m'attarder sur cette notion. Si cela vous pose vraiment problème, je vous conseille de vous référer à un manuel d'Algèbre linéaire.

Définition 3.3. *Matrice orthogonale.*
Une matrice carrée A est dite orthogonale si elle vérifie : ${}^tAA = I$, où I désigne la matrice identité.
L'ensemble des matrices orthogonales de taille n est noté $\mathcal{O}(n)$.

Les matrices orthogonales bénéficient des propriétés suivantes :

Propriété 3.3. *Soit A, une matrice de $\mathcal{O}(n)$. Alors :*
(i) Les vecteurs colonnes de A forment une base orthonormée de $\mathbb{R}^n$.
(ii) Les vecteurs lignes de A forment une base orthonormée de $\mathbb{R}^n$.

(iii) On a nécessairement $\det A = \pm 1$.
(iv) A est la matrice d'un changement de bases orthonormales.
(v) Soit $X \in \mathbb{R}^n$. *Alors :* $\|AX\| = \|X\|$.

Remarque : nous nous servirons de ces propriétés sans les démontrer car elles semblent un peu loin de ce cours.

Nous sommes donc parvenus à factoriser toute famille de vecteurs libres sous la forme QR, où $Q \in \mathcal{O}(n)$ et R est une matrice triangulaire supérieure. Nous généralisons ce résultat grâce au théorème qui suit :

Théorème 3.4. *Factorisation QR d'une matrice.*
Soit X, une matrice réelle de taille $n \times m$, avec $n \geq m$. Supposons de plus que X est de rang m. Alors, il existe une matrice $Q \in \mathcal{M}_{m\times n}(\mathbb{R})$ dont les vecteurs colonnes forment une famille orthonormée, et une matrice carrée triangulaire supérieure $R \in \mathcal{M}_n(\mathbb{R})$, telles que :

$$X = QR.$$

Remarque : cette factorisation est très utilisée en Analyse Numérique notamment, pour résoudre numériquement (c'est à dire par le biais de l'Informatique) des systèmes linéaires de grande taille ou encore pour calculer les valeurs propres de matrices de grande taille. Mais ça n'est pas là l'objet de ce cours.

Donnons, à titre anecdotique, un exemple de factorisation QR.

Exemple : on va réutiliser l'exemple étudié précédemment lorsque nous avons tenté d'appliquer le procédé d'orthonormalisation de Schmidt. Soit :

$$X = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & 3 & 2 \end{pmatrix}.$$

En fait, avec les notations de l'exemple précédent, $X = (e_1, e_2, e_3)$. Nous avons déjà fait tous les calculs nécessaires pour conclure. Il s'agit uniquement ici d'une présentation un peu différente du procédé d'orthonormalisation de Schmidt. On trouve :

$$X = QR, \text{ avec } Q = \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ -1 & 0 & 0 \end{pmatrix} \text{ et } R = \begin{pmatrix} -1 & -3 & -2 \\ 0 & -\sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & 0 & -\frac{\sqrt{2}}{2} \end{pmatrix}.$$

3.1.4 Notion d'orthogonalité

Je commence par donner quelques définitions relatives à cette notion.

Définition 3.4. *Soit E, un espace préhilbertien complexe (ou réel) muni d'une forme sesquilinéaire φ.*
(i) On dit que deux éléments x et y de E sont orthogonaux si, et seulement si $\varphi(x, y) = 0$.
*(ii) Soit $A \subset E$. On définit l'**orthogonal de** A, que l'on note $A^\perp$, l'espace :*

$$A^\perp = \{x \in E : \forall x \in A, \varphi(a, x) = 0\}.$$

(iii) Deux parties A et B de E sont dites ***orthogonales*** *si, et seulement si $A \subset B^{\perp}$, ce qui revient à dire que $B \subset A^{\perp}$.*

Les propriétés qui suivent résultent directement de cette définition. Je laisse au lecteur le soin de les démontrer.

Propriété 3.4. *Soit E, un espace préhilbertien complexe muni d'une forme sesquilinéaire φ. Soient A et B, deux sous espaces de E tels que $A \subset B$. Alors :*
(i) $\{0_E\}^{\perp} = E$ et $E^{\perp} = \{0_E\}$;
(ii) $A^{\perp} = (\text{vect } A)^{\perp}$.
(iii) $B^{\perp} \subset A^{\perp}$.
(iv) $A^{\perp}$ est un sous espace vectoriel de E.
(v) $A \subset (A^{\perp})^{\perp}$.
(vi) $A \cap A^{\perp} = \{0_E\}$.
(vii) Toute famille orthogonale de vecteurs non nuls est libre.

Le problème du supplémentaire orthogonal.
La question que nous nous posons à présent est la suivante. D'après la propriété (vi) de la proposition 3.4, la somme $F + F^{\perp}$ est directe. Mais dans quels cas a t-on $F \oplus F^{\perp} = E$, l'espace tout entier ?
Pour répondre à cette question, nous allons distinguer le cas où E est de dimension finie et celui où E est de dimension infinie. Il va de soit que le cas de la dimension finie est beaucoup plus simple, car on connaît (cf. chapitres précédents) de nombreuses propriétés topologiques et algébriques sur les espaces de dimension finie. Le théorème qui suit répond à cette question.

Théorème 3.5. *Soit E, un espace préhilbertien réel ou complexe,* ***de dimension finie****. Soit φ, un produit hermitien (ou scalaire) sur cet espace. Soit F, un sous espace vectoriel de E.*
Alors, on a : $F \oplus F^{\perp} = E$.

Démonstration. On écrit que E est un $\mathbb{K}$ espace vectoriel, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Si $F = \{0\}$, alors $F^{\perp} = E$, et le résultat est clair. Sinon, soit $(e_1, ..., e_p)$, une base de F. Considérons à présent l'application $\varpi : \begin{array}{ccc} E & \longrightarrow & \mathbb{K}^p \\ x & \longmapsto & (\varphi(x, e_1), ..., \varphi(x, e_p)) \end{array}$. ϖ est une forme linéaire. Quel est son noyau ? On a $x \in \ker \varpi \Longrightarrow \varphi(x, e_1) = 0$, ..., et $\varphi(x, e_p) = 0$, et puisque F est un espace vectoriel dont $(e_1, ..., e_p)$ est une base, on en déduit immédiatement que $x \in \ker \varpi$, si, et seulement si : $\forall y \in F$, $\varpi(x, y) = 0$, autrement dit, $\ker \varpi = F^{\perp}$. D'après le théorème du rang, $\dim F^{\perp} = \dim \ker \varpi = n - \dim \text{Im } \varpi \geq n - p$. Il s'ensuit, la somme $F + F^{\perp}$ étant directe, que $\dim(F + F^{\perp}) = \dim F + \dim F^{\perp} \geq p + (n - p) = n$, et par conséquent, $F + F^{\perp} = E$, car $F + F^{\perp}$ est un sous-espace vectoriel de E de même dimension que E. Ainsi, $F \oplus F^{\perp} = E$. □

Corollaire 3.1. *Soit E, un espace préhilbertien réel ou complexe,* ***de dimension finie*** *non nulle p. Soit φ, un produit hermitien (ou scalaire) sur cet espace. Alors, il existe une famille $(e_1, ..., e_p)$ de vecteurs de E formant une base orthonormée de E.*

Démonstration. On procède par récurrence. Si $n = 1$, le résultat est évident. Supposons que ce résultat est vrai pour un espace vectoriel de dimension $n - 1$, avec $n \geq 2$, et soit E, un espace vectoriel de dimension n. Choisissons alors e_1, un vecteur unitaire de E. $F = \mathbb{K}.e_1$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, est un sous espace vectoriel de E. Appliquons le théorème précédent. On en déduit que $E = \mathbb{K}.e_1 \oplus (\mathbb{K}.e_1)^{\perp}$, où $\dim(\mathbb{K}.e_1)^{\perp} = n - 1$ et il est clair que $\dim(\mathbb{K}.e_1)^{\perp} = n - 1$. On applique alors l'hypothèse de récurrence. Il existe donc $(e_2, ..., e_n)$, une base orthonormée de $(\mathbb{K}.e_1)^\perp$. On en déduit que $(e_1, ..., e_n)$, une base orthonormée de E. □

La question naturelle que l'on se pose à présent est : Que se passe t-il pour un espace E de dimension infinie ? Une réponse partielle va nous être apportée dans le théorème qui suit.

Théorème 3.6. *Soit E, un espace préhilbertien réel ou complexe,* ***de dimension infinie****. Soit φ, un produit hermitien (ou scalaire) sur cet espace.*
Si F est un sous espace vectoriel de E ***de dimension finie****, alors :*

$$F \oplus F^\perp = E.$$

Démonstration. F étant de dimension finie, on peut utiliser le théorème 3.5. Il existe donc $(e_1, ..., e_p)$, une base orthonormée de F. Soit $x \in E$. On va démontrer l'existence d'une famille $(\lambda_1, ..., \lambda_p)$ de scalaires, tels que $x - \sum_{k=1}^{p} \lambda_i.e_i \in F^\perp$. remarquons que $x - \sum_{k=1}^{p} \lambda_i.e_i \in F^\perp \iff \forall j \in \{1, ..., p\},\ \varphi\left(x - \sum_{k=1}^{p} \lambda_i.e_i, e_j\right) = 0 \iff \forall j \in \{1, ..., p\},\ \lambda_j = \varphi(x, e_j)$. Ainsi, on écrit que $x = \sum_{k=1}^{p} \lambda_i.e_i + y$, avec $y \in F^\perp$, ce qui prouve que $E = F \oplus F^\perp$. □

Remarque : On notera à l'avenir $p_F(x) = \sum_{k=1}^{p} \lambda_i.e_i = \sum_{k=1}^{p} \varphi(x, e_i).e_i$. $p_F(x)$ s'appelle la projection de x sur le sous espace vectoriel F.

3.1.5 Théorèmes de projection dans un espace préhilbertien

3.1.5.1 Introduction et aspects géométriques du problème

Si E est un espace préhibertien **réel**, alors il est muni d'un produit scalaire noté φ. Il est donc également muni d'une norme, induite par le produit scalaire. On la note $\|.\|$. De plus, E est muni d'une distance induite par la norme. On la note d. Elle est définie de la façon suivante

$$d: \begin{array}{ccc} E^2 & \longrightarrow & \mathbb{R}_+ \\ (x,y) & \longmapsto & \|x-y\| \end{array}.$$

Le problème de la projection, noté $(\mathcal{P}_\mathcal{A})$ peut alors se formuler de la façon suivante.

Soit $x \in E$ et $\mathcal{A}$, un sous espace de E. Peut-on trouver $y \in \mathcal{A}$ tel que $d(x, y) = \inf_{a \in \mathcal{A}} d(x, a)$?

Définition 3.5. *Distance à un sous espace vectoriel.*
Soit E, un espace préhilbertien réel et $\mathcal{A}$, un sous espace de E.
La quantité $\inf_{a \in \mathcal{A}} d(x, a)$ s'appelle la ***distance de x au sous espace $\mathcal{A}$, et est notée*** *$d(x, \mathcal{A})$.*

Dans un premier temps, examinons des cas simples, à l'aide de figures géométriques, que nous tenterons de généraliser dans les paragraphes qui suivent. Dans $\mathbb{R}^2$, espace euclidien muni de la norme euclidienne $\|.\|_2$, soit $X = (x_1, x_2)$, un point. La question devient donc :

$$\text{Peut-on trouver } Y = (y_1, y_2) \in \mathcal{A} \text{ tel que } \|X - Y\|_2 = \inf_{Z=(z_1,z_2)\in\mathcal{A}} \|X - Z\|_2\,?$$

- Si $\mathcal{A}$ désigne un cercle $\mathcal{C}$ de rayon R, on envisage deux cas.
  - Si X est en dehors du cercle $\mathcal{C}$. Dans ce cas, on se rend compte, à l'aide d'un petit dessin, de l'unicité de la projection y de x sur A.
  - Si X est le centre du cercle $\mathcal{C}$. Dans ce cas, on se rend compte qu'il existe une infinité de projections y de x sur $\mathcal{A}$.

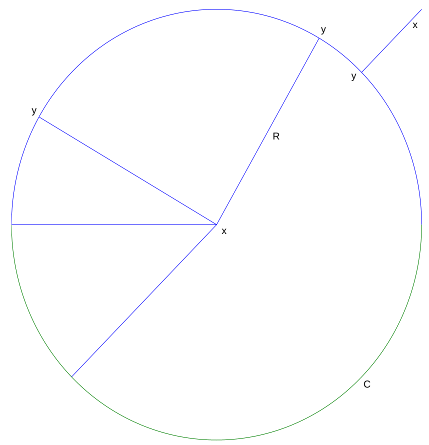

FIG. 3.1 – Quelques exemples de projections sur un cercle

- Si $\mathcal{A}$ désigne à présent un carré dont x est le centre. On se rend compte toujours à l'aide d'un dessin, qu'il existe un nombre fini de projections de x sur $\mathcal{A}$. (quatre, en l'occurrence...)

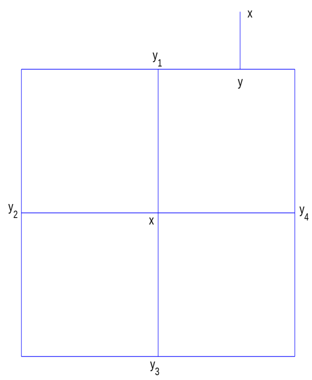

FIG. 3.2 – Un exemple de projection sur un carré

Un autre argument, dorénavant classique, nous permet d'affirmer que, lorsque $\mathcal{A}$ est compact, alors, pour tout $x \in E$, il existe une unique projection y de x sur $\mathcal{A}$. En effet, remarquons que $y \longmapsto d(x, y) = \|x - y\|$ est une application continue, en tant que composée d'applications continues ; il est alors clair que l'image de $\mathcal{A}$ par cette application est une partie compacte de

$\mathbb{R}_+$. On en déduit que la borne inférieure de ce compact est atteinte, et ainsi : $\exists y \in \mathcal{A} : \|x-y\| = \inf_{z\in\mathcal{A}} \|x-z\|$.

3.1.5.2 Le théorème de la projection orthogonale

Théorème 3.7. *Théorème de la projection orthogonale.*
Soit E, un espace préhilbertien réel muni d'un produit scalaire noté φ. Soit F, un sous espace vectoriel de E, et $x \in E$. Alors, les propositions suivantes sont équivalentes :
- *y est une projection de x sur F, autrement dit, y est une solution du problème $(\mathcal{P}_F)$.*
- *$y \in F$ et $x - y \in F^\perp$.*

Démonstration. La méthode est toujours la même. On va démontrer les deux implications qui composent l'équivalence.
- $(i) \Longrightarrow (ii)$: supposons que $y \in F$ soit une projection de x sur F. Alors, on a immédiatement, par définition de l'infimum d'une partie que $\forall z \in F,\ \|x-y\| \leq \|x-z\|$.
  Choisissons au lieu de z, le point $z_\lambda := y + \lambda.(z-y)$, avec $\lambda \in]0,1[$. Cette situation est résumée par le schéma qui suit. Puisque, F est un espace vectoriel, $z_\lambda \in F$. On en déduit que

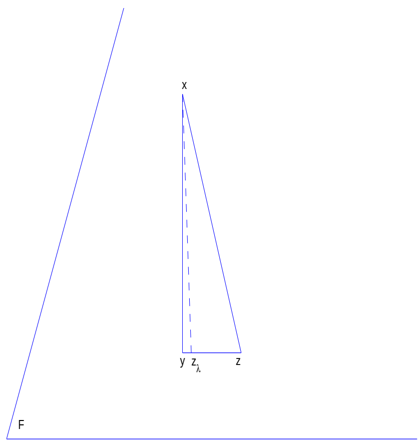

FIG. 3.3 – Visualisation de la preuve

  $\forall z \in F,\ \|x-y\| \leq \|x-y-\lambda.(z-y)\|$, et en élevant cette inégalité au carré, puis en divisant par $\lambda \neq 0$, on obtient :
$$\|x-y\|^2 \leq \|x-y\|^2 + \lambda^2\|z-y\|^2 - 2\lambda.\varphi(x-y, z-y).$$
  On obtient : $\varphi(x-y, z-y) \leq \lambda\|z-y\|^2$. Faisons à présent tendre λ vers 0. En remarquant que le vecteur $u = z-y$ décrit F, l'inégalité devient : $\varphi(x-y, u) \leq 0$. Or, cette inégalité étant vérifiée pour tout $u \in F$, elle est en particulier vérifiée pour tous les $-u \in F$, c'est à dire que : $\forall u \in F,\ \varphi(x-y,u) \geq 0$, et ainsi, $\varphi(x-y,u) = 0$. On en déduit que $x - y \in F^\perp$.
- $(ii) \Longrightarrow (i)$: supposons que $x-y \in F^\perp$ et $y \in F$. Alors, $\forall z \in F$, $\|x-(y+z)\|^2 = \|x-y\|^2 + \|z\|^2$, car $\varphi(x-y, z) = 0$. On remarque, comme précédemment, que lorsque y est fixé et que z décrit F, $y+z$ décrit F, car F est un espace vectoriel. On en déduit, en appelant $d(x,F)$, la distance de x à F, que :
$$d(x,F) = \inf_{z\in F} \|x-z\| = \inf_{z\in F} \|x-(y+z)\| = \sqrt{\inf_{z\in F}\{\|x-y\|^2 + \|z\|^2\}}.$$

Il est donc clair que l'infimum est atteint pour $z = 0$, et ainsi, $d(x, F) = \inf_{z\in F} \|x - y\|$. y est donc bien une projection de x sur F.

□

3.1.5.3 Version algébrique du théorème de la projection

Ces théorèmes sont des versions améliorées du théorème précédent dans des cas particuliers.

Théorème 3.8. *Théorème de la projection, version algébrique.*
Soit E, un espace préhilbertien réel muni d'un produit scalaire φ. On note $\|.\|$, la norme induite par φ. Soit F, un sous espace vectoriel **de dimension finie** *de E.*
Aors, pour tout $x \in E$, il existe une projection unique de x sur F. On la note en général $p_F(x)$, et si $(e_i)_{i\in\{1,\dots,\dim F\}}$ désigne une base orthonormée de F (on sait qu'elle existe d'après la section 3.1.3.1), on a :

$$p_F(x) = \sum_{i=1}^{\dim F} < x, e_i > .e_i.$$

Démonstration.

Preuve de l'unicité : comme à notre habitude, raisonnons par l'absurde. Supposons qu'il existe deux éléments de F, y_1 et y_2, tels que $d(x, F) = \|x - y_1\| = \|x - y_2\|$. Utilisons le théorème de la médiane. On obtient :

$$\begin{aligned} \|x - y_1\|^2 + \|x - y_2\|^2 &= 2\left\|x - \left(\frac{y_1 + y_2}{2}\right)\right\|^2 + \frac{\|y_1 - y_2\|^2}{2}. \\ \Longleftrightarrow 2\left(\|x - y_1\|^2 - \left\|x - \left(\frac{y_1 + y_2}{2}\right)\right\|^2\right) &= 3\frac{\|y_1 - y_2\|^2}{2}. \end{aligned}$$

Or, par définition de l'infimum d'un ensemble (qui est donc un minorant de tous les éléments de cet ensemble), il est clair que le membre de gauche de l'égalité ci-dessus est négatif. En effet, F étant un sous espace vectoriel, il est clair que $\frac{y_1 + y_2}{2}$ est encore un élément de F. Ainsi, on obtient que $\|y_1 - y_2\| \leq 0$, et il s'ensuit que $y_1 = y_2$.

- **Preuve de l'existence :** soit $(e_i)_{i\in\{1,\dots,n\}}$, une base orthonormée de F. On pose $y = \sum_{i=1}^{n} < x, e_i > .e_i$. Alors, $x - y = x - \sum_{i=1}^{n} < x, e_i > .e_i$. Soit alors $z = \sum_{i=1}^{n} z_i e_i \in F$. On a : $< x - y, z >=< x, z > - \sum_{i=1}^{n} < x, e_i > z_i = \left\langle x, \sum_{i=1}^{n} z_i e_i \right\rangle - \sum_{i=1}^{n} < x, e_i > z_i = 0.$
D'après le théorème 3.7, on en déduit que y est le projection de x sur F.

□

Exemple : on cherche à déterminer explicitement la quantité :

$$I = \inf_{(a,b)\in\mathbb{R}^2} \int_0^1 |e^x - ax - b|^2 dx.$$

Remarquons de prime abord que I existe, car d'après la positivité de l'intégrale, $I \geq 0$. Considérons à présent l'espace vectoriel normé **complet** $\mathcal{C}([0, 1], \mathbb{R})$. Cet espace est en particulier un espace préhilbertien réel, muni du produit scalaire $< .,. >$ défini par :

$$\forall (f,g) \in (\mathcal{C}([0,1],\mathbb{R}))^2, \ < f,g >:= \int_{-1}^{1} f(x)g(x)dx.$$

L'espace $\mathbb{R}_1[X]$ des polynômes à une indéterminée de degré au plus 1 est un sous espace vectoriel de $\mathcal{C}([0,1],\mathbb{R})$ de dimension finie. Appelons f, la fonction définie sur $\mathbb{R}$ par $f(x) = e^x$, le problème que l'on cherche à résoudre peut s'écrire :

Trouver $P \in \mathbb{R}_1[X]$ réalisant le minimum de la quantité $\|f - P\|^2$, où $\|.\|$ désigne la norme induite par $< .,. >$.

En effet, le théorème que nous venons d'énoncer (version algébrique du théorème de la projection) nous garantit l'existence et l'unicité d'un tel polynôme. P est en fait la projection de la fonction f sur le sous espace vectoriel $\mathbb{R}_1[X]$.
À présent que nous sommes fixés sur l'existence et l'unicité de P, nous allons utiliser le procédé d'orthonormalisation de Schmidt afin de déterminer une base orthonormée de $\mathbb{R}_1[X]$ pour le produit scalaire $< . >$. Appelons (e_1, e_2), cette base de $\mathbb{R}_1[X]$. On sait que $(1, X)$ est une base de $\mathbb{R}_1[X]$. C'est cette base que nous allons orthonormaliser.

- On peut poser $e_1 = \frac{1}{\sqrt{2}}$. Remarquons que $< e_1, e_1 >= 1$.
- Utilisons le procédé d'orthonormalisation de Schmidt. Posons : $e_2(x) = x + c$, avec $c \in \mathbb{R}$. On doit avoir $< e_2, e_1 >= 0$, donc $\displaystyle\int_{-1}^{1} e_2(x)dx = c = 0 \iff c = 0$, puis en notant $v(x) := x$, $e_2(x) = \dfrac{v(x)}{\sqrt{< v, v >}} = \sqrt{\dfrac{3}{2}}x$.

Conclusion : la base $\left(\dfrac{1}{\sqrt{2}}, \sqrt{\dfrac{3}{2}}x\right)$ de $\mathbb{R}_1[X]$ est orthonormée.

En utilisant le théorème de la projection (version algébrique), il est clair que, en appelant $P :=< f, e_1 > .e_1+ < f, e_2 > .e_2$, on a :

$$I = \|f - P\|.$$

Ainsi, il est immédiat que $< f, e_1 >= \sinh(1)$, où sinh désigne la fonction « sinus hyperbolique », et à l'aide d'une intégration par parties, on montre facilement que $< f, e_2 >= \sqrt{\dfrac{3}{2}} \displaystyle\int_{-1}^{1} xe^x dx = \sqrt{\dfrac{3}{2}}\dfrac{2}{e}$. Il s'ensuit que :

$$P = \sinh(1) + \frac{3}{e}X \text{ et ainsi } I = 1 - \frac{7}{e^2} \simeq 0,052653.$$

3.1.5.4 Matrice et déterminant de Gram

Nous venons de voir que le problème de la projection s'intéresse entre autres au calcul de la quantité $d(x, F)$, où x désigne un élément d'un espace préhilbertien réel E dont un sous espace vectoriel est F. Sans chercher à régler la question de l'existence voire l'unicité de la projection, il semble intéressant de se demander s'il est possible de calculer cette quantité. Et en effet, indépendamment de la détermination de la projection, on peut donner une réponse positive au calcul de $d(x, F)$, dans le cas où E est un **espace préhilbertien réel**. C'est l'objet des notions que j'introduis à présent.

Définition 3.6. *Matrice de Gram.*
Soit E, un espace préhilbertien réel et $(x_1, ..., x_n) \in E^n$, où n désigne un entier naturel non nul. On note $< .,. >$, le produit scalaire dont E est muni.
On appelle ***matrice de Gram*** *de $(x_1, ..., x_n)$, la matrice $(< x_i, x_j >)_{\substack{1\leq i\leq n \\ 1\leq j\leq n}}$.*
On note $G(x_1, ..., x_n)$, le déterminant de cette matrice. $G(x_1, ..., x_n)$ s'appelle le ***déterminant de Gram****.*

Nous allons à présent énoncer les propriétés les plus significatives de cette matrice, pour tenter d'aboutir à une formule, la plus simple possible, permettant de calculer $d(x, F)$, où F désigne un sous espace vectoriel de E. Je donnerai, en fin de partie, une application.

Propriété 3.5. *Soit E, un espace préhilbertien réel, et $(x_1, ..., x_n) \in E^n$, où n désigne un entier naturel non nul. On note $< .,. >$, le produit scalaire dont E est muni.*
(i) $(x_1, ..., x_n)$ et sa matrice de Gram ont même rang.
(ii) $(x_1, ..., x_n)$ est liée si, et seulement si $G(x_1, ..., x_n) = 0$.
(iii) $(x_1, ..., x_n)$ est libre si, et seulement si $G(x_1, ..., x_n) > 0$.
(iv) Soit $\sigma \in \mathfrak{S}_n$ (on rappelle que $\mathfrak{S}_n$ désigne l'ensemble des permutations sur l'ensemble $\{1, ..., n\}$). Alors, $G(x_{\sigma(1)}, ..., x_{\sigma(n)}) = G(x_1, ..., x_n)$.

Démonstration.
(*i*) Appelons A, la matrice de Gram de $(x_1, ..., x_n)$ et soit $Y = (y_i)_{1\leq i\leq n} \in \mathcal{M}_{n,1}(\mathbb{R})$. Alors, $AY = 0 \iff \forall i \in \{1, ..., n\},\ \sum_{j=1}^{n} y_j < x_i, x_j >= 0 = \left\langle x_i, \sum_{j=1}^{n} y_j x_j \right\rangle$. On en déduit que $\sum_{j=1}^{n} y_j x_j \in (\text{vect}(x_1, ..., x_n))^{\perp}$. Or, il est clair que $\sum_{j=1}^{n} y_j x_j \in \text{vect}(x_1, ..., x_n)$. On en déduit que $\sum_{j=1}^{n} y_j x_j \in (\text{vect}(x_1, ..., x_n))^{\perp} \cap \text{vect}(x_1, ..., x_n) = \{0_E\}$, donc $\sum_{j=1}^{n} y_j x_j = 0_E$. On désigne par ψ l'application :

$$\begin{array}{rccc} \psi: & \mathcal{M}_{n,1}(\mathbb{R}) & \longrightarrow & \text{vect}(x_1, ..., x_n) \\ & Y & \longmapsto & \sum_{j=1}^{n} y_j x_j. \end{array}$$

D'après ce que l'on vient d'écrire, il semble clair que $\ker(A) = \ker(\psi)$. D'après le théorème du rang, on en déduit immédiatement que $\text{rg}(\psi) = n - \dim(\ker(A))$, et d'après l'égalité des noyaux, il vient que : $\text{rg}(A) = \text{rg}(x_1, ..., x_n)$.

(*ii*) C'est immédiat, d'après (*i*). En effet, $(x_1, ..., x_n)$ est liée si, et seulement si $\det(x_1, ..., x_n) = 0$, autrement dit si, et seulement si $G(x_1, ..., x_n) = 0$, car $(x_1, ..., x_n)$ et sa matrice de Gram ont même rang.

(*iii*) D'après (*ii*), il suffit de montrer ici que $G(x_1, ..., x_n) \geq 0$ pour tous vecteurs x_1, ..., x_n. Supposons donc que $(x_1, ..., x_n)$ est libre. On appelle $(e_1, ...,_n)$, la famille orthonormale obtenue par le procédé d'orthonormalisation de Schmidt. On peut donc écrire, d'après le théorème de la projection, version algébrique, que $\forall j \in \{1, ..., n\},\ x_j = \sum_{k=1}^{n} < x_i, e_k > .e_k := \sum_{k=1}^{n} x_{i,k}.e_k$.
On appelle $\mathcal{B}$, la base $(e_1, ..., e_n)$. On a donc :

$$\forall (i,j) \in \{1, ..., n\}^2,\ < x_i, x_j >= \sum_{k=1}^{n} x_{k,i} x_{k,j} < e_k, e_k >=^t XX$$

, où $X = Mat\,((x_1, ..., x_n), \mathcal{B})$, la matrice de $(x_1, ..., x_n)$, exprimée dans la base $\mathcal{B}$. ${}^tXX$ est donc la matrice de Gram de $(x_1, ..., x_n)$; utilisons alors la propriété $\det X = \det^t X$ si X est carrée. On obtient que $G(x_1, ..., x_n) = \det^2(X) > 0$, car $\det X \neq 0$ par hypothèse. La conclusion s'ensuit.

(iv) Le résultat est trivial si $(x_1, ..., x_n)$ est liée. Supposons donc que $(x_1, ..., x_n)$ est non liée. Appelons $\varepsilon(\sigma)$, la signature de σ. L'idée est, comme dans la preuve précédente, en conservant ces notations, de décomposer $(x_{\sigma(1)}, ..., x_{\sigma(n)})$ dans la base $\mathcal{B}$. Posons $X_\sigma := Mat\,\big((x_{\sigma(1)}, ..., x_{\sigma(n)}), \mathcal{B}\big)$. Toujours en utilisant les propriétés classiques du déterminant, on a : $\det(X_\sigma) = \varepsilon\det(\sigma)X$. Il s'ensuit :

$$G(x_{\sigma(1)}, ..., x_{\sigma(n)}) = \varepsilon^2(\sigma)G(x_1, ..., x_n) = G(x_1, ..., x_n).$$

□

Le théorème qui suit nous intéresse particulièrement pour des raisons que nous avons explcitées en début de section.

Théorème 3.9. *On conserve les notations des définitions précédentes. Supposons que $(x_1, ..., x_n)$ est une famille libre de vecteurs de E. On appelle F, le sous espace vectoriel engendré par x_1, ..., x_n. Alors :*

$$\forall x \in E,\ d(x, F) = \sqrt{\frac{G(x_1, ..., x_n, x)}{G(x_1, ..., x_n)}}.$$

Démonstration. Nous sommes tout à fait dans le domaine d'application du théorème de la projection, version algébrique, car F est un sous espace vectoriel de E de dimension finie. On peut donc écrire $x = y + z$, avec $y = p_F(x) \in F$ et $z \in F^\perp$. On en déduit ainsi que $d(x, F) = \|z\|$. De plus, en élevant la relation précédente au carré, et puisque $< y, z >= 0$, on obtient l'égalité suivante qui rappelle le théorème de Pythagore en Géométrie : $\|x\|^2 = \|y\|^2 + \|z\|^2$. Nous en étudierons une version plus générale dans la section consacrée aux espaces de Hilbert. De plus, puisque $y \in F^\perp$, pour tout $i \in \{1, ..., n\}$, on a : $< x_i, x >=< x_i, y >$. Matriciellement, cela s'écrit :

$$\begin{pmatrix} < x_1, x_1 > & \dots & < x_1, x_n > & < x_1, x > \\ \vdots & & \vdots & \vdots \\ < x_{n-1}, x_1 > & \dots & < x_n, x_n > & < x_n, x > \\ < x, x_1 > & \dots & < x, x_n > & < x, x > \end{pmatrix} = \begin{pmatrix} < x_1, x_1 > & \dots & < x_1, x_n > & < x_1, y > \\ \vdots & & \vdots & \vdots \\ < x_{n-1}, x_1 > & \dots & < x_n, x_n > & < x_n, x > \\ < y, x_1 > & \dots & < y, x_n > & \|y\|^2 + \|z\|^2 \end{pmatrix}.$$

On a bien-sûr reconnu $G(x, x_1, ..., x_n)$ dans la matrice de gauche. Par ailleurs, le déterminant est une application multilinéaire. On en déduit que le membre de droite vaut :

$$\begin{vmatrix} < x_1, x_1 > & \dots & < x_1, x_n > & < x_1, y > \\ \vdots & & \vdots & \vdots \\ < x_{n-1}, x_1 > & \dots & < x_n, x_n > & < x_n, x > \\ < y, x_1 > & \dots & < y, x_n > & \|y\|^2 \end{vmatrix} + \begin{vmatrix} < x_1, x_1 > & \dots & < x_1, x_n > & 0 \\ \vdots & & \vdots & \vdots \\ < x_{n-1}, x_1 > & \dots & < x_n, x_n > & 0 \\ < y, x_1 > & \dots & < y, x_n > & \|z\|^2 \end{vmatrix}.$$

De plus, on a d'une part :

$$\begin{vmatrix} < x_1, x_1 > & \dots & < x_1, x_n > & < x_1, y > \\ \vdots & & \vdots & \vdots \\ < x_{n-1}, x_1 > & \dots & < x_n, x_n > & < x_n, x > \\ < y, x_1 > & \dots & < y, x_n > & \|y\|^2 \end{vmatrix} = G(x_1, ..., x_n, y)$$

et d'autre part, en effectuant un développement du déterminant par rapport à la dernière colonne :

$$\begin{vmatrix} < x_1, x_1 > & \dots & < x_1, x_n > & 0 \\ \vdots & & \vdots & \vdots \\ < x_{n-1}, x_1 > & \dots & < x_n, x_n > & 0 \\ < y, x_1 > & \dots & < y, x_n > & \|z\|^2 \end{vmatrix} = \|z\|^2 G(x_1, ..., x_n).$$

Remarquons enfin que $G(x_1, ..., x_n, y) = 0$, car y est combinaison linéaire des vecteurs x_1, ..., x_n. Ainsi :

$$d^2(x, F) = \|z\|^2 = \frac{G(x_1, ..., x_n, x)}{G(x_1, ..., x_n)}.$$

□

Exemple : reprenons l'exemple donné dans la section 3.1.5.3. Si l'on cherche à déterminer $d(f, \mathbb{R}_1[X])$, où f désigne la fonction exponentielle et $\mathbb{R}_1[X]$, l'espace des polynômes de degré au plus 1. $\mathbb{R}_1[X]$ étant un espace vectoriel de dimension finie, dont une base est $(1, X)$, nous sommes dans le domaine d'application du théorème précédent. En utilisant des notations abusives, mais probablement plus explicites aux yeux du lecteur, on peut écrire :

$$G(1, x, \exp(x)) = \begin{vmatrix} < 1, 1 > & < 1, x > & < 1, e^x > \\ < x, 1 > & < x, x > & < x, e^x > \\ < e^x, 1 > & < e^x, x > & < e^x, e^x > \end{vmatrix} = \begin{vmatrix} 2 & 0 & e - \frac{1}{e} \\ 0 & \frac{2}{3} & \frac{2}{e} \\ e - \frac{1}{e} & \frac{2}{e} & \frac{1}{2}(e^2 - \frac{1}{e^2}) \end{vmatrix} = \frac{4}{3}\left(1 - \frac{7}{e^2}\right).$$

Remarquons de plus que :

$$G(1, x) = \begin{vmatrix} < 1, 1 > & < 1, x > \\ < x, 1 > & < x, x > \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{vmatrix} = \frac{4}{3}.$$

On retrouve le résultat trouvé précédemment : $d(f, \mathbb{R}_1[X]) = 1 - \dfrac{7}{e^2}$.

Le théorème que j'énonce maintenant s'appelle l'**inégalité de Hadamard**. Je n'en donne pas immédiatement la preuve, car celle-ci sera proposée en exercice.

Propriété 3.6. *Inégalité de Hadamard.*
Soit E, un espace vérifiant les mêmes hypothèses que dans le théorème précédent.

$$\forall (x_1, ..., x_n) \in E^n,\ G(x_1, ..., x_n) = \overset{2}{\det}(x_1, ..., x_n) \leq \|x_1\|^2 ... \|x_n\|^2 = \prod_{i=1}^{n} \|x_i\|^2.$$

Remarque : on en déduit en particulier que :

$$\forall (x_1, ..., x_n) \in E^n,\ |\det(x_1, ..., x_n)| \leq \|x_1\| ... \|x_n\|.$$

Cette inégalité prouve que l'application « déterminant » est une application multilinéaire **continue**. Il suffit ici de revenir à la définition et aux propriétés d'une application multilinéaire continue, dans la chapitre précédent, section 2.5.3.

3.1.5.5 Version topologique du théorème de la projection

Avant de poursuivre en énonçant ce théorème, rappelons ce qu'est une partie convexe et comment cette notion se conçoit géométriquement.

Définition 3.7. *Partie convexe dans un EVN.*
Soit (E, N), un espace vectoriel normé. Soit $C \subset E$.
On dit que C est une ***partie convexe*** *de E si, et seulement si :*

$$\forall (x, y) \in C^2,\ \forall t \in [0, 1],\ t.x + (1 - t).y \in C$$

Remarque 1 : si x et y sont des réels, alors, lorsque t décrit le segment $[0, 1]$, $t.x + (1 - t).y$ décrit le segment $[x, y]$.

Remarque 2 : la remarque 1 tient encore en dimensions 2 et 3.

Remarque 3 : comment visualiser géométriquement un convexe ?
Géométriquement, un convexe est un ensemble tel que, si l'on considère deux points quelconques x et y de cet ensemble, alors tout point du segment $[x, y]$ est encore dans cet ensemble.

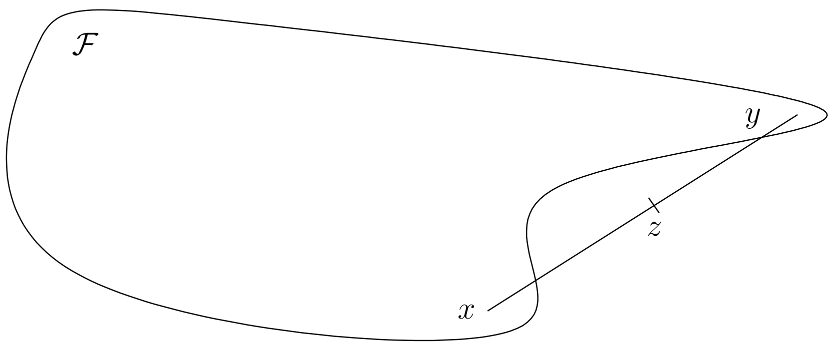

FIG. 3.4 – Un exemple d'ensemble non convexe $\mathcal{F}$

Dans l'exemple ci-dessus, le point z est un élément du segment $[x, y]$ qui n'appartient pourtant pas à $\mathcal{F}$. C'est pourquoi, juste en visualisant géométriquement la forme d'un ensemble, on pourra en général avoir une bonne intuition de sa convexité ou non-convexité.
La figure ci-dessous est, à l'inverse, un exemple de convexe.

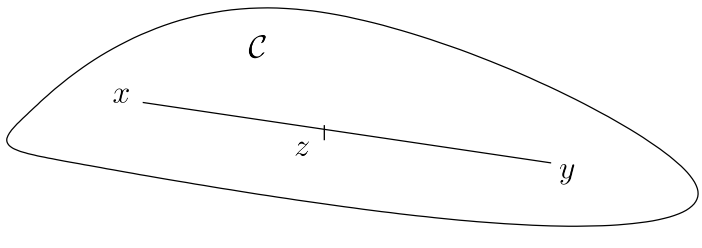

FIG. 3.5 – Un exemple d'ensemble convexe $\mathcal{C}$

Théorème 3.10. *Théorème de la projection, version topologique.*
Soit E, un espace préhilbertien réel muni d'un produit scalaire $< .,. >$. Soit F, une partie ***compacte*** *et* ***convexe*** *de E.*
Alors, pour tout élément x de E, il existe un unique $y \in F$ tel que :

$$d(x, F) = \|x - y\| = \inf_{f \in F} \|x - f\|.$$

On notera encore $y := p_F(x)$.

Démonstration.

- **Preuve de l'unicité :** raisonnons par l'absurde. Supposons qu'il existe deux éléments de F y_1 et y_2 tels que $d(x,F) = \|x-y_1\| = \|x-y_2\|$. Utilisons le théorème de la médiane. On obtient :
$$\begin{aligned} \|x-y_1\|^2 + \|x-y_2\|^2 &= 2\left\| x - \left(\frac{y_1+y_2}{2}\right) \right\|^2 + \frac{\|y_1-y_2\|^2}{2}. \\ \iff 2\left(\|x-y_1\|^2 - \left\| x - \left(\frac{y_1+y_2}{2}\right) \right\|^2 \right) &= 3\frac{\|y_1-y_2\|^2}{2}. \end{aligned}$$
Or, par définition de l'infimum d'un ensemble (qui est donc un minorant de tous les éléments de cet ensemble), il est clair que le membre de gauche de l'égalité ci-dessus est négatif. En effet, F étant un espace convexe, $\dfrac{y_1+y_2}{2}$ est encore un élément de F. Ainsi, on obtient que $\|y_1-y_2\| \leq 0$, et il s'ensuit que $y_1 = y_2$.
- **Preuve de l'existence :** si $d(x,F) = 0$, alors $x \in \overline{F} = F$, car F est supposé compact, donc en particulier fermé et on peut donc choisir $y = x$.
Si ça n'est pas le cas, posons : $\delta = d(x,F) = \inf_{f\in F} \|x-f\|$, puis utilisons une caractérisation bien connue de l'infimum. Cette technique a été maintes fois utilisée dans le chapitre précédent (caractérisation séquentielle) et ainsi, on peut construire une suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de F telle que :
$$\forall n \in \mathbb{N},\ \delta < \|x-f_n\| < \delta + \frac{1}{n}.$$
$\forall (p,q) \in \mathbb{N}^2,\ \dfrac{f_p + +f_q}{2} \in F$, car F est convexe. On en déduit que $\left\| x - \dfrac{f_p+f_q}{2} \right\| \geq \delta$. On utilise l'égalité de la médiane : $\|x-f_p\|^2 + \|x-f_q\|^2 = \dfrac{1}{2}\|f_p-f_q\|^2 + 2\left\| x - \dfrac{f_p+f_q}{2} \right\|^2$. On en déduit que :
$$\frac{1}{2}\|f_p-f_q\|^2 \leq \left(\delta+\frac{1}{p}\right)^2 + \left(\delta+\frac{1}{q}\right)^2 - 2\delta^2 \leq \frac{1}{p^2} + \frac{1}{q^2} + 2\delta\left(\frac{1}{p}+\frac{1}{q}\right) \xrightarrow[\substack{p\to+\infty \\ q\to+\infty}]{} 0.$$
Cela prouve donc que la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy dans F compacte, donc complète. Il existe donc $y \in F$ tel que $f_n \xrightarrow[n\to+\infty]{\text{dans } F} y$. Or, d'après l'inégalité $\forall n \in \mathbb{N},\ \delta < \|x-f_n\| < \delta + \frac{1}{n}$, et puisque la norme est une application continue (déjà vu maintes et maintes fois), la caractérisation séquentielle de la continuité nous permet un passage à la limite dans cette inégalité, ce qui prouve que $\|x-y\|$ réalise l'infimum et donc :
$$\|x-y\| = d(x,F).$$

□

3.2 Espaces de Banach et de Hilbert

3.2.1 Introduction et exemples

Définition 3.8. *Espace de Banach. Soit $(E, \|.\|)$, un espace vectoriel normé de dimension quelconque.*
*E est appelé **espace de Banach** lorsque E est **complet**.*

Nous avons déjà rencontré des espaces de Banach :
- $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$ et $\mathbb{C}^n$ munis des normes euclidiennes associées.
- $\ell^1(\mathbb{R}) := \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \sum_{k=1}^{+\infty} |u_n| < +\infty \right\}$, muni de la norme N_1 définie par : $\forall u \in \ell^1(\mathbb{R}),\ N_1(u) = \sum_{k=1}^{+\infty} |u_n|$.
- $\ell^2(\mathbb{R}) := \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \sum_{k=1}^{+\infty} |u_n|^2 < +\infty \right\}$, muni de la norme N_1 définie par : $\forall u \in \ell^2(\mathbb{R}),\ N_2(u) = \sum_{k=1}^{+\infty} |u_n|^2$.
- $\ell^\infty(\mathbb{R}) := \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \exists M > 0 : \forall n \in \mathbb{N},\ |u_n| < M \right\}$, muni de la norme N_∞ définie par : $\forall u \in \ell^\infty(\mathbb{R}),\ N_\infty(u) = \sup_{n\in\mathbb{N}} |u_n|$.
- $(\mathcal{C}(X,Y), \|.\|_\infty)$, lorsque X est compact et Y complet.
- $(L(E,F), |||.|||)$, l'espace des fonctions linéaires continues sur E, avec F complet.

Remarque : les normes $|||.|||$ et $\|.\|_\infty$ ont été introduites avec précision dans le chapitre précédent.

On reprend les notations du chapitre 1, section 1.1.2.

Définition 3.9. *Notion d'algèbre.*
Dans ce qui suit, E désigne un espace muni d'une norme $\|.\|$.

1. *Soit E, un ensemble muni de deux lois de composition internes, une addition notée $+$ et un produit noté $\times$, et d'une loi de composition externe notée ., tel que :*
   (i) $\times$ est distributive par rapport à $+$.
   (ii) $(E, +, .)$ est un $\mathbb{K}$-espace vectoriel.
   (iii) $(E, +, \times)$ est un anneau.
   (iv) $\forall (x,y) \in E^2,\ \lambda \in \mathbb{K},\ \lambda.(x \times y) = (\lambda.x) \times y = x.(\lambda.y)$.
   Une algèbre, également appelée $\mathbb{K}$-algèbre, se note en général $(E, \mathbb{K}, +, ., \times)$ ou tout simplement E s'il n'y a pas d'ambiguité.
   Si l'anneau $(E, +, \times)$ est commutatif (i.e. si la loi $\times$ est commutative), alors on parle d'algèbre commutative.
2. *Si de plus, E est une algèbre munie d'une norme $\|.\|$ telle que $\forall (x,y) \in E^2,\ \|x \times y\| \leq \|x\|.\|y\|$, on dit que $(E, \mathbb{K}, +, ., \times)$ est une **algèbre normée**. On dit alors que la norme est **sous-multiplicative**. On la note souvent $(E, \|.\|)$ pour simplifier. En général, on fait également l'hypothèse supplémentaire suivante : si 1_E désigne l'élément neutre pour la loi $\times$, on a $\|1_E\| = 1$.*
3. *On dit qu'une algèbre normée $(E, \|.\|)$ est une **algèbre de Banach** si E muni de $\|.\|$ est complet.*

Exemples :
- $(\mathbb{C}, \mathbb{R}, +, ., \times)$ est une $\mathbb{R}$-algèbre de Banach commutative normée par la norme euclidienne.
- $(\mathcal{M}_n(\mathbb{R}), \mathbb{R}, +, ., \times)$ est une $\mathbb{R}$-algèbre **non**-commutative normée pour la norme subordonnée $|||.|||$.
- Si E est un $\mathbb{K}$-espace vectoriel, $(\mathcal{L}(E), \mathbb{K}, +, ., \circ)$ est une $\mathbb{K}$ algèbre **non**-commutative normée pour la norme subordonnée $|||.|||$. On rappelle d'ailleurs (ce qui a déjà été montré dans le

chapitre précédent) que si u et v désignent deux éléments de $\mathcal{L}(E)$, on a : $|||u \circ v||| \leq |||u|||\,.\,|||v|||$. De plus, si E est complet, cette algèbre est de Banach.

- $(\mathbb{R}^3, \mathbb{R}, +, ., \wedge)$, où $\wedge$ désigne le produit vectoriel sur $\mathbb{R}^3$ est une $\mathbb{R}$-algèbre de Banach **non-commutative** normée pour la norme euclidienne. On remarque d'ailleurs que la propriété de sous-multiplicativité de la norme est assurée par le fait que, si u et v désignent deux éléments de $\mathbb{R}^3$, on a : $\|u \wedge v\|_2 = \|u\|_2.\|v\|_2 |\sin \widehat{(u,v)}| \leq \|u\|_2.\|v\|_2$.

Remarquons que l'on a la propriété suivante pour une algèbre normée :

Propriété 3.7. *Si $(E, \|.\|)$ est une algèbre normée. Alors, $\forall n \in \mathbb{N},\ \|u^n\| \leq \|u\|^n$.*

Démonstration. La démonstration est aisée. Elle se fait par récurrence, en considérant le fait que $\forall x \in E^2,\ \|x^n\| = \|xx^{n-1}\| \leq \|x\|.\|x^{n-1}\|$, en utilisant le fait que la norme est sous-multiplicative. □

3.2.2 Séries dans un espace de banach

Dans toute cette section, E désigne un espace de Banach, muni d'une norme notée $\|.\|$.

Définition 3.10. *Notion de série.*
Soit $(x_n)_{n\in\mathbb{N}}$, une suite d'éléments de E. On note $(\sum x_n)$ la suite $(S_n)_{n\in\mathbb{N}}$ définie par la relation : $\forall n \in \mathbb{N},\ S_n = \sum_{k=0}^{n} x_k$.
*On dit que $(\sum x_n)$ converge si, et seulement si la suite $(S_n)_{n\in\mathbb{N}}$ est de Cauchy dans E. Dans ce cas, on appelle S sa limite, également appelée **somme** de la série $(\sum x_n)$, et on note : $S = \sum_{k=0}^{+\infty} x_n$.*

Remarque : le critère de Cauchy pour la série $(\sum x_n)$ s'écrit ainsi :

$$\left(\sum x_n\right) \text{ converge } \iff \forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall (p,q) \in (\mathbb{N}^*)^2,\ q > p > N \Longrightarrow \|S_p - S_q\| < \varepsilon.$$

Comme dans le cas des séries à termes dans $\mathbb{R}$ ou $\mathbb{C}$, on définit la notion de convergence absolue.

Définition 3.11. *Convergence absolue d'une série.*
On dit qu'une série $(\sum x_n)$ d'éléments de E est absolument convergente lorsque la série à termes positifs $(\sum \|x_n\|)$ est convergente.

L'intérêt des séries absolument convergentes réside dans la propriété suivante :

Propriété 3.8. *Soit $(\sum x_n)$, une série absolument convergente d'éléments de l'espace de Banach E. Alors $(\sum x_n)$ est convergente.*
En revanche, la réciproque n'est pas vraie.

Contre-exemple : la critère spécial des séries alternées prouve que la série $\sum \frac{(-1)^n}{\sqrt{n}}$ est convergente. En revanche, le critère de Riemann pour les séries à termes positifs prouve que la série $\sum \frac{1}{\sqrt{n}}$ est divergente. Ceci montre que la réciproque n'est pas vraie. Mais passons à présent à la démonstration de cette proposition.

Démonstration. Soit $(\sum x_n)$, une série d'éléments de E. Soient p et q, deux entiers tels que $q > p$. Appelons $(S_n)_{n\in\mathbb{N}}$, la suite d'éléments de E définie pour tout entier naturel n par : $S_n = \sum_{k=0}^{n} x_k$. On a :

$$\|S_q - S_p\| = \left\| \sum_{k=p+1}^{q} x_k \right\| \leq \sum_{k=p+1}^{q} \|x_k\| = \sum_{k=0}^{q} \|x_k\| - \sum_{k=0}^{p} \|x_k\|.$$

Appelons $(T_n)_{n\in\mathbb{N}}$, la suite définie pour tout entier naturel n par : $T_n = \sum_{k=0}^{n} \|x_k\|$. Par hypothèse, $(T_n)_{n\in\mathbb{N}}$ converge car la série $(\sum x_n)$ est absolument convergente, donc $(T_n)_{n\in\mathbb{N}}$ est une suite de Cauchy de $\mathbb{R}$, et d'après l'inégalité ci-dessus, il s'ensuit que $(S_n)_{n\in\mathbb{N}}$ est une suite de Cauchy d'éléments de E, espace de Banach. On en déduit que $(\sum x_n)$ est convergente dans E. $\square$

Théorème 3.11. *Soit $(E, \|.\|)$, un espace de Banach et $u \in \mathcal{L}(E)$. Posons $e = Id_E$, et supposons que $|||u - e||| < 1$.*
Alors, u est inversible et u^{-1} est continue.

Démonstration. Souvenons-nous que $\mathcal{L}(E)$ est une algèbre normée, on a : $|||(u-e)^n||| \leq |||u-e|||^n$. Or, $|||u - e|||^n$ est le terme général d'une série géométrique convergente, car par hypothèse, $|||u - e||| < 1$. On en déduit que la série $\sum(u - e)^n$ converge absolument dans $\mathcal{L}(E)$. On pose donc $S(u) := \sum_{k=0}^{+\infty}(e-u)^n \in \mathcal{L}(E)$. Posons, de la même façon pour tout entier naturel n $S_n(u) := \sum_{k=0}^{n}(e-u)^k$ et $v := e - u$. On a :

$$u \circ S_n(u) = [e - (e - u)] \sum_{k=0}^{n}(e - u)^k = (e - v) \sum_{k=0}^{n} v^k = e - v^{n+1}.$$

La dernière égalité est obtenue à l'aide de la formule des anneaux. De plus, puisque $v^{n+1} = (e-u)^{n+1} \xrightarrow[n\to+\infty]{\text{dans } \mathcal{L}(E)} 0_E$, car $|||u-e||| < 1$, on en déduit que $\lim_{n\to+\infty} u \circ S_n(u) = e$. Or, considérons à présent l'application $\Lambda : \begin{array}{ccc} (\mathcal{L}(E))^2 & \longrightarrow & \mathcal{L}(E) \\ (u, v) & \longmapsto & u \circ v \end{array}$. Il est clair que Λ est bilinéaire. De plus, en vertu de l'inégalité $\forall (u, v) \in (\mathcal{L}(E))^2$, $|||u \circ v||| \leq |||u|||\,.\,|||v|||$, il est clair que Λ est continue, donc $\Lambda \in L(E)$. Utilisons à présent la caractérisation séquentielle de la continuité. On a :

$$\lim_{n\to+\infty} u \circ S_n(u) = e = u \circ \lim_{n\to+\infty} S_n(u) = u \circ S(u).$$

Enfin, on sait que u commute avec toute puissance de u, par conséquent, $S_n(u)$ et u commutent pour tout entier naturel n, et ainsi, de la même façon que précédemment, on démontre que : $S(u) \circ u = u \circ S(u) = e$, autrement dit u est inversible et son inverse est $S(u)$. $S(u)$ est continue. Cela se vérifie aisément. $\square$

Je dis à présent un mot sur la notion de convergence uniforme dans un espace de Banach.

Notation : soit $X \subset E$. On appelle $\mathcal{E} := \mathcal{B}(X, E)$, l'espace des fonctions définies sur X et à valeurs dans E. Supposons de plus que E est **complet**. $\mathcal{E}$ est muni de la norme $\|.\|_\infty$ définie pour $f \in \mathcal{E}$ par : $\|f\|_\infty := \sup_{x\in X} \|f(x)\|$.

Définition 3.12. *Différents modes de convergence dans un espace de Banach.*

1. ***Convergence uniforme dans un espace de Banach.***
   Soit $(f_n)_{n\in\mathbb{N}}$, une suite de fonctions de $\mathcal{E}$.
   *On dit que $(f_n)_{n\in\mathbb{N}}$ converge **uniformément** vers $f \in \mathcal{E}$ lorsque $\lim\limits_{n\to+\infty} \|f_n - f\|_\infty = 0$.*
2. ***Convergence uniforme d'une série de fonctions dans un espace de Banach.***
   Soit $(\sum f_n)_{n\in\mathbb{N}}$, une série de fonctions de $\mathcal{E}$.
   *On dit que $(\sum f_n)_{n\in\mathbb{N}}$ converge **uniformément** dans $\mathcal{E}$ lorsque la suite de fonctions $(S_n)_{n\in\mathbb{N}}$ définie pour tout entier naturel n par $S_n(x) = \displaystyle\sum_{k=0}^{n} f_k(x)$ converge uniformément dans $\mathcal{E}$.*
3. ***Convergence normale dans un espace de Banach.***
   Soit $(\sum f_n)_{n\in\mathbb{N}}$, une série de fonctions de $\mathcal{E}$.
   *On dit que $(\sum f_n)_{n\in\mathbb{N}}$ converge **normalement** dans $\mathcal{E}$ s'il existe une série convergente et à termes positifs $(\sum u_n)$ telle que : $\forall n \in \mathbb{N},\ \forall x \in E,\ \|f_n(x)\| \leq u_n$.*

La propriété qui suit est importante. Elle fournit un lien entre convergence normale et convergence uniforme.

Propriété 3.9. *Soit $(f_n)_{n\in\mathbb{N}}$, une suite de fonctions de $\mathcal{E}$.*
Si $(f_n)_{n\in\mathbb{N}}$ converge normalement dans $\mathcal{E}$, alors $(f_n)_{n\in\mathbb{N}}$ converge uniformément dans $\mathcal{E}$.

Ces définitions étant données, passons à présent à un théorème très important en Analyse. Il permet d'intervertir sous certaines hypothèses les symboles de limites d'une suite de fonctions : la limite par rapport à la variable de suite n, et la limite par rapport à la variable de fonction x.

Théorème 3.12. *Théorème de la double limite.*
Soit F, un espace vectoriel normé et $(f_n)_{n\in\mathbb{N}}$, une suite de fonctins de $X \subset E$ à valeurs dans F. Soit $a \in \overline{X}$. Supposons de plus que :
(i) $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur X.
(ii) $\forall n \in \mathbb{N},\ \lim\limits_{x\to a} f_n(x) = b_n$.
Alors, $(b_n)_{n\in\mathbb{N}}$ converge vers $b_\infty \in F$ et on a $\lim\limits_{x\to a} f(x) = b_\infty$, ce qui s'écrit de façon plus synthétique :

$$\lim_{n\to+\infty}\left(\lim_{x\to a} f_n(x)\right) = \lim_{x\to a}\left(\lim_{n\to+\infty} f_n(x)\right).$$

Remarque : ce résultat se transpose bien-sûr sans difficulté aux séries de fonctions. Sous les mêmes hypothèses, on pourra alors écrire que : si $(\sum f_n)_{n\in\mathbb{N}}$ converge uniformément sur X, et si pour $a \in \overline{X}$, et s'il existe $(b_n)_{n\in\mathbb{N}}$ telle que $f_n(x) \xrightarrow[x\to a]{\text{dans } X} b_n$, alors, on a :

$$\sum_{n=0}^{+\infty} \lim_{n\to a} f_n(x) = \lim_{n\to a} \sum_{n=0}^{+\infty} f_n(x).$$

Exemple : j'utilise ici un exemple très célèbre. Il s'agit de la fonction Zeta de Riemann. La fonction Zeta de Riemann est définie, d'après le critère de Riemann sur $]1, +\infty[$ par la relation :

$$\zeta(x) = \sum_{n=1}^{+\infty} \frac{1}{n^x}.$$

De plus, $x \geq 2 \Longrightarrow \forall n \in \mathbb{N}^*,\ \dfrac{1}{n^x} \leq \dfrac{1}{n^2}$. On en déduit que la série de fonction $\left(\sum \dfrac{1}{n^x}\right)$ converge normalement, donc uniformément sur $[2, +\infty[$. Nous sommes dans les conditions 'application du théorème de la double limite. Il s'ensuit que :

$$\lim_{x \to +\infty} \sum_{n=1}^{+\infty} \frac{1}{n^x} = \sum_{n=1}^{+\infty} \lim_{x \to +\infty} \frac{1}{n^x} = 0.$$

Remarque : on pourrait même montrer que $\zeta(x) \underset{x \to 1^+}{\backsim} \dfrac{1}{x-1}$.

3.2.3 Exponentielle d'endomorphismes dans un espace de Banach

Dans toute cette section, E désignera une algèbre de Banach. On note $\|.\|$, la norme d'algèbre dont on le munit. Elle vérifie donc en particulier $\forall (u, v) \in E^2,\ \|uv\| \leq \|u\|.\|v\|$ et en notant $e := Id_E$, on a : $\|e\| = 1$.

3.2.3.1 Définition et premières propriétés

Définition 3.13. *Exponentielle d'endomorphisme.*
Soit $u \in E$. La série $\left(\sum \dfrac{u^n}{n!}\right)$ est absolument convergente donc convergente.
Sa somme est notée $\exp u$ ou e^u.

Démonstration. Pour $u \in E$ donné, on a $\left\| \dfrac{u^n}{n!} \right\| \leq \dfrac{\|u\|^n}{n!}$. La série $\left(\sum \dfrac{u^n}{n!}\right)$ est donc absolument convergente dans E complet ; on en dédit qu'elle converge dans E, ce qui légitime la définition. □

Nous énonçons à présent des propriétés propres à l'exponentielle d'éléments d'une algèbre de Banach. La plupart de ces propriétés semblent assez naturelles, mais soyons tout de même méfiants. Il y a parfois des précautions à prendre.

Propriété 3.10. *Soient $(u, v) \in E^2$. On pose $e := Id_E$.*
(i) $\exp(1_E) = e.1_E$.
(ii) $\|\exp u\| \leq \exp(\|u\|)$.
(iii) Si u et v commutent dans E, alors $\exp(u + v) = \exp u. \exp v$.
(iv) Si $u \in E$, alors $\exp u$ *est inversible et a pour inverse* $\exp(-u)$.
(v) L'application $\Xi : \begin{array}{ccc} E & \longrightarrow & E \\ u & \longmapsto & \exp u \end{array}$ *est une application linéaire continue.*

Démonstration. *(i)* Appelons $(S_n)_{n \in \mathbb{N}}$, la somme partielle associée à la série $\exp(1_E)$. On a :

$$\forall n \in \mathbb{N},\ S_n := \sum_{k=0}^{n} \frac{(1_E)^k}{k!} = \left(\sum_{k=0}^{n} \frac{1}{k!}\right).1_E \xrightarrow[n \to +\infty]{\text{dans } E} e.1_E.$$

(ii) Appelons $(S_n(u))_{n \in \mathbb{N}}$, la somme partielle associée à la série $\exp u$. On a :

$$\|S_n(u)\| = \left\| \sum_{k=0}^{n} \frac{u^k}{k!} \right\| \leq \sum_{k=0}^{n} \frac{\|u\|^k}{k!} = \exp(\|u\|).$$

L'application « norme » étant continue, on peut utiliser la caractérisation séquentielle de la continuité. Puisque $S_n(u) \xrightarrow[n \to +\infty]{\text{dans } E} \exp u$. On en déduit l'inégalité.

(iii) E étant une algèbre de Banach, on peut, en partculier en déduire pour pour u et v, éléments de E qui commutent, on a : $\forall k \in \mathbb{N}^*,\ \dfrac{(u+v)^k}{k!} = \displaystyle\sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} \dfrac{u^i v^{k-i}}{k!}$. C'est la formule du binôme de Newton écrite dans un anneau commutatif. On en déduit donc que $\forall k \in \mathbb{N}^*,\ \dfrac{(u+v)^k}{k!} = \displaystyle\sum_{i=0}^{k} \frac{u^i}{i!} \frac{v^{k-i}}{(k-i)!} = \sum_{p+q=k} \frac{u^p}{p!} \frac{v^q}{q!}$. À ce stade, on introduit la suite $(W_n)_{n\in\mathbb{N}}$ définie par la relation explicite :

$$W_n = \sum_{k=0}^{2n} \frac{(u+v)^k}{k!} - \sum_{p=0}^{n} \frac{u^p}{p!} \sum_{q=0}^{n} \frac{v^q}{q!}.$$

D'après ce que l'on a écrit précédemment, il est clair que $\forall n \neq 0,\ W_n = \displaystyle\sum_{k=0}^{2n} \sum_{p+q=k} \frac{u^p}{p!} \frac{v^q}{q!} - \sum_{p=0}^{n} \frac{u^p}{p!} \sum_{q=0}^{n} \frac{v^q}{q!}$. Introduisons les ensembles d'indices :

$$A_n = \left\{(p,q) \in \mathbb{N}^2 : 0 \leq p \leq n, 0 \leq q \leq n\right\} \text{ et } B_n = \left\{(p,q) \in \mathbb{N}^2 : 0 \leq p+q \leq n\right\}.$$

Il est alors complètement clair que $B_n \subset A_n \subset B_{2n}$, et on peut écrire que $\forall n \in \mathbb{N}^*,\ W_n = \displaystyle\sum_{(p,q)\in B_{2n}\setminus A_n} \frac{u^p}{p!} \frac{v^q}{q!}$. Il s'ensuit que :

$$\forall n \in \mathbb{N}^*,\ \|W_n\| \leq \sum_{(p,q)\in B_{2n}\setminus A_n} \frac{\|u\|^p}{p!} \frac{\|v\|^q}{q!} = \sum_{k=0}^{2n} \frac{(\|u\| + \|v\|)^k}{k!} - \sum_{p=0}^{n} \frac{\|u\|^p}{p!} \frac{\|v\|^q}{q!}.$$

Posons $\forall n \neq 0,\ \alpha_n = \displaystyle\sum_{k=0}^{2n} \frac{(\|u\| + \|v\|)^k}{k!} - \sum_{p=0}^{n} \frac{\|u\|^p}{p!} \frac{\|v\|^q}{q!}$. Un passage à la limite fournit aisément que $\alpha_n \xrightarrow[n\to+\infty]{\text{dans } \mathbb{R}} e^{\|u\|+\|v\|} - e^{\|u\|}.e^{\|v\|} = 0$. Le théorème d'encadrement sur les suites nous indique alors que $\|W_n\| \xrightarrow[n\to+\infty]{\text{dans } \mathbb{R}} 0$, ce qui prouve exactement que $e^{u+v} = e^u.e^v$.

(iv) C'est un corollaire immédiat de (iii).

(v) Avant de commencer cette démonstration, établissons un lemme technique qui nous facilitera la tâche lorsque nous amorcerons la preuve.

Lemme 3.1. *Soient u et v, deux éléments de l'algèbre de Banach E. S'il existe $\lambda > 0$ tel que $\|u\| \leq \lambda$ et $\|v\| \leq \lambda$, alors, on a :*

$$\forall k \in \mathbb{N}^*,\ \|u^k - v^k\| \leq k.\lambda^{k-1}\|u - v\|.$$

Pour démontrer ce lemme, nous raisonnons par récurrence sur k. La cas $k = 1$ est immédiat. Montrons l'hérédité de la récurrence. Si pour un certain entier k, la relation $\|u^k - v^k\| \leq k.\lambda^{k-1}\|u - v\|$ est vérifiée, alors en écrivant que $u^{k+1} - v^{k+1} = (u^k - v^k)u - (u - v)v^k$, on a : $\|u^{k+1} - v^{k+1}\| \leq \|u\|.\|u^k - v^k\| + \|u - v\|.\|v\|^k \leq \lambda k.\lambda^{k-1}\|u - v\| + \|u - v\|\lambda^k = (k+1)\lambda^k\|u - v\|$, ce qui établit l'hérédité, donc le lemme.

À prèsent, pour démontrer (iv), nous allons utiliser la caractérisation séquentielle de la continuité. Soit $(u_p)_{p\in\mathbb{N}}$, une suite d'éléments de E convergeant vers $u \in E$. La suite $(u_p)_{p\in\mathbb{N}}$ est donc bornée car convergente. Il existe donc $\lambda > 0$ tel que, $\forall p \in \mathbb{N}^*$, $\|u_p\| \leq \lambda$, et à la limite, on a encore $\|u\| \leq \lambda$. On en déduit, d'après le lemme que nous venons d'établir, que :

$$\|\exp(u_p) - \exp(u)\| \leq \sum_{n=0}^{+\infty} \frac{\|u_p^n - u^n\|}{n!} \leq \sum_{n=1}^{+\infty} \frac{n\lambda^{n-1}}{n!} \|u_p - u\| = e^{\lambda}.\|u_p - u\| \xrightarrow[n\to+\infty]{\text{dans } \mathbb{R}} 0.$$

On en déduit que : $\exp(u_p) \xrightarrow[n\to+\infty]{\text{dans } E} \exp(u)$.

□

Remarque importante : le cours de première année vous a montré que si $u \in \mathcal{L}(E)$, avec E de dimension finie, alors, en appelant $\mathcal{B}$, une base de E, il existe une correspondance bijective entre u et sa matrice $M = Mat(u, \mathcal{B})$. De plus, $Mat(u \circ v, \mathcal{B}) = Mat(u, \mathcal{B}) \times Mat(v, \mathcal{B})$. C'est pourquoi on définit exactement de la même manière que pour un endomorphisme $u \in \mathcal{L}(E)$, l'exponentielle de la matrice $M = Mat(u, \mathcal{L}(E))$. De plus, on a : $Mat(\exp u, \mathcal{B}) = \exp\left(Mat(u, \mathcal{B})\right)$. Cela se prouve aisément.
Dans la section qui suit, nous nous intéressons à des calculs pratiques d'exponentielle de matrice.

3.2.3.2 Méthodes pratiques de calcul d'exponentielles

Dans cette partie, je vais donner quelques techniques fort utiles si l'on souhaite calculer l'exponentielle d'une matrice et *a fortiori* d'un endomorphisme. Ces techniques de calcul nécessitent de connaître des rudiments de réduction d'endomorphisme. Je redonnerai donc, mais sans rentrer dans les détails ces rudiments. Pour ceux qui souhaitent se rafraîchir les idées, je vous conseille de vous référer à un mannuel d'Algèbre linéaire.

<u>Cadre</u> : on se place dans une $\mathbb{R}$-algèbre normée E. Les matrices d'endomorphismes considérées seront donc à coefficients réels. J'appelle $\mathcal{B}$, une base de E.

1. **Cas où la matrice est diagonalisable :** ce cas est le plus simple. Je rappelle brièvement la définition :

   Définition 3.14. *Endomorphisme diagonalisable.*
   Soit E, un espace vectoriel et u, un endomorphisme. Alors :
   (i) On dit que u est diagonalisable si, et seulement si il existe une base de E formée des vecteurs propres de u.
   (ii) Notons $\sigma(u)$ l'ensemble des valeurs propres de u (on dit aussi le spectre de u), et pour $\lambda \in \sigma(u)$, on appelle E_λ, le sous-espace propre associé (i.e. l'ensemble des vecteurs x de E solutions de l'équation $u(x) = \lambda.x$).
   u est diagonalisable si, et seulement si, $E = \bigoplus_{\lambda\in\sigma(u)} E_\lambda$.
   (iii) Si E est de dimension finie et si u est diagonalisable, alors le polunôme caractérisque associé P_u défini par la relation $P_u(X) = \det(u - \lambda.I)$, où I désigne la matrice identité de E, est scindé, autrement dit, P_u se décompose en un produit de polynômes de degré 1.
   (iv) La caractérisation suivante est intéressante si E est de dimension finie : si P_u, le polynôme caractérisque de u est scindé à racines simples, alors u est diagonalisable.

   Si A est diagonalisable, il existe une base $\mathcal{B}'$ dans laquelle la matrice de l'endomorphisme u associé est diagonale. Ainsi, on peut écrire :

   $$A = Mat(u, \mathcal{B}) = Mat(\mathcal{B}', \mathcal{B}) \times Mat(u, \mathcal{B}') \times Mat(\mathcal{B}, \mathcal{B}').$$

$Mat(\mathcal{B}', \mathcal{B})$ désigne bien-sûr la matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$. On comprendra mieux, à l'aide d'un exemple, comment l'on calcule les coefficients de cette matrice. Notons : $P := Mat(\mathcal{B}, \mathcal{B}')$. On a alors :

$$A = P^{-1}DP.$$

Remarquons que, si n désigne un entier naturel non nul, alors :

$$A^n = (P^{-1}DP)^n = \underbrace{(P^{-1}DP)(P^{-1}DP)...(P^{-1}DP)}_{n \text{ fois}} = P^{-1}D^nP.$$

On en déduit que :

$$\exp A = \sum_{n=0}^{+\infty} \frac{A^n}{n!} = \sum_{n=0}^{+\infty} \frac{P^{-1}D^nP}{n!} = P^{-1} \left(\sum_{n=0}^{+\infty} \frac{D^n}{n!} \right) P = P^{-1} \exp DP.$$

Je donne tout de suite un exemple. Considérons la matrice A définie par :

$$A = \begin{pmatrix} 4 & 1 & -2 \\ 2 & 3 & -2 \\ 8 & 4 & -5 \end{pmatrix}.$$

Son polynôme caractérisque est donné par :

$$P(X) = (X-1)(X+1)(X-2).$$

Cherchons à déterminer les sous-espaces propres associés à :

- $\lambda_1 = 1$: on résout donc l'équation $AX = \lambda_1.X$. On trouve que les solutions de cette équation sont engendrées par le vecteur :

$$V_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

- $\lambda_2 = 2$: on résout donc l'équation $AX = \lambda_2.X$. On trouve que les solutions de cette équation sont engendrées par le vecteur :

$$V_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

- $\lambda_3 = -1$: on résout donc l'équation $AX = \lambda_3.X$. On trouve que les solutions de cette équation sont engendrées par le vecteur :

$$V_3 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

Ces petits calculs nous permettent alors de déterminer la matrice de passage P de la base (V_1, V_2, V_3) de vecteurs propres à la base initiale. On peut alors écrire A sous la forme $A = PDP^{-1}$, avec :

$$P = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & 1 \\ 2 & 0 & 3 \end{pmatrix}, \; P^{-1} = \begin{pmatrix} 2 & 1 & -1 \\ -\frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{4}{3} & -\frac{2}{3} & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Ainsi, et d'après ce qui a été expliqué précédemment, on calcule aisément l'exponentielle de la matrice A. On a donc :

$$\exp(A) = P \begin{pmatrix} e & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & \frac{1}{e} \end{pmatrix} P^{-1}.$$

2. **Cas où A^k se calcule aisément.**
On appelle A la matrice définie par :

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Une idée astucieuse, si l'on souhaite déterminer A^k, pour tout $k \geq 0$, est d'utiliser le théorème de Cayley-Hamilton. Ce théorème nous dit que le polynôme caractérisque de A est un polynôme annulateur de A. Or, dans notre cas, on vérifie aisément que le polynôme annulateur de A a pour expression :

$$P(X) = (X-1)^2(X+1) = X^3 - X^2 - X + 1.$$

On en déduit la relation suivante pour A : $A^3 - A^2 - A + I_2 = 0_{\mathcal{M}_2(\mathbb{R})}$, où I_2 désigne la matrice identité d'ordre 2. Remarquons de prime abord que A^{-1} se calcule aisément. En effet, il est clair que $A(I_2 + A - A^2) = (I_2 + A - A^2)A = I_2$ (A commute avec tout polynôme de A). Ainsi, $A^{-1} = I_2 + A - A^2$. L'idée, pour pouvoir utiliser cette méthode, est d'utiliser le polynôme annulateur de A. Traduisons la division euclidienne du polynôme X^k, pour $k \geq 3$ par le polynôme $X^3 - X^2 - X + 1$. On peut écrire qu'il existe un polynôme P_k de degré $k-3$, et trois suites $(a_k)_{k\geq 3}$, $(b_k)_{k\geq 3}$ et $(c_k)_{k\geq 3}$ tels que :

$$X^k = P_k(X)(X^3 - X^2 - X + 1) + a_k X^2 + b_k X + c_k.$$

Pour déterminer a_k, b_k et c_k en fonction de k, la bonne idée est de faire successivement $X = 1$ et $X = -1$ dans la relation ci-dessus. En remarquant que 1 est une racine double du polynôme annulateur de A, on peut être incité à dériver la relation ci-dessus, car si 1 est racine double du polynôme annulateur, 1 est encore racine de son polynôme dérivé. On en déduit successivement :

$$\forall k \geq 3, \left\{ \begin{array}{rcrcrcl} a_k & + & b_k & + & c_k & = & 1 \\ a_k & - & b_k & + & c_k & = & (-1)^k \\ 2a_k & + & b_k & & & = & k \end{array} \right. .$$

Un résolution classique de ce système donne pour solutions :

$$\forall k \geq 3, \left\{ \begin{array}{rcl} a_k & = & \frac{k}{2} - \frac{1}{4} + \frac{(-1)^k}{2} \\ b_k & = & \frac{1}{2} - +(-1)^{k+1} \\ c_k & = & \frac{3}{4} - \frac{k}{2} + \frac{(-1)^k}{2}. \end{array} \right. .$$

On en déduit que : $\forall k \geq 3,\ A^k = a_k.A^2 + b_k.A + c_k.I_2$. Il s'ensuit que :

$$\begin{aligned} \exp(A) &= I + A + \frac{A^2}{2!} + \left(\sum_{k=3}^{+\infty} \frac{a_k}{k!}\right).A^2 + \left(\sum_{k=3}^{+\infty} \frac{b_k}{k!}\right).A + \left(\sum_{k=3}^{+\infty} \frac{c_k}{k!}\right).I_2 \\ &= \left[\frac{1}{2} - \frac{5}{16}\left(e - \frac{1}{e}\right)\right] A^2 + \left[\frac{1}{4} + \frac{e}{2} - \frac{1}{e}\right].A + \left[\frac{21}{16}.\frac{e}{2} + \frac{21}{16}e - \frac{3}{4}\right] I_2. \end{aligned}$$

En effet, on a par exemple et en utilisant le fait que $\sum_{k=0}^{+\infty} \frac{1}{k!} = e$:

$$\sum_{k=3}^{+\infty} \frac{a_k}{k!} = \frac{1}{2}\sum_{k=3}^{+\infty} \frac{1}{(k-1)!} - \frac{1}{4}\sum_{k=3}^{+\infty} \frac{1}{k!} + \frac{1}{2}\sum_{k=3}^{+\infty} \frac{(-1)^k}{k!} = \frac{1}{2}(e-2) - \frac{1}{4}\left(e - \frac{5}{2}\right) + \frac{1}{2}\left(\frac{1}{e} - \frac{1}{2}\right).$$

Enfin, la formule écrite précédemment définit complètement l'exponentielle de la matrice A. On pourrait poursuivre les calculs pour trouver l'expression explicite de la matrice, mais cela n'a pas vraiment d'intérêt. Seul importe le principe de cette méthode.

3. **cas où la matrice A est trigonalisable.**
Rappelons brièvement ce que l'on entend par : A est trigonalisable.

Définition 3.15. *Endomorphisme trigonalisable.*
Soit E, un espace vectoriel et u, un endomiorphisme de E, et $\mathcal{B}$, une base de E. Alors :
(i) u est trigonalisable si, et seulement s'il existe une base $\mathcal{B}'$ de E telle que : $Mat(u, \mathcal{B}) = P^{-1}Mat(u, \mathcal{B}')P$, où $Mat(u, \mathcal{B}')$ est une matrice ***triangulaire supérieure*** *et P, la matrice de passage de la base $\mathcal{B}$ à $\mathcal{B}'$.*
(ii) Si E est de dimension finie, u est trigonalisable si, et seulement si son polynôme caractéristique est scindé à racines simples dans E, autrement dit s'il peut s'écrire comme un produit de polynômes de degré 1.

Considérons la matrice A définie par la relation :

$$A = \begin{pmatrix} 0 & 3 & -2 \\ 1 & 0 & 1 \\ 3 & -5 & 5 \end{pmatrix}.$$

Le polynôme caractéristique de A est : $P(X) = (1-X)(2-X)^2$. On cherche les sous espaces propres associés aux valeurs propres 1 et 2. Je laisse au lecteur le soin de vérifier les affirmations suivantes :
- Vecteur propre associé à 1 : $f_1(-1, 1, 2)$;
- Vecteur propre associé à 2 : $f_2(1, 0, -1)$.

(f_1, f_2) est une famille de dimension 2. Il faut donc **compléter** cette famille pour obtenir une base de $\mathbb{R}^3$. On choisit par exemple le vecteur $f_3(1, 1, 1)$. En dimension supérieure, il faudrait utiliser la notion de sous-espace caractéristique, que je' ne détaille pas ici, car cela nous enmènerait trop loin. On peut donc écrire que :

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 2 & -1 \\ -1 & 3 & -2 \\ 1 & -1 & 1 \end{pmatrix}.$$

Posons alors :

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{pmatrix}$$

On peut donc écrire que : $A = P^{-1}TP$, et en utilisant la même technique que dans le cas où A est diagonalisable, on montre aisément que : $\forall k \geq 1,\ A^k = P^{-1}T^kP$. Ensuite, T est une matrice triangulaire supérieure. Un la décompose de la façon suivante :

$$T = D + N, \text{ avec } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ et } N = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

D est une matrice diagonale et N est une matrice dite « nilpotente ».

Remarque 1 : une matrice N est dite nilpotente s'il existe un entier naturel k tel que $N^k = 0$.

Remarque 2 : la décomposition $T = N + D$, avec T, triangulaire supérieure, D diagonale et N nilpotente s'appelle la **décomposition de Dunford**.

Remarquons que $N^2 = 0$, ce qui nous confirme que N est nilpotente. Ensuite, puisque l'ensemble des matrices carrées est un anneau, et que les matrices diagonales commutent avec toute autre matrice, on peut écrire, d'après la formule du binôme de Newton que :

$$\forall k \geq 1,\ T^k = (D+N)^k = \sum_{i=0}^{k} \begin{pmatrix} k \\ i \end{pmatrix} N^i D^{k-i} = D^k + kND^{k-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^k & -k2^{k-1} \\ 0 & 0 & 2^k \end{pmatrix}.$$

Cette écriture nous permet de démontrer, puisque $\displaystyle\sum_{k=0}^{+\infty} \frac{2^k}{k!} = e^2$ et $\displaystyle\sum_{k=1}^{+\infty} \frac{k2^{k-1}}{k!} = e^2$, sur le même principe, que :

$$\exp(A) = P^{-1} \begin{pmatrix} e & 0 & 0 \\ 0 & e^2 & -e^2 \\ 0 & 0 & e^2 \end{pmatrix} P.$$

3.2.4 Espaces de Hilbert

3.2.4.1 Introduction et exemples

Dans cette partie, j'introduis un type particulier d'espace de Banach appelé *espace de Hilbert*. Il existe maints résultats relatifs aux espaces de Hilbert. Ces espaces sont très utilisés et très étudiés. Je ne présenterai donc ici qu'une infime partie de ce que l'on sait actuellement sur ces espaces.

Définition 3.16. *Espaces de Hilbert.*
Soit E, un espace préhilbertien réel ou complexe muni d'un produit $< .,. >$.
On dit que E est un ***espace de Hilbert*** *si E est complet pour la norme induite par $< .,. >$.*

Remarque 1 : je rappelle une enième fois que la norme induite par $< .,. >$ est l'application N définie par la relation :

$$\forall x \in E,\ N(x) = (< x, x >)^{\frac{1}{2}}.$$

Remarque 2 : les théorèmes de projection ou relatifs à l'existence d'un supplémentaire orthogonal, étudiés dans le cas des espaces préhilbertiens s'appliquent, *par définition* aux espaces de Hilbert.

Remarque 3 : un espace de Hilbert est en particulier un espace de Banach, puisqu'il est complet et muni d'une norme : la norme induite.

Je donne à présent quelques exemples d'espaces de Hilbert :

- Tous les exemples donnés dans la partie 3.1 sont encore valables.

- D'après les résultats énoncés dans la dernière section du chapitre précédent, on sait que tout espace préhilbertien de dimension finie est complet, donc définit un espace de Hilbert.
- L'exemple que je donne à présent est un peu plus compliqué. Si je souhaitais réellement présenter proprement les choses, je serais obligé de définir la notion d'intégration au sens de Lebesgue. C'est donc tout à fait délibérément que je resterai assez évasif sur la signification que je donne au symbole $\int$. Pour vous aider à vous représenter les choses, considérez les fonctions que vous avez l'habitude de manipuler, sans chercher de contre exemple « pathologique ». On appelle $\mathcal{F}(\mathbb{R})$, l'espace des fonctions $\mathbb{R} \longrightarrow \mathbb{R}$. On définit l'espace $L^2(\mathbb{R})$ par :

$$L^2(\mathbb{R}) = \left\{ f \in \mathcal{F}(\mathbb{R}) : \int_{-\infty}^{+\infty} |f(x)|^2 dx < +\infty \right\}.$$

Habituellement, on dit que deux fonctions f et g de $\mathcal{F}(\mathbb{R})$ sont égales si, et seulement si elles coïncident en tout point de $\mathbb{R}$, c'est à dire si $\forall x \in \mathbb{R},\ f(x) = g(x)$. Pour pouvoir définir une norme sur $L^2(\mathbb{R})$, il est nécessaire de définir une nouvelle notion d'égalité. Soient f et g, deux fonctions de $L^2(\mathbb{R})$. On dira alors que f et g sont égales **presque partout** si, en définissant $\Omega := \{x \in \mathbb{R} : f(x) \neq g(x)\}$, on a : $\int_\Omega dx = 0$. On notera alors $f \overset{p.p.}{=} g$. On appelle alors $\|.\|_{L^2(\mathbb{R})}$, l'application définie sur $L^2(\mathbb{R})$ par la relation :

$$\forall f \in L^2(\mathbb{R}),\ \|f\|_{L^2(\mathbb{R})} = \sqrt{\int_{-\infty}^{+\infty} |f(x)|^2 dx}$$

$\|.\|_{L^2(\mathbb{R})}$ est une norme sur $L^2(\mathbb{R})$. Remarquons que $\|f\|_{L^2(\mathbb{R})} = 0 \Longrightarrow f \overset{p.p.}{=} 0$. En effet, on sait qu'une intégrale « ne charge pas » les points, autrement dit que si une fonction est nulle en tout point de $\mathbb{R}$ excepté en un certain x_0, l'intégrale de cette fonction vaut 0. C'est pourquoi il nous a fallu être précautionneux dans le sens que l'on donne à l'égalité de deux éléments de $L^2(\mathbb{R})$. Cette précaution étant prise, remarquons que la norme $\|.\|_{L^2(\mathbb{R})}$ est induite par le produit scalaire $< .,. >$ défini par la relation :

$$\forall (f, g) \in \left(L^2(\mathbb{R})\right)^2,\ < f, g >:= \int_{-\infty}^{+\infty} f(x)g(x)dx.$$

Ainsi, $L^2(\mathbb{R})$ muni de ce produit scalaire est un espace de Hilbert. Cet exemple est essentiel en Analyse fonctionnelle.

3.2.4.2 Notion de base hilbertienne

Cette partie contient des résltats un peu théoriques sur les espaces de Hilbert. J'en montrerai l'utilité en détaillant l'exemple des séries de Fourier dans la section qui suit.

Dans toute cette section, H désignera un espace de Hilbert muni d'un produit $< .,. >$, dont la norme induite est notée $\|.\|$.

Définition 3.17. *Base hilbertienne.*
Soit I, un ensemble d'indices et $(e_i)_{i\in I}$, une famille d'éléments de H. On dit que $(e_i)_{i\in I}$ est une base ***hilbertienne*** *de H si, et seulement si :*
(i) La famille $(e_i)_{i\in I}$ est orthonormale.
(ii) Pour tout $x \in H$, $x = \sum_{i\in I} < e_i, x > .e_i$.

Remarque 1 : on déduit de (ii) et du théorème de Pythagore que, en posant $x_i =< x, e_i >$, on a : $\forall x \in H, \sum_{i\in I} |x_i|^2 = \|x\|^2$. Un résultat d'Analyse sur les familles sommables justifie que l'ensemble I est au plus dénombrable, c'est à dire dénombrable ou fini. Je rappelle qu'un ensemble I est dit dénombrable s'il existe une bijection de I dans $\mathbb{N}$. Cela signifie, de façon plus simple que le symbole $\sum_{i\in I}$ peut s'écrire $\sum_{n=1}^{+\infty}$.

Remarque 2 : le système orthonormal $(e_i)_{i\in I}$ est une base orthonormale si et seulement si le sous espace vectoriel engendré par $(e_i)_{i\in I}$ est dense dans H, autrement dit :

$$\overline{\{e_i, i \in I\}} = H.$$

Remarque 3 : <u>attention</u> ! Une base hilbertienne n'est pas nécessairement une base au sens des espaces vectoriels.

Remarque 4 : on peut faire le lien avec les théorèmes de projection étudiés précédemment. C'est ce que l'on traduit par l'inégalité de Bessel.

Propriété 3.11. *Inégalité de Bessel.*
Soit I, un ensemble d'indices fini, non vide. Soit $(e_i)_{i\in I}$, une famille orthonormale de vecteurs de H. Soit $x \in E$. On a l'inégalité :

$$\sum_{i\in I} | < x, e_i > |^2 \leq \|x\|^2.$$

Démonstration. On appelle V, le sous espace vectoriel engendré par $(e_i)_{i\in I}$. On désigne par p_V, la projection sur le sous espace vectoriel de dimension finie V. Le théorème de la projection, version algébrique nous assure que :

$$p_V(x) = \sum_{i\in I} < x, e_i > .e_i.$$

Utilisons alors le théorème de Pythagore. On en déduit que :

$$\sum_{i\in I} | < x, e_i > |^2 = \|p_V(x)\|^2 \leq \|x\|^2.$$

$\square$

L'égalité de Parseval que j'énonce à présent est proche de l'inégalité de Bessel. La différence est la suivante : dans le cas d'une famille orthonormale quelconque dans un espace de Hilbert, on a selement une inégalité, tandis que dans le cas où cette famille est une base hilbertienne, on a une égalité.

Théorème 3.13. *Théorème de Parseval.*
Soit I, un ensemble d'indices fini, non vide. Soit $(e_i)_{i\in I}$, une base hilbertienne de H. Soit $x \in E$. On a l'égalité :

$$\sum_{i\in I} | < x, e_i > |^2 = \|x\|^2.$$

La réciproque est exacte : si, pour tout $x \in H$,la relation de Parseval est vérifiée, alors, la famille $(e_i)_{i\in I}$ est une base hilbertienne de H.

La démonstration de ce théorème ressemble fortement à la démonstration de l'inégalité de Bessel. Il suffit d'utiliser le théorème de la projection, version algébrique, sur un espace de dimension finie n donnée. Ensuite, un passage à la limite fournit le résultat.
L'exercice qui suit est très détaillé. Il peut être utilisé comme un exemple.

Exercice : polynômes de Legendre.
Soit $n \in \mathbb{N}$. On définit la famille de polynômes $(P_n)_{n\in\mathbb{N}}$ par la relation :

$$\forall x \in [-1,1],\ P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n].$$

1. Calculer les cinq premiers termes de cette suite de polynômes.
2. On pose pour tout entier naturel n : $e_n = \sqrt{\dfrac{2n+1}{2}}P_n(x)$.
   Démontrer que $(e_n)_{n\in\mathbb{N}}$ est un système orthonormal de $L^2([-1,1])$.
3. En utilisant le théorème de Weierstrass, relatif à la densité des fonctions polynômiales dans $C([-1,1])$ muni de la norme $\|.\|_\infty$, démontrer que $(e_n)_{n\in\mathbb{N}}$ est une base hilbertienne de $L^2([-1,1])$.

Je ne démontre pas le théorème suivant, car il nécessite l'axiome du choix, ce qui nous entraînerait un peu loin. On s'attendait au résultat énoncé par ce théorème. Il s'agit d'un résultat d'existence.

Théorème 3.14. *Tout espace de Hilbert admet une base hilbertienne.*

Je reformule à présent le théorème de projection énoncé dans le cas d'un epace préhilbertien, appliqué cette fois à un espace de Hilbert. Je ne considère ici que les projections orthogonales.

Propriété 3.12. *Projection orthogonale dans un espace de Hilbert.*
Soit F, un sous-espace ***fermé*** *d'un espace de Hilbert H. Soit $x \in H$.*
Alors, il existe un unique $y \in F$ vérifiant l'une des assertions suivantes :
(i) $x - y \in F^\perp$. (autrement dit $< x-y, z >= 0,\ \forall z \in F$)
(ii) $\|x-y\| = d(x,y) = \inf\{\|x-z\|; z \in F\}$.

Le théorème qui suit est très utilisé en Analyse appliquée. Ce théorème s'appelle le *théorème de représentation de Riesz.*

Théorème 3.15. *Théorème de représentation de Riesz.*
Soit H, un espace de Hilbert. On désigne par H^*, le dual topologique de H, c'est à dire l'ensemble des formes linéaires continues sur H, ou encore l'ensemble des applications linéaires continues $\varphi : H \longrightarrow \mathbb{K}$, où H est un $\mathbb{K}$ espace vectoriel. $\mathbb{K}$ désigne en fait $\mathbb{R}$ ou $\mathbb{C}$.
Si ψ est une forme linéaire continue sur H, alors il existe un unique $a \in H$ tel que :

$$\forall x \in H,\ \psi(x) =< x, a > .$$

En particulier, l'application $a \in H \longmapsto \varphi_a$, où φ_a désigne la forme linéaire définie pour tout $x \in H$ par $\varphi_a(x) :=< x, a >$, est un isomorphisme de H sur H^*.

Démonstration. On peut supposer ψ non identiquement nulle. Posons alors $F = \ker\psi = \psi^{-1}(\{0\})$. Puisque ψ et continue, F est un sous espace vectoriel fermé de H. D'après le théorème de la projection, version algébrique, on en déduit que $F \oplus F^\perp = H$. Mais $F^\perp \neq \{0\}$, sinon, on aurait $F = H$, et ψ est non identiquement nulle. On peut donc choisir $b \in F^\perp$ tel que $\psi(b) = 1$. Soit

$x \in H$. Alors, le vecteur $y = x - \psi(x).b$ appartient à F. De plus, $\psi(y = \psi(x) - \psi(x)\psi(b) = 0$, et ainsi, $y \in F = \ker \psi$. Puisque $b \in F^{\perp}$, on a :

$$< y, b >= 0 =< x - \psi(x).b, b >=< x, b > -\psi(x) < b, b > .$$

Posons $a = \frac{b}{<b,b>}$. On en déduit que $\forall x \in H,\ \psi(x) =< a, x >$.
L'unicité s'obtient aisément, en raisonnant par l'absurde. Si a et a' sont deux éléments distincts de H tels que $\forall x \in H,\ < a, x >=< a', x >$, ainsi $\forall x \in H,\ < a - a', x >= 0$. Il suffit de choisir $x = a - a'$ pour obtenir le résultat souhaité.
La seconde partie de ce théorème découle de ce que l'on vient de montrer. □

3.2.4.3 Exemple : application aux séries de Fourier

Si vous avez déjà étudié les séries de Fourier, vous connaissez alors un exemple particulier d'espace de Hilbert. Je ne souhaite pas m'étendre sur ce thème qui commence déjà à s'éloigner un peu du thème qui nous intéresse : l'Analyse fonctionnelle. Je ne démontrerai donc aucun des théorèmes que je vais énoncer. Je souhaite présenter dans cette section un exemple d'utilisation de la notion de base hilbertienne.

Cadre de notre étude :
- on appelle $\mathcal{C}_{2\pi}$, l'espace des fonctions continues sur $\mathbb{R}$, 2π-périodiques et à valeurs complexes.
- Cet espace est muni d'un produit scalaire noté $< ., . >$, défini par :

$$\forall (f, g) \in \mathcal{C}_{2\pi}^2,\ < f, g >:= \int_0^{2\pi} \overline{f}(x)g(x)dx.$$

- On note $\|.\|$, la norme associée.
- On appelle $\mathcal{CM}_{2\pi}$, l'espace des fonctions continues par morceaux sur $\mathbb{R}$, 2π-périodiques et à valeursz complexes.
- **On admet que $\mathcal{C}_{2\pi}$ est dense dans $\mathcal{CM}_{2\pi}$.**
  Traduction : cela signifie que si $f \in \mathcal{CM}_{2\pi}$, il existe une suiste de fonctions $(f_n)_{n\in\mathbb{N}}$ telle que :

$$\|f_n - f\| \xrightarrow[n\to+\infty]{\text{dans } \mathbb{R}} 0.$$

- On définit la famille de fonctions $(e_n)_{n\in\mathbb{Z}}$ par la relation :

$$\forall x \in \mathbb{R},\ e_n(x) = e^{inx}.$$

Propriété 3.13. *La famille $(e_n)_{n\in\mathbb{Z}}$ est orthonormale dans l'espace $\mathcal{C}_{2\pi}$, autrement dit :*

$$\forall (n, p) \in \mathbb{Z}^2,\ < e_n, e_p >= \frac{1}{2\pi} \int_0^{2\pi} e^{it(n+p)}dt = \begin{cases} 0 & si\ p \neq n \\ 1 & sinon \end{cases} .$$

Définissons à présent les coefficients de Fourier.

Définition 3.18. *Coefficients de Fourier.*
Soit $f \in \mathcal{CM}_{2\pi}$. On définit la famille $(c_n(f))_{n\in\mathbb{Z}}$ par :

$$\forall n \in \mathbb{Z},\ c_n(f) :=< f, e_n >= \frac{1}{2\pi} \int_0^{2\pi} f(x)e^{-inx}dx.$$

Remarque 1 : $\forall f \in \mathcal{CM}_{2\pi}$, $|c_n(f)| \leq \|f\|_\infty$, en utilisant des propriétés ultra classiques des intégrales. On en déduit que l'application $f \longmapsto c_n(f)$ est linéaire et continue.

Remarque 2 : les coefficients c_n possèdent la propriété remarquable suivante : si $f \in \mathcal{C}_{2\pi}^{k-1:}$ et est de classe $\mathcal{C}^k$ par morceaux, alors :

$$c_n(f) = \underset{n\to+\infty}{o} \frac{1}{n^k}.$$

Nous allons à présent appliquer les résultats énoncés dans la section précédente aux séries de Fourier.

Appelons V, le sous espace vectoriel de $\mathcal{CM}_{2\pi}$ engendré par la famille de fonctions $(e_n)_{n\in\mathbb{Z}}$. Cet espace est appelé **espace vectoriel des polynômes trigonométriques.** Appelons V_n, pour $n \in \mathbb{Z}$, le sous espace défini par :

$$\forall n \in \mathbb{Z},\ V_n = \vec{(}e_k)_{-n\leq k\leq n}.$$

Soit $f \in \mathcal{CM}_{2\pi}$. Pour $n \in \mathbb{Z}$, on appelle S_n, la projection orthogonale de f sur V_n, autrement dit :

$$S_n(f) = \sum_{k=-n}^{n} < e_k, f > .e_k = \sum_{k=-n}^{n} c_n(f).e_k.$$

Définition 3.19. *Séries de Fourier.*
On appelle série de Fourier d'un élément f de $\mathcal{CM}_{2\pi}$, la série indexée par $\mathbb{Z}$ de terme général : $c_k(f).e_k$.
On dit alors que $S_n(f)$ est la somme partielle de rang n de la série de Fourier de f.

Traduisons l'inégalité de Bessel, étudiée dans la section précédente, au cas des séries de Fourier :

$$\forall f \in \mathcal{CM}_{2\pi},\ \forall n \in \mathbb{N},\ \sum_{k=n}^{n} |c_k(f)|^2 \leq \|f\|^2 < +\infty.$$

Remarque : cela nous montre en particulier que la famille $(c_k(f))_{k\in\mathbb{Z}}$ est de carré sommable, ou encore, avec les notations utilisées en début de chapitre que $(c_k(f))_{k\in\mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C})$.

Enfin, introduisons la notion de convergence en moyenne quadratique. On admet le théorème suivant :

Théorème 3.16. *L'espace $L^2([0, 2\pi])$ est un espace de Hilbert pour le produit scalaire $< .,. >$ défini précédemment. De plus, la famille $(e_n)_{n\in\mathbb{Z}}$ constitue une base hilbertienne de $L^2([0, 2\pi])$.*

Les espaces $\mathcal{C}_{2\pi}$ et $\mathcal{CM}_{2\pi}$ sont des espaces préhilbertiens. Ce sont des sous espaces de $L^2([0, 2\pi])$. On en déduit, en appliquant le théorème de projection, étudié dans la section précédente, à une fonction $f \in \mathcal{CM}_{2\pi} \subset L^2([0, 2\pi])$ le théorème :

Théorème 3.17. *Égalité de Parseval.*
Soit $f \in \mathcal{CM}_{2\pi}$. Alors :

$$\|f - S_n(f)\| \xrightarrow[n\to+\infty]{} 0 \text{ et on a } \|f\|^2 = \sum_{n\in\mathbb{Z}} |c_n(f)|^2.$$

Il s'agit là d'une conséquence directe de la notion même de base hilbertienne et de projection orthogonale dans un espace de Hilbert.
Enfin, j'énonce à présent un théorème fort célèbre sur les séries de Fourier. Il s'agit du théorème de Dirichlet, qui nous fournit des renseignements sur la notion de convergence simple d'une fonction vers sa série de Fourier.

Théorème 3.18. *Théorème de Dirichlet.*
Soit $f \in \mathcal{CM}_{2\pi}$, une fonction de classe $\mathcal{C}^1$ par morceaux. Alors sa série de Fourier converge simplement sur $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) et, pour tout $x \in \mathbb{R}$, on a :

$$\lim_{n \to +\infty} S_n(f)(x) = \frac{1}{2} \lim_{h \to 0^+} [f(x+h) + f(x-h)].$$

En particulier, si f est continue, on a :

$$\forall x \in \mathbb{R},\ f(x) = \sum_{n=0}^{+\infty} c_n(f).e_n.$$

Exemple : appelons f, la fonction définie sur $\mathbb{R}$ par :

$$f(x) = \begin{cases} 1 & \text{si } x \in [0, \pi[\\ -1 & \text{si } x \in [\pi, 2\pi[\\ f \in \mathcal{CM}_{2\pi} & \end{cases}$$

Calculons les coefficients de Fourier de f. Soit $n \in \mathbb{Z}$. Alors :

$$c_n(f) =< f, e_n >= \frac{i}{2\pi} \left(\int_0^{\pi} \sin(nx)dx - \int_{\pi}^{2\pi} \sin(nx)dx \right) = \begin{cases} \frac{4}{n\pi} & \text{si } n \text{ est impair ;} \\ 0 & \text{sinon.} \end{cases}$$

Sur l'intervalle $]0, 2\pi[$ par exemple, on peut donc écrire d'après le théorème de Dirichlet que :

$$f(x) = \frac{4}{\pi} \sum_{k=0}^{+\infty} \frac{\sin[(2n+1)]}{2n+1}.$$

De plus, on a convergence en moyenne quadratique de la série de Fourier de f vers f.

Index