Les Mathématiques pour l'Agrégation

C. Antonini
J.-F. Quint
P. Borgnat
J. Bérard
E. Lebeau
E. Souche
A. Chateau
O. Teytaud

21 mai 2002

Table des matières

Chapitre 1

Notations et définitions usuelles

$A \wedge B$ note la conjonction logique ; $A \wedge B$ si et seulement si A et B.
$A \vee B$ note la disjonction logique ; $A \vee B$ si et seulement si A ou B.
$A \rightarrow B$ signifie que A implique B.
$A \iff B$ signifie que A et B sont logiquement équivalents, c'est à dire que si on a A, on a nécessairement B.
Une relation sur E est un sous-ensemble de $E \times E$.
$x \in A$ signifie que x est un élément de l'ensemble A.
$A \subset B$ signifie que l'ensemble A est inclus dans (éventuellement égal à) B.
Une suite est dite **presque nulle** si elle est nulle sauf pour un nombre fini de valeurs de l'index.
Une relation $\mathcal{R}$ est dite **réflexive** si pour tout x $x\mathcal{R}x$.
Une relation $\mathcal{R}$ est dite **symétrique** si pour tout x et tout y $x\mathcal{R}y$ si et seulement si $y\mathcal{R}x$.
Une relation $\mathcal{R}$ est dite **antisymétrique** si pour tout x et tout y $x\mathcal{R}y \wedge y\mathcal{R}x \rightarrow x = y$.
Une relation $\mathcal{R}$ est dite **transitive** si pour tout x, tout y et tout z, $x\mathcal{R}y \wedge y\mathcal{R}z \rightarrow x\mathcal{R}z$.
Une relation $\mathcal{R}$ est une **relation d'équivalence** si elle est réflexive, symétrique et transitive.
Une relation $\mathcal{R}(x,y)$ est une **fonction** $\mathcal{R}$ est telle que pour tout x il existe au plus un y tel que $\mathcal{R}(x,y)$. C'est une **application** de A dans B si elle est incluse dans $A \times B$ et si pour tout x il existe un et un seul y dans B tel que $\mathcal{R}(x,y)$.
Etant donnée une relation d'équivalence $\mathcal{R}$ sur E, la **classe d'équivalence** de $x \in E$ est l'ensemble des y tels que $x\mathcal{R}y$. L'ensemble des classes d'équivalence réalise une partition de E.
Etant donné un ordre, on appelle **segment** d'extrémités a et b l'ensemble des x tels que $a \leq x \leq b \vee b \leq x \leq a$.
On note $liminf_{n \rightarrow +\infty} x_n$ la valeur $sup_{n \rightarrow +\infty} inf_{k \geq n} x_n$.
On note de même $limsup_{n \rightarrow +\infty} x_n$ la valeur $inf_{n \rightarrow +\infty} sup_{k \geq n} x_n$.
Etant donnés deux ensembles E et F avec $E \subset F$ on note χ_E ou 1_E et on appelle **fonction caractéristique de** E la fonction qui à x dans F associe 1 si $x \in E$ et 0 sinon.
On définit $\delta_{i,j}$ par $\delta_{i,j} = 1$ si $i = j$ et 0 sinon. Cette notation est appelée **fonction delta de Kronecker**.
On appelle **support** d'une fonction l'adhérence de l'ensemble des x tels que $f(x) \neq 0$.

signifie "Attention !"

signifie "Application" (sert à convaincre du fait (réel !) que les mathématiques servent à quelque chose !)

signifie "Pour y voir plus clair".

signifie "Remarque".

Chapitre 2

Ensembles ordonnés

Définition 1 *Soit E un ensemble. Un* **ordre** *(partiel) sur E est une relation $\leq$ telle que pour tout $(x, y, z) \in E^3$:*
• $x \leq x$
• $(x \leq y \wedge y \leq x) \rightarrow x = y$
• $(x \leq y \wedge y \leq z) \rightarrow x \leq z$
Ces trois propriétés sont respectivement la **réflexivité**, *la* **symétrie** *et la* **transitivité**.

E équipé d'un tel ordre est appelé «ensemble partiellement ordonné».

Un ordre $\leq$ donne naissance à une relation d'inégalité stricte $<$ par : $x < y \iff (x \leq y \wedge x \neq y$.

On définit aussi :
• $x \geq y \iff y \leq x$
• $x \not\leq y \iff \neg(x \leq y)$
• $x \parallel y \iff x \not\leq y \wedge y \not\leq x$ (x et y ne sont pas comparables)

Soit F un sous-ensemble de E, E étant muni d'un ordre partiel $\leq_E$; on définit l'ordre partiel $\leq_E$ **induit** sur F par $x \leq_F y \iff x \leq_E y$.

Définition 2 *Un ensemble E muni d'un ordre partiel E est dit* **totalement ordonné** *si et seulement si $\forall (x, y) \in E^2 x \leq y \vee y \leq x$. Un ensemble totalement ordonné est aussi appelé une* **chaîne**. *Un ensemble tel que $x \leq y \rightarrow x = y$ est appelé une* **antichaîne**.

Définition 3 *Une chaîne C est dîte maximale si et seulement si quel que soit l'élément x, l'ensemble $C \cup \{x\}$ n'est pas une chaîne.*
Une antichaîne C est dîte maximale si et seulement si quel que soit l'élément x, l'ensemble $C \cup \{x\}$ n'est pas une antichaîne.

On note n la chaîne $[0, n[$.
Dans la suite du texte $\leq$ désigne une relation d'ordre partiel.

Définition 4 *Etant donné $\leq$, on définit la* **relation de couverture** $\prec$ *par $x \prec y$ (y couvre x ou x est couvert par y) si et seulement si $x < y \wedge \forall z(x \leq z < y \rightarrow z = x)$. Ceci signifie qu'il n'y a pas de z tel que $x < y < z$.*

Si E est fini, la relation de couverture détermine la relation d'ordre partiel (et réciproquement).

Définition 5 *On définit maintenant le* **diagramme de Hasse** *pour un ensemble fini partiellement ordonné. A chaque élément de E on associe un point du plan, et on trace une ligne de x à y si $x \prec y$. On veille à ce que ces lignes n'intersectent pas les autres points, et on veille à ce que $x \prec y$ implique que l'ordonnée du point associé à x soit inférieure à l'ordonnée du point associé à y.*

Définition 6 *Une application $\phi : E \rightarrow F$ est dîte :*
• **monotone** *si $x \leq y \rightarrow \phi(x) \leq \phi(y)$.*
• *un* **morphisme** *si $x \leq y \iff \phi(x) \leq \phi(y)$.*
• *un* **isomorphisme d'ordre** *si c'est un morphisme bijectif.*

Quand ϕ est un morphisme, on écrit $\phi : E \hookrightarrow F$.
Quand ϕ est un isomorphisme on écrit $E \cong F$; E et F sont isomorphes.

Proposition 7 *Soit ϕ bijective de E dans F : alors les trois énoncés suivants sont équivalents :*
• *ϕ est un isomorphisme d'ordre*
• *$x < y$ dans E si et seulement si $\phi(x) < \phi(y)$ dans F*
• *$x \prec y$ dans P si et seulement si $\phi(x) \prec \phi(y)$*

Deux ensembles finis ordonnés sont isomorphes si et seulement si ils ont un diagramme de Hasse commun.

Définition 8 *Le* **dual** *d'un ensemble ordonné est le même ensemble mais muni de l'ordre $\leq^\delta$ tel que $x \leq^\delta y$ si et seulement si $y \leq x$. Le dual d'un énoncé ψ et l'énoncé ψ^δ obtenu en remplaçant $\leq$ par $\geq$ et réciproquement.*
Un énoncé est vrai pour tous les ensembles ordonnés si et seulement si son dual est vrai pour tous les ensembles ordonées.

Définition 9 *Soit F sous-ensemble de E tel que $F \subset E$, avec E ordonné. F est un* **idéal d'ordre** *si et seulement si $x \in F \wedge y \leq x \rightarrow y \in F$. F est un* **filtre d'ordre** *si et seulement si $(x \in F \wedge x \leq y) \rightarrow y \in F$.*
F est un filtre d'ordre si et seulement si le complémentaire de F est un idéal d'ordre.

On définit $\downarrow F$ par l'ensemble des x tel que pour un certain y dans F on a $x \leq y$. Par définition $\downarrow x$ est égal à $\downarrow x$. $\downarrow F$ se lit «section initiale engendrée par F».
On définit $\uparrow F$ par l'ensemble des x tel que pour un certain y dans F on a $y \leq x$. Par définition $\uparrow x$ est égal à $\uparrow x$. $\uparrow F$ se lit «section initiale engendrée par F».
$\downarrow F$ est donc le plus petit idéal d'ordre contenant F, et $\uparrow F$ est le plus petit filtre d'ordre contenant F.
On note $O(E)$ l'ensemble des idéaux d'ordre de l'ensemble ordonné E ; il est lui-même ordonné.

Les trois énoncés suivants sont équivalents :
• $x \leq y$
• $\downarrow x \subset \downarrow y$
• $(\forall F \in O(E)) y \in F \rightarrow x \in F$

Définition 10 *x est* **maximal** *si et seulement si $x \leq y \rightarrow x = y$*
x est le **maximum** *de E si et seulement si pour tout y on a $y \leq x$. On écrit $x = maxQ$.*
Les notions de **minimal** *et d'***élément minimum** *sont définies de manière duale, en renversant l'ordre.*
L'élément maximum d'un ensemble est généralement noté $\top$, et l'élément minimum est généralement noté $\perp$.

Lorsque l'ensemble est fini, l'ensemble des éléments maximaux et l'ensemble des éléments minimaux sont des anti-chaînes maximales.
Lorsqu'une chaîne $\{x_1 < x_2 < ... < x_n\}$ est maximale, alors $\forall i \; x_i \prec x_{i+1}$.

Définition 11 *On appelle généralement :*
• **graphe de la relation** *le graphe dans lequel on supprime les réflexivités.*
• **graphe de compatibilité** *l'ensemble des (x, y) avec x comparable à y.*
• **graphe de Hasse** *(ne pas confondre avec diagramme de Hasse) l'ensemble des (x, y) tels que $x \prec y$.*
• **graphe de couverture** *l'ensemble des x, y tels que $x \prec y$ ou $y \prec x$.*

On note dans la suite $E_\perp$ l'ensemble ordonné constitué de l'ensemble E auquel on ajoute une constante $\perp$ inférieure à tous les éléments de E. S'il y avait une relation d'ordre sur E, la relation sur $E_\perp$ contient cette relation. S'il n'y en avait pas, on obtient ce que l'on appelle un ordre plat.

Définition 12 *L'***union disjointe** *$E\dot{\cup}F$ de deux ensembles ordonnés disjoints E et F est l'ensemble union de E et F avec $x \leq_{E\dot{\cup}F} y \iff (x \leq_E y \vee x \leq_F y)$.*
La **somme linéaire** *$E \oplus F$ de deux ensembles ordonnés disjoints E et F est l'ensemble réunion de E et F muni de la relation $x \leq_{E \oplus F} \iff (x \leq_E y \vee x \leq_F y \vee (x \in E \wedge y \in F))$.*
On note $P \oplus_\perp Q$ la **somme séparée** *de P et Q, égale à $(P\dot{\cup}Q)_\perp$.*
On note $P \oplus_\vee Q$ la **somme coalescente** *de P et Q, obtenue en considérant $P\dot{\cup}Q$ et en identifiant les deux éléments $\perp$.*
Le **produit** *de $P_1, ..., P_n$ est défini sur l'ensemble produit cartésien par $(x_1, ..., x_n) \leq_{P_1 \times P_2 \times ... \times Pn} (y_1, ..., y_n) \iff \forall i\ x_i \leq_{P_i} y_i$.*

Soit $X = \{1, 2, ..., n\}$, et soit $\phi : P(X) \to \{0, 1\}^n$ défini par $\phi(A) = (\epsilon_1, ..., \epsilon_n)$ avec $\epsilon_i = 1$ si $i \in A$ et $\epsilon_i = 0$ sinon. Alors ϕ est un isomorphisme d'ordre.

L'ensemble Y^X des applications d'un ensemble X vers un ensemble ordonné Y sont naturellement ordonnées par $f \leq g \iff \forall x\ f(x) \leq g(x)$. Si X est lui-même ordonné, on peut considérer simplement l'ensemble des applications monotones, que l'on note $Y^{<X>}$.
On peut aussi considérer des fonctions au lieu de considérer des applications ; on considère alors que $f \leq g$ si et seulement si pour tout élément x du domaine de définition de f on a $f(x) \leq g(x)$.
Pour ordonner l'ensemble des fonctions de X dans Y on ajoute un élément $\perp$ dans Y inférieur à tous les éléments, et en remplaçant une fonction par l'application qui lui est égale sur son domaine de définition et qui est égale à $\perp$ en dehors de ce domaine. Cette fonction qui à une fonction de X dans Y associe une application de X dans $Y_\perp$ est un isomorphisme d'ordre.

Chapitre 3

Graphes

On s'intéresse ici au cas le plus général d'un ensemble X muni d'une relation binaire U. (X, U) est un graphe orienté.

On note $\Gamma^+(x)$ l'ensemble des y tels que $(x, y) \in U$; on l'appelle ensemble des **successeurs** de x.
On note $\Gamma^-(x)$ l'ensemble des y tels que $(y, x) \in U$; on l'appelle ensemble des **prédécesseurs** de x.
On note $d^+(x) = |\Gamma^+(x)|$ le **degré sortant** ou **degré externe** de x.
On note $d^-(x) = |\Gamma^-(x)|$ le **degré entrant** ou **degré interne** de x.
Si $d^-(x) = 0$ x est appelé une **source**.
Si $d^+(x) = 0$ x est appelé un **puits**.

Théorème 13 *Soit $G = (X, U)$ un graphe orienté fini. G est sans circuit si et seulement si les deux énoncés suivants sont vérifiés :*
• $\exists x \in X\ d^-(x) = 0$
• $\forall x \in X\ d^-(x) = 0 \Longrightarrow G \setminus \{x\}$ *est sans circuit.*

Théorème 14 *Soit $G = (X, U)$ un graphe orienté fini. G est sans circuit si et seulement si il existe une permutation $(x_1, x_2, ..., x_n)$ des sommets tels que $d^-_{G_i}(x_i) = 0$ où $G_i = G[\{x_i, ..., x_n\}]$.*

Ci-dessous un algorithme déterminant si oui ou non un graphe est sans circuit ou non.
Le codage machine employé consiste en une liste de successeurs pour chaque sommet.

Procédure sans-circuit(G)
Début : Pour tout $x \in X$ faire
    $d^-(x) = 0$
Pour tout $x \in X$ faire
    Pour tout $y \in \Gamma^+(x)$ faire

$d^-(y) \leftarrow d^-(y) + 1$
Source $\leftarrow \emptyset$
Nbsommets $\leftarrow 0$
Pour tout $x \in X$ faire
Si $d^-(x) = 0$ alors Source $\leftarrow$ Source $\cup\{x\}$.
Tant que Source $\neq \emptyset$ faire
$x \leftarrow$ choix(Source)
Source $\leftarrow$ Source privé de $\{x\}$
Nbsommets $\leftarrow$ Nbsommets $+1$
Pour chaque successeur y de x faire
$d^-(y) \leftarrow d^-(y) - 1$
Si $d^-(y) = 0$ alors
Source $\leftarrow$ Source $\cup\{y\}$.
Si (Nbsommets$=$ n) alors G est sans circuit sinon G a au moins un circuit.

La complexité de cet algorithme est $O(n + m)$ avec n le nombre de sommets et m le nombre d'arcs.
Source peut être implémentée sous forme de liste, avec pour fonction de choix la fonction simplissime qui choisit le premier élément.

Un **tri topologique** d'un graphe orienté sans circuit $G = (X, U)$ est une permutation $(x_1, x_2, ..., x_n)$ de X telle que $(x_i, x_j) \in U \Longrightarrow i < j$.

Notons que la permutation calculée par l'algorithme précédent (ie. l'ordre de sortie de Source) est un tri topologique.

L'algorithme suivant sert à engendrer *tous* les tris topologiques :

Procédure Tri-topologique(G)
Pour tout $x \in X$ calculer $d^-(x)$ (comme dans l'algorithme précédent)
$S \leftarrow \emptyset$
Pour tout $x \in X$ faire :
si $d^-(x) = 0$ alors $S \leftarrow S \cup \{x\}$
$\sigma \leftarrow \emptyset$
Tri-topo(G, S, σ)

avec la procédure récursive «Tri-topo» suivante :

Tri-topo(G, S, σ)
Si $S = \emptyset$ alors écrire σ sinon
Pour tout $x \in S$ faire
$S' \leftarrow S - \{x\}$
$\sigma \leftarrow \sigma.x$ (concaténation)
Pour tout $y \in \Gamma^+(x)$ faire
$d^-(y) \leftarrow d^-(y) - 1$
Si $d^-(y) = 0$ alors
$S' \leftarrow S' \cup \{y\}$
Tri-topo(G, S', σ)
Pour tout $y \in \Gamma^+(x)$ faire
$d^-(y) \leftarrow d^-(y) - 1$

$\sigma \leftarrow \sigma$ privé de x

La complexité est en $O((n+m)*|L(G)|)$.
Il existe des algorithmes de complexité $O(n*|L(G)|)$ (beaucoup plus compliqués).
$L(G)$ est le nombre de tri topologiques, n le nombre de sommets, m le nombre d'arcs.

On considère maintenant un graphe $G=(X,U)$ orienté et sans circuits. On pose $h(x)=0$ si $d^-(x)=0$ et $h(x)=max\{h(y)|(y,x)\in U\}+1$ sinon ; $h(x)$ est appelé **hauteur** de x.

Algorithme de construction des niveaux du graphe ; un **niveau** est de la forme $X_i=\{x\in X|h(x)=i\}$:
• Calculer les degrés internes.
• nb-sommets=0
• Pour tout $x\in X$ faire
• • Si $d^-(x)=0$ faire
• • • ajouter(x) au niveau 0.
• • • nb-sommets++
• $i\leftarrow 0$
• Tant que niveau(i) est non vide faire
• • Pour tout x dans niveau(i) faire
• • • pour tout y successeur de x faire • • • • $d^-(y)\leftarrow d^-(y)-1$
• • • • Si $d^-(y)=0$ alors
• • • • • niveau($i+1$) $\leftarrow$ niveau($i+1$) $\cup\{y\}$
• • • • • nb-sommets $\leftarrow$ nb-sommets $+1$
• • $i\leftarrow i+1$

Il y a des circuits si et seulement si à la fin de l'algorithme nb-sommets est différent de n.
La complexité est en $O(n+m)$.

Un chemin sur un graphe $G=(X,U)$ est un n-uplet $(x_1,...,x_k)$ tel que $\forall i(x_i,x_{i+1})\in U$.
Soit $G=(X,U)$ un graphe orienté, la **fermeture transitive** $G_f=(X,U_f)$ de G est définie par $(x,y)\in U_f$ si et seulement si il existe un chemin $(x_1,...,x_k)$ de G avec $x=x_1$ et $y=x_k$.
Soit $G=(X,U)$ un graphe orienté avec $X=\{1,...,k\}$ et $X_0=\emptyset$.
Soit $\mu=(x_1,...,x_l)$ un chemin de G, l'intérieur de μ, noté $I(\mu)$, est $\{x_2,...,x_{l-1}\}$.
$I(\mu)=\emptyset$ si et seulement si $l\leq 2$.
On définit une suite de graphes $G_k=(X,U_k)$ de la manière suivante : $(x,y)\in U_k$ si et seulement si il existe un chemin c entre x et y avec $I(c)\subset X_k$. En particulier $G_0=G\cup\{(x,x)|x\in X\}$, et $G_n=G_f$.

Lemme : $(i,j)\in U_k \iff ((i,j)\in U_{k-1}\vee((i,k)\in U_{k-1}\wedge(k,j)\in U_{k-1}))$

On déduit de ce lemme un algorithme destiné à déterminer la fermeture transitive d'un graphe, l'**algorithme de Roy-Warshall** :
• $a\leftarrow g$
(on initialise la matrice avec le graphe)
• Pour i variant de 1 à n faire $a_{i,i}\leftarrow 1$

• Pour k variant de 1 à n faire
• • Pour i variant de 1 à n faire
• • • Pour j variant de 1 à n faire
• • • • $a_{i,j} \leftarrow a_{i,j} \vee (a_{i,k} \wedge a_{k,j})$

On peut constater que l'on n'applique pas exactement le lemme, car les $a_{i,j}$ ne sont pas les bons, mais l'on peut aussi montrer que l'algorithme fonctionne tout de même.

La complexité de l'algorithme de Roy-Warshall est $O(n^3)$.

Soit $G = (X, U)$ un graphe orienté, $G_r = (X, U_r)$ est une **réduction transitive** de G si $U_r \subset U$, G_r n'a pas de transitivité et $(G_r)_f = G_f$.
La réduction transitive n'est pas toujours unique.
Si $G = (X, U)$ est un graphe orienté sans circuit, il a une unique réduction transitive appelé **graphe de Hasse**.

Algorithme général de calcul d'une fermeture transitive (pas seulement dans le cas d'un graphe sans circuit).
• Pour $x \in X$ faire
• • marque(x) $\leftarrow$ faux
• • $\Gamma_f^+(x) \leftarrow \emptyset$
• Pour tout $x \in X$ faire
• • file $\leftarrow \{x\}$
• • tant que file $\neq \emptyset$ faire
• • • $y \leftarrow$ premier(file)
• • • supprimer (y,file)
• • • Pour tout $z \in \Gamma^+(y)$ faire
• • • • Si marque(y)=faux alors
• • • • • marque(z) $\leftarrow$ vrai
• • • • • marque($\Gamma_f^+(x) \leftarrow \Gamma_f^+(x) \cup \{z\}$)
• • • • • file $\leftarrow$ file $\cup$ { z }
• • Pour tout $z \in Gamma_f^+(x)$ faire
• • • marque(x) $\leftarrow faux$

Le coût de cet algorithme est $O(n|U| + |U_f|)$.

On cherche maintenant à déterminer à la fois la fermeture et la réduction transitive dans un graphe sans circuit.
L'algorithme employé est l'**algorithme de Goralcicova-Koubek**.

• Pour tout $x \in X$ faire
• • $\Gamma_f^+(x) \leftarrow \emptyset$
• • $\Gamma_r^+(x) \leftarrow \emptyset$
• • marque(x) $\leftarrow$ faux
• On classe les points dans l'ordre d'un tri topologique $\tau = (x_1, x_2, ..., x_n)$
• Pour i variant de n à 1 faire
• • $S \leftarrow \Gamma^+(x_i)$
• • Tant que $S \neq \emptyset$
• • • $x \leftarrow min_\tau(S)$

• • • Si non(marque(X)) alors
(cas où (x_i, x) n'est pas un arc de transitivité)
• • • • $\Gamma_r^+(x_i) \leftarrow \Gamma_r^+(x_i) \cup \{x\}$
• • • • $\Gamma_f^+(x_i) \leftarrow \Gamma_f^+(x_i) \cup \{x\}$
• • • • marque(x) $\leftarrow$ vrai
• • • • Pour tout $y \in \Gamma_f^+(x)$ faire
• • • • • si non(marque(y)) alors
• • • • • • marque(y) $\leftarrow$ vrai
• • • • • • $\Gamma_f^+(x_i) \leftarrow \Gamma_f^+(x_i) \cup \{y\}$
• • • • $S \leftarrow S \setminus \{x\}$
• • Pour tout $y \in \Gamma_f^+(x_i)$ faire
• • • marque(y) $\leftarrow$ faux

La complexité est en $O(n + n|U_r| + |U_f|)$.

Problèmes ouverts : peut on réaliser la fermeture transitive de G en $O(n + |U_f|)$? La réduction transitive en $O(n + |U|)$? Comment détecter qu'un graphe est fermé transitivement ? Réduit transitivement ? (probablement faisable en temps linéaire)

Chapitre 4

Théorie des ensembles - autres systèmes axiomatiques - construction des ensembles usuels

4.1 Les axiomes de la théorie des ensembles de Zermelo-Fraenkel

Une **classe** est associée à une propriété d'une seul élément ; c'est à dire que l'on se donne une assertion comportant une et une seule variable libre ; un élément est dans la classe correspondante s'il vérifie l'assertion. Les formules comportant plusieurs variables libres sont appelées **relations**. Eventuellement on peut avoir une distinction entre des variables et des paramètres ; dans ce cas on a une classe pour chaque valeur possible des paramètres.
La théorie des ensembles est basée sur un ensemble d'axiomes. Les objets de cette théorie sont appelés **ensembles**, et la classe des ensembles est appelée **univers**. Les axiomes de la théorie des ensembles de Zermelo-Fraenkel sont les suivants :

Axiome 15 *Axiome d'extensionalité :*

$$\forall x \forall y (\forall z (z \in x \iff z \in y) \to x = y$$

(deux ensembles sont égaux si et seulement si ils contiennent exactement les mêmes éléments)

Axiome 16 *Axiome de l'union :*

$$\forall x \exists y \forall z (z \in y \iff \exists t (t \in x \wedge z \in t)$$

(une union d'ensembles est un ensemble)

Axiome 17 *Axiome de l'ensemble des parties :*

$$\forall x \exists y \forall z (z \in y \iff z \subset x)$$

(les parties d'un ensembles forment une partie. On note $x \subset y$ l'assertion $\forall z (z \in x \to z \in y)$)

Axiome 18 *Axiome du schéma de remplacement :*
Etant donné une formule $\mathcal{R}(x, y, z_0, ..., z_n)$ de paramètres $z_0, ..., z_n$, définissant pour toute valeur des z_i une fonction, alors :

$$\forall z_0 ... \forall z_n (\forall x \forall y \forall y' E(x, y, z_0, ..., z_n) \wedge E(x, y', z_0, ..., z_n) \to y = y')$$

$$\to \forall t \exists w \forall v (v \in w \iff \exists u (u \in t \wedge E(u, v, x_0, ..., x_{k-1})))$$

On ajoute usuellement un axiome supplémentaire à ces axiomes :

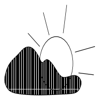
Axiome 19 *axiome de l'infini, qui affirme qu'il existe un ordinal infini. Nous verrons plus loin ce qu'est un ordinal, et ce qu'est un ordinal fini.*

Théorème 20 *La consistance de ces axiomes n'est pas changée si on remplace l'axiome de l'infini par sa négation.*

Démonstration : Voir [12]...□

On appelle **paire** l'ensemble $\{x, y\}$. Ne pas confondre avec le **couple** (x, y), qui désigne en fait l'ensemble $\{\{x\}, \{x, y\}\}$. On note de même (x, y, z) l'ensemble $(x, (y, z))$, et ainsi de suite pour les n**-uplets ordonnés**. La différence entre $\{x_0, ..., x_n\}$ et $(x_0, ..., x_n)$

est que dans le premier cas l'ordre des termes n'influe pas, alors que dans le second elle influe. On démontre l'associativité et la commutativité de l'union. On notera $\mathcal{P}(E)$ l'ensemble des parties de l'ensemble E.

 On notera que toutes les opérations intuitives sur les ensembles sont possibles, enfin presque. On peut en tout cas utiliser les intersections, définir l'ensemble des éléments d'un ensemble donné qui vérifient une propriété donnée, on peut travailler sur l'ensemble des parties d'un ensemble, on peut travailler sur un produit cartésien d'ensembles, bref toutes ces choses sans lesquelles les maths prendraient vraiment la tête. On peut aussi montrer l'existence et l'unicité de l'ensemble vide.

4.2 La "taille" des ensembles : ordinaux, cardinaux

4.2.1 Les ordinaux

Définition 21 (Définitions de base pour les ordinaux) *On dit qu'un ensemble muni d'une relation d'ordre est* **bien ordonné** *si et seulement si toute partie non vide de cet ensemble admet un élément minimum. L'ordre est alors appelé un* **bon ordre**. *On appelle* **segment initial** *d'une partie bien ordonnée un ensemble de cette partie tel que étant donné un élément de cette partie, tous les éléments qui sont inférieurs à cet élément sont aussi dans la partie. On appelle* **segment initial engendré par** x *l'ensemble des* y *plus petits que* x *; cette partie est clairement un segment initial.*
Un ensemble est dit **transitif** *si tout élément de cet ensemble est inclus dans cet ensemble. C'est à dire que si* $S \in E$*, alors* $S \subset E$ *(non, il n'y a pas de faute de frappe !). Un ensemble est un* **ordinal** *s'il est transitif s'il est bien ordonné par* $\in$*, cette relation étant une relation d'ordre strict. On note* On *l'ensemble des ordinaux.*

Par exemple les ensembles suivants sont des ordinaux :

$$\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}.$$

Proposition 22 • *Les segments initiaux d'un ordinal sont soit lui-même, soit ses éléments.*
• *Tout élément d'un ordinal est un ordinal.*
• *Un ordinal n'appartient pas à lui-même.*

Démonstration :
• Soit S un segment initial d'un ordinal O. Alors S est un segment initial engendré par un certain a (a est l'élément minimum de $O \setminus S$) ; l'ensemble des éléments qui sont plus petits que a étant les éléments qui appartiennent à a (puisque c'est ainsi que l'on a défini la relation d'ordre), a est donc le segment initial engendré par a.

• Facile...
• Il suffit de voir que l'on a imposé que $\in$ soit un ordre strict.□

Proposition 23 *Etant donné deux ordinaux O et P, une et une seule des trois assertions suivantes est vérifiée :* • $O \in P$
• $P \in O$
• $P = O$.

Démonstration : Il suffit de considérer l'intersection de O et P et d'examiner ses propriétés.□

La relation $\in$ est donc une relation d'ordre total sur la classe des ordinaux.

Proposition 24 • *La relation $\in$ est un bon ordre sur la classe des ordinaux.*
• *Le plus petit élément de la classe des ordinaux plus grands que E est $E \cup \{E\}$.*
• *L'union d'une classe d'ordinaux est un ordinal ; il est plus grand que tous les ordinaux de cette classe, et il est plus petit que tous les autres ordinaux plus grands que tous ces ordinaux.*

Démonstration :
• Il suffit de constater comme on l'a vu plus haut que le segment initial engendré par O est O.
• Il est clair que E doit appartenir à un tel élément, ainsi qu'être inclus dedans ; réciproquement l'ensemble $E \cup \{E\}$.
• Facile.□

Définition 25 *Etant donné E un ordinal, $E \cup \{E\}$ est appelé le* **successeur** *de E. On le note $E + 1$. E est dit le* **prédécesseur** *de $E + 1$.*

Propriété amusante :
La classe des ordinaux n'est pas un ensemble. En effet si un tel ensemble E existe, alors tout élément de E est un ordinal, et donc un ensemble d'ordinaux, et donc $E \in E$; ce qui n'est pas possible pour un ordinal puisque $\in$ est une relation d'ordre strict.

Définition 26 *On appelle* **morphisme d'ordre** *entre deux ensembles ou classes ordonnés A et B une application de A vers B telle que $f(a) \geq f(b) \iff a \geq b$. Un morphisme d'ordre bijectif est appelé* **isomorphisme d'ordre***. S'il existe un isomorphisme d'ordre entre deux ensembles ou classes alors on dit que ces ensembles ou classes sont isomorphes pour l'ordre.*

Théorème 27 *S'il existe un isomorphisme d'ordre entre deux ordinaux E et F, alors E et F sont égaux (alors l'isomorphisme d'ordre est l'identité).*

Théorème 28 *Pour tout ensemble ordonné E il existe un et un seul isomorphisme d'ordre de E vers un ordinal.*

Théorème 29 *Toute relation de bon ordre dont le domaine n'est pas un ensemble est nécessairement isomorphe pour l'ordre à la classe des ordinaux.*

Il est ensuite possible de montrer que s'il est vrai pour une certaine propriété P à une seule variable libre x que $P(x)$ est vrai pour tout ordinal E plus petit que F, alors P est vraie pour F, alors on peut conclure que la propriété est vraie pour tout ordinal.

Il reste de nombreuses choses à justifier pour expliquer toutes les petites choses que l'on s'autorise en maths sans se prendre la tête trop ; mais ces considérations dépassent mon propos...

4.2.2 Les cardinaux

⊡ Le théorème de Cantor-Bernstein

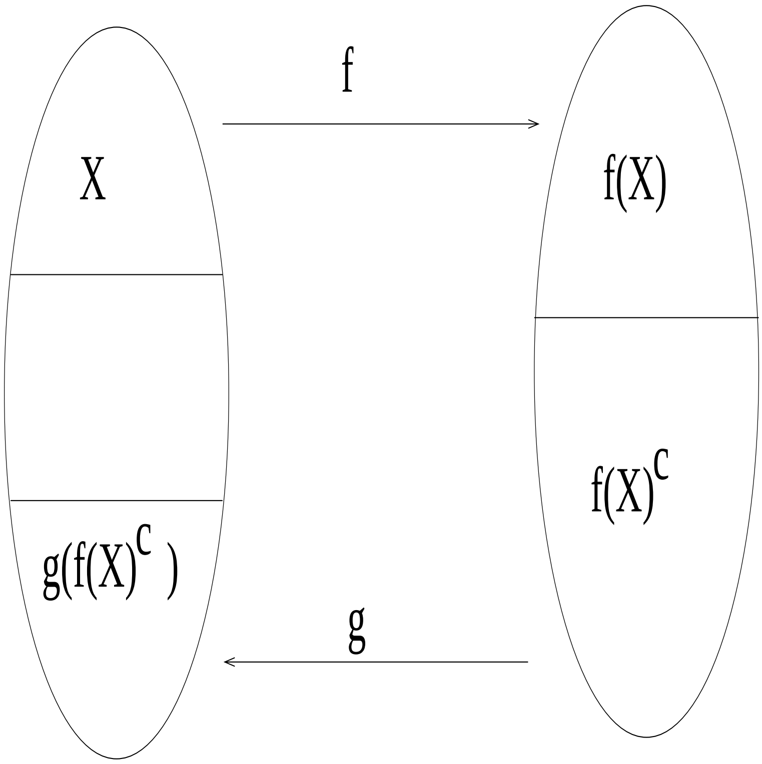


FIG. 4.1 – Démonstration du théorème de Cantor-Bernstein.

Théorème 30 (Théorème de Cantor-Bernstein) *Soit E et F deux ensembles, f une injection de E dans F, et g une injection de F dans E ; alors il existe une bijection de E dans F.*

Démonstration : • On considère l'ensemble des parties X de E telles que $g(f(X)^c) \cap X = \emptyset$.
• On montre que cet ensemble admet un élément maximal (car il est stable par réunion)
• On montre que le maximum X vérifie $g(f(X)^c) \cup X = E$
• On montre que la fonction qui à x associe $f(x)$ si $x \in X$ et l'unique y tel que $g(y) = x$ si $x \notin X$ est une bijection □

⊡ L'axiome du choix et ses dérivés

◇ Généralités - rappels

Définition 31 (Définitions sur les ordres) *Un* **ordre** *est une relation réflexive, antisymétrique, transitive.*
Une **relation d'ordre strict** *est une relation $<$ telle que $\leq$ définie par $x \leq y \iff (x = y \vee x < y)$ soit une relation d'ordre, et telle que pour tout x $\neg x < x$.*
Un élément x d'une partie E est un **minimum** *de cette partie E si et seulement si $x \in E$ et si $\forall e \in E\ e \geq x$.*
Un élément x d'une partie E est un élément **minimal** *de E si et seulement si $x \in E$ et si $e \in E$ et $e \leq x \rightarrow e = x$.*
Un élément x est dit **minorant** *d'une partie E si $\forall e \in E\ e \geq x$; il n'est pas nécessaire que x soit dans E.*
On définit de même **maximum, élément minimal, majorant**.
La figure 4.2 illustre ces notions.
Un **bon ordre** *est un ordre tel que toute partie non vide a un minimum.*

◇ L'axiome du choix

Axiome 32 (Axiome du choix, première version) *Etant donné un ensemble E, il existe une fonction f qui à une partie non vide de E associe un élément de cette partie.*

Axiome 33 (Axiome du choix, deuxième version) *Un produit d'ensembles non vides est non vide.*

On montre facilement que ces deux axiomes sont équivalents. Pour des applications

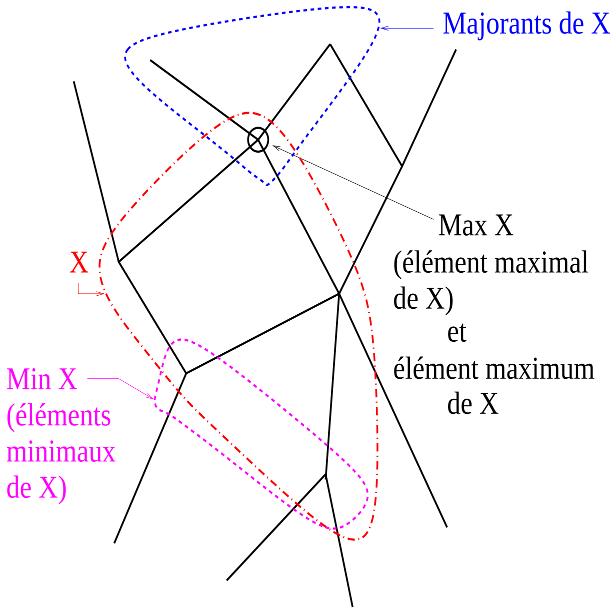


FIG. 4.2 – Illustration de quelques notions sur les ordres. X n'a pas de minorant, ni de plus petit élément ; par contre il a un maximum, c'est le même élément que l'élément maximal unique. Bien noter toutefois qu'un élément maximal, même lorsqu'il est unique, n'est pas nécessairement un maximum.

amusantes de l'axiome du choix on pourra consulter 6.3.

Théorème 34 (Théorème de Zermelo) *E non vide $\rightarrow$il existe une relation de bon ordre (i.e. telle que toute partie non vide admette un minimum).*

Il est difficile de montrer ce théorème à partir de l'axiome du choix. La réciproque est par contre facile.

Définition 35 (Ensemble inductif) *Deux éléments sont dits* **comparables** *si l'un des deux est inférieur ou égal à l'autre.*
On appelle **chaine** *un ensemble totalement ordonné, c'est à dire tel que deux éléments soient toujours comparables.*
Un ensemble ordonné est dit **inductif** *si toute chaine admet un majorant.*

Lemme 36 (Lemme de Zorn) *Tout ensemble non vide ordonné inductif admet un élément maximal.*

Le lemme de Zorn est équivalent au théorème de Zermelo, lui même équivalent aux deux versions de l'axiome du choix. On peut montrer Zermelo à partir de Zorn en considérant l'ensemble des bons ordres sur des parties de E, un couple $(X, \mathcal{R})$ étant inférieur à un couple $(X', \mathcal{R}')$, avec X et X' des parties de E et $\mathcal{R}$ et $\mathcal{R}'$ des bons ordres sur respectivement X et X', si $X \subset X'$, $\mathcal{R} \subset \mathcal{R}'$, et si $x \in X$ et $x' \in X'$ avec $x'\mathcal{R}'x$, alors $x' \in X$.

L'axiome du choix permet par exemple de démontrer l'existence d'une base pour tout espace vectoriel. L'axiome du choix est équivalent à l'existence d'une injection de A dans B ou de B dans A pour tous ensembles A et B ; la preuve de ce fait à partir de Zorn se fait facilement, en considérant les bijections entre des parties de A et des parties de B, par contre la réciproque est difficile.

L'axiome de fondation est l'assertion selon laquelle dans tout ensemble non vide il existe un élément d'intersection vide avec cet ensemble ; l'axiome de fondation sera plus détaillé en 4.5.

Théorème 37 (Consistance relative de AC et de $\neg AC$) *(AC désigne l'axiome du choix)*
Si la théorie axiomatique de Zermelo-Fraenkel avec axiome de fondation est consistante, alors la théorie de Zermelo-Fraenkel avec axiome de fondation et axiome du choix est consistante.
Si la théorie axiomatique de Zermelo-Fraenkel est consistante, alors la théorie de Zermelo-Fraenkel avec axiome du choix est consistante.
D'autre part si la théorie de Zermelo-Fraenkel est consistante, alors la théorie de Zermelo-Fraenkel avec la négation de l'axiome du choix est consistante.
Enfin, si la théorie de Zermelo-Fraenkel avec axiome de fondation est consistante, alors la théorie de Zermelo-Fraenkel avec axiome de fondation et avec la négation de l'axiome du choix (i.e. en supposant qu'il existe un ensemble sur lequel on ne peut pas construire une relation de bon ordre) est consistante.

Démonstration : Fortement non trivial. Je passe.□

Il est aussi possible de remplacer la négation de l'axiome du choix par le fait que $\mathcal{P}(\omega)$ ne puisse pas être bien ordonné ; une telle théorie est consistante si la théorie avec axiome de fondation est consistante.

◇ Quelques exercices

• On peut énoncer sans l'axiome du choix :
- un produit de groupes est non vide
- un produit dénombrable d'espaces métriques compacts est compact

⊡ Définition des cardinaux. Ordinaux finis et infinis

Définition 38 *Deux ensembles sont dits* **équipotents** *s'il existe une bijection de l'un dans l'autre.*

Il est évident qu'il s'agit là d'une relation d'équivalence.

L'axiome du choix permet de démontrer le théorème suivant :

Théorème 39 *Tout ensemble est équipotent à un ordinal.*

Définition 40 *Etant donné un ensemble, on sait qu'il existe au moins un ordinal auquel cet ensemble est équipotent. Eventuellement il peut y en avoir plusieurs ; le plus petit élément de ces ordinaux (au sens défini plus haut sur les ordinaux, c'est à dire la relation* $\in$*) est appelé le* **cardinal** *de l'ensemble. On note usuellement* $\overline{\overline{E}}$ *le cardinal de* E*, ou* $\#E$*. On note* $Card$ *la classe des cardinaux.*

Théorème 41 (Théorème de Cantor) *Pour tout ensemble* E*, on a* $\#E \leq \#\mathcal{P}(E)$*.*

Démonstration : Supposons le contraire ; alors il existe une surjection f de E dans $\mathcal{P}(E)$. Posons F l'ensemble des $x \in E$ tels que $x \notin f(x)$; il suffit alors de considérer le $x \in E$ tel que $f(x) = F$. On constate que si $x \in f(x)$, alors $x \notin f(x)$; et vice-versa.□

On notera que $Card$ n'est pas un ensemble ; sinon on pourrait construire un ensemble égal à On, ce qui est impossible.

Définition 42 (Somme de cardinaux) *On note* $A + B$ *le cardinal de l'union disjointe de deux ensembles respectivement équipotents à* A *et* B*.*

On notera que cette définition pose quelques petits problèmes de définition, pas difficiles à résoudre. L'addition de cardinaux est commutative et associative.
Une propriété importante est le fait que la somme des E_i pour $i \in I$ est le plus grand élément entre $\overline{\overline{I}}$ et les E_i, sous réserve que l'un au moins de ses ensembles (I ou l'un des E_i) soit infini.

Définition 43 (Produit de cardinaux) *On note* $A \times B$ *le cardinal du produit cartésien de* A *et de* B*.*

Là aussi il convient de vérifier que le produit de deux couples d'ensembles de mêmes cardinaux respectifs est le même. On peut en outre vérifier que la multiplication de cardinaux est associative et commutative, et distributive par rapport à l'addition. On notera que le produit de deux cardinaux est le plus grand de ces deux cardinaux.

Définition 44 (Exponentiation de cardinaux) *Etant donnés des cardinaux* A *et* B *on note* A^B *le cardinal de l'ensemble des applications de* B *dans* A*.*

On vérifiera facilement que la définition a bien un sens. On peut aussi montrer que $A^{B+C} = A^B \times A^C$ et que $A^{B^C} = A^{B \times C}$.

De nombreuses manipulations plus approfondies sur les cardinaux requièrent l'axiome du choix.

Définition 45 *Un ordinal est dit* **fini** *si tout ordinal inclus dans cet ordinal admet un prédécesseur. On appelle aussi* **entier naturel** *un ordinal fini.*

Moi ça m'amuse beaucoup de définir un entier naturel comme étant un ensemble E incluant chacun de ses éléments, tel que pour toute partie F de E il existe $x \in E \cap F$ tel que $x \in G$ pour tout $G \in F \setminus \{x\}$ et pour tout élément F de E il existe G incluant chacun de ses éléments, tel que pour toute partie H de G il existe $y \in G \cap H$ tel que $x \in I$ pour tout $I \in H \setminus \{y\}$, et tel que $G \cup \{G\} = F$. Si je me suis pas gouré.

On montre plein de choses bien agréables sur les ordinaux finis ; ils sont stable par union, produit, exponentiation. On montre aussi que tout ordinal fini est un cardinal.

Définition 46 *Un cardinal est dit* **fini** *s'il est fini en tant qu'ordinal. Dans le cas contraire il est dit* **infini***. On note $Card'$ la classe des cardinaux infinis.*

Nous supposons maintenant l'axiome de la théorie des ensembles selon lequel il existe un ordinal infini. Un **ordinal infini** est, par définition, un ordinal qui n'est pas fini. Cet axiome de la théorie des ensembles est équivalent à l'axiome selon lequel la classe des ordinaux finis est un ensemble ; ainsi, puisque la classe des ordinaux n'est pas un axiome, il existe un ordinal infini. On peut encore formuler cet axiome en disant qu'il existe un ordinal limite, au vu de la définition ci-dessous :

Définition 47 *Un ordinal différent du vide et sans prédécesseur est appelé un* **ordinal limite***. C'est donc un ordinal non vide tel que tout élément plus petit que lui a un successeur aussi plus petit que lui.*

Propriété : Un ordinal limite est l'union des ordinaux qui lui sont inférieurs.

Définition 48 *On appelle ω le minimum des ordinaux infinis. ω est donc un ordinal limite, c'est le plus petit, et c'est l'ensemble des ordinaux finis.*
Un ensemble est dit **fini** *si son cardinal est fini.*
Un ensemble est dit **dénombrable** *si son cardinal est inférieur ou égal à ω.*
ω est un cardinal ; on note $\aleph_0 = \omega$ et pour tout ordinal E n'étant pas un ordinal limite, alors avec F le prédécesseur de E, $\aleph_E$ est le plus petit ordinal plus grand que $\aleph_F$; et si E est un ordinal limite, alors $\aleph_E$ est l'union des $\aleph_F$ pour $F \in E$.

Propriété :
Un ensemble infini est un ensemble contenant une partie dénombrable infinie.
Un ensemble infini est un ensemble qui est en bijection avec l'une de ses parties

propres.
$\aleph_E$ est un cardinal.

4.3 L'axiome d'accessibilité

Définition 49 (Cardinal accessible et inaccessible) *Un cardinal E est dit* **inaccessible** *s'il est plus grand que ω, si pour tout F cardinal $< E$ on a $2^F < E$, et si toute famille de cardinaux $< E$, indexée par une famille de cardinal $< E$, a un sup plus petit que E.*
Un cardinal est dit **accessible** *s'il n'est pas inaccessible.*
*L'***axiome d'accessibilité** *affirme que tout cardinal est accessible.*

Théorème 50 *Si Zermelo-Fraenkel avec axiome du choix est consistant, alors Zermelo-Fraenkel avec axiome du choix et axiome d'accessibilité est consistant.*

Démonstration : Difficile...□

4.4 L'hypothèse du continu

Le théorème de Cantor nous dit que $\aleph_{E+1} \leq 2^{\aleph_E}$ (il est clair que $2^{\aleph_E}$ est le cardinal de l'ensemble des parties de E).

Définition 51 (Hypothèse du continu - hypothèse du continu généralisée)
On appelle hypothèse du continu l'assertion $\aleph_1 = 2^{\aleph}_0$.
On appelle hypothèse du continu généralisée l'assertion $\aleph_{E+1} = 2^{\aleph}_E$ pour tout E ordinal.

Propriété :
L'hypothèse du continu est équivalente à l'assertion selon laquelle les parties de ω peuvent être bien ordonnées de manière à ce que tout segment initial strict soit dénombrable.

Théorème 52 *Si la théorie de Zermelo-Fraenkel est consistante, alors la théorie de Zermelo-Fraenkel plus hypothèse du continu généralisée est consistante.*

Démonstration : Trop dure !□

4.5 L'axiome de fondation

Définition 53 (Axiome de fondation) *On appelle* **axiome de fondation** *l'axiome selon lequel pour tout ensemble E non vide il existe F tel que $F \in E$ et $F \cap E = \emptyset$.*

Cet axiome entraîne, par exemple, qu'il n'existe pas d'ensemble x tel que $x = \{x\}$, ni d'ensemble x tel que $x \in x$.

Théorème 54 *Si la théorie de Zermelo-Fraenkel est consistante, alors la théorie de Zaermelo-Fraenkel plus axiome de fondation est consistante.*

Démonstration : Dure !$\square$

Théorème 55 *Il n'existe pas de suite U_n d'ensembles telle que $U_{n+1} \in U_n$ pour tout n.*

Démonstration : La preuve, facile, nécessite l'axiome de fondation.$\square$

Théorème 56 *Si l'on utilise l'axiome de fondation, alors un ensemble E est un ordinal si et seulement si il est transitif et si deux éléments u et v de E vérifient au moins une des assertions suivantes :*
- $u \in v$
- $u = v$
- $v \in u$

Démonstration : La preuve est plus difficile, et je ne la donne pas ici car elle dépasse mon propos de simple brève introduction à la théorie des ensembles.$\square$

Bien sûr on peut montrer que si ces hypothèses sont vérifiées alors pour tout couple (u, v) c'est l'une *et une seule* des assertions qui est vérifiée.

Théorème 57 *Si l'on utilise l'axiome de fondation, alors pour tout ensemble E il existe un unique ensemble transitif contenant E et inclus dans tout ensemble transitif incluant E.*

Démonstration : Non triviale.$\square$

Définition 58 *On appelle* **fermeture transitive** *de E l'ensemble transitif dont l'existence est garantie par le théorème 57.*

Propriété :
La fermeture transitive de E est la réunion de E et de l'union des fermetures transitives des éléments de E.

Définition 59 *Une relation $\mathcal{R}$ est dite* **extensive** *si* $\forall(y,z)[\forall x(x\mathcal{R}y \iff x\mathcal{R}z) \rightarrow y = z]$.
Un ensemble est dit **extensif** *si $\in$ est une relation extensive sur cet ensemble. C'est à dire que si deux éléments ont même intersection avec l'ensemble, alors ils sont égaux.*

Propriétés :
Un ensemble transitive est extensif.
Un ensemble extensif est isomorphe à un ensemble transitif, et l'isomorphisme est unique (nécessitant l'axiome de fondation).

Il est possible de prouver que l'axiome de fondation est relativement consistant, c'est à dire que la théorie basée sur les axiomes de Zermelo-Fraenkel est consistante si et seulement si la théorie basée sur les mêmes axiomes plus l'axiome de fondation est consistante.

4.6 Résumé de théorie des ensembles

En résumé on a les implications de consistance du schéma 4.3.

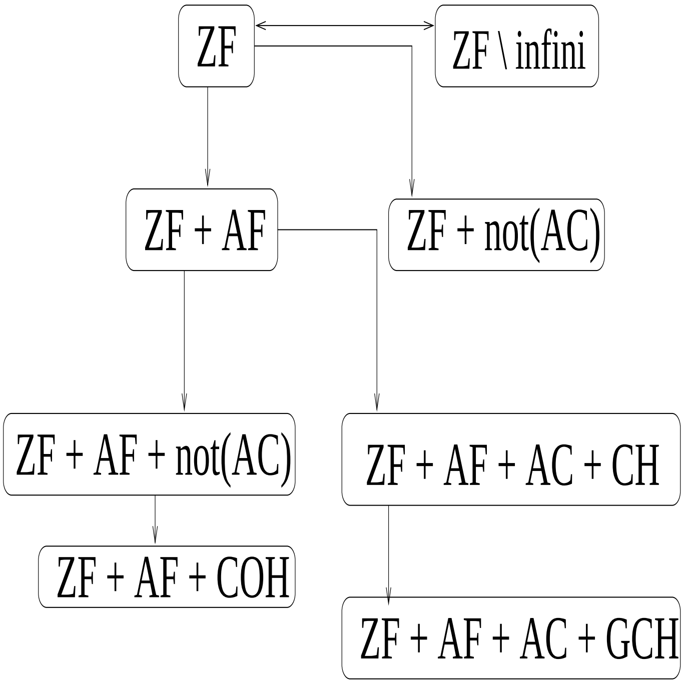


FIG. 4.3 – ZF désigne la théorie de Zermelo-Fraenkel. $ZF \setminus infini$ désigne la même théorie mais privée de l'axiome de l'infini et muni de sa négation. AC désigne l'axiome du choix. $not(AC)$ désigne la négation de AC. COH désigne l'axiome selon lequel les parties de ω ne peuvent pas être bien ordonnes. AF désigne l'axiome de fondation. ACC désigne l'axiome d'accessibilité. CH désigne l'hypothèse du continu, et GCH l'hypothèse du continu généralisée. Une flèche relie une théorie T à une théorie T' si T' est plus forte que T, c'est-à-dire que tous les théorèmes de T sont des théorèmes de T'. Notez bien que toutes les théories présentes sur la figure sont consistantes si et seulement si ZF est consistante. Notez bien aussi que si ZF est consistante, alors il est impossible de le prouver ; mais que par contre si elle ne l'est pas, on dispose d'un algorithme théorique permettant en temps fini de le prouver...

Chapitre 5

Topologie

5.1 Espaces topologiques

5.1.1 Cas le plus général d'espace topologique

Définition 60 (Topologie) *Une* **topologie** $\mathcal{T}$ *sur l'ensemble* X *est une partie* $\mathcal{T} \subset P(X)$ *vérifiant :*
• *L'ensemble vide* $\emptyset$ *et* X *sont dans* $\mathcal{T}$
• $\mathcal{T}$ *est stable par réunions arbitraires*
• $\mathcal{T}$ *est stable par intersections finies*

Un tel couple $(X, \mathcal{T})$ *est appelé espace topologique. Les éléments de* $\mathcal{T}$ *sont appelés les* **ouverts** *de la topologie.*
Une partie de X *est dite* **fermée** *si son complémentaire est ouvert.*

Exemples :
• La topologie discrète sur l'ensemble X est la topologie $\mathcal{T}_d = P(X)$
• La topologie grossière sur l'ensemble X est la topologie $\mathcal{T}_g = \{\emptyset, X\}$
• Sur $\overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$, la topologie usuelle est l'ensemble des U tels que $U \subset \mathbb{N}$ ou $+\infty \in U \wedge \mathbb{N} \setminus U$ est cofini.

On verra aussi d'autres exemples en parties 5.1.2 et 5.2.

Proposition 61 *Si* X *est un espace topologique alors*
• X *et* $\emptyset$ *sont des fermés de* X
• *Une intersection quelconque de fermés est un fermé*
• *Une union finie de fermés est un fermé*

Démonstration : Immédiat, par passage au complémentaire.□

Définition 62 (Séparation par des ouverts) *On dit que la partie A et la partie B sont* **séparées par des ouverts** *s'il existe deux ouverts U et V tels que $A \subset U$ et $B \subset V$ tels que $U \cap V = \emptyset$.*

5.1.2 Espaces métriques et espaces normés

Définition 63 (Métrique) *Une* **métrique** *ou* **distance** *sur l'ensemble X est une application $d : X \times X \to [0, +\infty[$ vérifiant :*
• $d(x, y) = 0 \iff x = y$
• $d(x, y) = d(y, x)$
• $d(x, y) \leq d(x, z) + d(z, y)$ *(propriété dite* **inégalité triangulaire***)*
On dit alors que (X, d) est un **espace métrique***.*

Exemples :
• $d_p = (\sum |x_i - y_i|^p)^{1/p}$
• $d_\infty = max|x_i - y_i|$

Propriété :
• $|d(x, z) - d(y, z)| \leq d(x, y)$

Définition 64 (Boules) *Si x est un point de l'espace métrique X et $r \in [0, +\infty[$, on appelle* **boule ouverte (resp. fermée)** *de centre x et de rayon r, l'ensemble des y tels que $d(x, y) < r$ (resp. $d(x, y) \leq r$).*
On appelle **sphère** *de l'espace métrique X de centre x et de rayon r l'ensemble des y tels que $d(x, y) = r$.*

Proposition 65 *Si X est un espace métrique, la famille de parties de X dont les éléments sont les réunions arbitraires de boules ouvertes est une topologie sur X. Cette topologie est appelée la topologie* **associée** *à la métrique.*
Une partie X d'un espace métrique E est dite **bornée** *si étant donné un point e dans E la distance de x à e pour x dans X est majorée par une certaine constante [a]. Cela équivaut aussi au fait que la distance entre deux points quelconques de X est bornée. C'est à dire que :*
- si pour un point x, $y \mapsto d(x, y)$ est bornée, alors pour tout point x, $y \mapsto d(x, y)$ est bornée.
- si pour tout point x $y \mapsto d(x, y)$ est bornée, alors $(x, y) \mapsto d(x, y)$ est aussi bornée.

[a] La notion est indépendante du point e choisi, grâce à l'inégalité triangulaire.

Démonstration : La vérification est fastidieuse et ne présente pas de difficulté.□

La notion de borné dépend de la métrique et pas de la topologie ! C'est à dire que même si deux métriques sont topologiquement équivalentes (voir définition 68) elles n'ont pas nécessairement les mêmes parties bornées. En fait pour toute métrique d, on peut construire une métrique équivalente d' par $d'(x,y) = ln(1 + \frac{d(x,y)}{1+d(x,y)})$, telle que toute partie soit bornée.

Propriétés :

• Dans un espace métrique, une partie est fermée si et seulement si elle contient la limite de toute suite convergente à valeurs dans cette partie.
• Une boule ouverte est ouverte, et donc un espace métrique est séparé
• Une boule fermée est fermée
• Une sphère est fermée
• Dans un espace métrique, une suite $(x_n)_{n\in\mathbb{N}}$ tend vers x si et seulement si $d(x_n, x)$ tend vers 0.

Exercice 66 • *La topologie usuelle sur $\mathbb{R}$ ou $\mathbb{C}$ est la topologie associée à la distance $d(x,y) = |x - y|$.*
• *La fonction qui à x et y associe 0 si $x = y$ et 1 sinon est une métrique. Cette métrique est associée à la topologie discrète, pour laquelle toute partie est à la fois un ouvert et un fermé.*
• *Si f est injective de X dans $\mathbb{R}$, alors la fonction qui à x et y associe $|f(x) - f(y)|$ est une distance sur X.*
• *La topologie usuelle sur $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ est définie par la distance $d(x,y) = |f(x) - f(y)|$, avec $f(x) = \frac{x}{1+|x|}$, $f(+\infty) = 1$ et $f(-\infty) = -1$.*

Définition 67 (Isométrie) *Etant donnés deux espaces métriques E et F, une application f de E dans F est une* **isométrie** *si $\forall (x,y)\ d_F(f(x), f(y)) = d_E(x,y)$.*

Définition 68 (Métrisable) *Une topologie est dite* **métrisable** *si et seulement si il existe une métrique telle que la topologie soit associée à cette métrique.*
Deux métriques d_1 et d_2 sont dites **équivalentes** *si il existe α et β tels que $\alpha d_1 < d_2 < \beta d_1$ [a], avec $\alpha, \beta > 0$.*
Deux métriques sont dites **topologiquement équivalentes** *si elles définissent la même topologie.*

[a] On dit aussi que d_1 et d_2 sont **Lipschitz-équivalentes**.

Soient deux distances d_1 et d_2 sur un espace E ; alors l'identité de (E, d_1) dans (E, d_2) est un homéomorphisme si et seulement si d_1 et d_2 sont topologiquement équivalentes, et elle est lipschitzienne et d'inverse lipschitzien [1] si et seulement si d_1 et

[1] Une application est dite **bilipschitzienne** si elle est lipschitzienne et d'inverse lipschitzien.

d_2 sont équivalentes.

Proposition 69 (Existence de topologie non métrisables) *Il existe des topologies, même séparées, non métrisables.*

Démonstration : Il est clair que toute topologie non séparée n'est pas métrisable.

Considérons, pour avoir un contre-exemple plus intéressant, une topologie séparée non métrisable. Ce contre-exemple fait appel à quelques notions qui seront définies ultérieurement, et peut donc être laissé de côté en première lecture.

Soit $\mathbb{R}^{\mathbb{R}}$, muni de la topologie produit.

Supposons que cet espace topologique soit métrisable.

Alors tout point est à base dénombrable de voisinage.

Soit (U_n) une base de voisinages de 0.

Alors pour tout n, U_n contient un voisinage de 0 (la fonction nulle de $\mathbb{R}$ dans $\mathbb{R}$) de la forme

$$V_n = \{f \in \mathbb{R}^{\mathbb{R}} / \forall i \in [1, N_n] |f(x_{n,i})| < \epsilon_n\}$$

On considère alors T l'ensemble des $x_{n,i}$ pour $i \leq N_n$ et $n \in \mathbb{N}$.

Cet ensemble est dénombrable comme union dénombrable d'ensemble finis.

Soit maintenant x dans $\mathbb{R}$ n'appartenant pas à T.

Alors $\{f \in \mathbb{R}^{\mathbb{R}} / |f(x)| < \epsilon\}$ est un ouvert, qui n'est manifestement inclus dans aucun V_n.

Il est à noter que $\{0,1\}^{\mathbb{R}}$ convient aussi.$\square$

Proposition 70 *Une topologie métrisable est entièrement caractérisée par les propriétés de convergence de suites.*
C'est à dire que si pour deux topologies métrisables, les suites convergentes sont les mêmes et ont mêmes limites, alors ces deux topologies sont égales.

Démonstration : Il suffit de voir que l'on caractérise un fermé F d'un métrique par le fait qu'il contient les limites de toute suite convergente d'éléments de F. Donc les fermés sont caractérisés par les propriétés de convergence de suite, et donc les ouverts aussi par passage au complémentaire.$\square$

Proposition 71 • *Si deux distance d_1 et d_2 sont équivalentes alors d_1 et d_2 définissent la même topologie.*
• *on peut avoir la même topologie sans avoir cette relation.*

Démonstration : Le premier • est facile, le second s'obtient en considérant $d'(x,y) = min(1, d(x,y))$, avec d une distance quelconque non bornée.$\square$
Il est intéressant de noter que même en ajoutant une condition à l'équivalence traduisant que l'on peut se limiter aux "petites" distances, on a un contre-exemple avec par exemple $d_{1/2}$ et d_1 qui définissent la même topologie sans être Lipschitz-équivalentes, même sur les petites distances.
On peut aussi noter que les d_p pour $p \geq 1$ sont Lipchitz-équivalentes entre elles, cela se montre par $d_\infty \leq d_p \leq n^{1/p} d_\infty$

Dans la suite $\mathbb{K}$ désigne un des deux corps $\mathbb{R}$ ou $\mathbb{C}$ muni de sa topologie usuelle.

Définition 72 (Norme) *Soit E un espace vectoriel sur le corps $\mathbb{K}$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Une* **norme** *sur E est une application $\| \, . \, \|$ de E dans $[0, +\infty[$ vérifiant :*
• $\| x \| = 0$ si et seulement si $x = 0$
• $\forall x, y \in E$, on a $\| x + y \| \leq \| x \| + \| y \|$
• $\forall \lambda \in \mathbb{K}\ \forall x \in E$ on a $\| \lambda . x \| = |\lambda| \, \| x \|$
S'il ne manque que la première propriété, on parle de **semi-norme**.
On appelle vecteur unitaire un vecteur x tel que $\| x \| = 1$.
Un espace muni d'une norme est appelé **espace normé** *ou* **espace vectoriel normé**.
Dans un espace normé une série $(\sum x_n)$ est dite **normalement convergente** *si $\sum_{i=1}^{n} \|x_i\|$ converge.*
Enfin une définition nécessitant la notion de continuité (définie ultérieurement) : on appelle **isomorphisme de l'espace vectoriel normé E dans l'espace vectoriel normé F** *une application linéaire continue bijective de réciproque continue (c'est à dire qu'il s'agit d'un morphisme algébrique (i.e. au sens des espaces vectoriels) et d'un homéomorphisme).*

Exemples :
• Sur $\mathbb{R}^n$, les applications suivantes sont des normes :
- $(x_1, ..., x_n) \mapsto \|x\|_\infty = max_{i \in \{1,\ldots,n\}} |x_i|$
- pour p réel ≥ 1, $x \mapsto \|x\|_p = (\sum_{i \in \{1,\ldots,n\}} |x_i|^p)^{1/p}$ • Un peu plus difficile : sur $\mathbb{R}[X]$, les applications suivantes sont des normes :
- $P \mapsto \|P\|_0 = sup_{x \in [0,1]} |P(x)|$
- $P \mapsto \|P\|_1 = \int_0^1 |P(t)dt$

Propriétés :
• La norme est convexe.

Définition 73 (Distance associée) *Etant donnée une norme on définit une* **distance associée** *par $d(x, y) = \| x - y \|$*

Définition 74 (Normes équivalentes) *Deux normes $\| \, . \, \|_1$ et $\| \, . \, \|_2$ sur un même espace vectoriel sont* **équivalentes** *si il existe $\alpha, \beta > 0$ tels que $\alpha. \| x \|_1 < \| x \|_2 < \beta. \| x \|_1$*

Théorème 75 *Deux normes sont équivalentes si et seulement si elles définissent la même topologie.*

Démonstration : L'une des deux implications résulte de 71. L'autre s'obtient facilement, l'une des deux inégalités après l'autre, en constatant qu'une boule de centre

0 et de rayon 1 pour l'une des normes contient une boule pour l'autre norme.□

5.1.3 Notion de voisinage

Définition 76 (Voisinage) *Soit X un espace topologique. Un* **voisinage** *V de $x \in X$ est un ensemble tel qu'il existe un ouvert U avec $x \in U \subset V$.*
On note par $\mathcal{V}(x)$ l'ensemble des voisinages de x.

Proposition 77 *Un sous-ensemble d'un espace topologique est ouvert si et seulement si il est un voisinage de chacun de ses points.*

Démonstration :
• Soit un ouvert U, et x dans U. On a $x \in U \subset U$... Donc U est voisinage de x.
• Soit U voisinage de chacun de ses points. A chaque point x associons l'ouvert U_x tel que $x \in U_x \in U$. La réunion des U_x est un ouvert, contient tous les x de U et est incluse dans U ; c'est donc U. Donc U est un ouvert. □

Proposition 78 • *Si $x \in X$, X espace topologique, et $V \subset V'$, et $V \in \mathcal{V}(x)$, alors $V' \in \mathcal{V}(x)$.*
• *pour tout $V, V' \in \mathcal{V}(x)$, alors $V \cap V' \in \mathcal{V}(x)$*

Démonstration : • V contient par définition un ouvert contenant x ; V étant inclus dans V', V' contient ce même ouvert. Donc V' est un voisinage de x.
• V et V' contiennent chacun un ouvert contenant x ; l'intersection de ces deux ouverts est un ouvert, contient x et est inclus dans $V \cap V'$; donc $V \cap V'$ est un voisinage de x. □

5.1.4 Fermeture, intérieur, extérieur, frontière

Définition 79 (Fermeture ou adhérence) *Si $A \subset X$,* **l'adhérence** *(dite aussi* **fermeture***) de A est l'intersection de tous les fermés contenant A, c'est donc le plus petit fermé contenant A. On note $\overline{A}$ l'adhérence de A.*

Propriété :
A, B parties de X ; alors $\overline{A \cup B} = \overline{A} \cup \overline{B}$ et $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$.

Définition 80 (Point d'accumulation d'une partie) *On appelle* **point d'accumulation d'une partie** *A un point x adhérent à $A \setminus \{x\}$.*
On appelle **ensemble dérivé de** *A l'ensemble des points d'accumulation de A.*

Propriété :
Un ensemble dérivé dans un espace séparé est toujours un fermé.

Lemme 81 *Si A est une partie de l'espace topologique X, on a l'équivalence suivante :*
$$x \in \overline{A} \iff \forall V \in \mathcal{V}(x), V \cap A \neq \emptyset$$

<u>Démonstration :</u> Il suffit de constater les points suivants :
• $y \notin \overline{A}$ si et seulement si on peut trouver un fermé F contenant A et ne contenant pas y.
• On considère le complémentaire de F.$\square$

Définition 82 (Ensemble dense) *Un sous-ensemble de X est* **dense** *dans X si son adhérence est X.*

La densité sera utilisée dans les théorèmes de prolongement, prolongement des identités, prolongement de fonctions uniformément continues (capital par exemple pour le théorème de Plancherel, cité dans la partie 5.1.8 et démontré dans [16]). Le prolongement de fonctions continues servira aussi à construire des solutions maximales d'équations différentielles (voir théorème de Cauchy-Lipschitz 515). On pourra aussi voir l'exercice ? ? référence selon lequel tout espace métrique complet connexe localement connexe est connexe par arcs.

La densité servira aussi pour prouver le théorème d'Arzéla-Ascoli 720, le théorème de Moore (voir livre ? ?), l'inégalité de Hardy (voir livre [2]).

De nombreux résultats de densité dans les Banach auront de vastes applications ; il y a déjà toutes les applications du théorème de Baire 249 (théorème de l'application ouverte, théorème du graphe fermé, théorème d'isomorphisme de Banach, que l'on trouvera tous à la suite du théorème de Baire 249). On pourra enfin consulter le théorème de Goldstine, dans le livre ? ?.

Par ailleurs, la séparabilité est par définition liée à la densité, voir la définition 96 et la liste d'applications qui y est donnée.

Enfin, certains résultats de densité seront fondamentaux pour de multiples applications pratique (approximation) : on pourra consulter le chapître 9. Cela servira par exemple pour la transformée de Fourier - en fait les bases hilbertiennes sont basées sur la densité.

N'oublions pas aussi de petits résultats dus à la densité de $\mathbb{Q}$ dans $\mathbb{R}$: le fait que tout ouvert de $\mathbb{R}$ s'exprime comme union dénombrable d'intervalles ouverts.

Proposition 83 *A est dense dans X si et seulement si tout ouvert non vide intersecte A.*

Démonstration : Cela résulte directement du lemme ci-dessus.□

Définition 84 (Intérieur) *L'***intérieur** *du sous-ensemble A de l'espace topologique X, noté $Int(A)$, est la réunion de tous les ouverts inclus dans A, c'est donc le plus grand ouvert contenu dans A.*

Propriétés :
• A, B inclus dans X ; alors $Int\ (A \cup B) = Int\ A \cup Int\ B$ et $Int\ (A \cap B) = Int\ A \cap Int\ B$. • Si deux ouverts sont disjoints, alors les intérieurs de leurs adhérences sont disjoints.
• $Int(X \setminus A) = X \setminus \overline{A}$ (ie $Int\ A = X \setminus \bar{(}X \setminus A)$).

Proposition 85 *Le point x est dans $Int(A)$ si et seulement si $A \in \mathcal{V}(x)$*
Le point x est dans $Int(A)$ si et seulement s'il existe $V \in \mathcal{V}(x)$ avec $V \subset A$

Démonstration : Je ne vous ferai pas l'injure de le démontrer.□

Définition 86 (Extérieur) *L'***extérieur** *de A, noté $Ext(A)$, est l'intérieur du complémentaire de A.*

Proposition 87 $Ext(A) = \{x | \exists V \in \mathcal{V}(x) / V \cap A = \emptyset\}$

Démonstration : Evident.□

Définition 88 (Frontière) *La* **frontière** *de A, notée $Fr(A)$ est son adhérence privée de son intérieur.*

Propriété : $Fr(A) = \overline{A} \cap \overline{X \setminus A}$.

Proposition 89 *Un ensemble est à la fois ouvert et fermé si et seulement si sa frontière est vide.*

Démonstration :
• Soit A cet ensemble. Comme A est fermé, il est égal à son adhérence, et comme il est ouvert, il est égal à son intérieur, donc sa frontière, égale à son adhérence privée de son intérieur, est vide.
• Réciproquement, si la frontière de A est vide, et s'il est non vide, cela signifie que son intérieur est au moins égal à A, donc qu'il est ouvert. Et si sa frontière est vide,

cela signifie que son adhérence ne peut pas être plus grande que lui, donc il est fermé.□

Théorème 90 • $Int(A) = \{x \in X | \exists V \in \mathcal{V}(x), V \cap X \setminus A \ = \emptyset\}$
• $Ext(A) = \{x \in X | \exists V \in \mathcal{V}(x), V \cap A = \emptyset\}$
• $Fr(A) = \{x \in X | \ \nexists V \in \mathcal{V}(x), V \cap A = \emptyset \vee V \cap (X \setminus A) = \emptyset\} = \{x \in X / \forall V \in \mathcal{V}(x), V \cap A \neq \emptyset \wedge V \cap (X \setminus A) \neq \emptyset\}$
X est réunion disjointe de son intérieur, son extérieur et sa frontière.

Démonstration : Chacune de ces propriétés se démontre en deux lignes, simplement en écrivant bien formellement ce que l'on cherche à démontrer.□

5.1.5 Base d'ouverts et base de voisinages

Définition 91 (Base d'ouverts) *Soit X un espace topologique. Une famille $\mathcal{B}$ d'ouverts de X est une* **base d'ouverts** *si tout ouvert est une réunion d'éléments de $\mathcal{B}$.*

Proposition 92 *Une famille $\mathcal{B}$ d'ouverts est une base d'ouverts si et seulement si quel que soit l'ouvert U et $x \in U$ il existe $V \in \mathcal{B}$ tel que $x \in V \subset U$.*

Démonstration : Si $\mathcal{B}$ est une base d'ouverts, alors étant donnés x et U, on considère un élément V de $\mathcal{B}$ qui contient x ; la réciproque se fait en considérant pour un ouvert donné la réunion des V obtenus par la propriété en considérant les différents x.□

Proposition 93 • *Dans un espace métrique, les boules ouvertes de rayon rationnel forment une base d'ouverts*
• *Dans le cas de $\mathbb{R}^n$ muni de la métrique usuelle, les boules ouvertes de rayon rationnel et à coordonnées toutes rationnelles forment une base <u>dénombrable</u> d'ouverts*
• *Dans $\mathbb{R}$ tout ouvert est en fait une réunion dénombrable d'intervalles ouverts deux à deux disjoints (et réciproquement).*
• *Dans $\mathbb{R}$ un fermé n'est pas nécessairement une réunion dénombrable d'intervalles fermés deux à deux disjoints, et une réunion dénombrable d'intervalles fermés deux à deux disjoints n'est pas nécessairement fermée.*

Démonstration :

• Soit U un ouvert d'un espace métrique, et x dans U ; on montre que U contient une boule de rayon rationnel contenant x. Pour cela on note que U est réunion de boules ouvertes, donc contient au moins une boule ouverte B de rayon r et de centre O contenant x ; on note alors r' la distance de x à O ; toute boule ouverte centrée en x de rayon rationnel inférieur à $r - r'$ convient (on peut aussi choisir de raisonner sur les boules centrées sur O de rayon adéquat...).

• Soit U un ouvert de $\mathbb{R}^n$, et x un point de U. On considère une boule ouverte contenant x et incluse dans U ; soit O son centre et r son rayon. Alors soit r' la distance de x à O, et y un point de coordonnées rationnelles situé à une distance d inférieure à $\frac{r-r'}{3}$ de O. Alors toute boule centrée sur y de rayon rationnel compris entre $r' + \frac{r-r'}{3}$ et $r' + 2.\frac{r-r'}{3}$ convient.

• En plusieurs points :

- Soit U un ouvert de $\mathbb{R}$; alors étant donné un rationnel de U on considère l'intervalle maximal le contenant. On parcourt ainsi tout U, et on a bien un ensemble dénombrable d'intervalles ouverts.

- Une réunion d'ouverts est toujours un ouvert.

• Deux contre-exemples :

- le cantor K^3 (voir partie 5.6.13) n'est pas une réunion dénombrable d'intervalles fermés disjoints.

- l'ensemble des $1/n$ est une réunion dénombrable d'intervalles fermés disjoints, mais n'est pas fermé. □

Définition 94 (Base dénombrable d'ouverts) *X est à* **base dénombrable d'ouverts** *si on peut trouver une base d'ouverts qui soit dénombrable.*

Proposition 95 *Un espace à base dénombrable d'ouverts contient un ensemble dénombrable dense*

Démonstration : Il suffit de considérer un point par ouvert non vide d'une base dénombrable.□

Définition 96 (Espace séparable) *Un espace est* **séparable** *si il contient un ensemble dénombrable dense.*

Cela sera notamment utile pour définir une métrique sur la boule unité fermée du dual d'un espace séparable (pour la topologie faible). Ceux qui veulent en savoir plus peuvent aller voir la proposition 194.

On note en particulier qu'un ensemble à base dénombrable d'ouverts est séparable (il suffit de prendre un point dans chaque ouvert) ; il s'agit de la proposition précédente. La réciproque est vraie dans le cas des espaces métriques :

Théorème 97 *Un espace métrique est séparable si et seulement s'il admet une base dénombrable d'ouverts.*

Ce résultat permettra de conclure que tout espace métrique compact admet une base dénombrable d'ouverts (voir résultat 196) et d'en déduire que tout espace métrique compact est de cardinal au plus la puissance du continu (voir résultat 194).

Démonstration : La remarque précédente donne l'un des deux sens. Réciproquement supposons que X soit métrique séparable. Soit $\{x_n/n \in \mathbb{N}\}$ un ensemble dense dénombrable. Alors l'ensemble des boules de centre x_i et de rayon $1/j$ avec $(i,j) \in \mathbb{N} \times \mathbb{N}^*$ est une base dénombrable d'ouverts.□

Définition 98 (Base de voisinages) *Soit $x \in X$, une famille $\mathcal{B}(x)$ de voisinages de x est une* **base de voisinages** *de x si pour tout $V \in \mathcal{V}(x)$ il existe $V' \in \mathcal{B}(x)$ avec $V' \subset V$.*

Définition 99 *Un espace est* **à base dénombrable de voisinages** *si chacun de ses points admet une base dénombrable de voisinages.*

Exercice 100 *Tout espace métrique est à base dénombrable de voisinages.*

Démonstration : Il suffit de considérer les boules de rayon $1/i$ de centre x pour avoir une base dénombrable de voisinages de x.$\square$

5.1.6 Continuité et limite

Définition 101 (Continuité ponctuelle) *Soit f une application entre espaces topologiques. f est* **continue en** *x si et seulement si quel que soit $V \in \mathcal{V}(f(x))$, l'image réciproque $f^{-1}(V)$ est un voisinage de x (ie si $\exists U \in \mathcal{V}(X)/f(U) \subset U$).*
f est **continue** *si f est continue en tout point.*

Exemples :
• La distance est continue (en vertu de la propriété $|d(x,z) - d(y,z)| \leq d(x,y)$).
• La norme est continue (comme composée d'applications continues, puisque $x \mapsto (x,x)$ est continue, et $(x,y) \mapsto d(x,y)$ est continue, avec d la distance associée à la norme.
• La multiplication par un scalaire et l'addition sont continues pour la topologie associée à la norme.

Définition 102 (Semi-continuité) *Une application f de X dans $\mathbb{R}$ est* **semi-continue inférieurement** *si pour tout c on a $f^{-1}(]c,+\infty[)$ ouvert.*
Une application f de X dans $\mathbb{R}$ est **semi-continue supérieurement** *si pour tout c on a $f^{-1}(]-\infty,c[)$ ouvert.*

Proposition 103 • *Une fonction à valeurs dans $\mathbb{R}$ est continue si et seulement si elle est à la fois semi-continue inférieurement et semi-continue supérieurement.*
• *La borne sup d'une famille de fonctions semi-continues inférieurement est semi-continue inférieurement.*
• *La fonction caractéristique d'un ouvert (resp. fermé) est semi-continue inférieurement (resp. supérieurement).*

Théorème 104 (Stabilité de la continuité par composition) *Si f est continue en x et si g est continue en $f(x)$, alors $g \circ f$ est aussi continue en x.*

Démonstration : L'image réciproque d'un voisinage de $g(f(x))$ est un voisinage de $f(x)$, l'image réciproque d'un voisinage de $f(x)$ par f est un voisinage de x, donc l'image réciproque d'un voisinage de $g \circ f(x)$ par $g \circ f$ est un voisinage de x. D'où la continuité de $g \circ f$ en x.□

Corollaire 105 *Si f et g sont continues, alors $g \circ f$ est continue.*

Démonstration : L'image réciproque d'un ouvert par f est un ouvert, l'image réciproque d'un ouvert par g est un ouvert, donc l'image réciproque par $g \circ f$ est un ouvert.□
(on peut aussi simplement utiliser le théorème 107)

Proposition 106 *Soit $\mathcal{B}$ une base de voisinages de $f(x)$.*
f est continue en x si et seulement si quel que soit $V \in \mathcal{B}$, $f^{-1}(B) \in V(x)$.

Démonstration : Soit un voisinage U de $f(x)$, il contient un certain V appartenant à $\mathcal{B}$. L'image réciproque de V étant un voisinage de x, l'image réciproque de U contient V et est donc aussi un voisinage de x.□

Théorème 107 *Les assertions suivantes sont équivalentes :*
• *f est continue*
• *Pour tout ouvert U, $f^{-1}(U)$ est un ouvert de X.*
• *Pour tout fermé F, $f^{-1}(F)$ est un fermé de X.*
• *Pour tout ouvert $V \in \mathcal{B}$, avec $\mathcal{B}$ une base d'ouverts, $f^{-1}(V)$ est ouvert*
• *Pour tout A, $f(\overline{A}) \subset \overline{f(A)}$*

Démonstration : L'équivalence entre les 4 premières assertions est claire. La cinquième assertion est une conséquence facile de la continuité de F (il suffit de voir qu'elle équivaut à $\overline{A} \subset f^{-1}(\overline{f(A)})$ et de rappeler que l'adhérence de A est l'intersection de tous les fermés contenant A). Réciproquement, en supposant la cinquième assertion vraie, on montre facilement que tout fermé F de l'image de f vérifie $f^{-1}(F)$ fermé. Il suffit de voir alors que f est continue de X vers Y si et seulement si elle est continue en tant que restriction de X sur $f(X)$.$\square$

On peut noter alors que si f est une application de X dans Y, alors si X est muni de la topologie discrète (topologie égale à l'ensemble des parties de X) ou si Y est muni de la topologie grossière (topologie limitée à $\{\emptyset, Y\}$) alors f est nécessairement continue.

Définition 108 (Limite) *Soit $f : X \setminus \{x_0\} \to Y$, avec $x_0 \in X$. On dit que y est une* **limite** *de f en x_0, si pour tout voisinage V de y dans Y, la réunion $f^{-1}(V) \cup \{x_0\}$ est un voisinage de x_0.*

Proposition 109 *Les propriétés suivantes sont équivalentes au fait que l soit limite de x_n :*
- *pour tout voisinage V de l, il existe un nombre fini de x_n en dehors de V.*
- *Dans le cas où l'espace est métrique : la distance de x_n à c tend vers 0.*

Lemme 110 *f est continue en x_0 si et seulement si $f(x_0)$ est limite de $f_{|X \setminus \{x_0\}}$ en x_0.*

Démonstration : Faisable sans trop de difficultés.$\square$

Définition 111 (Point isolé) *x_0 est* **isolé** *si et seulement si $\{x_0\}$ est ouvert.*
Un espace topologique est dit **discret** *si tous ses éléments sont des points isolés.*

Lemme 112 *Le point x_0 n'est pas isolé si et seulement si $V \setminus \{x_0\} \neq \emptyset$, pour tout $V \in \mathcal{V}(x_0)$, et encore si et seulement si $x_0 \in \overline{X \setminus \{x_0\}}$.*

Démonstration : Clair.$\square$

Un problème est la non-unicité de la limite, a priori. Nous avons donc besoin de la notion d'espace séparé, que l'on définira un peu plus loin.

Définition 113 (Homéomorphisme) *Un* **homéomorphisme** *est une application bijective continue et de réciproque continue.*

Exercice 114 (Quelques propriétés des homéomorphismes.) • *L'identité est un homéomorphisme.*
• *Une composition d'homéomorphisme est un homéomorphisme.*
• *Sur un espace normé, la translation et l'homothétie de rapport non nul sont des homéomorphismes.*
• *L'ensemble des homéomorphismes de X vers X est un sous-groupe de l'ensemble des bijections de X vers X.*

Démonstration : Rien de difficile dans tout ça ; notons que la réciproque d'une homothétie est une homothétie, et qu'une homothétie est continue parce que les opérations algébriques sont continues (voir proposition 163). □

5.1.7 Espace séparé

Définition 115 (Espace séparé) *Un espace est* **séparé** *si pour toute paire de points (x, y) on peut trouver un voisinage de x et un voisinage de y disjoints.*

Exercice 116 • *Un espace métrique est séparé.*
• *Une topologie discrète est séparée.*

Exercice 117 *Tout ensemble fini d'un espace séparé est fermé.*

Démonstration : Dans le cas d'un singleton il est clair que le complémentaire est voisinage de chacun de ses points, donc ouvert, par 77. Le passage à un ensemble fini se voit par les propriétés immédiates des fermés données en 61.□

Théorème 118 *Soit $f : X \to Y$.*
Si x_0 n'est pas isolé et si Y est séparé, alors l'application f a au plus une limite en x_0.

Démonstration : On considère les voisinages respectifs de deux limites, et on considère l'intersection de leurs images inverses respectives ; cette intersection est réduite à un singleton ; or c'est un voisinage de x_0.□

Théorème 119 *Soient f_1 et f_2 deux applications continues ayant même ensemble de départ et même ensemble séparé d'arrivée. Alors $\{x | f_1(x) = f_2(x)\}$ est fermé.*

L'hypothèse de séparation est nécéssaire (de même que dans le théorème suivant, même contre-exemple) ; considérer par exemple f_1 et f_2 deux applications de

$\mathbb{R}$ (muni de sa topologie usuelle) dans $\{0, 1\}$ muni de la topologie grossière. f_1 est l'application nulle, f_2 est nulle sauf en 0 ; $f_2(0) = 1$.

Démonstration : On montre que l'ensemble complémentaire est ouvert. Pour cela on considère x dans ce complémentaire, et deux voisinages disjoints de $f_1(x)$ et $f_2(x)$; l'intersection des images réciproques de ces voisinages est un voisinage de x qui montre que notre complémentaire est bien un voisinage de x.□

Corollaire 120 *Si f_1 et f_2 coïncident sur un ensemble dense et ont valeurs dans un espace séparé, alors elle coïncident partout.*

Démonstration : Il suffit de se rappeler qu'un fermé est égal à son adhérence, et que l'adhérence d'un ensemble dense est l'espace tout entier.□

Lemme 121 *Si f est continue et injective, et si l'espace d'arrivée est séparé, alors l'espace de départ est aussi séparé.*

Ce lemme servira à montrer le théorème 160

Démonstration : On considère deux points distincts de l'espace de départ, leurs images sont distinctes par l'injectivité, on peut les séparer par deux ouverts, d'images réciproques ouvertes. La suite est triviale. □

5.1.8 Continuité et limite dans les espaces métriques ou normés

Définition 122 (Continuité séquentielle) *f est* **séquentiellement continue en** *x si et seulement si pour toute suite x_n convergeant vers x les $f(x_n)$ convergent vers $f(x)$.*

Théorème 123 *Soit X à base dénombrable de voisinages, alors toute fonction séquentiellement continue est continue.*

Démonstration : On considère une suite de voisinages décroissants (V_n) de x. Soit W un voisinage de $f(x)$. Si $f^{-1}(W)$ n'est pas un voisinage de x, alors on peut trouver $x_n \in V_n \setminus f^{-1}(W)$; x_n tend vers x ; or $f(x_n) \notin W$, et donc $f(x_n)$ ne peut pas tendre vers $f(x)$.□

Corollaire 124 *Si f est séquentiellement continue sur un espace métrique, alors f est continue.*

Démonstration : Il faut simplement considérer l'exercice 100□

Ce corollaire servira notamment pour le théorème 369.

Proposition 125 (Définition $\epsilon - \delta$ de la continuité) *Soit f application entre espaces métriques ; f est continue en x si pour tout ϵ il existe δ tel que $d(x, x') < \delta \to d(f(x'), f(x)) < \epsilon$*

Démonstration : Il suffit de remarquer que la famille des boules ouvertes de rayon ϵ et de centre $f(x)$ est une base de voisinages de $f(x)$, et que la famille des boules ouvertes de rayon δ et de centre x est une base de voisinages de x.$\square$

Définition 126 (Continuité uniforme) *Une application f d'un espace métrique dans un autre espace métrique est dite* **uniformément continue** *si, pour tout $\epsilon > 0$ il existe $\alpha > 0$ tel que, pour tout $(x, y) \in X^2$, $d(x, y) < \alpha \to d(f(x), f(y)) < \epsilon$.*

La continuité uniforme n'est pas une notion topologique mais une notion métrique ; i.e. deux distances équivalentes ont la même notion de continuité uniforme (que l'on change la distance dans l'espace de départ ou dans l'espace d'arrivée), mais le fait que deux métriques soient associées à la même topologie ne suffit pas pour qu'elles aient la même notion de continuité uniforme.$\square$

La continuité uniforme est une notion très importante ayant de nombreuses applications.

Pour montrer la continuité uniforme, on dispose des outils suivants :
- une fonction Lipschitzienne entre métriques est uniformément continue
- une fonction bornée de $\mathbb{R}$ dans $\mathbb{R}$ et monotone est uniformément continue
- une fonction continue sur un compact est uniformément continue (théorème de Heine 198, voir le dit théorème pour d'innombrables applications)
- si p et q sont conjugués et si f et g appartiennent à L^p et L^q de $\mathbb{R}^n$ respectivement, alors $f * g$ (convoluée) est uniformément continue.

Une propriété essentielle est le théorème 246.

Définition 127 *On dit qu'une suite f_n d'applications de X dans Y avec Y un espace métrique* **converge uniformément** *vers f si pour tout ϵ positif il existe N tel que pour tout $n \geq N$ et tout x dans X $d(f(x), f_n(x)) < \epsilon$.*

Les applications et des exemples classiques :

Tout d'abord, quelques résultats célèbres de densité pour la topologie de la convergence uniforme : voir le théorème de Runge 1210, le théorème de Stone 415 (avec son corollaire le théorème de Stone-Weierstrass ; voir en particulier les polynômes de Bernstein $B_n(f)(x) = \sum_{k=1}^{n} f(\frac{k}{n}) C_n^k x^k (1-x)^{n-k}$ qui convergent uniformément vers f sur $[0, 1]$, voir théorème 417).

Il faut absolument se rappeler la convergence uniforme d'une série entière sur tout disque de rayon strictement inférieur au rayon de convergence.

Quelques résultats célèbres utilisant la convergence uniforme : 634,637 (intégration de fonctions réglées), 674 (sur la limite uniforme d'une suite de fonctions holomorphes). Quelques variantes à notre convergence uniforme ci-dessus définie, et d'autres résultats (notamment métrisabilité) : voir définition 710, et les résultats qui suivent ; voir aussi Ascoli et ses conséquences, 20.2.

Il convient enfin de signaler quelques applications de la convergence uniforme aux espaces L^p et à l'intégration :
- théorème de Plancherel : il existe un unique isomorphisme de L^2 dans L^2 appelé transformation de Fourier L^2 notée $f \mapsto \hat{f}$ telle que pour tout f dans $L^1 \cap L^2$ $\hat{f}$ est la transformée de Fourier L^1 de f, $\|\hat{f}\|_2 = \|f\|_2$ (voir par exemple le livre [16])
- théorème de Sard : voir [6].
- Intégration au sens de Riemman : voir partie 15.3.

Définition 128 (Applications lipschitzienne) *Une application h est dite* **lipschitzienne** *s'il existe $K \in [0, +\infty[$ tel que*

$$d(h(x), h(x')) \leq K.d(x, x')$$

On dit aussi qu'elle est K-lipschitzienne.
On définit la **constante de Lipschitz** *par*

$$Lip(h) = sup\{\frac{d(h(x), h(x'))}{d(x, x')} | x, x' \in X, x \neq x'\}$$

Proposition 129
• *Les fonctions lipschitziennes sont continues, et même uniformément continues.*
• *Les fonctions C^1 d'un compact de $\mathbb{R}$ dans un espace vectoriel normé sont Lipschitziennes, ainsi que les fonctions dérivables de $\mathbb{R}$ dans un espace vectoriel normé à dérivée bornée (voir le théorème 458).*

Exemple : la distance $x \mapsto d(x, x_0)$ sur un espace métrique E avec x_0 appartenant à E est 1-lipschitzienne de E dans $\mathbb{R}$. La distance de $E \times E$ dans $\mathbb{R}$ est lipschitzienne, pour toutes les normes usuelles.

Définition 130 (Norme d'une application linéaire) *Si ϕ est une application linéaire entre espaces normés, on définit sa* **norme** $\| \phi \|$ *par*
$\| \phi \| = sup\{\| \phi(x) \| \ / \ \| x \| \leq 1\}$
Cette norme peut a priori être infinie - ce qui signifie donc que l'appellation "norme", bien que classique, est abusive. Il ne s'agit d'une norme qu'en se restreignant à l'ensemble des applications pour lesquelles cette "norme" est finie.

Lemme 131

$$\| \phi \|= sup\{\| \phi(x) \| \ / \ \| x \|= 1\} = sup\{\frac{\| \phi(x) \|}{\| x \|}/x \neq 0\}$$

Démonstration : Il suffit d'avoir la patience de le vérifier...□

Théorème 132 *Une application linéaire entre espaces normés est continue si et seulement si sa norme est $< \infty$. Elle est continue si et seulement si elle est lipschitzienne et son coefficient de Lipschitz est égal à sa norme.*

Démonstration : Si ϕ est continue en zéro, il est clair que pour r suffisamment petit, $\| x \|< r$ implique $\| \phi(x) \|< 1$; on constate alors par linéarité que $\| \phi \|\leq r^{-1}$.

Réciproquement si ϕ a une norme finie, alors ϕ est lipschtzienne $\| \phi(x)-\phi(y) \|=\| \phi(x-y) \|\leq\| \phi \| . \| (x-y) \|$, et $Lip(\phi) \leq\| \phi \|$; en considérant x de norme 1, on constate que $Lip(\phi) =\| \phi \|$; d'où le résultat.□

Exercice 133 (Critère de continuité pour une forme linéaire sur un espace normé) *ϕ fonction de E dans son corps $K = \mathbb{R}$ ou $K = \mathbb{C}$ est continue si et seulement si son noyau $\phi^{-1}(0)$ est fermé.*

Démonstration : Si ϕ est continue, il est clair que l'image réciproque d'un singleton est un fermé. Réciproquement, par contraposée, supposons que ϕ n'est pas continue, alors f n'est pas non plus séquentiellement continue (voir le corollaire 124), donc il existe une suite x_n tendant vers 0 telle que $\phi(x_n)$ ne tend pas vers 0. La suite $y_n = \frac{x_n}{\phi(x_n)}$ (définie pour les n tels que $\phi(x_n)$ soit $> \epsilon > 0$ après extraction d'une sous-suite) tend vers 0. On considère alors un certain a tel que $\phi(a) = 1$ (si ϕ est nulle elle est continue), et on constate que la suite $z_n = y_n - a$ tend vers $-a \notin \phi^{-1}(0)$, alors que $z_n \in \phi^{-1}(0)$.□

Définition 134 (Borné) *Soit E un espace normé. Un sous-ensemble $A \subset E$ est dit* **borné** *si $sup\{\| x \| \ |x \in A\} < +\infty$.*
On dit que l'application f est **bornée** *sur B si et seulement si $f(B)$ est borné.*

Exercice 135 *Soit ϕ une application linéaire entre espaces normés. Les assertions suivantes sont équivalentes :*
- *ϕ est continue*
- *ϕ est continue en 0*
- *ϕ est bornée sur une boule de rayon > 0*
- *ϕ est bornée sur une sphère de rayon > 0*

Démonstration : Ces preuves sont faciles, je me contente de rappeler quelques faits qui permettent de les rédiger proprement.

La topologie est invariante par translation (puisque toute translation est un homéomorphisme), donc la continuité en 0 équivaut à la continuité en un point quelconque.

Le fait que ϕ soit bornée sur une boule équivaut trivialement au fait que ϕ soit bornée sur une sphère (par linéarité).

Si ϕ est bornée sur une boule, par linéarité il est clair qu'elle tend vers 0 en 0.

Enfin si ϕ est continue, on a montré un peu plus tôt que sa norme est finie, ce qui se voit facilement au fait que pour x suffisamment petit, on doit avoir $\phi(x)$ petit, et donc pour $\|x\| < 1$, $\|\phi(x)\| \leq 1/r$. $\square$

5.1.9 Valeur d'adhérence

Définition 136 (Valeur d'adhérence) *Soit $f : X \setminus \{x_0\} \to Y$, avec X et Y des espaces topologiques ; on dit que $y \in Y$ est une* **valeur d'adhérence** *de f en x_0 si et seulement si pour tout $V_{x_0} \in \mathcal{V}(x_0)$ et tout $V_y \in \mathcal{V}(y)$ on a $V_y \cap f(V_{x_0} \setminus \{x_0\}) \neq \emptyset$.*

Lemme 137 *L'ensemble des valeurs d'adhérence de f en x_0 est donné par l'intersection des $\overline{f(V_{x_0} \setminus \{x_0\})}$, pour V_{x_0} voisinage de x_0 ; en particulier c'est un fermé.*

Démonstration : Soit y une valeur d'adhérence, alors par définition y appartient à l'adhérence de $\overline{f(V \setminus \{x_0\})}$ pour tout V voisinage de x_0. La réciproque, tout aussi simple, est laissée de côté.$\square$

Corollaire 138 *Si x_0 n'est pas isolé, alors les limites sont des valeurs d'adhérence.*

Démonstration : Clair.$\square$

Proposition 139 (Le cas des suites) *Soit x_n une suite dans un espace topologique X.*
• Les limites de suites extraites sont des valeurs d'adhérence
• Si une valeur d'adhérence a une base dénombrable de voisinages, alors c'est la limite d'une suite extraite.

Démonstration :

• l'infini n'est pas isolé pour la topologie usuelle de $\mathbb{N}$. Donc les limites d'une suite sont des valeurs d'adhérence. Et les valeurs d'adhérence d'une suite extraite sont clairement des valeurs d'adhérence de la suite.

• Soit (V_n) une suite de voisinages de l, valeur d'adhérence de x_n ; soit $\phi(1)$ tel que $x_{\phi(1)}$ soit inclus dans V_1, $\phi(2)$ tel que $\phi(2)$ soit inclus dans V_2 et $\phi(1) < \phi(2)$,

$\phi(3)$ tel que $\phi(3)$ soit inclus dans V_3 et $\phi(2) < \phi(3)$, et ainsi de suite...$\square$

Corollaire 140 *Dans un espace métrique, les valeurs d'adhérence d'une suite sont exactement les limites des sous-suites extraites.*

Attention à l'hypothèse métrique ! Dans le cas général, ce n'est pas vrai, voir 5.6.7.

5.2 Construction de topologies

Définition 141 *Etant donné $A \subset X$, on appelle* **topologie induite** *par la topologie de X sur A l'ensemble des intersections d'ouverts de X avec A.*

Il est facile de vérifier qu'il s'agit bien d'une topologie.

Exercice 142 • *Si X est séparé, alors A est séparé pour la topologie induite.*
• *A est ouvert (resp. fermé) dans X si et seulement si tout $B \subset A$ est ouvert (resp. fermé) pour la topologie induite si et seulement si B est ouvert (resp. fermé) pour la topologie de X*
• *Si A est ouvert (resp. fermé) dans X, alors l'intérieur (resp. l'adhérence) de $B \subset A$ est le même dans X et dans A*

5.2.1 Topologie quotient

On suppose X muni d'une relation d'équivalence $\mathcal{R}$. On note π la projection canonique de X sur l'ensemble quotient.

Définition 143 (Topologie quotient) *La* **topologie quotient** *est définie comme suit :*
$U \subset X/\mathcal{R}$ est ouvert si et seulement si $\pi^{-1}(U)$ est ouvert.

On peut vérifier facilement qu'il s'agit bien d'une topologie.

Proposition 144 *Soit X un espace topologique, et $\mathcal{R}$ une relation d'équivalence sur X. On note Π la projection canonique de X sur $X/\mathcal{R}$.*
Les propriétés suivantes de la topologie quotient sur $X/\mathcal{R}$ sont fondamentales :
- la projection canonique est continue (c'est à dire que l'image réciproque de tout ouvert est un ouvert)
- la projection canonique est ouverte (c'est à dire que l'image de tout ouvert est un ouvert) si la relation d'équivalence est associée à un groupe agissant par homéomorphismes sur X (voir partie 21.3).

Démonstration : Il est clair par définition que la projection canonique est continue. Pour la réciproque il suffit de voir que si U est un ouvert de X, $\Pi^{-1}(\Pi(U))$ est la réunion des $g(U)$ pour g dans le groupe d'homéomorphismes agissant sur X.$\square$

La topologie quotient sert un peu partout, par exemples elle définit une topologie sur un espace projectif et le rend compact pour cette topologie (voir le théorème 1275).

5.2.2 Topologie sur un espace d'applications linéaires

On note $\mathcal{L}(E, F)$ l'espace vectoriel normé des applications linéaires continues de l'espace normé E dans l'espace normé F. Cet espace est normé par

$$\| \phi \| = sup\{\| \phi(x) \| \mid \| x \| \leq 1\} = sup\{\frac{\| \phi(x) \|}{\| x \|} \mid \| x \| \neq 0\}$$

On peut vérifier facilement qu'il s'agit bien d'un espace vectoriel normé.

Définition 145 (Dual topologique) *L'espace* **dual topologique** *du $\mathbb{K}$-espace vectoriel normé E est l'espace $E' = \mathcal{L}(E, \mathbb{K})$ des formes linéaires continues.*

Définition 146 (topologie forte) *On appelle* **topologie forte** *la topologie définie sur le dual par la norme usuelle.*

On va voir un peu plus loin des topologies plus ludiques sur le dual. La topologie usuelle sur le dual est la topologie faible, et pas la topologie forte (voir définition plus loin...). Notamment la partie 20 est plus fournie en la matière.

5.2.3 Topologie définie par une famille de parties d'un ensemble

Lemme 147 *Une intersection quelconque de topologies est une topologie.*

Démonstration : Evident en revenant à la définition d'une topologie.$\square$

Définition 148 *Si une topologie $\mathcal{T}$ est incluse dans une topologie $\mathcal{T}'$, on dit que T' est* **plus fine** *que $\mathcal{T}$, ou que $\mathcal{T}$ est moins fine que $\mathcal{T}'$.*

Proposition 149 *Soit A une famille de parties de X ; l'intersection de toutes les topologies contenant A est une topologie, c'est la plus petite topologie contenant A. On la note $\mathcal{T}(A)$, et on dit que c'est la* **topologie engendrée par** *A. $\mathcal{T}(A)$ est la famille des réunions arbitraires d'intersections finies de parties de $A \cup \{\emptyset, X\}$. Les intersections finies de parties de $A \cup \{\emptyset, X\}$ forment une base d'ouverts pour cette topologie.*

Démonstration : Il suffit de considérer le lemme 147 pour avoir l'existence de la plus petite topologie contenant A. Le reste est un petit exercie pas trop dur...$\square$

5.2.4 Topologie définie par une famille d'applications

Proposition 150 *Etant donné Z un ensemble, et X_i une famille d'espaces topologiques, avec $f_i : Z \to X_i$, il existe une plus petite topologie sur Z rendant toutes les f_i continues ; c'est la topologie engendrée par les $f_i^{-1}(U)$ avec U ouvert. Une base de cette topologie est donc l'ensemble des intersections finies d'images réciproques d'ouverts par des f_i.*

Démonstration : Facile avec la proposition 149$\square$

Remarquons que pour $A \subset X$ la topologie engendrée par la fonction (dite injection canonique) qui à x dans A associe x dans X est la topologie induite sur A par celle de X.

Théorème 151 *Dans la situation ci-dessus, une application f de Y dans Z est continue si et seulement si toutes les composées $f_i \circ f$ sont continues.*

On verra une application pour la continuité lorsque l'espace d'arrivée est un espace produit ; théorème 159. Ce théorème permet aussi de montrer la proposition 155.

Démonstration : Application immédiate des définitions.$\square$

Définition 152 *On dit que la famille d'applications f_i est* **séparante** *si et seulement si pour tout (x, y) il existe i tel que $f_i(x) \neq f_i(y)$.*

Proposition 153 *Si les f_i sont séparantes et si les topologies sur les X_i sont séparées, alors la topologie engendrée est séparée.*

Ce lemme permettra de montrer qu'un produit d'espaces séparés est séparé, théorème 160.

Démonstration : Supposons que x et y soient distincts ; alors puisque la famille d'applications est séparante il existe f_i telle que $f_i(x) \neq f_i(y)$; et puisque X_i est séparé, il existe un ouvert U contenant x et un ouvert V contenant y tels que U et V sont disjoints. Les ensembles $f_i^{-1}(U)$ et $f_i^{-1}(V)$ sont ouverts, puisque f_i est continue (par définition de la topologie engendrée !), et disjoints. Le résultat en découle.□

Définition 154 (Topologie faible et topologie faible *) *On appelle* **topologie faible** *sur l'espace normé E la topologie engendrée par l'ensemble des formes linéaires continues de E dans K. On appelle* **topologie faible *** *sur le dual de l'espace normé E la topologie engendré par l'ensemble des applications qui à ϕ associent $\phi(x)$, étant donné $x \in X$.*

Proposition 155 *La topologie forte définie en 146 est plus fine que la topologie faible *.*

Démonstration : En vertu du théorème 151, il suffit de voir que pour tout x la fonction qui à ϕ associe $\phi(x)$ est continue pour la norme, ce qui est facile à prouver (en se ramenant en zéro, une application linéaire étant continue si et seulement si elle est continue en zéro).□

Proposition 156 *La topologie forte d'un espace vectoriel normé est plus fine que la topologie faible.*

Démonstration : En vertu du théorème 151, il suffit de voir que toute ϕ dans E' est continue pour la norme, ce qui est évident.□

Proposition 157 *La topologie forte sur le dual E' est plus fine que la topologie faible, elle même plus fine que la topologie faible *.*

Démonstration : La première partie étant déjà montrée, il suffit de voir que la topologie faible est plus fine que la topologie faible *. Or ceci découle simplement du fait que si deux familles d'applications sont incluses l'une dans l'autre, alors les topologies engendrées sont plus fines l'une que l'autre.□

5.2.5 Topologie produit

Définition 158 (Topologie produit) *On appelle* **topologie produit** *sur le produit des X_i la topologie engendrée par les projections canoniques de X sur X_i.*

Théorème 159 *Avec π_i les projections canoniques, une application f de Y dans X est continue si et seulement si pour tout i $\pi_i \circ f$ est continue.*

Démonstration : Il suffit d'utiliser le théorème 151. □

Théorème 160 *Un produit d'espaces topologiques non vides est séparé si et seulement si chacun des facteurs l'est.*

Démonstration : Les π_i sont séparantes, donc si chaque X_i est séparé, X est séparé, par la proposition 153.
Réciproquement, il suffit de considérer un élément du produit, grâce à l'axiome du choix ; grâce à cet élément, on peut facilement construire une application de X_i dans X qui soit continue et injective ; donc X_i est séparé par le lemme 121.$\square$

Proposition 161 *La topologie sur $X_1 \times X_2$ avec X_i métrique est la topologie associée à la métrique $d((x_1, x_2), (y_1, y_2)) = max(d(x_1, y_1), d(x_2, y_2))$; on pourrait aussi prendre la somme.*

Démonstration : On rappelle simplement que les boules constituent une base d'ouverts dans un espace métrique$\square$

⚠ Cette proposition se généralise à un produit fini, et même à un produit dénombrable ; la distance entre $(x_1, x_2, ...)$ et $(y_1, y_2, ...)$ est donnée par $\sum_n \frac{min(1, d_n(x_n, y_n))}{2^n}$, avec d_n la distance sur X_n.

Exercice 162 *Le lemme précédent se généralise à un produit fini quelconque.*

Proposition 163 *Sur un espace normé la somme (opération entre deux élément des l'espace) et la multiplication (d'un élément du corps par un élément de l'espace) sont continues.*

Démonstration : L'addition est continue grâce à l'inégalité triangulaire. La multiplication est continue grâce à $\|\lambda x\| = |\lambda| \|x\|$.$\square$

Proposition 164 *Soit $E_1, ..., E_n$ et F des espaces vectoriels normés . Soit f multilinéaire de $E_1 \times ... \times E_n$ dans F, alors f est continue si et seulement si* $\| \phi \| = sup\{\| \phi(x_1, ..., x_n) \| \mid \| x_1 \| \leq 1, ..., \| x_n \| \leq 1\} < +\infty$

Démonstration : Facile.$\square$
Contrairement au cas des applications linéaires, notons qu'une application multilinéaire continue n'est pas nécessairement lipschitzienne.

Exercice 165 *Une application multilinéaire continue entre un produit d'espaces vectoriels normés et un espace vectoriel normé est lipschitzienne sur chaque sous-ensemble borné.*

Démonstration : Facile.$\square$

5.3 Compacité - liens entre complétude et compacité

5.3.1 Généralités

Définition 166 (Recouvrement ouvert) *Un* **recouvrement ouvert** *de l'espace topologique X est une famille d'ouverts U_i avec $X = \cup U_i$.*

Définition 167 (Compact) *X est* **compact** *s'il est séparé et si de tout recouvrement ouvert on peut extraire un sous-recouvrement fini.*
Un sous-ensemble K de l'espace X est dit **compact** *s'il est compact pour la topologie induite.*
Une partie A de X est dite **relativement compacte** *si sa fermeture $\overline{A}$ est compacte.*

On verra plus tard (voir lemme 177) que tout compact d'un espace séparé est fermé, et que tout compact d'un métrique est borné (s'il n'était pas borné on extrairait une sous-suite convergente d'une suite non bornée, par le théorème 199).

Un compact, dans le cas général, n'est absolument pas nécessairement fermé ! Considérer par exemple un point, dans un ensemble X contenant au moins deux points et dont la topologie est réduite à $\{\emptyset, X\}$.

Définition 168 *Un espace vérifie la* **propriété de Borel-Lebesgue** *si de tout recouvrement ouvert on peut extraire un recouvrement fini.*

Un espace est donc compact s'il est séparé et s'il vérifie la propriété de Borel-Lebesgue.

La compacité : éclaircissements, utilisation.

On verra d'autres caractérisations de la compacité que la définition par "séparé+Borel-Lebesgue". Néanmoins cette définition servira par exemple pour le théorème 381 (résultats de régularité sous le signe somme). Elle permettra aussi, en partie 5.6.12, de montrer que la compactifié d'Alexandrov est compact. Les deux premiers points de l'exercice 170, la proposition 171 (l'image continue d'un compact dans un séparé est compact), le théorème 176 de séparation des compacts, le théorème 718 (semblable au théorème de Heine dans le cas de familles équicontinues), le résultat selon lequel tout métrique compact est homéomorphe à une partie du cube de Hilbert en partie 5.6.9, le théorème de Stone 415, le corollaire du champ rentrant dans la sphère 313, le théorème d'Ascoli 719 utilisent cette même caractérisation.

Les méthodes usuelles pour montrer la compacité d'un ensemble sont le fait qu'un sous-ensemble fermé d'un compact est compact, le fait qu'un produit (quelconque) de compacts est compact (voir le théorème de Tykhonov 186[2], le théorème d'Arzéla-

[2]Le théorème de Tykhonov, conjoint au fait qu'un fermé d'un compact est compact, implique d'ailleurs que la sphère unité de $\mathbb{R}^n$ est compacte, et donc notamment l'équivalence des normes en dimension finie - voir théorème 188

Ascoli 720 (aux multiples applications), et le fait que l'image continue d'un compact dans un séparé est compacte (par exemple, dans le cas des espaces projectifs).

Des théorèmes incontournables en matière de compacité sont le théorème de Banach-Alaoglu 193 (utilisant Tykhonov), le théorème de Heine 198 ; le théorème de Baire (sous une forme moins connue que la forme classique basée sur la complétude) 249 s'applique aux espaces localement compacts. Citons aussi le théorème de Riesz 192, le théorème de Krein-Milman (soit E un espace vectoriel normé de dimension finie, K un compact convexe de E non vide, alors K est l'enveloppe convexe de ses points extrémaux : on trouvera une preuve dans [13]), le théorème de Montel 682.

La compacité dans le cas métrique offre des résultats fondamentaux :

- théorème de Bolzano-Weierstrass 199
- un espace métrique compact est séparable
- une isométrie d'un espace métrique compact dans lui-même est une isométrie[3]
- un espace métrique compact est complet (voir corollaire 235)[4]
-

> **Théorème 169** *Un espace métrique* **précompact**[a] *et complet est compact.*
>
> [a]Un espace métrique E est dit précompact si quel que soit $\epsilon > 0$ il existe un recouvrement fini de E par des boules de rayon $< \epsilon$.

Démonstration : On a déjà vu qu'un espace compact métrique est complet (corollaire un peu plus haut). Il est clair qu'il est aussi précompact. C'est donc la réciproque qui pose problème.

Supposons donc E précompact et complet. Pour montrer sa compacité, nous allons utiliser le théorème de Bolzano-Weierstrass 199. Considérons donc une suite (x_n) de E. Nous allons en chercher une sous-suite convergeante.

Il existe, par définition, pour i entier ≥ 1, $y_{i,1}, y_{i,2}, ..., y_{i,N_i}$ tels que les boules centrées sur les $y_{i,j}$ et de rayon $\frac{1}{2^j}$ recouvrent E. Construisons par récurrence sur i $1 \leq j_i \leq N_i$ tel qu'une infinité de points x_n soit dans l'intersection des boule de rayon $\frac{1}{2^l}$ centrée sur x_{l,j_l} pour $l \geq i$. On choisit alors $a_i \in \mathbb{N}$, construit aussi par récurrence, tel que la suite des a_i soit croissante, et x_{a_i} soit dans l'intersection des boule de rayon $\frac{1}{2^l}$ centrée sur x_{l,j_l} pour $l \geq i$.

Ceci définit une suite extraite de la suite des x_n, dont on montre facilement qu'elle est de Cauchy. Elle converge donc, par complétude de E. Donc, E est compact.

Une belle application est la proposition 284. $\square$

Un ensemble discret[5] dans un compact est fini ; on en déduit en particulier qu'une fonction holomorphe non nulle a un nombre fini de zéros dans un compact convexe.

Enfin il est capital que l'image d'un compact par une application continue à valeurs dans un espace séparé est compacte (voir théorème 171). Cela entraine en particulier qu'une fonction continue sur un intervalle fermé de $\mathbb{R}$ atteint ses bornes (d'où le théorème de Darboux 577, le théorème de Rolle 575, et certains critères de recherche de minima - voir partie 12.2).

---

[3]On en trouvera une preuve en application de Bolzano-Weierstrass.

[4]On en déduit notamment que le théorème du point fixe 474 s'applique dans un compact métrique et donc que la boule unité fermée $l^2(\mathbb{N})$ n'est pas compacte ; en cas contraire, l'application $(x_n)_{n\geq 0} \mapsto (y_n)_{n\geq 0}$ avec $y_i = x_{i-1}$ si $i > 0$ et $y_0 = 0$ serait bijective car une isométrie d'un espace complet compact sur lui-même est une bijection comme dit ci-dessus.

[5]I.e. tout point est isolé.

Dans les ouvrages en anglais, "compact space" est simplement un espace vérifiant la propriété de Borel-Lebesgue. L'équivalent de notre espace compact est "compact Hausdorff space".

Exercice 170 • *Toute partie finie d'un espace séparé est compacte.*

• *Tout intervalle fermé borné* $[a, b]$ *de* $\mathbb{R}$ *est compact.*

• *Soit* $(x_n)_{n\in\mathbb{N}}$ *une suite d'éléments d'un espace topologique* X *séparé tendant vers une limite* x*. Alors* $\{x_n/n \in \mathbb{N}\} \cup \{x\}$ *est un compact (preuve facile, en considérant un recouvrement par des ouverts, puis en considérant un des ouverts contenant* x*, et en voyant qu'un nombre fini des éléments de la suite est en dehors de cet ouvert.*

• $O_n(\mathbb{R})$*,* $SO_n(\mathbb{R})$ *sont des compacts (en tant que fermés bornés de* $\mathbb{R}^n$*, qui est de dimension finie).*

• *Les espaces projectifs sont compacts (voir 33.2.4).*

• *Le cube de Hilbert (voir 5.6.9) est compact.*

• *Le compactifié d'Alexandrov d'un espace séparé non compact localement compact est compact (voir 5.6.12)*

Démonstration : La première assertion est triviale. Pour la deuxième, on se donne un recouvrement ouvert $\mathcal{U}$ on considère le plus grand x tel que $[a, x]$ peut être recouvert par un recouvrement fini extrait de $\mathcal{U}$. La suite est facile ou comporte une référence vers une preuve complète.□

Proposition 171 *Si* f *est une application continue d'un espace compact* K *dans un espace séparé* Y*, alors* $f(K)$ *est compact.*

Démonstration : $f(K)$ est évidemment séparé. Etant donné un recouvrement ouvert de $f(K)$ on peut considérer le recouvrement ouvert de K constitué des images réciproques de ces ouverts ; on en extrait un recouvrement fini, et il n'y a plus qu'à repasser dans Y.□

Cette propriété servira notamment pour le théorème de Rolle 575, ou pour montrer qu'un espace projectif est compact (théorème 1275). Elle permettra aussi de montrer que tout compact métrique est isomorphe à un sous-espace topologique du cube de Hilbert (voir partie 5.6.9). Enfin, elle permet de montrer que toute isométrie d'un métrique compact sur lui-même est une bijection (corollaire 260).

Il faut noter qu'une propriété plus fine sera parfois utile :

Proposition 172 *Soit* f *une application semi-continue supérieurement d'un compact dans* $\mathbb{R}$*. Alors* f *est majorée et atteint sa borne sup.*

Cela servira notamment pour le théorème 682.

Démonstration :

• Soit K un compact, et f semi-continue supérieurement de K dans $\mathbb{R}$. Soit x la borne sup de $f(t)$ pour t dans K (à priori x peut être égal à $+\infty$).

• Soit $(x_n)_{n\in\mathbb{N}}$ croissante tendant vers x avec x_n élément de l'image de f pour tout n. Supposons que la borne sup ne soit pas atteinte (soit elle est infinie, soit x_n tend vers x sans jamais l'atteindre).

• On a alors $K = \cup_{n \in N} f^{-1}(]-\infty, x_n[)$. On peut extraire de ce recouvrement de K un recouvrement fini (en fait, un recouvrement par un seul des $f^{-1}(-\infty, x_n[$ puisque ces ensembles sont croissants) ; donc f est bien majorée.

• K est alors égal à $f^{-1}(]-\infty, x_n[)$ pour un certain n, ce qui contredit le fait que x_n croisse vers x sans jamais l'atteindre - en effet $x_n < x$ implique qu'il existe t dans K tel que $f(t) > x_n$.□

Définition 173 (Propriété d'intersection finie non vide) *Une famille $\mathcal{A}$ de parties de X a la* **propriété d'intersection finie non vide** *si et seulement si tout sous-ensemble fini de $\mathcal{A}$ a une intersection non vide.*

Proposition 174 *Un espace topologique est compact s'il est séparé et si toute famille de fermés qui a la propriété d'intersection finie non vide a une intersection non vide.*

Démonstration : Il suffit de considérer les complémentaires des fermés, qui ont le bon goût d'être ouverts.□

Outre les corollaires qui suivent, on pourra voir la proposition 183, ou le lemme 410.

Corollaire 175 *Un fermé d'un compact est compact.*

Voir (par exemple...) 410.

Démonstration : Un fermé d'un compact est évidemment séparé ; il suffit ensuite de voir qu'un fermé de notre fermé est un fermé de notre espace et d'utiliser la proposition précédente.□

Théorème 176 *Deux compacts disjoints d'un espace séparé peuvent être séparés par des ouverts.*

Démonstration : On montre tout d'abord le lemme suivant :

Lemme 177 *Si X est séparé, et K compact inclus dans X, alors K est fermé.*

Cela servira à chaque fois qu'on voudra montrer que compact équivaut à fermé borné dans un espace donné, par exemple 724.

Démonstration : On considère x dans le complémentaire de K ; pour tout y appartenant à K on peut séparer x et y par des ouverts U_y et V_y. On peut alors considérer le recouvrement de K par les ouverts V_y et en extraire un recouvrement fini. En prenant l'intersection des U_y correspondants à notre recouvrement fini, on a un ouvert autour de x, n'intersectant pas K. Donc le complémentaire de K est ouvert, donc K est fermé.□

On peut donc terminer la preuve de notre théorème, en considérant un deuxième compact K', et pour tout x de K', on peut trouver un ouvert U_y autour de x et un ouvert V_x contenant K ; on applique la compacité de K', et on obtient facilement deux ouverts disjoints séparant K' de K.

Corollaire 178 *Dans un espace compact, les sous-ensembles fermés sont les sous-ensembles compacts.*

Démonstration : Il suffit de considérer le corollaire 175 et le lemme 177.

Corollaire 179 *Tout point d'un compact possède une base de voisinage compacts.*

Démonstration : (voir figure 5.1) Soit W un voisinage ouvert de x dans l'espace compact X. Le fermé $X \setminus W$ est compact. On peut donc séparer les compacts $\{x\}$ et $X \setminus W$ par deux ouverts U et V. Alors $X \in U \subset X \setminus V \subset W$; et donc $X \setminus V$ est un voisinage compact de x inclus dans W.$\square$

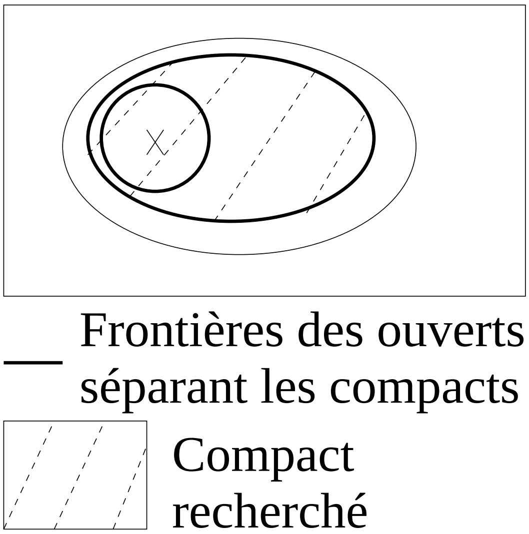


FIG. 5.1 – Construction d'une base de voisinages compacts dans un compact.

Corollaire 180 *Une fonction continue bijective d'un compact dans un espace séparé est un homéomorphisme.*

On peut citer en applications les résultats 279 et 267 (propriétés du cube de Hilbert et du Cantor triadique).

Démonstration : Il suffit de voir que l'image d'un fermé (donc compact) est compacte dans l'espace image, et donc elle est aussi fermée. Donc l'image réciproque de tout fermé par la fonction inverse est un fermé.$\square$

On peut utiliser ce résultat pour montrer que tout compact métrique est homéomorphe à une partie du cube de Hilbert, partie 5.6.9.

Théorème 181 *Les compacts de $\mathbb{R}$ sont les fermés bornés.*

Démonstration : Il suffit de considérer un interval fermé borné autour d'une partie bornée pour montrer facilement ce résultat à partir des résultats précédents et de 170.$\square$

Corollaire 182 *Etant donnée une fonction continue d'un compact dans $\mathbb{R}$, ses bornes supérieures et inférieures sont atteintes.*

Ce résultat sert dans la vie de tous les jours, mais on peut par exemple citer le joli théorème 290, le theorème de Rolle 575, la recherche de points extrémaux sur un compact (voir 12.2). Citons aussi le résultat 292 sur les billards strictement convexes du plan. Enfin, il servira pour le théorème 290 (point fixe commun à un sous-groupe compact d'automorphismes d'un espace de Hilbert).

Démonstration : Trivial au vu du résultat précédent et de la proposition 171.$\square$

Proposition 183 *Toute suite à valeurs dans un compact admet une valeur d'adhérence.*

Démonstration : La suite des $\overline{\{x_m/m \geq n\}}$ a la propriété d'intersection finie non vide ; il ne reste plus qu'à appliquer la proposition 174.$\square$

Définition 184 (Localement compact) *Un espace topologique est* **localement compact** *s'il est séparé et si tout point possède un voisinage compact.*

Proposition 185 *Tout point d'un espace localement compact possède une base de voisinages compacts.*

Démonstration : Si $x \in Int(K)$ avec K compact, alors x possède une base de voisinages compacts dans K muni de la topologie induite (par 179). Comme $x \in Int(K)$, cette base de voisinages est aussi une base de voisinages de x dans X.$\square$

5.3.2 Le théorème de Tykhonov

Théorème 186 (Théorème de Tykhonov) *Soit X_i une famille d'espaces tous non vides. Le produit est compact si et seulement si chacun des facteurs l'est.*

Démonstration : On a déjà montré que le produit est séparé si chacun des facteurs l'est (voir 160). La compacité du produit X entraîne la compacité de chacun des facteurs comme on peut s'en rendre compte en considérant la projection canonique sur chacun des facteurs. Il reste donc à voir la réciproque, c'est à dire que X est compact, si chacun des facteurs l'est. On trouvera une démonstration dans Bourbaki, ou bien dans [13]. La démonstration utilise le lemme de Zorn 36.

Il est important de noter que l'on peut prouver Tykhonov dans le cas d'un produit dénombrable de compacts métriques (X_i, d_i) sans faire appel à l'axiome du choix. Cela se fait simplement en considérant :
• La métrique d_i' associée à la métrique d_i, avec $d_i' = min(d_i, 1)$.
• La métrique sur le produit des compacts définie par $D(x, y) = \sum \frac{1}{2^i} d_i'(x_i, y_i)$.
• La topologie de cette métrique est la topologie produit.
• Il ne reste plus qu'à utiliser la caractérisation des compacts métriques par les sous-suites (théorème de Bolzano-Weierstrass, théorème 199).□

Dans le cas d'un produit fini de compacts métriques, la preuve est évidente.

Corollaire 187 *Les compacts de $\mathbb{R}^n$ sont les fermés bornés.*

Démonstration : Etant donnée une partie bornée, on considère un produit d'intervalles fermés bornés dans lequel cette partie est incluse, et le résultat vient tout seul.□

5.3.3 Application aux espaces vectoriels normés

Théorème 188 *Toutes les normes sur un $\mathbb{R}$- ou $\mathbb{C}$- espace vectoriel de dimension finie sont équivalentes.*

Démonstration : On considère une base, et la norme qui a un élément de E associe la somme des valeurs absolue de ses composantes. On montre qu'une norme quelconque est équivalente à cette norme. Il suffit pour cela de noter que la sphère unité (pour notre norme) est compacte, par compacité de la même sphère dans R^n et continuité des opérations algébriques, et de vérifier que toute norme est continue et donc atteint sur cette sphère un minimum et un maximum (NB : toute norme est continue car K-lipschitzienne pour K le max des normes d'images d'éléments d'une base orthonormale).□

Corollaire 189 *Un sous-espace vectoriel (de dimension finie) d'un espace normé est fermé.*

Une application se trouve juste après le théorème de Baire 249 : un espace de Banach de dimension infinie ne possède pas de base dénombrable.

Démonstration : Nous avons tout d'abord besoin d'un lemme :

Lemme 190 *Un sous-espace vectoriel de dimension finie d'un espace vectoriel de dimension finie est fermé.*

Démonstration : On considère la même norme que dans le théorème précédent. Pour cette norme notre espace est clairement fermé (au vu des équations le définissant). Plus précisément, on considère une base de notre espace vectoriel E, telle que F soit engendré par les k premiers éléments de cette base (c'est possible grâce au théorème de la base incomplète). Alors F est l'intersection d'hyperplans fermés d'équations $x_i = 0$.$\square$
On peut maintenant finir notre preuve ; soit $x \in \overline{F}$, avec F de dimension finie ; alors on se place dans l'espace généré par une base de F plus le vecteur x, et on utilise le lemme ci-dessus.$\square$

Exercice 191 *Toute application linéaire d'un espace normé de dimension finie dans un espace normé est continue.*

Démonstration : Il suffit de considérer une base et la norme définie plus haut.$\square$

Théorème 192 (Théorème de Riesz) *Un espace normé est de dimension finie si et seulement si sa boule unité fermée est compacte.*

On verra une application amusante avec le corollaire 725, une autre (utilisant aussi le théorème d'Arzéla-Ascoli et le théorème d'isomorphisme de Banach) avec le théorème 747.

Démonstration : Supposons E de dimension finie, alors toutes les normes sont équivalentes, on peut se ramener à $E = \mathbb{R}^n$; comme la boule unité est fermée bornée, elle est compacte. Réciproquement (voir figure 5.2), supposons la boule unité fermée compacte, alors on peut la recouvrir par des boules ouvertes de diamètre 0.5 en nombre fini. On considère alors l'espace F engendré par les centres de ces boules, et on montre que l'on peut approcher tout point de la boule arbitrairement bien avec des points de F ; ensuite on utilise le fait que F est de dimension finie et donc est fermé.$\square$

Théorème 193 (Théorème de Banach-Alaoglu) *Soit E' le dual d'un espace normé, alors sa boule unité fermée est compacte pour la topologie faible * (ie la topologie engendrée par les applications qui à $\phi \in E'$ associent $\phi(x)$ pour un certain $x \in E$.*

La boule unité fermée est l'ensemble des formes linéaires ϕ telles que $\|\phi(x)\| \leq \|x\|$.

Démonstration : (voir figure 5.3) On identifie E' à une partie du produit $\mathbb{K}^E$, en identifiant ϕ à $(\phi(x))_{x \in E}$. La topologie faible * est alors la topologie induite sur E' par la topologie produit sur $\mathbb{K}^E$. La boule unité $\overline{B_{E'}}$ est contenue dans $\Gamma = \Pi_{x \in E} \overline{B}(0, \|$

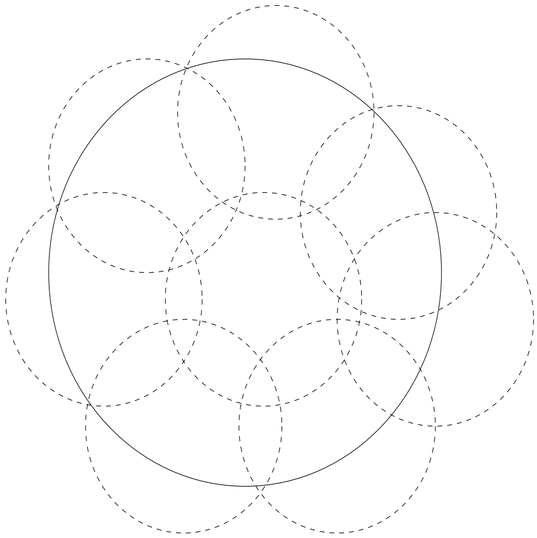


FIG. 5.2 – Le théorème de Riesz. On approxime x de la boule par le centre du cercle le plus proche, et on réitère avec le double de la distance entre x et ce centre.

$x \parallel) \subset \mathbb{K}^E$. Par le théorème de Tykhonov ce produit est compact. Il suffit donc maintenant de montrer que $\overline{B}_{E'}$ est fermé comme sous-ensemble de Γ muni de la topologie produit, ce qui se fait aisément en considérant les équations définissant $\overline{B}_{E'}$ (qui sont simplement les équations définissant les fonctions linéaires).□

Voir la proposition 742 par exemple.

Proposition 194 *La boule unité fermée du dual d'un espace séparable est métrisable pour la topologie faible *.*

Démonstration : On considère une suite x_n dense dans E, à valeurs non nulles ; la topologie faible sur la boule unité fermée peut être définie par la métrique

$$d(\phi, \psi) = \sum_{n \geq 0} \frac{|\phi(x_n) - \psi(x_n)|}{\parallel x_n \parallel .2^n}$$

Cette (courte) vérification étant faite, le résultat est acquis.□

Corollaire 195 *On peut en outre extraire de toute suite de la boule unité fermée du dual d'un espace séparable une suite convergeant faiblement.*

Démonstration : Laissée au lecteur...□

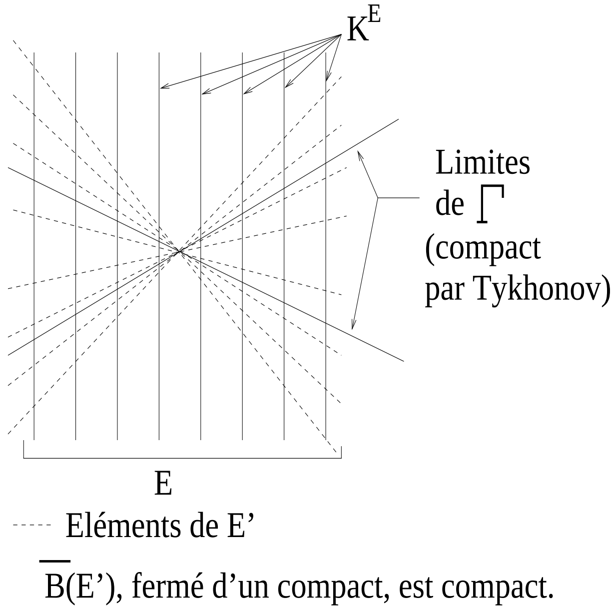


FIG. 5.3 – Schéma explicatif de la preuve du théorème de Banach-Alaoglu.

5.3.4 Espaces métriques compacts

Théorème 196 *Un espace métrique compact est séparable. Il possède donc une base dénombrable d'ouverts.*

__**Démonstration :**__ Soit X métrique compact. Pour tout n on peut trouver une suite finie de points telle que les boules centrées sur ces points et de rayon $\frac{1}{n}$ recouvrent X. La suite obtenue en mettant bout à bout toutes ces suites finies est dense dans X.$\square$

Corollaire 197 *Un espace métrique compact est de cardinal inférieur ou égal à celui de* $\mathbb{R}$.

__**Démonstration :**__ Un espace métrique compact est séparable ; donc il admet une base dénombrable d'ouverts. En prenant un x_i dans chaque ouvert, on obtient donc que tout point est limite d'une suite de x_i. Il suffit alors de voir que l'ensemble des suites d'un ensemble au plus dénombrable est de cardinal au plus la puissance du continu, ce qui se voit facilement, en considérant par exemple la fonction qui à un réel $x \in [0,1]$ dont le développement binaire comporte une infinité de 1 associe la suite $(u_n)_{n\in\mathbb{N}}$ telle que u_n est égal au nombre de 0 entre le n-ième 1 et le $n+1$-ième 1.$\square$

Théorème 198 (Théorème de Heine) *Une application continue d'un espace métrique compact vers un espace métrique est uniformément continue.*

Ce théorème servira par exemple pour le théorème 423. Il peut aussi servir à montrer qu'une application continue de $\mathbb{R}$ dans $\mathbb{R}$ tendant vers une limite finie en plus et moins l'infini est uniformément continue.

Démonstration : On considère, pour $\epsilon > 0$, pour chaque $x \in X$, $\alpha_x > 0$ tel que $d(x,y) < \alpha_x \to d(f(x), f(y)) < \epsilon/2$. Par compacité, on peut recouvrir X par un nombre fini de boules de centre x et de rayon $\alpha_x/2$. On prend alors $\alpha = inf \alpha_i$, et le résultat vient tout seul...$\square$

Théorème 199 (Théorème de Bolzano-Weierstrass) *Un espace métrique est compact si et seulement si toute suite à valeurs dans X contient une sous-suite convergente.*

Voir par exemple le théorème de Brouwer 311, le théorème de Tykhonov dans le cas d'un produit dénombrable d'espaces métriques (voir juste après le théorème 186) sans utiliser l'axiome du choix. Le théorème est aussi utilisé dans le lemme 1208, qui servira à démontrer le théorème de Runge. Le corollaire 201 est une autre application : toute isométrie d'un espace métrique compact dans lui-même est une bijection.

Démonstration : Si X est métrique compact, alors toute suite (x_n) a une valeur d'adhérence (considérer la suite décroissante de parties de X constituées des éléments $X_n = \overline{\{x_k/k \geq n\}}$; la suite des adhérences de ces parties a la propriété d'intersection finie), et X étant métrique, une sous-suite converge vers cette valeur d'adhérence.
Réciproquement, considérons tout d'abord les deux lemmes suivants :

Lemme 200 (Lemme de Lebesgue) *Soit (X,d) un espace métrique tel que toute suite contienne une sous-suite convergente. Si V_i est un recouvrement ouvert de X, alors il existe $\epsilon > 0$ tel que pour tout $x \in X$, il existe i tel que $B(x,\epsilon) \subset V_i$.*

Démonstration : Dans le cas contraire, on peut pour tout entier n trouver un x_n tel que la boule de centre x_n et de rayon $1/n$ ne soit contenue dans aucun V_i. Alors on extrait de cette suite une sous-suite convergente. On obtient que pour n assez grand les boules en question seront incluses dans le V_i qui contient x.$\square$

Corollaire 201 *Une isométrie d'un espace métrique compact sur lui-même est une bijection.*

Démonstration : Supposons E un tel espace, et f une isométrie de E dans E.
Supposons que x n'appartienne pas à l'image de f. Alors, x est à distance $> \epsilon >$

0 de l'image de f (en effet l'image de f est compacte comme image continue d'un compact, voir proposition 171, or la distance entre un compact et un fermé disjoint de lui est > 0, voir corollaire 260).

Considérons alors $u_n = f^n(x)$, et supposons que u_{k_n} converge, pour (k_n) une certaine suite strictement croissante. Si l'on aboutit à une contradiction, alors le théorème de Bolzano Weierstrass permettra de conclure que l'espace ne peut être compact.

$d(u_{k_n}, u_{k_{n+1}}) = d(u_{k_{n+1}-k_n}, x)$ puisque f est une isométrie. Or $d(u_{k_{n+1}-k_n}, x) > \epsilon$ par définition de x et puisque les u_n appartiennent à l'image de f pour $n > 0$. D'où la contradiction recherchée.$\square$

Lemme 202 *Sous les mêmes hypothèses que le lemme 200, pour tout $\epsilon > 0$, il existe une suite finie x_i telle que les boules $B(x_i, \epsilon)$ recouvrent X.*

Démonstration : Si le lemme est faux pour un certain ϵ, alors on peut construire par récurrence une suite telle que chaque point soit à une distance au moins ϵ des autres points, ce qui contredit l'hypothèse.$\square$

Avec ces deux lemmes on conclut facilement ; si toute suite contient une sous-suite convergente, alors étant donné un recouvrement ouvert (V_i), on peut construire par le premier lemme un ensemble de boules recouvrant X et tel que chaque boule est incluse dans l'un des V_i ; ensuite par le deuxième lemme, on se ramène à un nombre fini de points, et il ne reste plus qu'à cueillir le bon sous-ensemble des V_i.$\square$

5.4 Connexité

Définition 203 *Un espace topologique est dit* **connexe** *si les seuls sous-ensembles de X à la fois ouverts et fermés sont $\emptyset$ et X. Une partie d'un espace topologique est connexe si elle est connexe pour la topologie induite.*

On utilisera la connexité pour montrer :
- certaines formes du théorème des valeurs intermédiaires 209.
- le corollaire 470 sur la dérivabilité d'une limite d'une suite de fonctions.
- les lemmes 495 et 496, utile pour une démonstration du théorème de Jordan
- la proposition 955 utilisera la connexité pour définir une distance dans un ouvert connexe d'un espace vectoriel normé
- théorème de Runge, 1210.
- Tous les résultats basés sur l'indice, par exemple le théorème de Cauchy 661, et beaucoup de résultats sur les fonctions holomorphes.
- L'exercice de la partie 5.6.18, montrant qu'une fonction f C^∞ de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\forall x\ \exists n\ f^{(n)}(x) = 0$ est polynomiale.

On trouvera diverses autres applications de la connexité plus loin dans ce chapitre.

Proposition 204 *Les assertions suivantes sont équivalentes :*
(i) X est connexe
(ii) Une application ϕ de X dans $\{0,1\}$ continue est constante, avec $\{0,1\}$ muni de la topologie discrète.
(iii) Pour tout couple d'ouverts A et B de X, si $X = A \cup B$ et $A \cap B = \emptyset$, alors $A = \emptyset$ ou $B = \emptyset$
(iv) Pareil avec des fermés
(v) Toutes les parties de X non triviales (i.e. autres que X et $\emptyset$) ont une frontière non vide.

Démonstration : Facile :

(i) $\rightarrow$ (ii) Si X est connexe, montrons que toute application continue de X dans $\{0,1\}$ est constante.

En effet, si une telle application f n'était pas constante, on partitionnerait X en deux ouverts non vides ($f^{-1}(0) = f^{-1}(]-\frac{1}{2}, \frac{1}{2}[)$ et $f^{-1}(1) = f^{-1}(]\frac{1}{2}, \frac{3}{2}[)$) ; chacun d'eux serait alors à la fois ouvert, fermé, et non trivial.

La réciproque (ii) $\rightarrow$ (i) est non moins simple (raisonner par l'absurde : si A ouvert et fermé non vide et différent de X, alors prendre la fonction caractéristique de A dans X).

(ii) $\rightarrow$ (iii) Facile, en voyant que si A et B contredisent l'hypothèse, A est ouvert et fermé et non trivial.

Le reste est du même niveau de difficulté, je le passe sous silence... $\square$

Proposition 205 • *Si $A \subset X$ est connexe et si $A \subset B \subset \overline{A}$, alors B est connexe.*
• *Si les A_i sont des parties connexes de X et $\cap A_i \neq \emptyset$, alors $\cup A_i$ est connexe.*
• *Si les A_i sont des parties connexes de X et pour tout couple A_i, A_j il existe $i_0, ..., i_k$ avec $i_0 = i$ et $i_k = j$ tels que A_{i_l} intersecte $A_{i_{l+1}}$, alors $\cup A_i$ est connexe.*

Démonstration : Pour montrer la première assertion on utilise la deuxième des caractérisations des connexes données en 204.
La deuxième assertion n'est qu'un cas particulier de la troisième.
La troisième assertion là aussi se montre en utilisant la seconde des caractérisations des connexes données en 204.$\square$

Théorème 206 *Les connexes de $\mathbb{R}$ sont les intervalles.*

Démonstration : Facile.$\square$

Théorème 207 *L'image d'un connexe par une fonction continue est un connexe.*

Démonstration : Facile toujours en utilisant la même caractérisation des connexes.□

Théorème 208 *Soit f une application continue définie sur un connexe et à valeurs dans $\mathbb{R}$. Alors l'image de f est un intervalle.*

Démonstration : Facile au vu des deux théorèmes précédents.□

Corollaire 209 *Le théorème des valeurs intermédiaires (dans le cas d'une fonction continue, pas dans le cas d'une fonction dérivée) découle immédiatement du théorème ci-dessus.*

Théorème des valeurs intermédiaires pour une fonction dérivée, dit aussi théorème de Darboux, 577.

Théorème 210 (passage à la douane) *Soit X un espace topologique et $A \subset X$ connexe. Si A intersecte à la fois B et son complémentaire, alors A intersecte la frontière de B.*

Démonstration : Il suffit de voir que les deux ouverts $Int(B)$ et $Ext(B)$ ne peuvent recouvrir A.□

Théorème 211 *Un produit d'ensembles non vides est connexe si et seulement si chacun des facteurs l'est.*

Démonstration : Il est facile de voir, via les projections canoniques, que si le produit est connexe, chacun des facteurs l'est. La réciproque est plus difficile. On commence par le cas où le produit est un produit de deux espaces (voir figure 5.4). Pour cela on montre que tous les couples $(x, y) = (x_1, x_2), (y_1, y_2)$ sont contenus dans un sous-ensemble connexe de $X_1 \times X_2$; on utilisera ensuite la proposition 205. Il suffit pour ce résultat intermédiaire de considérer l'union de $X_1 \times \{y_2\}$ et de $\{x_1\} \times X_2$. Par récurrence, on généralise ce résultat à tout produit fini de connexes.

On considère maintenant un produit quelconques X de facteurs X_i connexes non vides. On considère un élément y de X, en utilisant l'axiome du choix. Pour A fini inclus dans I (I est l'index de X_i), on définit alors le sous-ensemble X_A de X défini par $(x_i) \in X_A$ si et seulement si $x_i = y_i$ pour tout i tel que $i \notin A$. X_A est connexe

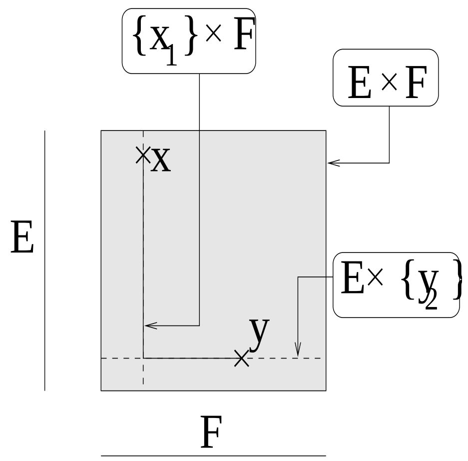


FIG. 5.4 – Un produit fini d'ensembles est connexe si et seulement si chacun des facteurs l'est. La généralisation à un produit infini se fait par un argument de connexité de l'adhérence d'une partie connexe convenablement choisie (voir le texte).

puisqu'homéomorphe à un produit fini de X_i. On peut vérifier que la réunion des X_A est dense dans X (en se rappelant qu'une base d'ouverts d'une topologie produit est l'ensemble des intersections finies d'images d'ouverts par les projections inverses) et connexe (par le deuxième point de la proposition 205), et on conclut par le premier point de la proposition 205.□

> **Théorème 212** *Une fonction localement constante sur un connexe est constante.*

Démonstration : Il suffit de voir que l'image réciproque d'un singleton est à la fois ouverte et fermée.□

> **Définition 213 (Composante connexe)** *Avec $x \in X$, la* **composante connexe** *de x, notée $C(x)$, est la réunion de tous les connexes contenant x.*

> **Proposition 214** • *Tout point appartient à sa composante connexe*
> • *La composante connexe d'un point est le plus grand connexe contenant ce point*
> • *Les composantes connexes sont fermées*
> • *Deux composantes connexes sont disjointes ou confondues. En particulier, la famille des composantes connexes forment une partition de l'espace.*

Démonstration : Le premier point est trivial, le deuxième aussi par 205, le troisième découle de la connexité de $\overline{C(x)}$, le quatrième point découle du fait que la réunion de deux connexes non disjoints est un connexe (deuxième point de la proposition 205).□

Définition 215 (Arc ou chemin, ligne brisée) *Un* **arc ou chemin** *est une application continue de* $[0,1]$ *dans* X. *L'image de* 0 *et l'image de* 1 *sont les extrémités de l'arc.*
Une **ligne brisée entre** a **et** b *est une suite finie de segments* $[x_i, x_{i+1}]$ *avec* $i \in [0, n-1]$, $x_0 = a$ *et* $x_n = b$.
On appelle **longueur d'une ligne brisée** *la somme des longueurs de ses segments.*

Exercice 216 • *Dans un espace normé, l'application qui à* t *associe* $(1-t).x + t.y$ *est un arc d'extrémités* x *et* y *(on dit aussi un arc entre* x *et* y*). L'image de cet arc est appelée segment, noté* $[x, y]$.
• *D'un arc entre* x *et* y *et un arc entre* y *et* z *on peut déduire un arc entre* x *et* z.

Définition 217 (Connexe par arcs) *Un espace topologique est dit* **connexe par arcs** *si il existe un arc entre toute paire de points.*
Une partie d'un espace topologique est dite connexe par arcs si elle est connexe par arcs pour la topologie induite.

Exemples : Un convexe est connexe par arcs.
Démonstration : Cela découle des exemples ci-dessus.□

Proposition 218 *Un connexe par arcs est connexe. La réciproque est fausse.*

Démonstration : On fixe x dans un espace connexe par arcs. Chaque arc est un connexe, car image d'un connexe ($[0,1]$) par une fonction continue ; la réunion des arcs partant de x est connexe (par la proposition 205), or par définition cette réunion est l'espace tout entier. Pour la réciproque, considérer la figure 5.5.□

Exercice 219 *Soit l'application* $f :]0,1] \to \mathbb{R}$, *qui à* x *associe* $1/sin(x)$. *Montrer que la fermeture de son graphe est connexe mais pas connexe par arcs.*

Démonstration : On suppose qu'il existe une fonction ϕ continue qui à 0 associe $(0,1)$ et à 1 associe $(1, sin(1))$, et telle que pour tout x on ait $\phi(x)$ appartienne au graphe de f. On considère x_0 le sup de l'ensemble des x tels que la première composante de $\phi(x)$ soit nulle. Il suffit ensuite de considérer la limite de la deuxième composante pour x tendant vers x_0.□

Proposition 220 *Soit* C_i *une famille de parties connexes par arcs. Si pour toute paire* i, j *il existe une suite finie* $C_{a_0}, ..., C_{a_k}$ *avec* $C_{a_h} \cap C_{a_{h+1}} \neq \emptyset$ *et* $a_0 = i$ *et* $a_k = j$, *alors la réunion est connexe par arcs.*

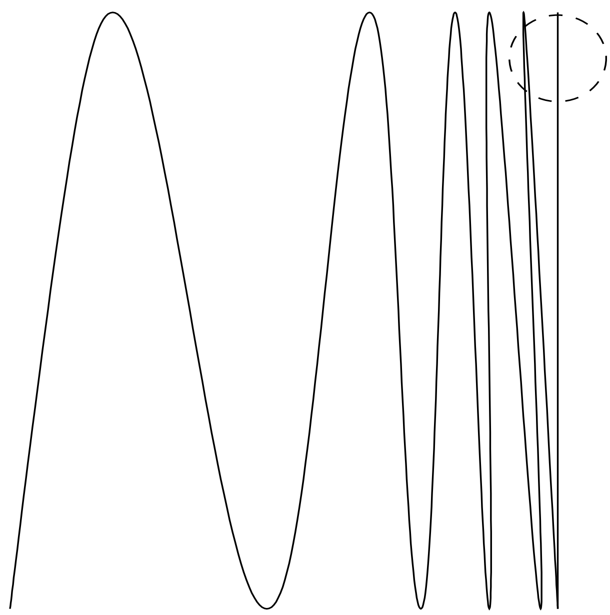


FIG. 5.5 – Un connexe qui n'est pas connexe par arcs. La même figure fournit un exemple de connexe qui n'est pas localement connexe. Il s'agit de la courbe des $(x, sin(1/x))$ vers 0 par valeurs inférieures, plus la frontière $\{0\} \times [-1, 1]$. On voit que la figure n'est pas localement connexe en considérant ce qu'il se passe au voisinage du point $(0, 1)$.

Démonstration : Facile.□

Définition 221 (Composante connexe par arcs) *La* **composante connexe par arcs** *de x est la réunion de tous les connexes par arcs passant par x ; on la note $C_a(x)$.*

Proposition 222 • *La composante connexe par arcs d'un point est connexe par arcs.*
• *Deux composantes connexes par arcs sont soit disjointes soit confondues.*
• *$C_a(x) \subset C(x)$, car $C_a(x)$ est un connexe contenant x, et $C(x)$ est le plus grand connexe contenant x par définition.*

Définition 223 (Localement connexe (par arcs)) *Un espace est* **localement connexe (resp. par arcs)** *si tout point de l'espace possède une base de voisinage connexes (resp. par arcs).*

⚠ Attention ; un espace peut être connexe sans être localement connexe. Voir par exemple la figure 5.5.

Notamment, alors qu'un espace dont tout point possède un voisinage compact (par exemple un espace compact !) est localement compact, un espace dont tout point possède un voisinage connexe n'est pas nécessairement localement connexe.

Théorème 224 *Dans un espace localement connexe (resp. localement connexe par arcs), les composantes connexes (resp. par arcs) des ouverts sont ouvertes.*

Démonstration : Facile.□

Corollaire 225 *Dans un espace localement connexe (resp. localement connexe par arcs) tout point possède une base de voisinages ouverts et connexes (resp. connexes par arcs).*

Démonstration : Il suffit de considérer, étant donné x et un voisinage V de x, un ouvert inclus dans V et contenant x,et une composante connexe (resp. par arcs) de x dans cet ouvert.□

On peut noter le théorème suivant :

Théorème 226 *Dans un espace localement connexe par arcs, les ouverts connexes sont connexes par arcs. Notamment, les ouverts connexes de $\mathbb{R}^n$, ou de tout espace vectoriel normé [a] sont connexes par arcs.*

[a] Ou même de tout espace vectoriel topologique.

5.5 Complétude

5.5.1 Suites de Cauchy. Espace complet

Définition 227 (Suite de Cauchy) *Une suite (x_n) dans un espace métrique est dite* **suite de Cauchy** *si pour tout $\epsilon > 0$ il existe un $N \in \mathbb{N}$ tel que $\forall n, m > N$ on a $d(x_n, x_m) < \epsilon$.*

La notion de suite de Cauchy est une notion métrique et non une notion topologique. Même si deux distances sont équivalentes, on ne peut être sûr que les suites de Cauchy soient les mêmes pour les deux métriques. Par exemple avec $d(x, y) = |arctan(x) - arctan(y)|$, la topologie sur $\mathbb{R}$ est la même que pour la topologie usuelle, mais la suite $u_n = n$ n'est pas de Cauchy pour la métrique usuelle, alors qu'elle est de Cauchy pour cette métrique.

Proposition 228 • *Etant donnée une suite x_n, notons $X_n = \{x_k / k \geq n\}$; alors la suite x_n est de Cauchy si et seulement si le diamètre de X_n tend vers* 0.
• *Dans un espace métrique toute suite convergente est de Cauchy.* • *L'image d'une suite de Cauchy par une application uniformément continue est une suite de Cauchy.*

Définition 229 (Espace complet) *Un espace métrique X est* **complet**, *si toute suite de Cauchy de X a une limite dans X.*

Quelques exemples d'espaces complets :
• les exemples de Banach donnés un peu plus loin.
• $C^k(\Omega)$ avec Ω un ouvert de $\mathbb{R}^n$, voir partie 20.3.

Une propriété fondamentale des espaces complets est le théorème du point fixe 474.

Définition 230 (Espace de Banach) *Un* **espace de Banach** *est un espace vectoriel normé complet.*
Un **isomorphisme entre l'espace de Banach** E **et l'espace de Banach** F *est un isomorphisme des espaces vectoriels normés sous-jacents.*

Quelques exemples d'espaces de Banach :
• $\mathbb{R}$
• $\mathbb{R}^n$ muni d'une des normes suivantes :
- $(x_1, ..., x_n) \mapsto \sum_{i=1}^n |x_i|$
- $(x_1, ..., x_n) \mapsto \sqrt{\sum_{i=1}^n x_i^2}$
- $(x_1, ..., x_n) \mapsto max_{i=1}^n |x_i|$
• L'ensemble des applications continues bornées d'un espace topologique X dans $\mathbb{R}$ ou $\mathbb{C}$, muni de la norme $f \mapsto sup_{x \in X} |f(x)|$
• Les espaces L^p, comme on le verra en 8
• Si F est un Banach et E un espace vectoriel normé , alors $\mathcal{L}(E, F)$ (ensemble des fonctions linéaires continues de E dans F) est un Banach (pour la norme $f \mapsto sup_{\|x\|=1} \|f(x)\|$).

On rappelle que deux normes sont dites équivalentes si elles définissent la même topologie.

Tout d'abord quelques propriétés des Banachs issues directement de la partie 5 :
• Un isomorphisme algébrique (i.e. un isomorphisme au sens des espaces vectoriels) continu entre espaces de Banach est un isomorphisme d'espaces vectoriels normés .
• Toutes les normes sont équivalentes sur des $\mathbb{R}$-espaces vectoriels de dimension finie.
• Deux normes sont équivalentes si et seulement si chacune d'elle est inférieure à une certaine constante multipliée par l'autre

• Un espace vectoriel normé de dimension finie est complet, et donc est un Banach.
• Les compacts d'un espace vectoriel normé de dimension finie sont les fermés bornés.

Proposition 231 *Etant donnés des espaces métriques E_i en nombre fini, le produit $E_0 \times ... \times E_n$ peut être équipé d'une métrique définie par $d((x_i),(y_i))$ de l'une des formes suivantes (entre autres) :*
• $\sum_i d(x_i, y_i)$
• $\sqrt{\sum_i d(x_i, y_i)^2}$
• $\sqrt[p]{\sum_i d(x_i, y_i)^p}$
• $max_i d(x_i, y_i)$
Ce sont bien des distances et elles sont équivalentes entre elles. La topologie ainsi définie est la topologie produit, que l'on a définie plus tôt.

Proposition 232 *Un espace métrique est complet si et seulement si l'intersection de toute suite décroissante de fermés non vides de diamètre tendant vers 0 est non vide et donc réduite à un point.*

Démonstration : Si l'espace métrique est complet, alors on considère x_n appartenant au n-ième fermé ; la suite est de Cauchy, et converge donc vers un point ; quel que soit n, ce point est limite d'une suite de points de X_n ; donc il appartient à X_n puisque X_n est fermé. En outre, le diamètre tendant vers 0, le diamètre de l'intersection est 0 ; donc il s'agit d'un seul point.
Réciproquement, étant donnée une suite de Cauchy x_n, on considère la suite des X_n avec $X_n = \overline{\{x_k / k \geq n\}}$; cette suite vérifie les hypothèses, donc l'intersection des X_n est réduite à un point. On montre facilement que ce point est limite des x_n.□

Proposition 233 *Un produit fini d'espaces métriques complets, muni d'une métrique comme défini ci-dessus, est complet.*
Réciproquement un produit fini d'espaces métriques, muni d'une métrique comme défini ci-dessus, est complet si et seulement si chacun des facteurs l'est.

Démonstration : La démonstration (pas très difficile) est laissée au lecteur.□

Proposition 234 *Si une suite de Cauchy a une valeur d'adhérence, elle est convergente.*

Démonstration : On considère une suite extraite qui converge, et on montre facilement que la suite tend vers la même limite.□

Corollaire 235 *Un espace métrique compact est complet.*

Démonstration : S'il est compact, toute suite a une valeur d'adhérence (par le théorème de Bolzano-Weierstrass 199) ; il suffit alors d'appliquer la proposition précédente.□

Théorème 236 *Le corps $\mathbb{R}$ est complet pour sa métrique ; de même $\mathbb{R}^n$ muni d'une norme est complet pour cette norme. Plus généralement un espace normé de dimension finie est complet.*

Démonstration : On considère une norme sur E de dimension finie et une suite de Cauchy x_n. On montre que pour un certain N la suite est à valeurs dans la boule de centre x_N et de rayon 1 à partir du rang N, directement par la définition d'une suite de Cauchy ; on a donc une suite dans un compact, et donc la suite de Cauchy converge vers un élément de cette boule.□

On trouvera par exemple une utilisation de ce théorème dans 244.

Proposition 237 *Un sous-ensemble d'un métrique complet est complet si et seulement si il est fermé.*

Démonstration : Soit A un sous-ensemble fermé de X complet. Si x_n est une suite de Cauchy dans A, c'est aussi une suite de Cauchy dans X, donc elle converge. Si A est fermé la limite est dans A. Réciproquement, on suppose x dans $\overline{A}$, et on choisit une suite x_n qui tend vers x ; et on remarque que x_n est de Cauchy et donc converge vers une limite dans A puisque A est complet.□

Définition 238 (Série absolument convergente) *Soit E un espace vectoriel normé. (x_n) dans E est appelée une* **série absolument convergente** *si $\sum_{n\geq 0} \| x_n \| < +\infty$.*

Théorème 239 *Un espace vectoriel normé E est complet si et seulement si toute série absolument convergente (x_n) est convergente dans E.*

Démonstration : Supposons E complet.
Soit une série x_n absolument convergente. Pour $m > n$ on a

$$\begin{aligned} d(\sum_{i=0}^{n} x_i, \sum_{i=0}^{m} x_i) &= \| \sum_{i=n+1}^{m} x_i \| \\ &\leq \sum_{i=n+1}^{m} \| x_i \| \\ &\leq \sum_{i=n+1}^{+\infty} \| x_i \| \\ &\to 0 \end{aligned}$$

Donc la suite $y_n = \sum_{i=0}^{n} x_i$ est de Cauchy, et donc converge.

Réciproquement supposons maintenant que toute série absolument convergente converge. On se donne x_n une suite de Cauchy. On en extrait une sous-suite, et $\| x_m - x_{n_k} \| \leq \frac{1}{2^k}$ pour $m \geq n_k$; la série correspondante est absolument convergente ; il est facile d'en déduire que la suite a une valeur d'adhérence, et donc qu'elle converge.$\square$

> **Théorème 240** *Si E est normé et si F est de Banach, alors l'espace normé $\mathcal{L}(E,F)$ est aussi de Banach.*

Démonstration : Soit f_n une suite de Cauchy dans $\mathcal{L}(E,F)$. Pour tout $x \in E$, on a $\| f_n(x) - f_m(x) \| \leq \| f_n - f_m \| . \| x \|$, donc la suite $f_n(x)$ est de Cauchy dans F ; elle converge vers un élément que l'on note $f(x)$. Il est clair que f est linéaire. On fixe alors $\epsilon > 0$. On choisit N tel que $\| f_n - f_m \| \leq \epsilon$, pour $n, m > N$, et on considère x de norme < 1. En faisant tendre m vers l'infini on obtient que $\forall n > N$ $\| f_n(x) - f(x) \| \leq \epsilon$; donc f est bornée sur la boule unité, et donc f est continue. On obtient avec la même formule la convergence de f_n vers f au sens de la norme. En résumé, la preuve s'obtient en montrant la convergence simple (facile), puis en montrant que la limite est linéaire (trivial), puis qu'elle est continue (y'a qu'à l'écrire et ça roule). $\square$

> **Corollaire 241** *Le dual E' d'un espace normé E est un espace de Banach.*

Démonstration : Par application immédiate du théorème précédent.$\square$

Voir le corollaire 704.

> **Théorème 242** *Si K est un espace compact et Y un espace complet, alors l'espace des applications continues de K dans Y $C^0(K,Y)$ est métrique complet pour la distance $d(f,g) = sup_x d(f(x), g(x))$.*

Démonstration : La compacité de K permet de vérifier que la fonction d est bien définie ; elle est clairement effectivement une métrique. Etant donnée f_n une suite de Cauchy dans l'espace considéré, on montre facilement que cette suite converge simplement vers une certaine fonction f ; en utilisant la continuité de f_n et la convergence uniforme on conclut facilement à la continuité de f. La convergence uniforme des f_n découle facilement du critère de Cauchy dans l'espace considéré.$\square$

Voir par exemple le théorème 634. On peut citer aussi le fait que l'espace des applications continues d'un espace K compact dans un espace E de Banach est de Banach pour la norme $\|f\|_\infty = sup_x \|f(x)\|_E$.

Exercice 243 *Si $E_1, ...E_n$ sont des espaces vectoriels normés, avec F de Banach, montrer que l'espace normé $L(E_1, ..., E_n); F)$ est aussi de Banach.*

5.5.2 Complété d'un espace métrique

Théorème 244 *Tout espace métrique (X, d) se plonge isométriquement dans un espace complet $(\tilde{X}, \tilde{d})$ avec X dense dans $\tilde{X}$.*
Si on se donne deux tels plongements, alors l'identité sur X s'étend de manière unique en une isométrie de $\tilde{X}_1$ et $\tilde{X}_2$.

Définition 245 *Un tel espace métrique complet est appelé* **complété** *de X.*

Démonstration :
Existence :
• on commence par introduire une relation d'équivalence entre les suites de Cauchy : on dit de deux suites qu'elles sont équivalentes si la distance entre l'une et l'autre tend vers 0 (ie (x_n) est équivalente à (y_n) si $\lim x_n - y_n \to 0$).
• On considère $\tilde{X}$ l'ensemble des classes d'équivalences pour cette relation.
• On note par $[x_n]$ la classe d'équivalence de la suite x_n.
• On remarque que la distance entre x_n et y_n tend vers une limite donnée pour n tendant vers $+\infty$ (noter que pour ce point on utilise la complétude de $\mathbb{R}$ montrée un peu plus tôt).
• On peut donc prendre pour distance sur $\tilde{X}$ la limite de la distance entre deux suites pour n tendant vers l'infini ; on vérifie facilement qu'il s'agit bien d'une distance.
• On peut prendre pour plongement la fonction qui à x associe la suite constante égale à x.
• On constate facilement que ce plongement est une isométrie.
• On peut voir facilement que l'image du plongement est dense dans $\tilde{X}$ en considérant pour une suite donnée x_n la suite des suites de Cauchy constante égales à x_n.
• On peut maintenant identifier X et son image.
• On considère maintenant une suite de Cauchy dans $\tilde{X}$, notée y_n.
• Pour tout n on peut choisir x_n, suite de Cauchy, telle que la distance (dans $\tilde{X}$) entre x_n et y_n soit inférieure à $1/n$.
• La suite x_n est de Cauchy dans $\tilde{X}$, et donc aussi dans X. Par définition de $\tilde{X}$, la limite de la suite x_n est la classe des suites dont la distance à x_n est nulle.
• Il ne reste plus qu'à voir que y_n tend aussi vers cette limite.

L'unicité résulte du corollaire 247.□

Théorème 246 *Soient deux espaces métriques A et B avec B complet. Si D est une partie dense de A et $f : D \to B$ est uniformément continue, alors il existe un et un seul prolongement continu $\tilde{f} : A \to B$. Cette fonction $\tilde{f}$ est de plus uniformément continue.*

Démonstration : • L'unicité est évidente, par unicité de la limite et par la densité de D dans A.
• L'existence découle immédiatement du critère de Cauchy, grâce à la complétude

de B.

• Il reste à montrer l'uniforme continuité : soient x et y distincts dans A. Soit $x_n \to x$ et $y_n \to y$, avec x_n et y_n dans D. $d(\tilde{f}(x), \tilde{f}(y)) \leq 2d(f(x_n), f(y_n))$ si n est assez grand. L'uniforme continuité de $\tilde{f}$ en découle immédiatement.

Cela servira à montrer quelques propriétés simples des espaces de Hölder, voir 738, et le théorème de Plancherel. L'escalier de Cantor utilise aussi ceci (voir partie 5.6.13). Cela permet aussi de voir que si E est métrique complet connexe localement connexe, alors E est connexe par arcs

Corollaire 247 *Une isométrie $i : D \to B$ d'un sous-ensemble dense de l'espace métrique A sur une partie de l'espace métrique complet B s'étend de manière unique en une isométrie $\tilde{i} : A \to B$ de A sur une partie de B. L'extension $\tilde{i}$ est une bijection de A sur B si et seulement si $i(D)$ est dense dans B.*

Démonstration : Evident, même preuve que pour le théorème 246.$\square$

5.6 Zoologie de la topologie

5.6.1 Séparation de fermés par des ouverts dans un métrique

Proposition 248 *Soient F_1 et F_2 deux fermés disjoints d'un espace métrique (E, d). Alors il existe deux ouverts U_1 et U_2 tels que $F_1 \subset U_1$ et $F_2 \subset U_2$, U_1 et U_2 étant disjoints.*
En outre il existe une fonction continue de E dans $[0, 1]$ dont la restriction à F_1 est égale à 0 et dont la restriction à F_2 est 1 (c'est à dire que $\chi_{F_2} \leq f \leq \chi_{F_1^c}$).

Démonstration : Considérer la fonction f définie par

$$f(x) = \frac{\sup(d(x, F_1) - d(x, F_2), 0)}{d(x, F_1)}$$

si $x \notin F_1$ et $f(x) = 0$ sinon. puis $U_1 = f^{-1}([0, 0.5[)$ et $U_2 = f^{-1}(]0.5, 1])$.$\square$

5.6.2 Théorème de Baire

Théorème 249 (Théorème de Baire) *Soit X un espace topologique. Si X est localement compact, ou s'il est métrique complet, alors*
• Toute intersection dénombrable d'ouverts denses est dense
• Une réunion dénombrable de fermés recouvrant X comporte un fermé d'intérieur non vide

Démonstration : Il suffit de montrer la première assertion, la seconde étant équivalente par passage au complémentaire.
On se donne U_i une famille dénombrable d'ouverts avec $\overline{U_i} = X$. Soit V un ouvert

non vide. On veut montrer que l'intersection des U_i a une intersection non vide avec V. On pose $V_0 = U_0 \cap V$ (ouvert non vide par densité de U_0). Ensuite, par récurrence :
• $\overline{V_n} \subset V_{n-1} \cap U_n$ pour $n \geq 1$
• Cas métrique complet : on impose $diam(V_n) \leq \frac{1}{2n}$
• Cas localement compact : $\overline{V_n}$ compact
L'intersection des $\overline{V_n}$ est non vide, car :
• Dans le cas localement compact, il s'agit d'une suite décroissante de compacts non vides.
• Dans le métrique complet, il s'agit d'une suite décroissante de parties non vides de diamètres tendant vers 0.
Il suffit alors de choisir un élément dans l'intersection des V_i.□

Les lignes qui suivent fournissent de nombreuses applications du théorème de Baire. Il y a aussi par exemple l'application 5.6.18.

Proposition 250 *Un espace de Banach de dimension infinie n'a pas de base algébrique dénombrable.*

Démonstration :

Supposons que E, espace de Banach de dimension infinie, ait une base dénombrable (e_n) pour $n \in \mathbb{N}$.

Définissons alors F_n, espace vectoriel engendré par les e_i pour $i \leq n$.

F_n est alors un fermé (car de dimension finie, résultat 189) et d'intérieur vide (facile). Or l'union des F_n est égale à E ; donc E devrait être d'intérieur vide grâce au théorème de Baire, ce qui est absurde.□

Corollaire 251 $\mathbb{R}[X]$ *n'est de Banach pour aucune norme.*

Théorème 252 (Théorème de Banach-Steinhaus) *Soit $T_\alpha : E \to F$ une famille d'applications linéaires continues de l'espace de Banach E dans l'espace normé F. Si $\forall x\ sup_\alpha \parallel T_\alpha x \parallel < \infty$, alors $sup_\alpha \parallel T_\alpha \parallel < \infty$.*

On verra une application à la transformation de Toeplitz (proposition 610), qui fournit une preuve élégante de la moyenne de Césaro (corollaire 611).

Démonstration : On pose B_n l'ensemble des x tels que $\forall \alpha$ on a $\parallel T_\alpha(x) \parallel \leq n$
B_n est fermé. L'hypothèse permet de dire que l'union des B_n est E. Par le théorème de Baire, l'un des B_n est d'intérieur non vide. On en déduit facilement le résultat.□

Corollaire 253 *Soit $T_n : E \to F$ une suite d'applications linéaires continues de l'espace de Banach E dans l'espace normé F. Si $T.x = lim_{n\to+\infty} T_n(x)$ existe pour tout x, alors T est une application linéaire continue.*

Démonstration : Facile.□

Définition 254 (Application ouverte) *Une application est dite* **ouverte** *si l'image de tout ouvert est un ouvert.*

Théorème 255 (Théorème de l'application ouverte) *Une application linéaire continue surjective entre espaces de Banach est ouverte.*

Démonstration : Donnons nous une telle application f, entre deux espaces de Banach E et F. f est donc supposée linéaire, continue, et surjective. On montre qu'elle est ouverte.

• Soit U un ouvert de E. Il suffit de montrer que $f(U)$ est ouvert dans F.

• Soit x dans $f(U)$. Il suffit de montrer que $f(U)$ est un voisinage de x.

• Par translation, on peut supposer $x = 0$.

• Il suffit donc de montrer qu'une certaine boule $B_F(0,r)$ dans F centrée en 0 de rayon un certain $r > 0$ est incluse dans l'image par f d'une boule arbitraire $B_E(0,r')$ dans E, centrée en 0, de rayon r', avec r' tel que $B_E(0,r') \subset U$. Par linéarité de f, on peut se limiter à $r' = 1$. Il convient donc de montrer qu'il existe r tel que $B_F(0,r) \subset f(B_E(0,1))$.

• Définissons, pour $n \in \mathbb{N}$, $F_n = \overline{f(B_E(0,n))}$.

• D'après le théorème de Baire (ci-dessus), l'union des F_n étant égale à E, il existe un F_n d'intérieur non vide. Du coup, F_1, par linéarité, est lui-même d'intérieur non vide.

• Il existe donc une boule $B_F(y,\epsilon)$ centrée en y de rayon $\epsilon > 0$ incluse dans F_1.

• Par symétrie $-y \in F_1$. Finalement, $B_F(y,\epsilon) - y \subset F_1 + F_1$, donc $B_F(0,\epsilon) \subset F_2$ ($F_1 + F_1 = F_2$ comme on s'en convaincra aisément).

• $B_F(0,\frac{\epsilon}{2}) \subset F_1$, d'où le résultat. □

Corollaire 256 (Théorème d'isomorphisme de Banach) *Une application continue linéaire bijective entre espaces de Banach a un inverse continu (→est un isomorphisme).*

Démonstration : Conséquence immédiate du théorème de l'application ouverte.□

Voir par exemple le théorème 747, utilisant aussi le théorème de Riesz et le théorème d'Arzéla-Ascoli.

Corollaire 257 *Si E muni de la norme N_1 est un Banach et si E muni de la norme N_2 est un Banach, alors si $N_1 \leq \lambda.N_2$ pour un certain λ implique $N_2 \leq \mu N_1$ pour un certain μ. Donc si une des deux normes est plus fine que l'autre, alors elles sont équivalentes.*

Démonstration : L'identité de (E,N_2) dans (E,N_1) est continue, bijective, linéaire ; donc c'est un isomorphisme.□

Corollaire 258 (Théorème du graphe fermé) *Soit $T : E \to F$, linéaire entre les Banach E et F. L'application T est continue si et seulement si le graphe de T est fermé dans $E \times F$.*

Démonstration : Un sens ne pose pas de problème ; le graphe d'une application continue est toujours fermé.
Réciproquement, supposons le graphe fermé, voir figure 5.6.
• L'espace $E \times F$ est en fait un espace de Banach.
• Le graphe est en fait un Banach (la linéarité de T permet de conclure que le graphe est en fait un espace vectoriel, qui est fermé par hypothèse).
• La projection du graphe sur E restreinte au graphe est une application linéaire bijective du graphe sur E ; par le corollaire 256, son inverse est aussi continue. La projection du graphe sur F est aussi continue.
• La fonction considérée, composition de deux fonctions continues, est donc continue.□

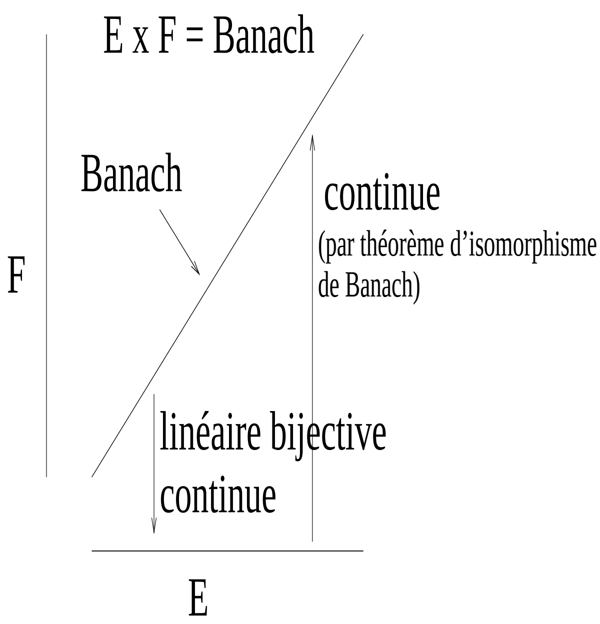


FIG. 5.6 – Schéma explicatif de la preuve du théorème du graphe fermé.

5.6.3 Distance d'un point à une partie

Proposition 259 *Soit A une partie non vide d'un espace métrique (E, d). L'application $\tilde{d}$ qui à x dans E associe $d(x, A) = inf\{d(x, a)/a \in A\}$ est continue et 1-lipschitzienne.*

Cette proposition servira un peu partout, par exemple pour le lemme 413 (très utile pour approximer des fonctions par des fonctions C^∞), ou pour le lemme 261, ou pour voir que les ϵ-voisinages sont ouverts.

Démonstration : Soit x dans E, montrons que $\tilde{d}$ est continue en x. Considérons t dans E, et donnons nous $\epsilon > 0$ (figure 5.7).

Par définition de $\tilde{d}$ et de l'inf, il existe $a \in A$ tel que $d(x,a) \leq d(x,A) + \epsilon$. Alors $d(t,A) \leq d(t,x) + d(x,a) \leq d(t,x) + d(x,A) + \epsilon$, donc $\tilde{d}(t) \leq \tilde{d}(x) + d(t,x) + \epsilon$. En faisant tendre ϵ vers 0 on obtient $\tilde{d}(t) \leq \tilde{d}(x) + d(x,t)$. De même on aurait $\tilde{d}(x) \leq \tilde{d}(t) + d(x,y)$. Donc

$$|\tilde{d}(x) - \tilde{d}(t)| \leq d(x,t)$$

On en déduit la continuité et le caractère 1-lipschitzien de $\tilde{d}$.$\square$

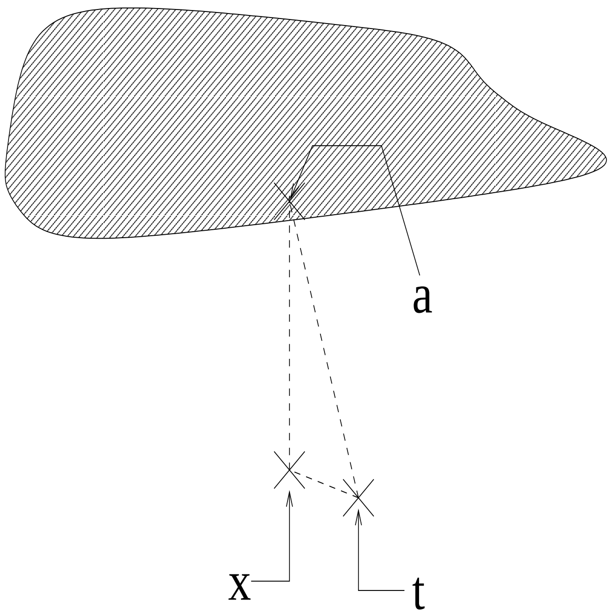


FIG. 5.7 – Continuité de la distance à une partie : une simple application de l'inégalité triangulaire.

Corollaire 260 *La distance entre un compact et un fermé disjoints est* > 0.

par distance entre un compact et un fermé on entend l'$\inf$ de la distance entre un point du compact et un point du fermé.

Démonstration : La distance à un ensemble étant continue, la distance d'un point du compact au fermé atteint son minimum sur le compact (voir 182). Si la distance est nulle, alors elle est nulle en un certain point x du compact. On prend alors une suite x_n du fermé tendant vers x (par exemple $d(x_n, x) < 1/n$) ; sa limite est nécessairement dans le fermé, donc x est à l'intersection du fermé et du compact, donc ces deux ensembles ne sont pas disjoints. D'où la contradiction, et le résultat.$\square$

La distance entre deux fermés disjoints n'est pas nécessairement non nulle !

Considérer dans $\mathbb{R}$ les fermés F_1 et F_2 définis par

$$F_1 = \mathbb{N}$$

$$F_2 = \{\frac{n+1}{n} / n \in \mathbb{N} \wedge n \geq 2\}$$

5.6.4 Approximation d'ouverts par des compacts

Lemme 261 *Soit U un ouvert de $\mathbb{R}^n$. Pour $m \geq 1$, notons K_m l'intersection de*

$$\{x \in U / d(x, U^c) \geq \frac{1}{m}\}$$

et de

$$\overline{B}(0, m)$$

Alors :
• Pour tout $m > 0$ K_m est compact
• $K_m \subset Int(K_{m+1})$
• $U = \cup_i K_i$
• $\forall K$ compact de U $\exists m / K \subset K_m$

Ce résultat servira par exemple pour le corollaire 725, ou pour l'utilisation de la définition 729, ou pour la proposition 714.

Démonstration : • K_m est borné (clairement), K_m est une intersection de deux fermés (rappelons que pour une partie non vide donnée l'application qui à un point associe sa distance à cette partie est continue, voir proposition 259). Un fermé borné de $\mathbb{R}^n$ est compact (corollaire 187).

• Il suffit de voir que l'intérieur d'une intersection finie est l'intersection des intérieurs.

• La distance d'un point x de U au complémentaire de U est > 0 car le complémentaire de U est fermée (un point à distance nulle d'un fermé est dans ce fermé).

• Soit K un compact inclus dans U.

L'inf de la distance d'un point de K au complémentaire de U est > 0 grâce à un corollaire précédent. Donc cette distance est supérieure à $1/m$ pour m assez grand. Il suffit ensuite de prendre m assez grand pour que K soit inclus dans la boule fermé $\overline{B}(0, m)$.□

Lemme 262 (Approximation d'ouverts du plan par des compacts pas trop troués) *Soit Ω un ouvert de $\mathbb{C}$ (on pourrait dire $\mathbb{R}^2$). Alors il existe une suite de compacts K_n inclus dans Ω tels que :*
• $K_n \subset Int(K_{n+1})$
• Tout compact de Ω est inclus dans un certain K_n
• Toute composante connexe de $(\mathbb{C} \cup \{\infty\}) \setminus K_n$ contient une composante connexe de $(\mathbb{C} \cup \{\infty\}) \setminus \Omega$

La dernière condition signifie simplement qu'il n'y a pas de "trous" superflus dans les K_n. La seconde condition implique que la réunion des K_n est Ω.

Démonstration : On utilise les mêmes K_n que dans la partie précédentes.

Le seul problème est de vérifier que la dernière condition est vérifiée.

On se donne donc C une composante connexe de $\hat{C} \setminus {K_n}$[6], et x appartenant à cette composante connexe.

Alors nécessairement $|x| > n$ ou $|x - f| < 1/n$ pour un certain f dans F, avec F le complémentaire de Ω.

Dans le cas $|x| > n$, alors les $\lambda.x$, pour λ réels ≥ 1, forment une demi-droite, qui unie à $\{\infty\}$, forme un connexe, inclus dans C, et intersectant une composante connexe de $\hat{C} \setminus \Omega$ (puisque $\infty \notin \Omega$!).

Dans le cas $|x - f| < 1/n$, le segment $[x, f]$ est inclus dans C, donc C contient f, et donc intersecte $\hat{C} \setminus \Omega$, au moins en f.

D'où le résultat.□

5.6.5 Homéomorphisme entre une boule fermée et un compact convexe d'intérieur non vide

Proposition 263 *Soit K un compact convexe d'intérieur non vide de $\mathbb{R}^n$. Alors K est homéomorphe à la boule unité fermée.*

Démonstration :

• On peut supposer que 0 appartient à l'intérieur de K.

• On peut agrandir K jusqu'à ce qu'il contienne la boule unité fermée.

• On définit la fonction f de la boule dans K qui à x associe $T.x$ avec T le sup des $t \in \mathbb{R}^+$ tels que $t.u$ appartient à K, avec u le vecteur directeur de x $(x/\|x\|)$. On définit $f(0) = 0$.

• Montrons tout d'abord que f est bien définie. Si x est non nul, K étant borné, le sup des t en question est bien défini si x est non nul (le cas $f(0)$ étant séparé).

• $t.u$ appartient à K pour tout $t < T$, par convexité de K. Le fait que K est fermé fait que $T.u$ appartient à K. Si x est de norme 1, le problème est donc résolu. Si x est de norme plus petite que 1, a fortiori, $T.x$ appartient à K par convexité de K.

• Il faut maintenant montrer que f est continue.

• f est continue en 0. En effet il est clair que $f(x)$ tend vers 0 quand x tend vers 0.

• Il convient maintenant de montrer que f est continue en x autre que 0. Pour cela il suffira de montrer que la fonction qui à u associe T le sup des $t \in \mathbb{R}^+$ tels que $t.u$ appartient à K est continue sur la sphère (ensuite il est clair que la multiplication par un scalaire est continue, que la fonction qui à un vecteur associe son vecteur directeur est continue (par quotient $x/\|x\|$).

• La figure 5.8 parle d'elle même.□

Cela permet d'appliquer des triangulations sur la boule unité fermée (enfin sur un simplexe homéomorphe à la boule), voir théorème 312.

5.6.6 Intersection vide d'une suite décroissante de fermés convexes non vides bornés d'un espace vectoriel normé

Sur $\mathbb{R}$ l'intersection d'une suite décroissante de convexes fermés bornés non vides ne saurait être vide. Dans le cas général il en est tout autrement.

[6] $\hat{C}$ est le compactifié d'Alexandrov de $\mathbb{C}$ $(\mathbb{C} \cup \{\infty\})$

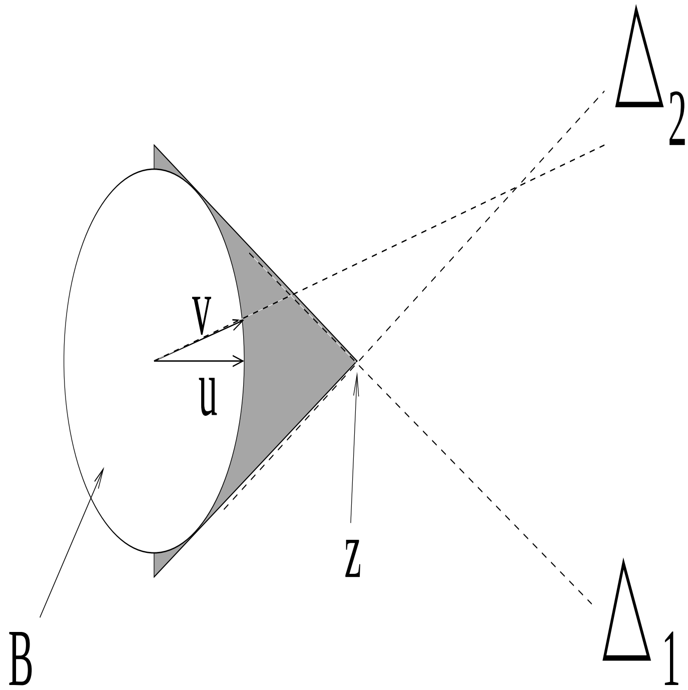


FIG. 5.8 – Par hypothèse, la boule B est incluse dans K. On se donne z le point $f(u)$, c'est à dire le "bord" de K dans la direction u. Alors la zone grisée appartient nécessairement à K par convexité. $f(v)$ est alors nécessairement au delà de Δ_1 par convexité de K, et en deça de Δ_2, par définition de z. D'où la continuité de f.

• Soit E l'espace des fonctions continues de $[0,1]$ dans $\mathbb{R}$.

• C'est un espace vectoriel normé pour la norme infinie. C'est même un espace de Banach.

• Soit $(x_n)_{n\in\mathbb{N}}$ une suite de rationnels dense dans $[0,1]$.

• Soit C_n l'ensemble des applications de E nulles en x_i pour tout i dans $[0,n]$, bornées par 2 et d'intégrale 1 sur $[0,1]$.

• Les C_n sont non vides, convexes, fermés, bornés, décroissants.

• L'intersection des C_n ne peut contenir que des fonctions nulles sur tous les rationnels, et continues, donc l'intersection des C_n ne peut pas contenir de fonction non nulle. Or L'intersection des C_n ne peut contenir que des fonctions d'intégrale 1.□

5.6.7 Valeurs d'adhérence $\neq$ limites de suites extraites

Proposition 264 *L'ensemble des valeurs d'adhérence d'une suite n'est pas nécessairement égal à l'ensemble des limites de sous-suites extraites.*

Démonstration : En effet, soit E l'ensemble des applications continues de $[0,1]$ dans $[0,1]$; on le munit de la topologie produit, c'est à dire de la topologie de la convergence simple (il est facile de vérifier que c'est pareil...).

• Un voisinage de la fonction nulle dans E est de la forme $\{f/\forall i \in [1,n] |f(x_i)| \leq \epsilon_i\}$ $(*)$, avec les ϵ_i positifs, et les x_i dans $[0,1]$.

• On considère les applications en dents de scie, égales à 0 en x_0, en x_1, en x_2, ... , en x_n, avec $x_i < x_{i+1}$, $x_0 = 0$ et $x_1 = 1$; et affine entre x_i et $\frac{x_i+x_{i+1}}{2}$ et $\frac{x_i+x_{i+1}}{2}$ et x_{i+1}, avec $f(\frac{x_i+x_{i+1}}{2}) = 1$, avec les x_i rationnels.

• On peut clairement les énumérer, et donc il s'agit d'une suite.

• la suite nulle est clairement dans l'adhérence de cette suite (prendre un voisinage de la fonction nulle écrit sous la forme $(*)$, et regarder ce qu'il se passe.

• aucune suite extraite ne peut tendre simplement vers la fonction nulle, sinon le théorème de convergence dominée de Lebesgue permettrait de dire que l'intégrale de la fonction nulle est la limite de l'intégrale des fonctions de la suite - or toute fonction de notre suite a une intégrale $\frac{1}{2}$.$\square$

5.6.8 Les espaces projectifs

Proposition 265 *Les espaces projectifs sont compacts et connexes par arcs.*

On trouvera plus d'informations sur ce sujet dans la partie 33.2.4.

5.6.9 Le cube de Hilbert

Définition 266 *On appelle* **cube de Hilbert** *l'espace produit $[0,1]^{\mathbb{N}}$, muni de la topologie produit ($[0,1]$ étant muni de la topologie usuelle sur les segments de réels).*

Propriétés :

• Le cube de Hilbert est connexe par arcs (considérer, étant donnés deux suites de réels $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$, l'application qui à t associe $(x_n + t.(y_n - x_n))_{n\in\mathbb{N}}$). Chaque composante étant continue, cette application est continue.

• Le cube de Hilbert est métrisable (considérer l'application qui à (x_n) et (y_n) associe $\sum_{n\in\mathbb{N}} \frac{|x_n-y_n|}{2^n}$.

• Le cube de Hilbert est compact ; par application du théorème de Tykhonov.

Théorème 267 *Tout espace métrique compact K est homéomorphe à un sous-espace topologique du cube de Hilbert.*

<u>Démonstration :</u>

• Etant donné $n \in \mathbb{N}$, on considère un recouvrement de K par des boules $b_{n,i}$ pour $i \in [1, t(n)]$, en nombre fini et de rayon $1/n$ (on peut toujours construire un recouvrement fini, en extrayant un recouvrement fini du recouvrement comportant TOUTES les boules de rayon $1/n$, via la compacité de K).

• On note $B_{n,i}$ la boule de même centre que $b_{n,i}$, mais de rayon double $(2/n)$

• On peut, par le lemme d'Urysohn 411, trouver une fonction continue $f_{n,i}$ égale à 1 sur $b_{n,i}$, comprise entre 0 et 1, et nulle en dehors de la boule $B_{n,i}$.

• On peut alors construire l'application f qui à un point x de K associe

$$(f_{1,1}(x), ..., f_{1,t(1)}(x), f_{2,1}(x), ..., f_{2,t(2)}(x), ..., f_{m,1}(x), ..., f_{m,t(m)}(x)...),$$

qui est un élément du cube de Hilbert.

• Cette application est continue, puisque toutes ses composantes sont continues.

• Elle est injective, on l'a d'ailleurs un peu beaucoup construite pour ça.

• f de K dans $f(K)$ est alors une application continue bijective d'un espace compact dans un espace séparé ; ceci implique que f est un homéomorphisme, d'après le corollaire 180.$\square$

5.6.10 Fonction non continue vérifiant la propriété des valeurs intermédiaires

Il suffit de considérer la fonction qui à un réel x associe $sin(1/x)$ si x est non nul et 0 sinon.

5.6.11 Tous les fermés de $\mathbb{R}^n$ s'expriment comme zéros de fonctions indéfiniment dérivables

$\boxdot$ **Des fermés particuliers, le cas de la dimension** 1

> **Lemme 268** *Il existe une fonction C^∞ de $\mathbb{R}$ dans $\mathbb{R}$ qui s'annule exactement sur $]-\infty, 0]$.*

Démonstration :

• On pose $f(x) = exp(-1/x)$ pour $x > 0$, $f(x) = 0$ sinon.

• Il est clair que f est C^∞ sur $\mathbb{R}^*$.

• En 0 on peut facilement voir que toutes les dérivées sont nulles, car leurs limites sont nulles, puisqu'elles s'expriment comme produit d'une fraction rationnelle par un $exp(-1/x)$. $\square$

> **Lemme 269** *Tout intervalle ouvert de $\mathbb{R}$ s'exprime comme complémentaire de l'ensemble des zéros d'une fonction C^∞.*

Démonstration : • $]a, +\infty[$ ou $]-\infty, a[$: voir lemme précédent.

• $]a, b[$ est l'ensemble des zéros de $x \mapsto f(x-a).f(b-x)$.$\square$

> **Lemme 270** *Tout fermé de $\mathbb{R}$ s'exprime comme ensemble des zéros d'une fonction C^∞.*

Démonstration :

• On note U le complémentaire du fermé à étudier.

• U est ouvert.

• U est réunion dénombrable d'intervalles ouverts disjoints (preuve en vérifiant qu'il y a un rationnel dans toute composante connexe d'un ouvert)

• On note ϕ_n une fonction (voir lemme précédent) C^∞ qui s'annule exactement sur le complémentaire du n-ième intervalle de la partition ci-dessus.

• On note ϕ la somme des ϕ_n. Cette somme est bien définie car il y a au plus un des ϕ_n qui est non nul en un point donné.

• ϕ est indéfiniment dérivable, comme on s'en rend facilement compte en regardant ce qu'il se passe au voisinage d'un point donné - qui n'appartient qu'à un seul support de ϕ_n.□

⊡ Des fermés particuliers, le cas général : dimension quelconque

> **Lemme 271** *Soit B une boule ouverte de $\mathbb{R}^n$; il existe une fonction C^∞ nulle partout sauf dans cette boule, où elle est > 0.*

Démonstration : On montre le résultat pour la boule unité ouverte, la généralisation étant évidente.

• Soit $f(x) = exp(-\frac{1}{1-\|x\|^2})$ pour x tel que $\|x\| < 1$, et $f(x) = 0$ sinon. La norme ici évoquée est la norme euclidienne.

• la situation étant invariante par rotation, on se contente de montrer que la fonction est C^∞ sur le premier axe (i.e. l'ensemble des $(x, 0, 0, ...0)$ pour x dans $\mathbb{R}$).

• Pour cela on montre que chaque dérivée partielle est C^∞.

• Tout d'abord dans le cas d'un point autre que 0 ou 1 :

- La dérivée partielle suivant un autre axe que le premier est clairement nulle, par symétrie du problème.

- La dérivée partielle suivant le premier axe est clairement C^∞, comme composée d'applications C^∞, voir le lemme 269.

• Et en zéro, il suffit de voir que le carré de la norme euclidienne est une fonction polynômiale, donc C^∞.

• Le cas 1 se traite facilement, comme dans le lemme 268.□

> **Lemme 272** *Tout ouvert s'écrit comme réunion dénombrable de boules ouvertes.*

Démonstration :

• On considère la suite x_n des points à coordonnées rationnelles de U, un ouvert.

• Pour tout x_n, on définit r_n le sup des r tels que $B(x_n, r) \subset U$.

• r_n est bien positif strictement, puisque U est ouvert.

• Il est clair que tout rationnel de U est inclus dans la réunion des $B(x_n, r)$.

• Tout point x est inclus dans une boule centrée sur x de rayon ϵ incluse dans U ; donc une boule centrée sur un rationnel situé à une distance au plus $\epsilon/3$ de x et de rayon maximal va contenir x. En effet

• Des deux • précédents, on déduit donc que notre ouvert s'exprime comme réunion dénombrable de boules ouvertes.□

> **Théorème 273 (Le résultat tant attendu)** *Tout fermé de $\mathbb{R}^n$ s'exprime comme zéro d'une fonction C^∞.*

Démonstration : • Soit F un fermé. On considère U son complémentaire. On écrit U comme une réunion dénombrable de boules ouvertes B_n.

• On fait alors une somme pondérée de fonctions comme définies dans le lemme 271. Avec ϕ la fonction donnée par le lemme 271 pour la boule unité, on peut écrire cette somme comme $\sum_{i=1}^{\infty} c_i \phi(2\frac{x-x_i}{r_i}$.

• Il faut maintenant parvenir à sommer la n-ième fonction pondérée par le terme (strictement positif) nécessaire pour ramener toutes ses dérivées en dessous de $1/2^n$ fois une constante ne dépendant que de l'ordre de la dérivée ; ainsi on aura convergence normale de toutes les dérivées et donc la somme sera C^∞. La difficulté est que contrairement au cas de la dimension 1, les boules ne sont pas disjointes.

• Il est suffisant pour cela que la somme des c_i/r_i^k soit convergente. En effet, dans ce cas la dérivée k-ième de $c_i \phi(\frac{x-x_i}{r_i}$ sera majorée par la borne sur la dérivée k-ième de ϕ, divisée par r_i^k.

• Il suffit de choisir $c_i = e^{-\frac{1}{r_i}}.2^{-i}$; ainsi $\sum c_i/r_i^k = \sum 2^{-i}.e^{-\frac{1}{r_i}}/r_i^k \leq \sum 2^{-i}.M_k$ avec M_k le sup de $x^k.e^{-\frac{1}{x}}$.$\square$

5.6.12 Le compactifié d'Alexandrov

On se donne X un espace topologique séparé, non compact, localement compact. L'objectif va être de construire un espace $\tilde{X}$ à peine plus gros que X, qui lui sera compact, et qui contient un sous-espace topologique homéomorphe à X.

On pose $\tilde{X} = X \cup \{\infty\}$. On définit $\mathcal{T}$ l'ensemble constitué :

- des ouverts de X

- des $\tilde{X} \setminus K$, où K est un compact de X.

• Il est facile de vérifier que $\mathcal{T}$ est une topologie. L'ensemble des ouverts de X est bien stable par intersection finie et par réunion quelconque ; et l'ensemble des complémentaires de compacts de X dans $\tilde{X}$ est bien lui aussi stable par intersections finies et réunion quelconques (rappelons qu'une réunion finie de compacts est compacte et qu'une intersection quelconque de compacts est compact - comme fermé d'un compact) ; il suffit donc de vérifier que la réunion (resp. l'intersection) d'un ouvert de X et d'un complémentaire de compact de X est bien un ouvert de X ou un complémentaire de compact de X.

Pour cela soit U un ouvert de X, et K un compact de X, de complémentaire V. $U \cap V$ est l'intersection d'un ouvert avec $V \setminus K$ qui est un ouvert ; donc c'est un ouvert de X. Et $U \cup V$ est le complémentaire de $K \cap U'$, avec U' le complémentaire de U dans $\tilde{X}$.

• Montrons que X est dense dans $\tilde{X}$. Pour le voir il suffit de voir que tout voisinage de ∞ intersecte X ; ce qui est clair car <u>X n'est pas compact</u>[7].

• On va maintenant montrer que X est homéomorphe à un sous-espace de $\tilde{X}$. L'identité de X dans $\tilde{X}$ est injective. Les ouverts de $\tilde{X}$ inclus dans X étant exactement les ouverts de X, il est clair qu'il s'agit bien d'un homéomorphisme.

• Montrons maintenant que $\tilde{X}$ est séparé (premier pas pour montrer qu'il est compact). On peut séparer deux points de X par des ouverts, puisque <u>X est séparé</u>. Montrons maintenant qu'on peut séparer un point $x \in X$ de ∞. On se donne pour cela K un voisinage compact de x, ce qui peut se faire puisque <u>X est localement compact</u>. $IntK$ et $\tilde{X} \setminus K$ sont des ouverts séparant x et ∞.

• Montrons maintenant que $\tilde{X}$ vérifie la propriété de Borel-Lebesgue, c'est à dire que de tout recouvrement d'ouverts de $\tilde{X}$ on peut extraire un recouvrement fini. Soit $X = \cup_{i \in I} U_i$, avec les U_i ouverts. Un certain U_{i_0} contient ∞. Son complémentaire est

[7] Je souligne de temps à autre les endroits où s'appliquent les hypothèses

compact, et recouvert par les U_j, pour $j \neq i_0$; on peut donc le recouvrir par les U_j, pour $j \in J$ fini. L'ensemble des U_i pour $i \in J \cup \{i_0\}$ est un recouvrement fini de $\tilde{X}$.

Exemple : $\widetilde{\mathbb{R}^n}$ est homéomorphe à la sphère unité de $\mathbb{R}^{n+1}$ [8].

Théorème 274 (Compactifié d'Alexandrov) *Si X est un espace topologique non compact et localement compact, il existe un espace topologique $\tilde{X}$ compact appelé* **compactifié d'Alexandrov** *de X et tel que :*
• X est dense dans $\tilde{X}$. • X est homéomorphe à un sous espace topologique de $\tilde{X}$.

5.6.13 Le cantor K_3

Définition 275 (Cantor K_3) *On note C_0 l'ensemble $[0, 1]$.*
On note C_1 l'ensemble $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$.
On note C_2 l'ensemble $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \ \cup \ [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, \frac{9}{9}]$
...
On note C_n l'ensemble $\frac{1}{3}.C_{n-1} \cup (C_{n-1} + 2).\frac{1}{3}$.

On note K_3 l'intersection des C_n, pour $n \in \mathbb{N}$. On appelle cet ensemble **ensemble triadique de Cantor**.
On le munit d'une topologie en considérant la restriction de la distance usuelle à K_3.

FIG. 5.9 – Construction de l'ensemble de Cantor. Les lignes successives représentent C_1, C_2, C_3, C_4.

[8]Cette sphère est de dimension topologique n.

Proposition 276 *L'ensemble triadique de Cantor K_3 est aussi l'ensemble des réels de $[0, 1]$ dont le développement 3-adique ne comporte que des 0 ou des 2.*

Démonstration : Cela se prouve facilement en considérant l'intersection des C_i jusqu'à un certain rang, et en prenant la limite en l'infini.□

Proposition 277 *L'ensemble triadique de Cantor K_3 est compact.*

Démonstration : K_3 est fermé, car c'est une intersection de fermés, et borné car inclus dans $[0, 1]$. Donc il est compact, comme tout fermé borné de $\mathbb{R}$.□

Proposition 278 *L'ensemble triadique de Cantor K_3 est de mesure nulle et d'intérieur vide.*

Démonstration : K_3 est mesurable, comme intersection dénombrable de fermé. La mesure de K_3 est inférieure à la mesure de C_n, pour tout n ; donc K_3 est de mesure nulle. K_3 est d'intérieur vide, sinon il ne serait pas de mesure nulle.□

Proposition 279 *L'ensemble triadique de Cantor K_3 est homéomorphe à $\{0, 1\}^{\mathbb{N}}$, ensemble des suites de $\{0, 1\}$, muni de la topologie produit de la topologie discrète sur $\{0, 1\}$.*

Démonstration : Soit la fonction f qui à une suite u_n de $\{0, 1\}$ associe la somme des $2.u_n/3^n$. Cette fonction est injective, clairement. Elle est surjective (proposition 276). Voyons maintenant la continuité de f ; en fait on va considérer la continuité de f^{-1}. Pour cela on considère l'image réciproque d'un ouvert non vide de la base d'ouverts de la topologie produit constituée des produits d'ouverts tels qu'un nombre fini d'ouverts seulement soient différents de $\{0, 1\}$. Il est suffisant pour que l'image réciproque de x soit dans cet ouvert que les premiers chiffres soient les mêmes, et donc que la distance soit suffisamment petite.
Enfin toute fonction continue bijective d'un compact dans un séparé est un homéomorphisme (d'après le corollaire 180), ce qui permet de conclure.□

Proposition 280 *L'ensemble triadique de Cantor K_3 est* **totalement discontinu**, *ce qui signifie que la composante connexe d'un point est réduite à ce point.*

Il suffit de montrer qu'étant donnés x et y dans K_3, il existe deux ouverts fermés disjoints contenant l'un x et l'autre y. En effet ainsi la composante connexe de x sera différente de la composante connexe de y.
Pour cela on peut considérer indifférement K_3 comme le produit $\{0, 1\}^n$ ou comme

l'intersection des C_n ; dans le premier cas il suffit de considérer le premier rang auquel les deux suites diffèrent, dans le deuxième cas, le premier chiffre dans le développement triadique pour lequel x et y diffèrent.$\square$

Proposition 281 *L'ensemble triadique de Cantor K_3 ne comporte pas de point isolé.*

On note qu'un ensemble **parfait** est un ensemble fermé et dépourvu de point isolé. K_3 sera donc un ensemble parfait.

Démonstration : Facile, soit en considérant un ouvert de la base d'ouverts dans le cas du produit $\{0,1\}^{\mathbb{N}}$, soit en considérant l'intersection d'une boule ouverte avec K_3 dans le cas de l'intersection des C_n (bien entendu, une seule de ces deux preuves suffit !).$\square$

Proposition 282 *K_3 et $\emptyset$ sont les deux seuls compacts K inclus dans $[0,1]$ qui vérifient*

$$K.\frac{1}{3} \cup (K.\frac{1}{3} + \frac{2}{3}) = K$$

Démonstration : Il est facile de vérifier que $\emptyset$ et K_3 sont des solutions de l'équation donnée.
On considère maintenant l'ensemble $K([0,1])$ des compacts non vides inclus dans $[0,1]$, et l'application f qui à un compact K associe $K.\frac{1}{3} + K.\frac{1}{3} + \frac{2}{3}$.
Cette application associe bien un compact inclus dans $[0,1]$ à un compact inclus dans $[0,1]$. On va considérer un compact A donné, non vide, et on va montrer que $f^n(A)$ tend vers K_3 pour la distance de Hausdorff.

Définition 283 (Définition de la distance de Hausdorff) *On définit tout d'abord :*

$$V_\epsilon(A) = \{x | d(x,A) < \epsilon\}$$

*$V_\epsilon(A)$ est appelé ϵ-**voisinage ouvert de** A. Il est ouvert par la proposition 259. Ensuite on note $D(A,B)$ et on appelle **distance de Hausdorff** le réel*

$$D(A,B) = inf\{x / A \subset V_x(B) \wedge B \subset V_x(A)\}$$

défini sur l'ensemble $K(E)$ des compacts non vides d'un espace métrique E donné.

Il s'agit bien d'une distance ;
• $D(A,B) >= 0$ et $D(A,B) < \infty$ est clair
• $D(A,B) = 0 \rightarrow A = B$ est clair
• l'inégalité triangulaire se vérifie facilement

Proposition 284 *Si E est un espace métrique complet, alors l'ensemble des compacts non vides de E muni de la distance de Hausdorff est complet.*

<u>**Démonstration :**</u> Soit K_n une suite de Cauchy dans l'ensemble des compacts non vides de E.
Alors il existe une suite $\epsilon_N \to 0$ telle que

$$\forall k, n > N \ D(K_k, K_n) < \epsilon_N$$

et donc

$$\forall k, n > N \ K_k \subset V_{\epsilon_N}(K_n)$$

On considère alors K l'ensemble des x tels qu'il existe une suite x_n telle que $x_n \in K_n$ et x_n admet x pour valeur d'adhérence.
K est fermé. En effet :
- soit y_∞ dans $\overline{K}$. Il existe (y_m) suite dans K tendant vers y_∞.
- y_m est limite d'une certaine suite d'éléments x_n tels que $x_n \in K_n$. On considère une suite extraite x_{n_m} telle que $d(x_{n_m}, y_m) \to 0$ comme $m \to \infty$. On définit x_n pour n n'appartenant pas à l'ensemble des n_m, en le choisissant de manière quelconque dans K_n.
Alors

$$d(x_{n_m}, y_\infty) \leq d(y_m, y_\infty) + d(x_{n,m}, y_m) \to 0$$

Donc y_∞ est valeur d'adhérence des x_n, donc $y_\infty \in K$. D'où $\overline{K} \subset K$, et donc K est fermé. Fermé d'un complet, il est donc aussi complet.

K est aussi précompact. En effet : - Soit $\epsilon > 0$.
- Il existe clairement N tel que K soit inclus dans $V_\epsilon(K_N)$.
- pour tout y dans K, il existe x_y dans K_N tel que $d(y, x_y) < \epsilon$.
- On peut construire sur K_n (puisqu'il est compact) un recouvrement fini par des boules centrés sur les z_i de rayon ϵ.
- Les boules centrées sur les z_i de rayon 2ϵ recouvrent donc K. On peut supprimer les z_i inutiles, ie tels que $B(z_i, 2\epsilon) \cap K = \emptyset$. Il reste les z_i, pour, disons, $i \in [1, M]$.
- On peut alors déterminer, pour tout $i \in [1, M]$, un élément z'_i de K à distance $< 2\epsilon$ de z_i (puisqu'on a supprimé les z_i inutiles).
- Les boules centrées sur les z'_i et de rayon 4ϵ montrent alors que K est précompact.
Précompact et complet, K est donc compact (voir théorème 169).

Il convient de montrer que K est non vide, ce qui sera fait en même temps que la preuve de la convergence des K_n ci-dessous (en effet K_n sera inclus dans $V_\epsilon(K)$).
K est limite des K_n pour la distance de Hausdorff ; pour le montrer il faut voir que pour tout ϵ il existe un N tel que pour $n \geq N$ on ait
• $K \subset V_\epsilon(K_n)$ (trivial)
• $K_n \subset V_\epsilon(K)$: pour cela on considère x dans K_n, avec n tel que $\epsilon_n \leq \epsilon$.
On considère alors
• $n_0 > n$ tel que $\epsilon_{n_0} \leq \epsilon/2^1$, et x_{n_0} dans K_{n_0}, tel que $d(x, x_{n_0}) \leq \epsilon/2^0$
• $n_1 > n_0$ tel que $\epsilon_{n_1} \leq \epsilon/2^2$, et x_{n_1} dans K_{n_1}, tel que $d(x_{n_0}, x_{n_1}) \leq \epsilon/2^1$
• $n_2 > n_1$ tel que $\epsilon_{n_2} \leq \epsilon/2^3$, et x_{n_2} dans K_{n_2}, tel que $d(x_{n_1}, x_{n_2} \leq \epsilon/2^2$
...
• $n_p > n_{p-1}$ tel que $\epsilon_{n_p} \leq \epsilon/2^{p+1}$, et x_p dans K_{n_p}, tel que $d(x_{n_{p-1}}, x_{n_p}) \leq \epsilon/2^p$
...

La suite des x_{n_p} est de Cauchy, donc elle converge vers un certain y ; en sommant les distances ont montre facilement que $d(x, y) \leq 2.\epsilon$. Pour compléter la suite des x_n pour $n_p \leq n \leq n_{p+1}$, il suffit de prendre un point quelconque dans K_n.□
On peut maintenant terminer notre démonstration sur le fait que K_3 est le seul compact non vide tel que $K_3 = f(K_3)$.
On montre facilement que f est contractante de rapport $\frac{1}{3}$ pour la distance de Hausdorff. On peut donc conclure par le théorème du point fixe 474 ; K est le seul compact non vide satisfaisant à l'équation.□

Une autre propriété est le fait que pour E métrique compact, $K(E)$ est compact.

On peut utiliser le Cantor triadique K_3 pour construire une fonction continue, croissante, dérivable presque partout, de dérivée nulle presque partout, et poutant non constante (égale à 0 en 0 et à 1 en 1).

5.6.14 Une distance sur les sous-espaces vectoriels d'un espace vectoriel de dimension finie

On se donne E un espace vectoriel de dimension finie n. On appelle S l'ensemble des sous-espaces vectoriels de E. Alors on définit la distance d sur S par

$$d(F, G) = dim(F + G) - dim(F \cap G)$$

Définition 285 *Cette distance est appelée* **distance de Grassman**.

Proposition 286 *La distance de Grassman bien une distance.*

Démonstration : En effet :
• d est bien positive (facile)
• d est symétrique (encore plus facile)
• $d(F, G) = 0$ implique $dim\ F \cap G = dim\ F + G$ or $F \cap G \subset F \subset F + G$ et donc $F \cap G = F$ et donc $F \subset G$; de même on obtiendrait $G \subset F$; et donc au final $F = G$.
• Il reste à montrer l'inégalité triangulaire. Pour cela on aura besoin de la définition suivante :

Définition 287 *On appelle* **chaîne entre les sous-espaces vectoriels** F **et** G *une suite finie de sous-espaces vectoriels , le premier étant F, le dernier étant G, et chaque élément de la chaîne étant un hyperplan du sous-espace vectoriel suivant, ou au contraire le sous-espace vectoriel suivant étant un hyperplan de ce sous-espace vectoriel ; formellement cela signifie qu'il existe $F_1, ..., F_p$ tels que $F = F_1$, $G = F_p$, et pour tout $i \in [1, p-1]$, F_i est un hyperplan de F_{i+1} ou F_{i+1} est un hyperplan de F_i.*
p est appelé **longueur de la chaîne**.

On procède par étapes :
- Si $F \subset G$, il y a entre F et G une chaîne de longueur $dim\ G - dim\ F$.

- Dans le cas général il y a entre F et G une chaîne de longueur $d(F, G)$ (facile en passant par l'espace $F \cap G$ - il suffit de se rappeler que $dim\ F + G = dim\ F + dim\ G - dim\ F \cap G$). On va maintenant se préoccuper de montrer que cette chaîne est de longueur minimale.

- Si on a une chaîne A, B, C, avec A et C hyperplans de B (c'est-à-dire que B est le plus grand de nos 3 éléments A, B et C), alors on a aussi une chaîne $A, A \cap C, C$, à moins que $A = C$.

- Si on a une chaîne, on peut la remplacer par une chaîne de même longueur entre les deux mêmes sous-espaces vectoriels et de manière à avoir des inclusions décroissantes puis croissantes.

- la longueur d'une chaîne entre F et G est au moins $d(F, G)$. On sait donc, avec le résultat obtenu plus haut, que $d(F, G)$ est la longueur minimale d'une chaîne entre F et G. On peut remarquer qu'on aurait pu raisonner de même en utilisant une chaîne croissante puis décroissante en passant par $F+G$ au lieu de décroissante puis croissante en passant par $F \cap G$.

- l'inégalité triangulaire en résulte aisément.$\square$

Proposition 288 *Pour la distance de Grassman, tout isomorphisme algébrique est une isométrie.*

Démonstration : Facile, puisque la distance de Grassman ne dépend que de dimension d'espaces, qui est préservé par isomorphisme algébrique.$\square$

5.6.15 Topologie et approximation de fonctions caractéristiques

On consultera la partie 9.1 (et les parties suivantes pour des applications).

5.6.16 Points fixes

Point fixe d'un endomorphisme dans un compact convexe

Lemme 289 *Soit K un compact convexe d'un espace vectoriel normé E, et f un endomorphisme continu de E tel que $f(K) \subset K$; alors il existe $x \in K$ tel que $f(x) = x$.*

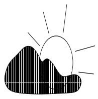
Démonstration :

- On se donne $x_0 \in K$, et on définit la suite (x_n) par $x_{n+1} = f(x_n)$.
- On définit alors $y_n = \frac{1}{n}\sum_{i=0}^{n-1} x_n$; $y_n \in K$ par convexité de K.
- Par compacité de K, on peut extraire une suite convergente de la suite des (y_n) (puisque E est un espace vectoriel normé , donc un espace métrique, le théorème de Bolzano-Weierstrass 199 s'applique).
- Soit y la limite de cette suite.
- $f(y_n) - y_n = \frac{x_n - x_0}{n}$; la suite (x_n) étant bornée (puisque $\underline{K \text{ est compact dans un métrique}}$) et en passant à la limite $\underline{\text{puisque } f \text{ est continue}}$ $f(y) = y$.$\square$

Un théorème de point fixe dans un espace de Hilbert

Ce résultat est directement inspiré de la note "Un théorème de point fixe en dimension finie : application aux sous-groupes compacts de $\mathbb{R}^n$", de Richard Antetomaso, dans la 104ème intégrale de la Revue de Maths Spé, 1993-1994.

Théorème 290 *On se donne H un espace de Hilbert, et K un compact convexe non vide de H, et un sous-groupe compact G de l'ensemble des automorphismes de H (qui sont aussi des homéomorphismes). On suppose que pour tout g dans G, $g(K) \subset K$. Alors il existe a appartenant à H point fixe commun, ie tel que $\forall g \in G, g(a) = a$.*



Il faut bien comprendre pour quelle topologie G est compact. Il s'agit de la topologie produit de H^H.

Démonstration :

• On définit sur H une norme $\|.\|'$ définie par

$$\|x\|' = \sup\{\|g(x)\|/g \in G\}$$

Cette norme est bien définie et à valeurs finies, car G est compact ; utiliser le corollaire 182 avec la fonction qui à g associe $\|g(x)\|$.

• Montrons qu'il s'agit bien d'une norme.

- $\|\lambda x\|' = \sup\{\|g(\lambda x)\|/g \in G\}$

$$= \sup\{\||\lambda| g(\lambda x)\|/g \in G\}$$

$$= |\lambda| \sup\{\|g(\lambda x)\|/g \in G\}$$

$$= |\lambda| \|g(x)\|'$$

- $\|x\|' = 0$ implique clairement $x = 0$.

- Enfin,

$$\|x + y\|'$$

$$= \sup\{\|g(x + y)\|/g \in G\}$$

$$= \|g_0(x + y)\|$$

$$\leq \|g_0(x)\| + \|g_0(y)\|$$

$$\leq \|x\|' + \|y\|'$$

car par hypothèse G est compact et une fonction continue sur un compact atteint ses bornes (voir corollaire 182).

Le cas d'égalité est atteint si $\|g_0(x)\| + \|g_0(y)\| = \|g_0(x + y)\|$, donc si $g_0(x)$ et $g_0(y)$ sont positivement liés (car H est un espace de Hilbert), donc si x et y sont positivement liés puisque g_0 est un automorphisme (les éléments de G sont des automorphismes). Cela signifie précisément que notre norme est strictement convexe.

• Supposons maintenant que $\forall x \exists g \in G/g(x) \neq x$.

• Considérons, pour $g \in G$, l'ensemble Ω_g des éléments de x de K tels que $g(x) \neq x$.

• $\forall g \Omega_g$ est ouvert, puisque g, l'identité, et l'addition sont continues (g est continue par hypothèse, l'identité est continue[9], et voir la proposition 163 pour vérifier que l'addition est bien continue).

[9] L'image réciproque de tout ouvert est bien un ouvert !

• Les Ω_g recouvrent K (c'est ce qu'on a supposé ci-dessus).
• Par définition des compacts, et puisque K est compact, on peut extraire un recouvrement fini $K = \cup_{i\in[1,n]}\Omega_{g_i}$. On note que $n \geq 1$, car K est non vide.
• On applique alors le lemme 289 à l'endomorphisme continu $x \mapsto \frac{1}{n}g_1(x)+\dots+g_n(x)$ (de K dans K, bien défini par convexité de K), continue par continuité des g_i. On en déduit qu'il existe x tel que $nx = \sum_i g_i(x)$.
• On a alors $n\|x\|' \leq \sum_i \|g_i(x)\| = n\|x\|$ (car les g_i sont des isométries et car les isomorphismes d'espaces de Hilbert sont des isométries.).
• On a montré plus haut que la norme était strictement convexe ; donc pour avoir le cas d'égalité ci-dessus il faut que les $g_i(x)$ soient positivement liés ; or ils ont tous même norme, puisque les g_i sont des isométries ; donc tous les $g_i(x)$ sont égaux.
• Du coup pour tout i dans $[1,n]$, $ng_i(x) = nx$; donc $g_i(x) = x$. D'où la contradiction ; x n'appartient pas à l'union des Ω_{g_i}, alors qu'il appartient à X, et que ces Ω_{g_i} recouvrent K.□

5.6.17 Cas particulier des espaces vectoriels normés de dimension finie

Voir 11.5.5.

5.6.18 $f : \mathbb{R} \to \mathbb{R}$ est C^∞ et $\forall x \exists n / f^{(n)}(x) = 0$, alors f est polynomiale

Proposition 291 *$f : \mathbb{R} \to \mathbb{R}$ est C^∞ et $\forall x \exists n / f^{(n)}(x) = 0$, alors f est polynomiale.*

Cet exercice est extrait du livre [5] ; nous avons tâché de préciser un peu plus les détails de la preuve.

Démonstration : • On considère Ω l'ensemble des points au voisinage desquels f est polynomiale
• Ω est ouvert, car si $x \in \Omega$, alors il existe V ouvert contenant x dans lequel tout point y admet un voisinage sur lequel f est polynomiale : il s'agit simplement de l'ouvert V.
• Soit $]u,v[$ inclus dans Ω, alors il existe un polynôme P tel que f est égale à P sur $[u,v]$.
Pour le prouver on considère x_0 dans $]u,v[$, un couple (x_1,x_2) tel que $f = P$ sur $]x_1,x_2[$ et $x_1 < x_0 < x_2$, et l'ensemble J des $x \in]x_0,v[$ tels que $f = P$ sur $]x_0,x[$; J est non vide car contenant x_2 ; il est fermé dans $]x_0,v[$ par continuité ; on montre facilement que si $x \in J$ alors $]x-\epsilon,x+\epsilon[\subset J$ pour un certain ϵ ; donc J est ouvert dans $]x_0,v[$; donc $J =]x_0,v[$ par connexité. On a donc $f = P$ sur $]x_0,v[$, et de même on aurait $f = P$ sur $]u,x_0[$.
• Soit F le complémentaire de Ω. F est fermé. Montrons qu'il ne comporte pas de point isolé. S'il comporte un point isolé a, on applique l'hypothèse et le développement de Taylor en a, et on en déduit une contradiction.
• On suppose F non vide pour arriver à une contradiction
• On définit F_n l'ensemble des $x \in F$ tels que $f^{(n)}(x) = 0$.
• On montre qu'il existe un intervalle ouvert non vide dont l'intersection avec F est incluse dans un certain F_{n_0}.
Les F_n sont fermés. Donc on applique le théorème de Baire 249 dans F ; il existe F_{n_0}

d'intérieur non vide. On peut alors choisir un intervalle I ouvert d'intersection non vide avec F, et cette intersection est incluse dans F_{n_0}.
• $I \cap F \subset F_n \ \forall n \geq n_0$
En effet soit $a \in I \cap F$; a n'est pas isolé et est donc limite d'une suite a_p d'éléments de F différents de a. Il suffit alors d'écrire la dérivée pour constater que $a \in F_{n_0+1}$.
Par récurrence $I \cap F \subset F_n$ pour $n \geq n_0$.
• On montre maintenant que f_{n_0} est nulle sur toute composante connexe de $I \cap \Omega$.
Tout d'abord $I \cap \Omega$ est non vide, sinon $I \subset F$, puis $I \cap F = I$.
$I \cap \Omega$ est ouvert ; donc c'est une réunion disjointe d'intervalles ouverts. Soit $]u, v[$ une telle composante connexe. $f = P$ sur $[u, v]$. $]u, v[\neq I$ sinon $F \cap I = \emptyset$
Donc $u \in I$ ou $v \in I$; supposons $u \in I$. Alors $u \in I \cap F$, et $u \in F_n$ pour tout $n \geq n_0$. Le degré de P est donc inférieur à n_0. Donc $f^{(n_0)} = 0$ sur $[u, v]$. Ca marche sur toutes les composantes connexes, donc $f^{(n_0)} = 0$ sur $I \cap \Omega$ mais aussi sur $I \cap F$. Donc $f^{(n_0)}$ est nulle sur I, donc f est un polynôme sur I donc $I \subset \Omega$, ce qui est impossible puisque $I \cap F \neq \emptyset$.□

5.6.19 Les billards strictement convexes

Proposition 292 *Soit K un ensemble strictement convexe de $\mathbb{R}^2$; on suppose que par tout point de la frontière de K il passe une unique tangente à K. On définit une trajectoire périodique de période n par la donnée de n points $a_0, \ldots, a_{n-1}$ de la frontière de K tels que pour tout $i \in [0, n-1]$ $\theta_i = \theta_{i+1}$ (voir figure 5.10) en notant $a_n = a_0$.*
Alors pour tout $n \geq 2$ il existe des trajectoires périodiques de période n.

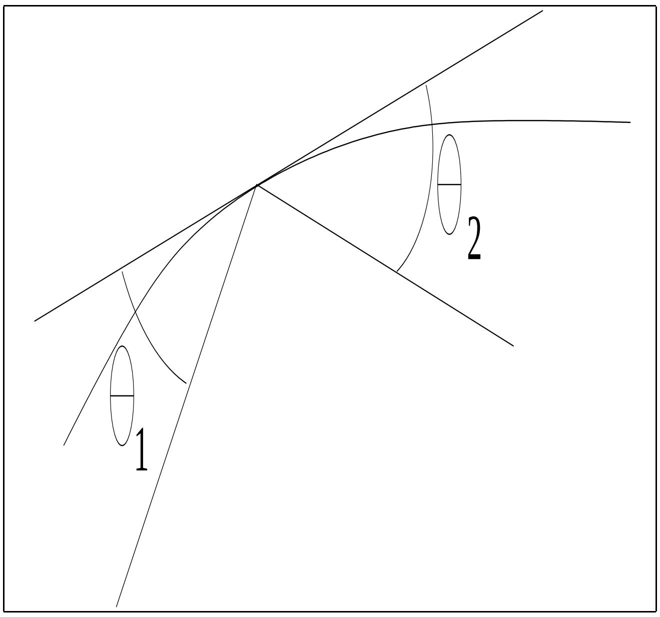


FIG. 5.10 – Illustration de la définition d'une trajectoire périodique

Démonstration :
• On se donne $n \geq 2$.
• Notons δK la frontière de K.
• On considère l'ensemble

$$\tilde{K} = \{(a_0, \ldots, a_{n-1}) / \forall i a_i \in \delta K\}$$

Il est égal à $(\delta K)^n$.
• On définit $f : \tilde{K} \to \mathbb{R}$ défini par

$$f(a) = \sum_{i=0}^{n-1} |a_i - a_{i-1}|$$

($|.|$ désigne ici la norme euclidienne)
• f est C^0 (par inégalité triangulaire)
• $\tilde{K}$ est compact (comme produit de compacts - s'agissant d'un produit fini d'espaces métriques il n'est pas nécessaire d'invoquer Tykhonov, voir le paragraphe qui suit le théorème 186).
• f atteint son maximum sur $\tilde{K}$.
• Tout point où f atteint son maximum vérifie la propriété énoncée, comme le lecteur le vérifiera aisément.$\square$

5.6.20 Deux jolies inégalités géométriques

On s'inspire ici de [4].

5.6.21 Inégalité isopérimétrique

Lemme 293 *On se donne Γ une fonction C^1 de $[0, 1]$ dans $\mathbb{C}$. On suppose $\int_0^1 \Gamma = 0$. Alors*

$$4\Pi^2 \int_0^1 |\Gamma|^2 \leq \int_0^1 |\Gamma'|^2$$

et il n'y a égalité que si Γ est une combinaison linéaire de $e^{2i\Pi x}$ et $e^{-2i\Pi x}$.

Démonstration : On procède comme suit :
• On considère Γ sur $[0, 1[$, et on considère sa série de Fourier.

$$\Gamma(x) = \sum_{n \in \mathbb{Z}} c_n e^{2i\Pi n x}$$

• On applique Parseval (théorème 1108) ;

$$\int_0^1 |\Gamma|^2 = \sum_n c_n^2$$

• On applique Parseval à la dérivée de Γ, $\Gamma'(x) = 2i\Pi \sum_n n c_n e^{2i\Pi n x}$

$$\int_0^1 |\Gamma'|^2 = 4\Pi^2 \sum_n (n c_n)^2$$

• On sait que $c_0 = 0$, car $\int_0^1 \Gamma = 0$.
• On a donc l'inégalité souhaitée, et le cas d'égalité.$\square$

Théorème 294 (Inégalité isopérimétrique) *La courbe C^1 fermée du plan qui à longueur donnée délimite une aire maximale est le cercle.*

Démonstration :
• On se donne une courbe fermée C^1 Γ de $[0,1]$ dans $\mathbb{C}$.
• Quitte à reparamétrer, on suppose que Γ' est constant.
• Quitte à translater Γ on suppose que $\int \Gamma = 0$.
• On applique alors le théorème de Green-Riemann, qui affirme que l'aire A est donnée par la formule

$$A = \frac{1}{2}\int xy' - x'y$$

avec $\Gamma = x + iy$, et x et y à valeurs réelles.
• Or

$$\int \Gamma\overline{\Gamma'} = \int (x+iy).(x'-iy')$$
$$= \int xx' + yy' + iyx' - iy'x$$

or $\int xx' = \int yy' = 0$ par périodicité donc

$$\int \Gamma\overline{\Gamma'} = i\int yx' - y'x$$

et donc

$$2A \leq \sqrt{\int |\Gamma|^2}\sqrt{\int |\Gamma'|^2}$$
$$\leq \frac{1}{2\Pi}\int |\Gamma'|^2$$

grâce au lemme précédent. Or Γ' étant constant, on obtient

$$A \leq \frac{l^2}{4\Pi}$$

avec l la longueur de l'arc.
• Il y a égalité si et seulement si l'inégalité de Cauchy-Schwartz est en fait une égalité, et donc si et seulement si on a non seulement tous les c_i nuls sauf c_1 et c_{-1}, mais aussi Γ et Γ' liés ; la famille $(x \mapsto e^{2i\Pi x}, x \mapsto e^{-2i\Pi x})$ étant libre, on constate que les solutions se trouvent pour un des deux coefficients c_1 et c_{-1} nul, c'est-à-dire que Γ est un cercle.$\square$

5.6.22 Inégalité isodiamétrique

On considère l'espace $\mathbb{R}^n$ muni de sa structure usuelle d'espace euclidien, et de la mesure de Lebesgue.

Théorème 295 *Quel que soit K compact de $\mathbb{R}^n$ de mesure finie, $\mu(K) \leq \mu(B(0, \delta(K)/2))$, avec $\delta(E)$ pour E une partie de $\mathbb{R}^n$ le diamètre de E, c'est à dire le $\sup$ des distances entre deux points de E.*

Cela revient à dire que le plus grand volume possible à diamètre donné est celui d'une boule.

Démonstration :

Nous aurons besoin de la définition suivante :

Définition 296 *Etant donné K un compact de $\mathbb{R}^n$, on appelle* **symétrisé de Steiner** *de K par rapport à l'hyperplan P l'ensemble*

$$S_P(K) = \{x = p + tu/p \in P \wedge D(p,u) \cap K \neq 0 \wedge |t| \leq \frac{1}{2}\mu'(K \cap D(p,u)\}$$

où u désigne un vecteur directeur unitaire de la droite orthogonale à P, et où $D(p,u)$ désigne la droite de vecteur unitaire u passant par p.
μ' désigne la mesure de Lebesgue sur la droite $D(p,u)$.

Lemme 297 (Première propriété fondamentale du symétrisé de Steiner) *Quel que soit K compact et P hyperplan, $S_P(K)$ a même mesure que K.*

Démonstration : Il suffit de considérer le théorème de Fubini, appliqué à la fonction caractéristique de K, pour voir que l'intégrale est l'intégrale sur $p \in P$ de la mesure $\mu'(K \cap D(p,u))$ (cette dernière quantité étant la mesure de $\{t \in \mathbb{R}/|t| \leq \frac{1}{2}\mu'(K \cap D(p,u))\} \subset \mathbb{R}$.$\square$

Lemme 298 (Deuxième propriété fondamentale du symétrisé de Steiner) *Quel que soit K compact et P hyperplan, $S_P(K)$ a un diamètre inférieur ou égal à celui de K.*

Démonstration : On considère deux points x et y de $S_P(K)$, à une distance d l'un de l'autre ; on cherche à montrer qu'il existe deux points x' et y' de K à distance supérieure ou égale à d.

- On note x_P et y_P les projetés orthogonaux de x et y sur P.
- On note l_x et l_y les mesures de $K \cap D(x,u)$ et $K \cap D(y,u)$.
- On note d' la distance entre x_P et y_P.
- d^2 est majorée par $d'^2 + (l_x/2)^2 + (l_y/2)^2$.
- On note L_x et L_y les <u>diamètres</u> de $S_P(K) \cap D(x,u)$ et $S_P(K) \cap D(y,u)$.

• Il est clair que $L_x \geq l_x$ et que $L_y \geq l_y$. Une étude de cas montre rapidement qu'en considérant les points extrémaux x_1, x_2, y_1 et y_2 de $S_P(K) \cap D(x,u)$ et $S_P(K) \cap D(y,u)$ respectivement, l'une des distances $d(x_i, y_i)$ est supérieure ou égale à $\sqrt{d'^2 + (L_x/2)^2 + (L_y/2)^2}$.$\square$

On a encore besoin d'un nouveau lemme :

Lemme 299 (Symétrisation d'un compact de $\mathbb{R}^n$) *On note $(e_1, ..., e_n)$ la base canonique de $\mathbb{R}^n$, et P_1, ..., P_n les hyperplans orthogonaux aux e_i passant par 0. On se donne K un compact de $\mathbb{R}^n$.*
Alors $S_{P_1} \circ S_{P_2} \circ S_{P_3} \circ \cdots \circ S_{P_n}(K)$ est stable par $x \mapsto -x$.

Démonstration :

Il suffit de procéder tranquillement, par récurrence ; $S_{P_n}(K)$ est clairement invariant par symétrie par rapport à P_n, $S_{P_{n-1}}$ est clairement invariant par symétrie par rapport à P_{n-1}, et par rapport à P_n aussi car la symétrisation de Steiner par rapport à un hyperplan P ne perturbe pas les symétries par rapport à des hyperplans orthogonaux à P (vérification immédiate sur la formule !), et ainsi de suite... L'invariance par rapport aux symétries par rapport aux n hyperplans donnent aussi l'invariance par $x \mapsto -x$.$\square$

On peut maintenant finir la preuve du théorème, en se donnant un compact K quelconque, en lui associant un compact K' égal à $S_{P_1} \circ S_{P_2} \circ S_{P_3} \circ \cdots \circ S_{P_n}(K)$, qui, par les lemmes précédents, est invariant par symétrie par rapport à l'origine, et donc est inclus dans la boule $B(0, \delta(K')/2)$. Il ne reste qu'à appliquer les différents lemmes pour conclure...$\square$

5.6.23 Triangulations d'un simplexe - Lemme de Sperner - conséquences

On s'inspire ici du livre [4], en tâchant de donner une preuve plus détaillée.

Définition 300 *On appelle* **simplexe** *de dimension n l'enveloppe convexe de $n+1$ points formant un repère affine.*
On appelle **face d'un simplexe** *l'enveloppe convexe d'un nombre fini (quelconque) de ses points. Sa* **dimension** *est par définition le nombre de points de cette face moins 1. On appelle g-***face** *une face de dimension g.*

Lemme 301 *Tout élément P appartenant à un simplexe Δ est décrit de manière unique par ses coordonnées barycentriques dans le repère des sommets de ce simplexe, si l'on impose que la somme des dites coordonnées est 1. Chaque coordonnée est ≥ 0.*

Démonstration : Voir la proposition-définition 1234.$\square$

On se donne pour la suite un simplexe Δ de $\mathbb{R}^n$ de sommets x_0, x_1, ... , x_n.

Tout point x de Δ est donc repéré par ses coordonées barycentriques $c_0(x)$, ... ,

$c_n(x)$, avec $\sum_{i=0}^n c_i(x) = 1$, et $\sum_{i=0}^n c_i(x).\overrightarrow{xx_i} = \overrightarrow{0}$.

Définition 302 *Soit $\sigma \in \sigma_{n+1}$ une permutation de $[0,n]$.*
On note Δ_σ l'ensemble des points x de Δ tels que

$$c_{\sigma(0)}(x) \geq c_{\sigma(1)}(x) \geq ... \geq c_{\sigma(n)}(x)$$

Proposition 303 • *Δ est l'union des Δ_σ.*
• *Pour tout $\sigma \in \sigma_{n+1}$, Δ_σ est un simplexe.*
• *Les intérieurs des Δ_σ sont disjoints.*

Démonstration : Le premier • ne mérite pas notre attention.

Pour le second • , c'est plus difficile, et nous allons donc détailler :

Un point x est dans Δ_I avec I la permutation identité, si ses coordonnées c_0, c_1,...,c_n vérifient $c_0 \geq c_1 \geq ... \geq c_n$. En écrivant

$$t_0 = c_0$$

$$t_1 = c_1 - c_0$$

$$t_2 = c_2 - c_1 - c_0$$

$$t_i = c_i - c_{i-1} - c_{i-2} - ... - c_0$$

$$t_n = c_n - c_{n-1} - c_{n-2} - ... - c_0$$

on voit que x est dans Δ_I si et seulement si il est dans l'enveloppe convexe des points de coordonnées barycentriques

$$(1, 0, 0, 0, ..., 0)$$

$$(1, 1, 0, 0, ..., 0)$$

$$(1, 1, 1, 0, ..., 0)$$

$$...$$

$$(1, 1, 1, 1, ..., 1)$$

Je n'ai pas normalisé pour ne pas alourdir la notation, sinon on obtient

$$(1, 0, 0, 0, ..., 0)$$

$$(\frac{1}{2}, \frac{1}{2}, 0, 0, ..., 0)$$

$$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, ..., 0)$$

$$...$$

$$(\frac{1}{n}, \frac{1}{n}, \frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n})$$

Donc il s'agit bien d'un simplexe. Il est non vide car les sommets définis ci-dessus forment bien un repère affine et donc l'intérieur est un voisinage de leur isobarycentre.

Il en va de même pour autre chose que l'identité ; Δ_σ est l'enveloppe convexe de n points comportant respectivement un 1 et seulement des 0 ailleurs, deux 1 et seulement des 0 ailleurs, trois 1 et seulement des 0 ailleurs, et ainsi de suite, chaque fois les 1 étant conservés, et un nouveau 1 étant ajouté.

Le troisième $\bullet$ est démontré dans le lemme qui suit.$\square$

Définition 304 *On appelle* **triangulation d'un simplexe** Δ *un ensemble fini de simplexes* Δ_i *tels que :*
$\bullet$ $\cup_i \Delta_i = \Delta$
$\bullet$ *Si* $i \neq j$, $Int(\Delta_j) \cap Int(\Delta_i) = \emptyset$
$\bullet$ *L'intersection d'une face de* Δ_i *et d'une face de* Δ_j *(pour* $i = j$ *ou* $i \neq j$*) est soit vide soit une face de* Δ_i *et une face de* Δ_j*.*

Lemme 305 *L'ensemble des* Δ_σ *pour* $\sigma \in \sigma_{n+1}$ *est une triangulation de* Δ*.*

Démonstration : $\bullet$ Il est bien clair que la réunion des Δ_σ est bien égale à Δ.

$\bullet$ L'intersection des intérieurs de Δ_σ et Δ_τ avec $(\sigma, \tau) \in \sigma_{n+1}^2$ est incluse dans l'intérieur des intersections, et donc incluse dans un hyperplan ; donc elle est vide.

$\bullet$ Regardons ce qu'est une face de simplexe, par exemple le simplexe Δ_I, avec I la permutation identité.

Il s'agit du barycentre d'un nombre fini de sommets, de la forme

$$(1/p, 1/p, ..., 1/p, 1/p, 0, 0, ..., 0, 0).$$

C'est à dire d'une somme pondérés de

$$(1, 0, 0, ..., 0)$$
$$(\frac{1}{i}, ..., \frac{1}{i}, 0, ..., 0)$$
$$(\frac{1}{i}, ..., \frac{1}{j}, 0, ..., 0)$$

On constate donc qu'une face est entièrement décrite par des équations du type $c_0(x)r_0c_1(x)r_1c_2(x)r_2...r_{n-1}c_n(x)r_n0$, avec r_i opérateur $=$ ou $\geq$.

Une intersection de faces va encore être du même type, car au plus elle va remplacer des $\geq$ par des $=$. D'où le résultat.$\square$

Lemme 306 *Dans la triangulation de* Δ *en les* Δ_σ *où* σ *appartient à* σ_{n+1}*, les* Δ_σ *ont un diamètre inférieur à* $\frac{n}{n+1}$ *fois le diamètre de* Δ*.*

Démonstration :

$\bullet$ le centre de gravité x (ou isobarycentre) de Δ appartient à tout simplexe Δ_σ (car tous les $c_i(x)$ sont égaux, égaux à $\frac{1}{n+1}$).

$\bullet$ la distance de x à un point de Δ est inférieur ou égale aux distances aux sommets de Δ, donc la distance d'un point de Δ_σ à x est toujours inférieure ou égale à $\frac{n}{n+1}$ fois la longueur de la médiane.

$\bullet$ le diamètre de Δ_σ, qui est la longueur maximal d'un de ses côtés, est donc majoré par la longueur max des côtés sur la face opposée à x, et par la longueur max des arêtes débouchant sur x... Dans tous les cas, cette longueur est majorée par celle d'une médiane de Δ ou d'une face du simplexe (une face de dimension quelconque, éventuellement une arête).$\square$

Lemme 307 *Pour tout ϵ on peut obtenir des triangulations de Δ en simplexes de diamètre inférieurs à ϵ.*

Démonstration :

On va utiliser par récurrence le lemme précédent. Les deux premiers $\bullet$ de la définition d'une triangulation sont faciles à obtenir, le problème est de montrer qu'une partition de chaque élément d'une partition donne encore une partition vérifiant le troisième point, c'est à dire le fait que l'intersection de deux faces de deux éléments distincts de la partition est soit vide soit égale à une face de chacun des deux éléments. Intuitivement, le problème est de voir que les faces se "recollent" bien.

Pour cela il suffit de voir que la triangulation faite selon le lemme précédent induit une triangulation de chacune des faces du dit simplexe - triangulation égale à celle qu'aurait donné le même lemme sur cette face. Cela se voit facilement en voyant qu'une face est une partie du simplexe où l'on annule une des composantes.$\square$

Définition 308 (Numérotation standard d'un simplexe) *Etant donnée une triangulation Δ_i d'un simplexe Δ, on note S l'ensemble des sommets des éléments de cette triangulation. On appelle* **numérotation standard** *d'une triangulation d'un simplexe de sommets $x_0, ..., x_n$ une application f de S dans $\{0, n\}$ telle que $f(x_i) = i$ et si pour tout s dans S $f(s) = i$ pour un certain i tel que x_i est un sommet de la face de Δ de dimension minimale contenant x.*

NB : la caractérisation "pour tout s dans S $f(s) = i$ pour un certain x_i tel que x_i est un sommet de la face de Δ de dimension minimale contenant x" inclue la condition $f(x_i) = i$, car une face peut très bien avoir une dimension 0.

On constate donc que dans une triangulation comme celles que l'on a construites plus haut, l'isobarycentre est autorisé à prendre n'importe quelle valeur puisque la seule face qui le contienne est Δ lui-même.

Proposition 309 *Une numérotation standard f d'une triangulation du simplexe Δ enveloppe convexe de $(x_0, ..., x_n)$ induit une numérotation standard de la triangulation induite sur le simplexe Δ' enveloppe convexe de $(x_0, ..., x_{n-1})$.*

Démonstration :

$\bullet$ Il est clair que tout sommet de la triangulation de Δ' a bien un numéro $< n$.

$\bullet$ Soit x sommet de la triangulation de Δ' appartenant à une face F minimale de Δ.

$\bullet$ F est forcément incluse dans Δ'.

$\bullet$ F est donc la même face que la face minimale de x dans Δ'.

• Donc la numérotation induite est bien standard.□

> **Lemme 310 (Lemme de Sperner)** *Toute triangulation d'un simplexe de dimension n munie d'une numérotation standard possède un élément numéroté $(0, ..., n)$.*

Démonstration :

• Soit Δ notre simplexe, supposé muni d'une numérotation standard f sur une triangulation T de Δ.

• On note $P(U)$ la propriété pour un simplexe U de dimension r d'avoir un sommet numéroté i et un seul pour tout i dans $[1, r]$.

• Soit E l'ensemble des simplexes de T ayant la propriété P.

• Soit F l'ensemble des simplexes de T qui ne sont pas dans E mais dont un numéro et un seul est numéroté i pour tout i dans $[0, n-1]$ (attention ils ne vérifient donc pas la propriété P).

• Soit G l'ensemble des $n-1$-faces de simplexes de T inclus dans Δ' et vérifiant la propriété P (rappelons qu'une face de simplexe est un simplexe).

• Soit H l'ensemble des $n-1$-faces de simplexes de T qui ne sont pas contenues dans Δ' et qui vérifient la propriété P.

• Chaque élément de E contient une et une seule $n-1$-face ayant la propriété P.

• Chaque élément de F contient exactement deux faces ayant la propriété P (facile, il y a exactement deux éléments numérotés pareil, donc on bâtit deux simplexes ayant la numérotation requise).

• Un élément de G est face d'un et d'un seul simplexe (car il est inclus dans Δ').

• Un élément de H est face d'exactement deux simplexes, car il n'est pas inclus dans Δ', et car il n'est pas non plus sur une autre face puisqu'il ne contient pas de sommet numéroté n.

(par simplicité dans la suite et pour alléger les notations je note I le cardinal $|I|$ d'un ensemble I)

• On en déduit $E + 2F = G + 2.H$ en comptant le nombre de $n-1$-faces ayant la propriété P, avec leurs multiplicités (c'est à dire en comptant deux fois les faces communes à deux simplexes).□

• En comptant modulo 2, on en déduit que E et G on la même parité.

• Il est clair que G est "le" E correspondant à Δ'.

• Par récurrence sur la dimension, on en déduit donc que G a toujours la même parité. Or dans le cas de la dimension 1, on constate facilement que ce nombre est impair ; on a une alternance de 0 et 1, 0 en premier, 1 à la fin, donc on a changé un nombre impair de fois de 0 à 1 ou de 1 à 0.

• On en déduit donc le résultat tant attendu ; E ne peut être nul puisqu'impair...□

> **Théorème 311 (Théorème de Brouwer)** *Soit f une application continue d'un simplexe Δ dans Δ. Alors f admet au moins un point fixe.*

Démonstration : • On suppose que f n'a pas de point fixe.

• Soit Δ_i, pour $i \in [0, n]$, l'ensemble des $x \in \Delta$ tels que $c_i(x) > c_i(f(x))$ (intuitivement f "éloigne" x du sommet i - attention, pas au sens de la distance, mais au sens du poids barycentrique du sommet i ; les points les plus "loin" étant les points de la face opposée, le point le plus proche étant le sommet lui-même).

• $\Delta = \cup_i \Delta_i$. En effet, soit $x \in \Delta$.

- Supposons $c_i(x) \leq c_i(f(x))$ pour tout i.
- $\sum c_i(x) = \sum c_i(f(x)) = 1$
- donc $c_i(x) = c_i(f(x))$ pour tout i
- alors on a f admettant un point fixe en x.
- c'est une contradiction, donc il existe i tel que $c_i(x) > c_i(f(x))$.

• x_i appartient à Δ_i (évident ; f ne peut qu'"éloigner" un point de lui-même, puisque f n'a pas de point fixe)

• x_i n'appartient pas à Δ_j si $j \neq i$ (non moins évident ; x est déjà "loin" autant que possible de x_j, puisqu'il appartient à la face opposée)

• Si x appartient à la face de Δ engendrée par les x_i pour $i \in I$ pour un certain sous-ensemble I de $[0, n]$, alors x n'appartient pas aux Δ_i si $i \notin I$ (toujours évident - si x appartient à la face engendrée par les x_i pour $i \in I$, il appartient à la face opposée à x_j pour tout $j \notin I$, et ne peut donc en être éloigné).

• Soit T une triangulation de Δ, ayant pour ensemble de sommets l'ensemble S. Soit g l'application qui à $s \in S$ associe $g(s)$ avec $s \in \Delta_{g(s)}$; on cherche à montrer qu'il s'agit d'une numérotation standard.

• g est bien définie, puisque l'on a montré que les Δ_i recouvraient Δ.

• le fait que $g(x_i) = i$ a déjà été prouvé ($x_i \in \Delta_i$, $x_i \notin \Delta_j$ quand $j \neq i$).

• soit $s \in S$, et F une face minimale de Δ contenant s. Il faut montrer que $g(s)$ est le numéro d'un sommet de F.

• si $F = \Delta$, le résultat est clair.

• si s appartient à une face stricte de Δ, alors s n'appartient pas aux Δ_j pour j ne désignant pas un numéro de sommet de F (prouvé un peu plus haut) ; donc $g(s)$ est forcément le numéro d'un sommet de F.

• g est donc bien une numérotation.

• On a montré qu'on pouvait construire des triangulations aussi fines que l'on voulait, au sens où chaque simplexe de la triangulation pouvait être imposé de diamètre inférieur à $1/n$. On se donne T_n une telle triangulation, avec g_n la numérotation correspondante, donnée par les questions précédentes.

• D'après le lemme de Sperner, il existe un élément de la triangulation T qui a la propriété P évoquée plus tôt, c'est à dire qu'il comporte un sommet numéroté i pour tout i dans $[0, n]$.

• On peut considérer alors la suite de sommets numérotés 0 dans des simplexes ayant la propriété P de la triangulation T_n.

• Puisque l'on est dans un compact métrique, on peut en extraire une sous-suite convergente, par le théorème de Bolzano-Weierstrass (voir théorème 199). Soit x la limite.

• x est aussi la limite des suites de sommets numérotés i dans des simplexes ayant la propriété P, pour $i \in [1, n]$, car le diamètre des simplexes tend vers 0.

• Par continuité de f, $c_i(f(x)) \geq c_i(x)$ pour tout i.

• Or $\sum_i c_i(x) = \sum_i c_i(f(x)) = 1$, donc $c_i(f(x)) = c_i(x)$ pour tout i.

• On en déduit alors que $x = f(x)$. D'où la contradiction...□

Corollaire 312 (Brouwer) *Toute application continue d'une boule unité fermée de $\mathbb{R}^n$ dans elle-même comporte un point fixe.*

<u>Démonstration :</u>

En fait il suffit de montrer que la boule unité fermée est homéomorphe à un simplexe. Cela est fait dans la proposition 263.□

Corollaire 313 (Champ rentrant dans la sphère) *Soit V un champ de vecteur continu de la boule unité fermée de $\mathbb{R}^n$ dans $\mathbb{R}^n$. Si pour tout x, $\tilde{V}(x) =< V(x), x >$ est négatif strictement [a], alors V s'annule au moins en un point de la boule unité.*

[a] On dit que le champ est rentrant.

Démonstration :

• On note S la sphère unité fermée, et C_r la couronne constituée par la boule unité fermée privée de la boule ouverte de rayon $1 - r$.

• On considère pour tout ϵ l'application f_ϵ de $\overline{B}$ (la boule unité fermée) dans $\mathbb{R}^n$ qui à x associe $x + \epsilon.\overrightarrow{V}(x)$.

• f_ϵ est continu.

• $\tilde{V}$ étant continue sur un compact $\overline{B}$, on peut lui trouver un majorant à $\|V\|$. Notons M ce maximum.

• V étant continue sur un compact S, elle y atteint son maximum, qui est négatif. Notons m ce maximum ; on a $m < 0$.

• En tout point x de S, on peut centrer une boule ouverte de rayon r_x sur laquelle $< x, V(x) >$ est inférieur à $m/2$. La sphère S est recouverte par les boules centrées en x de rayon $r_x/2$; on peut donc extraire de ce recouvrement un recouvrement fini. Les boules de rayon r_x recouvrent elle aussi S, et elles recouvrent aussi une couronne C_r pour un r assez petit.

• Sur C_r, on a donc $< x, V(x) >$ inférieur à $m/2$.

• $\|f_\epsilon(x)\|^2 = \|x\|^2 + \epsilon^2 \|V(x)\|^2 + 2\epsilon < V(x), x >$.

• Donc $\|f_\epsilon(x)\|^2 \leq \|x\|^2 + \epsilon^2.M^2 + 2\epsilon < V(x), x >$

• Sur C_r, on a alors $\|f_\epsilon(x)\|^2 \leq \|x\|^2 + \epsilon^2.M^2 + \epsilon.m$

• Pour ϵ suffisamment petit, on a donc $\|f_\epsilon(x)\|^2 < \|x\|^2$

• Pour ϵ suffisamment petit et $\|x\| < 1 - r$ on a aussi $\|f(x)\| \leq 1$ (puisque V est borné).

• On déduit de tout cela que f_ϵ pour ϵ assez petit est une application de la boule unité fermée dans la boule unité fermée. Par le résultat 312, on en déduit donc que f_ϵ admet un point fixe, et que donc $\tilde{V}$ s'annule quelque part.□

Chapitre 6

Intégration

6.1 σ-algèbre, mesure

6.1.1 Définitions, généralités

Soit X un ensemble, $\mathcal{A} \in \mathcal{P}(x)$.

> **Définition 314 (Algèbre)** *$\mathcal{A}$ est une* **algèbre** *(on dit aussi parfois* **clan***) si elle vérifie :*
> • $X \in \mathcal{A}$
> • *Stabilité par union finie*
> • *Stabilité par passage au complémentaire*

Propriétés :
• $\emptyset \in \mathcal{A}$
• Stabilité par intersection finie
• Stabilité par différence
• Une intersection quelconque d'algèbres est une algèbre.

En fait une algèbre est stable par n'importe quelle suite finie d'opérations sur les ensembles.

Exemples :
a) $\{\emptyset, X\}$
b) $\mathcal{P}(X)$
c) $\{A \subset X | A$ fini ou cofini $\}$
d) $\{A \subset X | A$ ou A^c au plus dénombrable$\}$
e) $X, E \subset X$; $\{A/A \subset E \vee X - E \subset A\}$
f) $X, E \subset X$; $\{E \subset A \vee A \cap E = \emptyset\}$

Définition 315 (σ-algèbre) *$\mathcal{A}$ est une σ-**algèbre (ou tribu)** si elle vérifie :*
• $X \in \mathcal{A}$
• *stabilité par union dénombrable*
• *stabilité par passage au complémentaire*

Propriété :
• $\emptyset \in \mathcal{A}$
• stabilité par intersection dénombrable.
• Une σ-algèbre est une algèbre.
• stabilité par différence.
• L'image réciproque d'une tribu par une application est une tribu.
• Etant donnée une tribu $\mathcal{A}$ sur X, et une application f de X dans Y, alors l'ensemble des $U \subset Y$ tels que $f^{-1}(U)$ appartient à $\mathcal{A}$ est une tribu sur Y.
• Une intersection quelconque de tribus est une tribu.

En fait une σ-algèbre est stable par n'importe quelle suite dénombrable d'opérations sur les ensembles.

Notez que l'ensemble des parties mesurables au sens de Riemann d'un segment $[a, b]$ au sens de Riemann est un clan, mais pas une tribu ; puisque certaines parties dénombrables ne sont pas mesurables.

Définition 316 (Espace mesurable) *$(X, \mathcal{A})$ est un **espace mesurable** si $\mathcal{A}$ est une tribu sur X*
*Une partie de X est dite $\mathcal{A}$-**mesurable** si elle appartient à $\mathcal{A}$.*

Dans les exemples plus haut, tous sont des σ-algèbres, sauf c), à moins que X soit fini.

6.1.2 σ-algèbre engendrée

Afin de pouvoir définir la notion de σ-algèbre engendrée, nous avons besoin d'un petit lemme (évident !) :

Lemme 317 *X ensemble, (A_α) famille d'algèbres (resp. de σ algèbres), alors $\cap A_\alpha$ est une algèbre (resp. une σ-algèbre.*

Définition 318 *X un ensemble, $\mathcal{M} \subset \mathcal{P}(X)$ famille de parties de X, l'***algèbre engendrée par** *$\mathcal{M}$ (resp. la σ-***algèbre engendrée par** *$\mathcal{M}$) est l'intersection de toute les algèbres (resp. σ-algèbre) contenant $\mathcal{M}$.*
*X un ensemble, la σ-***algèbre engendrée par une famille de fonctions** *de X vers des espaces mesurables est la σ-algèbre engendrée par les images réciproques d'ensembles mesurables par ces fonctions.*

C'est bien une algèbre (resp. σ-algèbre) et c'est la plus petite qui contienne $\mathcal{M}$.

la σ-algèbre engendré par l'image réciproque d'une famille $\mathcal{F}$ de sous-ensembles de F par une application $f : E \to F$ est l'image réciproque de la σ-algèbre engendrée par $\mathcal{F}$.

Exemple fondamental :

Définition 319 *X muni d'une topologie $\mathcal{T}$; la σ-algèbre engendrée par $\mathcal{T}$ s'appelle la σ-***algèbre borélienne***. Ses éléments sont appelés les* **boréliens***.*

Dans $\mathbb{R}$ ou $\mathbb{R}^n$, c'est la plus petite σ-algèbre contenant les boules ouvertes. C'est aussi la plus petite σ-algèbre contenant les boules fermées.

Proposition 320 *La σ-algèbre engendrée par une base d'ouverts est la σ-algèbre engendrée par la topologie ; donc lorsqu'un ensemble engendre une topologie, il engendre aussi les boréliens.*
Les ensembles suivants engendrent les boréliens de $\mathbb{R}$:
• *les ouverts*
• *les fermés*
• *les intervalles ouverts*
• *les intervalles fermés*
• *les intervalles fermés bornés*
• *Les intervalles ouverts bornés*
• *Les $[a, b[$ avec $(a, b) \in \mathbb{R}^2$*
• *Les $]a, +\infty[$*
• *Les $[a, +\infty[$*

Les ensembles suivants engendrent les boréliens de $\overline{\mathbb{R}}$:
• *Les $[a, +\infty]$*
• *Les $]a, +\infty]$*
• *Les $[-\infty, a[$*
• *Les $[-\infty, a]$*

Les ensembles suivants engendrent les boréliens de $\mathbb{R}^n$:
• *Les ouverts*
• *Les pavés ouverts*
• *Les boules ouvertes*
• *Les bandes ouvertes :*
$\{\mathbb{R}^{i-1} \times]a, b[\times \mathbb{R}^{n-i}\}$
• *Les pavés compacts*
• *Les bandes fermées :*
$\{\mathbb{R}^{i-1} \times [a, b] \times \mathbb{R}^{n-i}\}$
• *Les bandes comme suit :*
$\{\mathbb{R}^{i-1} \times]a, b] \times \mathbb{R}^{n-i}\}$
• *Ou les ensembles de la forme suivante :*
$\{\mathbb{R}^{i-1} \times]a, +\infty[\times \mathbb{R}^{n-i}\}$

Disposer de parties génératrices petites est pratique pour certaines propriétés des boréliens.

Exemple : on munit $\overline{\mathbb{R}}$ d'une distance comme suit :
$d(x, y) = |arctan x - arctan y|$ avec $arctan(\infty) = \Pi/2$ et $arctan(-\infty) = -\Pi/2$.
Les boréliens pour cette distance sont engendrés par les $\{]a, +\infty], a \in \mathbb{R}]$. Toute suite monotone dans $\overline{\mathbb{R}}$ admet une limite, et toute suite dans $\overline{\mathbb{R}}$ admet une valeur d'adhérence.

6.1.3 Mesures

Définition 321 (Mesure) *Etant donné $(X, \mathcal{A})$ mesurable, on appelle* **mesure** *une application $\mu : \mathcal{A} \to [0, +\infty]$ telle que :*
• $\mu(\emptyset) = 0$
• *Si les A_i sont disjoints, au plus dénombrables, alors $\mu(\cup A_i) = \sum \mu(A_i)$ (σ* **additivité***)*

$(X, \mathcal{A}, \mu)$ est appelé **espace mesuré***.*
Etant donné $(X, \mathcal{A})$ mesurable, on appelle **mesure complexe** *une application $\mu : \mathcal{A} \to \mathbb{C}$ telle que :*
• $\mu(\emptyset) = 0$
• *Si les A_i sont disjoints, au plus dénombrables, alors $\mu(\cup A_i) = \sum \mu(A_i)$ (σ* **additivité***) quel que soit l'ordre de la sommation - c'est à dire que la somme est absolument convergente.*

On notera bien qu'une mesure peut prendre $+\infty$ pour valeur, et pas une mesure complexe. Une mesure complexe n'est pas un cas particulier de mesure, et une mesure n'est pas un cas particulier de mesure complexe.

On note aussi que dans le cas des mesures complexes, le deuxième • de la définition suffit à imposer le premier (pas dans le cas réel, à cause de la possibilité $+\infty$).

Définition 322 (Propriété vraie presque partout) *Une propriété P est dite vraie* **presque partout** *si l'ensemble des éléments pour lesquels elle est fausse est inclus dans un ensemble de mesure nulle. Une partie est dite* **négligeable** *si elle est incluse dans une partie de mesure nulle, c'est à dire si sa fonction caractéristique est nulle presque partout.*
Un espace mesuré est dit **complet** *si tout ensemble négligeable est mesurable (et donc de mesure 0).*

⚠ Un ensemble négligeable n'est pas nécéssairement mesurable !

Une réunion dénombrable d'ensembles négligeables est négligeable.

Proposition 323 • *μ est finiment additive (outre qu'elle est σ-additive)*
• *μ est croissante ($\mathbb{R}$ muni de l'ordre usuel, les ensembles mesurables munis de l'inclusion)*
• *La mesure d'une union dénombrable est inférieure ou égale à la somme des mesures.*
• *Avec A et B $\mathcal{A}$-mesurables, $\mu(A \cup B) \leq \mu(A) + \mu(B)$*
• *$\mu(U_{i \leq n} F_i) \leq \sum_{1 \leq i \leq n} \mu(F_i)$, avec F_i $\mathcal{A}$-mesurable.*
• *Si $\mu(X)$ est finie, alors $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$*
• **Formule d'inclusion exclusion** *; avec $F_i \in \mathcal{A}$, on a $\mu(\cup_{i \leq n} F_i) = \sum_{i \leq n} \mu(F_i) - \sum_{1 \leq i < j \leq n} \mu(F_i \cap F_j) + \sum_{1 \leq i < j < k \leq n} \mu(F_i \cap F_j \cap F_k) ... + (-1)^{n-1} \mu(\cap_{1 \leq i \leq n} F_i)$*

Exemples :

Définition 324 • *Considérons l'espace mesurable* $(X, \mathcal{P}(X))$.
On considère sur X *la* **mesure de dénombrement** *définie pourtout* $A \subset X$ *par*
$\mu(\mathcal{A}) = card(A)$ *si* A *fini*
$\mu(\mathcal{A}) = +\infty$ *sinon*
• $(X, \mathcal{P}(X))$, *soit* $a \in X$, *on appelle* **mesure de dirac en** a *la fonction* δ_a *telle que pour toute partie* A *inclue dans* X, $\delta_a(A) = 1$ *si* $a \in A$ *et* 0 *sinon.*

Théorème 325 *Toute espace mesuré* $(X, \mathcal{A}, \mu)$ *peut être remplacé par un espace mesuré* $(X, \overline{\mathcal{A}}, \overline{\mu})$ *complet, avec* $\mathcal{A} \subset \overline{\mathcal{A}}$ *et* $\overline{\mu}_{|\mathcal{A}} = \mu$. *On peut même garantir que tout ensemble* C *contenant* A *et inclus dans* B *avec* $(A, B) \in \mathcal{A}^2$ *et* $B - A$ *négligeable appartient à* $\overline{\mathcal{A}}$.

Démonstration : On considère l'ensemble des parties C décrites dans le théorème ; on vérifie facilement qu'il s'agit bien d'une σ-algèbre . Ensuite on montre que l'on peut définir la mesure de C comme égale à la mesure de A, et que la définition est bien correcte. $\square$

Définition 326 (Lebesguiens) *La tribu obtenue à partir de la tribu des boréliens en appliquant le théorème 325 s'appelle* **tribu des lebesguiens**. *Les éléments de cette tribu sont appelés les* **lebesguiens**.

Donc lorsque l'on travaille avec des boréliens, certains ensembles négligeables ne sont pas mesurables, alors qu'avec les lebesguiens, tous les ensembles négligeables sont mesurables.

Théorème 327 (Théorème fondamental) *Il existe une unique mesure sur* $\mathbb{R}$ *muni des boréliens classique telle que* $\mu([a, b]) = b - a$ *pour* $b > a$. μ *s'appelle* **mesure de Lebesgue sur** $\mathbb{R}$.
Il existe une unique mesure sur $\mathbb{R}^n$ *muni des boréliens classiques telle que* $\mu(\Pi_i[a_i, b_i]) = \Pi_i(b_i - a_i)$ *pour* $b_i > a_i$. μ *s'appelle* **mesure de Lebesgue sur** $\mathbb{R}^n$.
La mesure de Lebesgue vérifie en outre les propriétés suivantes :
• *à une constante de proportionnalité près, c'est la seule mesure sur les boréliens invariante par translations et finie sur les intervalles bornés.*
• *Tout ensemble au plus dénombrable est de mesure nulle.*
• *Etant donnée* E *une partie mesurable, la mesure de* E *est égale à l'*inf *des mesures des parties ouvertes contenant* E.
• *Etant donnée* E *une partie mesurable, la mesure de* E *est égale au* sup *des mesures des parties compacts inclues dans* E.

Démonstration : Admise.$\square$

Proposition 328 $(X, \mathcal{A}, \mu)$ *espace mesuré,* $\forall i \in \mathbb{N}\ A_i \in \mathcal{A}$.
i) Si $A_i \subset A_{i+1}$*, alors* $\mu(\cup_{i\in\mathbb{N}} A_i) = lim_{i\to+\infty}\mu(A_i)$
ii) Si $A_{i+1} \subset A_i$ *et si* $\mu(A_0) < +\infty$ *alors* $\mu(\cap A_i) = lim_{i\to+\infty}\mu(A_i)$

NB : ne pas oublier la seconde condition pour la deuxième assertion ; contre-exemple avec $A_i = [i, +\infty[$.

Démonstration :
i) $B_0 = A_0$, $B_i = A_i - A_{i-1}$, suite facile.
ii) $\mu(D - C) = \mu(D) - \mu(C)$ si $C \subset D$ et $\mu(D) < +\infty$
$B_i = A_{i-1} - A_i$
$\cap A_i = A_0 - \cup B_k$
$\mu(\cap A_i) = \mu(A_0) - \mu(\cup B_k) = \mu(A_0) - \sum[\mu(A_{k-1}) - \mu(A_k)]$
$\square$

Définition 329 (mesure finie ou σ -finie) $(X, \mathcal{A}, \mu)$ *mesuré,* μ *est* **finie** *si* $\mu(X) < +\infty$.
$(X, \mathcal{A}, \mu)$ *mesuré,* μ *est* σ**-finie** *si*
$\exists (X_k \in \mathcal{A}) / \cup_k X_k = X \wedge \mu(X_k) < \infty$
Si $\mu(X) = 1$ *alors* μ *est appelée une* **mesure de probabilité**.

Donc $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ avec μ la mesure de Lebesgue est σ-finie car $\mathbb{R} = \cup_{k\in\mathbb{N}}] -k, k[$

6.2 Π-systèmes, d-systèmes, et théorème de Carathéodory

Définition 330 (Π-systèmes) *Un* **Π-système** *sur X est un sous-ensemble de $P(X)$ stable par intersections finies.*

Définition 331 (d-système, alias classe monotone) *D est un* **d-système** *(on dit aussi une* **classe monotone** *si*
• $S \in D$
• *D est stable par soustraction ($A \in D$, $B \in D$, alors $A \cap B^c \in D$).*
• *Pour toute suite A_n croissante, $A_n \in D$, alors $\cup A_n \in D$*

Proposition 332 *Une intersection de d-systèmes est un d-système.*

Définition 333 (d-système engendré) *On appelle* **d-système engendré** *par un ensemble de parties de X l'intersection de tous les d-systèmes contenant X.*

Proposition 334 *Un ensemble inclus dans $P(X)$ est une σ-algèbre si et seulement si c'est un Π-système et un d-système.*

Lemme 335 (Lemme de Dynkin) *Soit I un Π-système, alors la σ-algèbre engendrée par I, notée $\sigma(I)$, est égale au d-système engendré par I.*

Démonstration : Pour le prouver il suffit de montrer que $d(I)$ est un Π système, vu la proposition 334. Pour cela on montre tout d'abord que le sous-ensemble D_1 de $d(I)$ constitué des éléments de $d(I)$ dont l'intersection avec tout élément de I appartient à $d(I)$, est égal à $d(I)$. Le raisonnement est le suivant :
• $d(I) \subset D_1$ car :
- $I \subset D_1$
- D_1 est un Π-système
• $D_1 \subset d(I)$ trivialement

On montre ensuite que le sous-ensemble D_2 de $d(I)$ constitué des éléments de $d(I)$ dont l'intersection avec tout élément de $d(I)$ appartient à $d(I)$, est égal à $d(I)$; en effet : • $d(I) \subset D_2$ car
- $I \subset D_2$ car $D_1 = d(I)$
- D_2 est un d-système
• $D_2 \subset d(I)$ trivialement

Or $d(I) = D_2$ est exactement l'énoncé du fait que $d(I)$ est un Π-système.$\square$

Lemme 336 • *Soit μ_1 et μ_2 deux mesures sur $(X, \mathcal{A})$ coïncidant sur un Π-système engendrant $\mathcal{A}$ et telles que $\mu_1(X) < +\infty$ et $\mu_2(X) < +\infty$. Alors μ_1 et μ_2 sont égales.*
• *La même propriété est vraie si X est de mesure σ-finie pour μ_1 et μ_2.*

Démonstration : On considère le d-système des F tels que $\mu_1(F) = \mu_2(F)$. Il contient un Π-système, donc il ne reste qu'à conclure via le lemme 335.□

Théorème 337 (Théorème de Carathéodory) *Soit $\mathcal{A}$ la σ-algèbre engendrée par $\mathcal{B}$ une algèbre, et μ σ-additive de $\mathcal{B}$ dans $[0, +\infty]$. alors il existe une mesure μ' sur $\mathcal{A}$ dont la restriction à $\mathcal{B}$ est μ. Si $\mu(X) < +\infty$, alors cette extension est unique.*

Démonstration : L'existence est ici admise.
L'unicité résulte simplement de 336.□

Proposition 338 *Soit μ une mesure sur un espace mesurable X, et f une fonction mesurable de X dans Y un autre espace mesurable ; alors l'application qui à une partie mesurable E de Y associe $\mu(f^{-1}(E))$ est une mesure sur Y. On note μ^f cette mesure.*

Démonstration : Facile.□
• Si $Y = \mathbb{R}^n$ muni des boréliens, alors si μ est positive on a

$$\int_{\mathbb{R}^d} f d\mu^f = \int_X \phi \circ f \, d\mu$$

• Si $Y = \mathbb{R}^n$ et si f est à valeurs quelconques, alors ϕ est L_1 pour μ^f si et seulement si $\phi \circ f$ est L_1 pour μ, et on a alors l'égalité

$$\int_{\mathbb{R}^d} f d\mu^f = \int_X \phi \circ f \, d\mu$$

6.3 Parties non mesurables

Théorème 339 (Banach et Tarski) • *Il existe un ensemble F inclus dans la sphère unité S_2 de $\mathbb{R}^3$ tel que pour tout $k \geq 3$ (éventuellement $k = \infty$), S_2 est la réunion disjointe de k images de F par des rotations. Façon amusante de constater qu'on ne peut pas mesurer n'importe quoi... Cette preuve nécessite l'axiome du choix.*
• *On peut même aller plus loin et étant donnés A et B d'intérieurs non vides de $\mathbb{R}^3$, on peut décomposer A en une réunion de A_i finie, et B en une réunion de B_i de même cardinal, avec A_i et B_i égaux via une rotation et une translation.*

• Si l'on n'utilise pas l'axiome du choix, alors on peut utiliser à sa place un autre axiome, qui affirme que toute partie de $\mathbb{R}$ est mesurable.

6.4 Exercices sur les σ-algèbre et les mesures

• Soit $\mathcal{A} \subset P(E)$.
- Supposons $E \in \mathcal{A}$ et $\mathcal{A}$ stable par soustraction. Montrer que $\mathcal{A}$ est une algèbre.

Démonstration : Il suffit de vérifier que E et $\emptyset$ sont dans $\mathcal{A}$, que l'on a bien stabilité par passage au complémentaire, et stabilité par union de deux éléments car $E - (E - A - B) = A \cup B$.□
- Supposons $E \in \mathcal{A}$, $\mathcal{A}$ stable par passage au complémentaire, et stable par union disjointe. Montrer que $\mathcal{A}$ n'est pas nécessairement une σ-algèbre .

□Considérer l'ensemble des parties de cardinal pair de $\{a, b, c, d\}$.□

• Montrer que la réunion d'une suite croissante d'algèbres est une algèbre.

• Montrer que la réunion d'une suite croissante de σ-algèbres n'est pas nécessairement une σ-algèbre .

Démonstration : Considérer les σ-algèbre sur $\mathbb{N}$ engendrées respectivement par $\{\{1\}\}, \{\{1\}, \{2\}\}, \{\{1\}, \{2\}, \{3\}\}$; la réunion n'est pas une σ-algèbre car par exemple l'ensemble des entiers pairs n'en fait pas partie.□

6.5 Fonctions mesurables

Définition 340 (Fonction mesurable) *Etant donnés $(X, \mathcal{A})$ et $(X, \mathcal{B})$ des espaces mesurables, $f : X \to Y$ est dite* **mesurable** *si $\forall B \in \mathcal{B}\ f^{-1}(B) \in \mathcal{A}$*
On définit parfois aussi la notion de fonction mesurable d'un espace mesurable vers un espace topologique ; la condition est alors le fait que l'image réciproque d'un ouvert soit une partie mesurable.

Une fonction caractéristique d'un ensemble est mesurable si et seulement si l'ensemble est mesurable.

Si Y est topologique et qu'on n'a rien précisé, la tribu $\mathcal{B}$ est celle des boréliens.

Proposition 341 *Avec les conditions de la définitions, si $\mathcal{B}$ est la σ-algèbre engendrée par $\mathcal{M}$, alors f est mesurable si et seulement $\forall \mathcal{B} \in \mathcal{M}\ f^{-1}(B) \in \mathcal{A}$*

Démonstration : Le sens $\rightarrow$ est trivial.
Pour le sens $\leftarrow$, considérons l'ensemble des $B \in \mathcal{B}$ tels que $f^{-1}(B) \in \mathcal{A}$, c'est une σ-algèbre de Y ; d'où le résultat.□

Corollaire 342 *• Si $f : (X, \mathcal{A}) \to Y$ avec Y topologique, f mesurable si et seulement si $\forall U$ ouvert , $f^{-1}(U) \in \mathcal{A}$*
• $f_k : (X, \mathcal{A}) \to \overline{\mathbb{R}}$ ou $[0, +\infty]$ mesurable
alors la fonction f qui à x associe $sup_k f_k(x)$ est mesurable.

Démonstration : Le premier point est clair.
Pour le second on considère $f^{-1}(]a, +\infty]) = \{x/sup\{f_k(x)\} > a\}$
$= \{x/\exists k \in \mathbb{N} f_k(x) > a\} = \cup_{k \in \mathbb{N}} \{x/f_k(x) > a\}$
$= \cup f_k^{-1}(]a, +\infty[)$ qui appartient à $\mathcal{A}$ car f_k mesurable et $\mathcal{A}$ σ-algèbre .□

Proposition 343 *La composition de deux fonctions mesurables est mesurable. f mesurable et g continue alors $g \circ f$ est mesurable.*

Corollaire 344 *Si f est mesurable de X dans $\mathbb{C}$ ou $\mathbb{R}$, alors $Re(f)$, $Im(f)$, $|f|$ sont mesurables.*

Proposition 345 *f mesurable de X dans Y avec Y topologique, g mesurable de X dans Z avec Z topologique, Y et Z à bases dénombrables d'ouverts. alors $(f, g) : X \to Y \times Z$ est mesurable pour la topologie produit.*

Corollaire 346 *• f et g mesurables de $(X, \mathcal{A})$ dans $\mathbb{R}$ ou $\mathbb{C}$, alors $f+g$, $f-g$, fg, f/g (g ne s'annulant pas), sont mesurables.*
• f de $(X, \mathcal{A})$ dans $\mathbb{C}$ est mesurable si et seulement si sa partie réelle et sa partie imaginaire sont mesurables.
• f de $(X, \mathcal{A})$ dans $\mathbb{C}$ est mesurable si et seulement si la fonction $|f|$ et la fonction $x \mapsto arg(f(x))$ sont mesurables
• f de $(X, \mathcal{A})$ dans $\mathbb{R}$ ou $\mathbb{R}$ est mesurable si et seulement si f^+ et f^- sont mesurables.

6.6 Suites de fonctions mesurables

Théorème 347 *$(X, \mathcal{A})$ mesurable, f_k de X quelconque dans Y métrique, telle que pour tout x $f_k(x) \to f(x)$. Alors f est mesurable.*

Démonstration : Soit U un ouvert de Y, et U_n l'ensemble des u tels que $d(x, Y \setminus U) > 1/n$, et soit F_n l'ensemble des u tels que $d(x, Y \setminus U) \geq 1/n$.
On montre facilement que U_n est ouvert et que F_n est fermé.
L'union des U_n est U car U est ouvert.

L'union des F_n est donc U aussi. (voir figure 6.1)

$$f^{-1}(U)$$
$$= \cup_n f^{-1}(U_n)$$
$$= \cup_n \{u/lim_{k\to+\infty} f_k(u) \in U_n\}$$
$$\subset \cup_n \{u/\exists K/k \geq K \to f_k(u) \in U_n\}$$

car U_n est un ouvert.

$$\subset \cup_n \cup_K \cap_{k\geq K} f_k^{-1}(U_n) \in \mathcal{A}$$

Considérons maintenant

$$\cup_n \cup_K \cap_{k\geq K} f_k^{-1}(U_n)$$
$$\subset \cup_n \cup_K \cap_{k\geq K} f_k^{-1}(F_n)$$

or $u \in \cup_K \cap_{k\geq K} f_k^{-1}(F_n)$ si et seulement si

$$\exists K/\forall k \geq K, f_k(u) \in F_n$$

donc

$$\subset \cup_n f^{-1}(F_n) = f^{-1}(\cup F_n) = f^{-1}(F_n) = f^{-1}(U)$$

donc $f^{-1}(U) \in \mathcal{A}$, donc f est mesurable.□

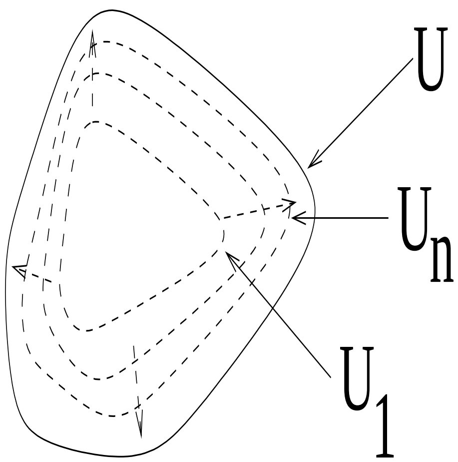


FIG. 6.1 – Illustration de la preuve de la mesurabilité de la limite simple d'une suite de fonctions mesurables.

Corollaire 348 *f_k suite de fonctions mesurables de $(X, \mathcal{A})$ dans $[0, +\infty]$ ou $\overline{\mathbb{R}}$, alors $lim_{n\to+\infty} sup_{k\geq n} f_k$ et $lim_{n\to+\infty} inf_{k\geq n} f_k$ existent et sont mesurables.*

Démonstration : Rappelons simplement qu'une fonction monotone a nécéssairement une limite dans $\mathbb{R}$.□

Proposition 349 *Si pour tout $n \in \mathbb{N}$ f_n est mesurable de X dans $\overline{\mathbb{R}}$, alors $sup_{n \in \mathbb{N}} f_n$ est mesurable.*

Démonstration : Il suffit de considérer les images réciproques d'ensembles de la forme $]a, +\infty]$ par les f_n. Plus précisément, fixons a, et considérons $E_n = f_n^{-1}(]a, +\infty[)$. Avec $f = sup\ f_n$, $E = f^{-1}(]a, \infty[)$ est égal à $\cup E_n$, et donc est mesurable. Les boréliens sont engendrés par les intervalles $]a, +\infty[$; donc f est bien mesurable.$\square$

6.7 Intégration - théorème de convergence dominée de Lebesgue

$(X, \mathcal{A}, \mu)$ un espace mesuré.

6.7.1 Fonctions étagées et fonctions simples

Définitions et généralités

Définition 350 *Une fonction est dite* **étagée** *si elle ne prend qu'un nombre fini de valeurs.*
Une **fonction simple** *est une application s de X dans $[0, \infty[$ mesurable et étagée.*

Si s est simple, α_1,...,α_n ses valeurs, $A_i = \{x / s(x) = \alpha_i\} = s^{-1}(\{\alpha_i\})$, alors $s = \sum_i \alpha_i 1_{|A_i}$.

Notons que les A_i sont des boréliens, en tant qu'image inverse de borélien, pour peu que la fonction s soit mesurable.

L'écriture de s est unique si on a :
- $\alpha_i \neq \alpha_j$ si $i \neq j$
- $A_i \neq \emptyset$
- $A_i \cap A_j = \emptyset$ si $i \neq j$
- $\cup_i A_i = X$

Proposition 351 *Si f est une fonction mesurable de X dans $[0, +\infty[$, alors il existe une suite croissante s_n de fonctions simples qui converge simplement vers f.*

Démonstration :
Etant donné $n \in \mathbb{N}$, on considére les segments $S_i = [\frac{i}{2^n}, \frac{i+1}{2^n}[$, pour i variant de 0 à $n2^n$. Il suffit alors de considérer la somme f_n des $\frac{i}{2^n}\chi_{f^{-1}(S_i)}$ pour i dans $[0, n2^n]$.$\square$

Pour la suite de notre propos, on devra utiliser une multiplication dans $[0,+\infty[$, prolongeant la multiplication usuelle, et vérifiant $0.\infty = 0$. Attention, ce produit n'est pas continu.

⊡ Intégration des fonctions simples

Définition 352 $s = \sum \alpha_i 1_{|A_i}$ *avec la condition d'unicité donnée ci-dessus, alors on appelle* **intégrale** *sur* $E \in \mathcal{A}$ *de* s *pour la mesure* μ

$$\int_E s.d\mu = \sum \alpha_i \mu(A_i \cap E)$$

Lemme 353 *Supposons* s *et* t *deux fonctions simples vérifiant les conditions d'unicité. Alors :*
• *Si* $s \leq t$, *alors* $\int_E sd\mu \leq \int_E td\mu$
• *additivité :* $\int_E (s+t)d\mu = \int_E sd\mu + \int_E td\mu$
• $\int_E sd\mu = \int_E sd\mu$
• $c < +\infty \rightarrow \int c.sd\mu = c \int sd\mu$
• *La fonction de* $\mathcal{A}$ *dans* $[0,+\infty]$ *qui à* E *associe l'intégrale de* s *sur* E *pour la mesure* μ *est une mesure.*

6.7.2 Fonctions positives

Définition 354 $f : (X, \mathcal{A}) \rightarrow [0,+\infty]$ *mesurable*
On appelle **intégrale** *de* f *sur* E *de* f *pour* μ

$$\int_E f d\mu = sup_{s \leq f} \int_E s.d\mu$$

où $s \leq f$ *désigne l'ensemble des fonctions simples plus petites (ou égale) que* f.

Rmq : Si f est simple, alors cette définition coïncide avec la précédente.

La notation $\int_E f d\mu$ désigne $\int f.\chi_E.d\mu$.

Propriétés, avec f et g positives mesurables :
• $\int f + g = \int f + \int g$
• $f \leq g \rightarrow \int f \leq \int g$
• $\int |a|.f = |a| \int f$

Théorème 355 (Théorème de convergence monotone, dit aussi théorème de Beppo-Levi)
Soit f_n une suite croissante de fonctions mesurables positives.
Alors $f = sup f_n$ est mesurable et $\int f = lim\ \int f_n$.

Démonstration : On procède par étapes comme suit :
• On montre tout d'abord que f est mesurable (considérer $f^{-1}(]a, +\infty))$.
• On montre que $\int f_n$ est croisante ; notons F sa limite (éventuellement infinie).
• $F \leq \int f$ par passage à la limite.
L'autre inégalité est un peu plus difficile et est illustrée par la figure 6.2.
• On considère une fonction simple $\leq f$, notée s, et un réel $a < 1$.
• On considère E_n l'ensemble des x tels que $a.s(x) \leq f_n(x)$.
• $\int_{E_n} a.u \leq \int_{E_n} f_n \leq \int_X f_n \leq F$.
• La réunion des E_n est X, et la réunion est dénombrable ; donc $\int_{E_n} u \to \int_X u$.
• Par passage au sup on a $\int f \leq F$.$\square$

On trouvera de multiples applications à ce théorèmes, citons le lemme de Fatou 357, le théorème justifiant la définition des mesures images 358, le théorème importantissime de la convergence dominée de Lebesgue 366, pour le théorème de Fubini 378. Cela servira aussi en probabilité, avec la proposition 1326, la proposition 1327, les resultats sur les variables aléatoires indépendantes 1334 ; on trouvera aussi une application au théorème 1367 sur l'extinction d'un processus de branchement. Par ailleurs, on utilisera ce résultat aussi pour montrer que les espaces L^p sont des Banach, voir corollaire 404.

On prendra garde à ne pas utiliser des arguments plus lourds (notamment la convergence dominée de Lebesgue) quand un résultat plus simple comme celui-ci suffit.

Corollaire 356 *(**Permutation des signes somme - intégrales**) Avec f_n pour $n \in \mathbb{N}$ mesurable positive, $\int \sum f_n = \sum \int f_n$*

Lemme 357 (Lemme de Fatou) *Avec f_n de X vers $[0, +\infty]$ mesurable, on a $\int liminf\ f_n \leq liminf \int f_n$*

Démonstration : Il suffit de définir $g_k = inf_{n \geq k} f_n$; g_k est croissante ; donc par le théorème de convergence monotone $\int liminf\ f_n = lim \int g_n = liminf \int g_n \leq liminf\ \int f_n$.$\square$

Théorème 358 *Pour toute fonction f mesurable de X dans $[0, +\infty]$, la fonction ν qui à un borélien E associe $\int_E f(x).d\mu$ est une mesure sur X.*
En outre pour toute fonction g mesurable de X dans $[0, +\infty]$, $\int g.d\nu = \int f.g.d\mu$.

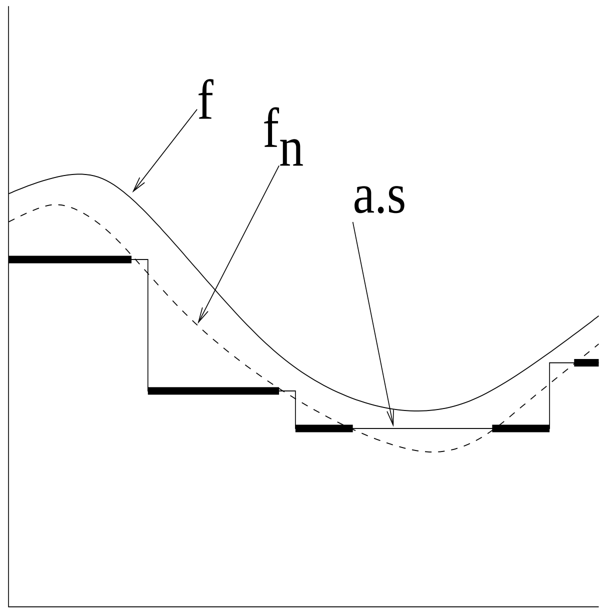


FIG. 6.2 – Les ensembles sur lesquels f_n dépasse la fonction simple multipliée par $a < 1$ ont pour réunion X et sont en nombre dénombrable.

Démonstration : Il est facile de voir que ν est une mesure avec les outils que l'on s'est donnés plus haut. La suite est un peu plus laborieuse, mais se résoud en utilisant le théorème de convergence monotone. □

6.7.3 Le cas général

Définition 359 (Fonction intégrable) *Une fonction f de X dans $\mathbb{C}$ est dite* **intégrable** *si elle est mesurable et si $\int |f|$ est finie. On note $\mathcal{L}^1(X,\mathbb{C})$ l'ensemble des fonctions intégrables de X dans $\mathbb{C}$, et $\mathcal{L}^1(X,\mathbb{R})$ l'ensemble des fonctions intégrables de X dans $\mathbb{R}$. $\mathcal{L}^1(X)$ tout court désigne généralement $\mathcal{L}^1(X,\mathbb{R})$ (voir selon le contexte).*

$\mathcal{L}^1(X)$ est un $\mathbb{C}$-espace vectoriel (on le munit de l'addition et de la multiplication par un scalaire). L'application qui à une fonction associe son intégrale est une forme linéaire sur cet espace vectoriel.
$\mathcal{L}^1(X)$ est un $\mathbb{R}$-espace vectoriel (on le munit de l'addition et de la multiplication par un scalaire). L'application qui à une fonction associe son intégrale est une forme linéaire sur cet espace vectoriel.

Notez que l'intégrabilité dépend de la mesure, alors que la mesurabilité ne dépend que de l'espace mesurable.

Définition 360 (Intégrale (au sens de Lebesgue) d'une fonction intégrable) *Etant donnée f une fonction intégrable de X dans $\mathbb{R}$, on appelle* **intégrale** *de f et on note $\int f$ le réel $\int f^+ - \int f^-$.*
Etant donnée $f = g + i.h$ une fonction intégrable de X dans $\mathbb{C}$, on appelle **intégrale** *de f et on note $\int f$ le complexe $\int g + i. \int h$.*

On notera bien que cette définition ne permet de définir d'intégrale que lorsque $\int |f|$ est finie. Ainsi l'intégrale d'une fonction mesurable à valeurs dans $[0, +\infty]$ est toujours définie mais la fonction n'est pas nécéssairement intégrable (lorsque l'intégrale est infinie). Et dans le cas d'une fonction Riemann-intégrable on doit résister à la tentation d'utiliser une intégrale d'une fonction pas intégrable, puisque l'intégrale de Lebesgue n'est définie que dans le cadre de fonctions intégrables.

Propriétés :
• Si f est intégrable alors $|f|$ est intégrable et $|\int f| \leq \int |f|$.
• Une fonction inférieure en module à une fonction g intégrable, est intégrable.
• Si deux fonctions f et g sont intégrables et à valeurs dans $\mathbb{R}$ et si $f \leq g$ alors $\int f \leq \int g$.

Définition 361 (Intégrale sur une partie mesurable) *Soit E une partie mesurable de X, et f de X dans $\mathbb{R}$ ou $\mathbb{C}$, alors f est dite* **intégrable sur** *E si $f.\chi_E$ est intégrable, avec χ_E la fonction caractéristique de E. On définit alors $\int_E f = \int f.\chi_E$.*

On peut vérifier que $\int_E f = \int f_{|E}$.

Exemple important : les sommes de séries.
L'ensemble des parties de $\mathbb{N}$ est une σ-algèbre sur $\mathbb{N}$. On peut munir l'espace mesurable ainsi défini d'une mesure μ telle que $\mu(A) = card(A)$ si A est fini et $\mu(A) = +\infty$ sinon.

On se donne alors une fonction f de $\mathbb{N}$ dans $[0, +\infty[$, c'est à dire une suite de réels positifs.
Cette fonction est évidemment mesurable.

On peut alors considérer les fonctions $f.\chi_[0, n]$; la suite de ces fonctions converge vers f, donc par le théorème de convergence monotone, l'intégrale de f sur $\mathbb{N}$ est la limite pour $\mathbb{N}$ tendant vers $+\infty$ de l'intégrale de $f.\chi_[0, n]$. $f.\chi_[0, n]$ étant une fonction étagée, son intégrale est facile à calculer ; il s'agit de la somme des f_i pour $i \in [0, n]$.

On peut retrouver ainsi divers résultats classiques du calcul de séries, par exemple le changement d'ordre des termes dans une série absolument convergente. On peut aussi considérer le cas des séries complexes.

Par contre, on ne peut rien faire au niveau des séries non-absolument convergentes.

Proposition 362 • *Une fonction mesurable de X dans $[0,+\infty]$ est d'intégrale nulle si et seulement si elle est nulle presque partout.*
• *Si une fonction mesurable de X dans $[0,+\infty]$ est d'intégrale finie alors elle est finie presque partout.*
• *f mesurable de X dans $\mathbb{C}$; alors $|\int f| \leq \int |f|$, et si $|\int f| = \int |f|$, alors il existe a tel que $f = a.|f|$ presque partout.*

Démonstration :
• le premier point est facile, il suffit de considérer les ensembles sur lesquels f est supérieure à $1/n$, et leur réunion dénombrable.
• Considérer l'ensemble des x tels que $f(x) = +\infty$ et sa mesure.
• Considérer l'argument de l'intégrale de f, et la fonction $a.f$ avec a un complexe de module 1 tel que $\int a.f \in \mathbb{R}^+$. La suite est facile...□

On peut considérer différentes structures à l'intérieur de l'espace vectoriel $\mathcal{F}$ des fonctions de X dans $\mathbb{R}$:
• Le sous-espace vectoriel des fonctions mesurables
• Le sous-espace vectoriel des fonctions intégrables
• Le sous-espace vectoriel des fonctions nulles presque partout

La dernière propriété permet notamment de définir un espace quotient de l'espace des fonctions, par la relation d'équivalence $\mathcal{R}$ définie par $f\mathcal{R}g \iff f(x) = g(x)$ presque partout . On considère alors l'espace constitué par les classes contenant au moins une fonction intégrable. La forme linéaire qui à f associe son intégrale induit une forme linéaire sur cet espace quotient (notez que deux fonctions intégrables appartenant à la même classe ont même intégrale). On peut normer cet espace par la norme suivante :

$$\| f \|_1 = \int |f|$$

La forme linéaire qui à f associe son intégrale est continue pour cette norme, et de norme ≤ 1, c'est à dire que $|\int f| \leq \| f \|_1 = \int |f|$. Dans la plupart des cas, c'est à dire dès qu'il existe une partie de mesure finie non nulle, cette norme est 1 (on pourra s'en convaincre en considérant la fonction caractéristique d'une telle partie).

Définition 363 *On note L^1 le sous-espace des classes contenant au moins une fonction intégrable ainsi normé (ne pas confondre avec $\mathcal{L}^1$). Cette notation est dépendante du contexte ; formellement il faudrait préciser l'espace de départ et l'espace d'arrivée (ce dernier étant généralement $\mathbb{C}$, quelquefois $\mathbb{R}$).*

Théoriquement il n'est pas possible d'écrire $\| f \|_1$ pour une fonction de X dans $[0,+\infty]$, $+\infty$ inclus ; néanmoins on verra souvent cette notation pour $\int |f|$. On se permettra ainsi de parler de la classe d'une fonction dont l'intégrale est finie, même s'il s'agit par exemple d'une fonction à valeurs dans $[0,+\infty]$.
On assimilera souvent une fonction et sa classe, l'intégrale d'une fonction et l'intégrale

d'une fonction intégrable de sa classe, etc.

6.7.4 Fonctions vectorielles

Une **fonction vectorielle** est une fonction de X dans $\mathbb{R}^n$.

Définition 364 *Une fonction vectorielle est dite* **intégrable** *si toutes ses composantes sont intégrables. Son intégrale est alors le vecteur dont chaque coordonnée est l'intégrale de la coordonnée correspondante de f.*

Proposition 365 *Une fonction vectorielle est intégrable si et seulement si elle est intégrable et si sa norme est intégrable (indépendamment du choix de cette norme).*

Démonstration : On majore chaque composante par la norme multipliée par une certaine constante, et réciproquement ; le résultat est ensuite facile.$\square$

6.7.5 Théorème de la convergence dominée de Lebesgue. Corollaires

Théorème 366 (Théorème de la convergence dominée de Lebesgue) *Soit f_n de X dans $\mathbb{C}$ pour tout $n \in \mathbb{N}$.*
Hypothèses :
• *f_n mesurable*
• *Pour presque tout x, $f_n(x)$ converge*
• *Il existe une fonction g intégrable de X dans $[0, +\infty]$ majorant toutes les fonctions f_n.*

Alors :
• *une certaine fonction f est limite simple des f_n ; cette fonction est intégrable.*
• *$\int |f - f_n| \to 0$ (convergence L_1).*
• *$\int f_n \to \int f$ pour $n \to +\infty$*

Démonstration : On passe par les étapes suivantes :
• Tout d'abord le cas d'une suite de fonctions tendant monotonement vers la fonction nulle (démonstration en utilisant le théorème de convergence monotone).
• Ensuite suite de fonctions tendant vers la fonction nulle (on se ramène au cas précédent en considérant $g_n(x) = sup_{k \geq n} f_k(x)$.
• Ensuite le cas général se traite en considérant $f - f_n$, majorée par $2.g$.
• L'assertion $\int f_n \to \int f$ pour $n \to +\infty$ découle simplement du fait que $|\int a - \int b| \leq \int |a - b|$.$\square$

On utilisera ce résultat très souvent, citons par exemple le corollaire ci-dessous, le théorème 369 de continuité sous le signe intégral, le théorème 370 de dérivation sous le signe intégral, le lemme de Scheffé 371, le résultat de densité 424, le contre-exemple du paragraphe 5.6.7, le théorème de Doob 1354.

Corollaire 367 *Soit une suite de fonctions f_n intégrables de X dans $\mathbb{C}$, telle que $\sum \| f \|_1$ converge.*
• Alors pour presque tout x on a $\sum_n f_n(x)$ convergente vers un certain $f(x)$.
• En outre f est intégrable et $\int f = \sum_n \int f_n$.
• La suite $\sum_{i=0}^n f_n$ converge vers f pour L_1.

Démonstration : On considère $g(x) = \sum_n |f_n(x)|$. Par le théorème de convergence monotone, $\int g = \sum \int |f_n|$. On peut donc définir $h(x)$ comme limite des fonctions $h_n(x) = sum_{i=0}^n f_n(x)$. On a $|h(x)| \leq g(x)$; par le théorème de convergence dominée, l'intégrale de f est égale à la limite de l'intégrale des h_n, donc à la limite de la somme des $\int f_i$ pour $i \in [0, n]$ et la dernière assertion découle de la convergence pour L_1 dans le théorème de convergence dominée.□

Corollaire 368 *Les espaces vectoriels normés L^1 et $L^1(\mathbb{C})$ sont complets.*

Démonstration : Un espace vectoriel normé est complet si et seulement si toute série normalement convergente est convergente ; donc d'après le corollaire précédent, L^1 et $L^1(\mathbb{C})$ sont complets.□

Corollaire 369 *Soit Y un espace métrique et $f : X \times Y \to \mathbb{C}$ une application telle que :*
• pour tout y l'application qui à x associe $f(x, y)$ est intégrable
• pour tout x appartenant à $X - N$ avec N négligeable l'application qui à y associe $f(x, y)$ est continue
• il existe g intégrale de X dans $\mathbb{R}^+$ telle que pour tous x et y $|f(x, y)| \leq g(x)$.
Alors l'application qui à y associe $\int f(x, y).dx$ est continue.

Démonstration : On considère une suite y_n tendant vers y, la continuité séquentielle impliquant la continuité dans un espace métrique. Il suffit d'appliquer le théorème de convergence dominée à la suite de fonctions de la forme $x \mapsto f(x, y_n)$.□

Corollaire 370 *Soit Y un intervalle ouvert de $\mathbb{R}$ et $f : X \times Y \to \mathbb{C}$ une application telle que :*
• pour tout y l'application qui à x associe $f(x,y)$ est intégrable
• pour $x \in X - N$, avec N négligeable, l'application qui à y associe $f(x,y)$ est dérivable de dérivée $f_2'(x,y)$
• Pour une certain fonction g positive L^1, pour tout y, $|f_2'(x,y)| \leq g(x)$
Alors pour tout y $x \mapsto f_2'(x,y)$ est intégrable et l'application qui à y associe $\int f(x,y).dx$ est dérivable, de dérivée $\int f_2'(x,y).dx$.

Démonstration : On considère y_n une suite tendant vers y.
On définit $k_n(x) = \frac{f(x,y_n)-f(x,y)}{y_n-y}$. Pour tout n, k_n est intégrable.
D'après l'inégalité des accroissements finis, $|k_n(x)|$ est majoré par le sup de $|f_2'(x,.)|$ à son tour majoré par $g(x)$. On peut donc appliquer le théorème de convergence dominée, et déduire le résultat.$\square$

Lemme 371 (Lemme de Scheffé) *Supposons que f_n soit une suite de fonctions $\mathcal{L}^1$ de (S,μ) dans $\mathbb{R}$, et supposons que pour presque tout x $f_n(x) \to f(x)$ quand $n \to +\infty$. Alors*

$$\int |f_n|.d\mu \to \int |f|.d\mu$$

si et seulement si

$$\int |f_n - f|.d\mu \to 0$$

Démonstration : La partie "si" est triviale ; voyons maintenant la partie "et seulement si".
On montre d'abord le résultat pour des fonctions positives.
On suppose donc que $\int f_n(x).d\mu(x) \to \int f(x).d\mu(x)$. Notons $g_n = f_n - f$, et g_n^+ et g_n^- les parties positives et négatives de g_n. Alors :
• $g_n^+(x) \to 0$ et $\int g_n^-(x) \to 0$ presque partout
• $g_n^- \leq f$, donc par le théorème de convergence dominée $\int g_n^-(x).d\mu(x) \to 0$.
• Par ailleurs $\int g_n^+(x) = \int f_n(x).d\mu(x) - \int f(x).d\mu(x) - \int g^-(x).d\mu(x)$ qui tend donc vers 0.
• On peut donc déduire par les deux points précédents que $\int g_n.d\mu(x) \to 0$.
• On a donc prouvé le résultat pour des fonctions positives.
• On passe au cas général. $\int |f_n(x)|.d\mu(x) \to \int |f(x)|.d\mu(x)$ signifie que

$$\int f_n^+(x).d\mu(x) + \int f_n^-(x).d\mu(x) \to \int f^+(x).d\mu(x) + \int f^-(x).d\mu(x).$$

• Le lemme de Fatou implique que

$$\int liminf\ f_n^+.d\mu \leq liminf\ \int f_n^+ d\mu$$

et $\int liminf\ f_n^-.d\mu \leq liminf\ \int f_n^- d\mu$
et donc pour n assez grand

$$\int f_n^+(x).d\mu(x) + \int f_n^-(x).d\mu(x) \geq \int f^+(x).d\mu(x) + \int f^-(x).d\mu(x).$$

• Les résultats des deux points précédents permettent de conclure que

$$\int f_n^+(x).d\mu(x) \to \int f^+(x).d\mu(x)$$

et

$$\int f_n^-(x).d\mu(x) \to \int f^-(x).d\mu(x)$$

• Il ne reste plus qu'à appliquer le résultat dans le cas des fonctions positives.$\square$

6.8 Intégration dans les espaces produits. Changement de variable

On rappelle qu'une tribu ou σ-algèbre sur E est un sous-ensemble de $P(E)$, contenant E, stable par passage au complémentaire et par union dénombrable (et du même coup par intersection dénombrable).
On rappelle aussi qu'un clan ou algèbre sur E est un sous-ensemble de $P(E)$ stable par passage au complémentaire et par union finie (et du même coup par intersection finie).

Définition 372 *On se donne X et Y deux espaces mesurables, munis respectivement de la tribu $\mathcal{X}$ et de la tribu $\mathcal{Y}$.*
Un ensemble $A \times B$ est dit **rectangle mesurable** *de $X \times Y$ si $A \in \mathcal{X}$ et $B \in \mathcal{Y}$.*
On appelle **tribu produit** *ou* σ**-algèbre produit** *de $\mathcal{X}$ et $\mathcal{Y}$ et on note $\mathcal{X} \otimes \mathcal{Y}$ la tribu engendré par les rectangles mesurables. Ce sera la σ-algèbre par défaut par la suite ; un ensemble mesurable de $X \times Y$ est en particulier un élément de cette σ-algèbre .*
Un sous-ensemble de $X \times Y$ est dit **ensemble élémentaire** *si il est réunion finie de rectangles mesurables.*
Etant donné E inclus dans $X \times Y$ et $x \in X$, on appelle **première coupe suivant x de** *E l'ensemble des y dans Y tels que $(x, y) \in E$.*
Etant donné E inclus dans $X \times Y$ et $y \in Y$, on appelle **deuxième coupe suivant y de** *E l'ensemble des x dans X tels que $(x, y) \in E$.*
Etant données μ_X et μ_Y des mesures sur $(X, \mathcal{X})$ et $(Y, \mathcal{Y})$ respectivement on appelle mesure produit de μ_X et μ_Y une mesure sur $(X \times Y, \mathcal{X} \otimes \mathcal{Y})$ telle que la mesure d'un rectangle mesurable $A \times B$ soit $\mu_X(A).\mu_Y(B)$.

Proposition 373 • *L'ensemble des ensembles élémentaires est un clan.*
• *Un ensemble élémentaire peut sécrire comme réunion disjointe d'un nombre fini de rectangles mesurables.*
• *La première coupe suivant x d'un ensemble mesurable de $\mathcal{X} \otimes \mathcal{Y}$ est mesurable.*
• *La deuxième coupe suivant y d'un ensemble mesurable de $\mathcal{X} \otimes \mathcal{Y}$ est mesurable.*
• *Pour E mesurable de $X \times Y$, si on se donne deux mesures sur X et Y qui soient σ-finies, l'application de X dans $\mathbb{R}$ qui à x associe la mesure de la première coupe suivant x de E est mesurable*

Démonstration : Les deux premiers • sont faciles.
• Le troisième est plus délicat :
- Soit E dans $\mathcal{X} \otimes \mathcal{Y}$ et x dans X.
- Si E est un rectangle mesurable le résultat est clair.
- Il suffit donc de montrer que l'ensembles des F dans $\mathcal{X} \otimes \mathcal{Y}$ tels que la coupe suivant x de F est mesurable est une tribu, ce qui est facile.□
• Ce • est évidemment équivalent au précédent.
• Le cinquième • est admis.□

Théorème 374 *Etant donnés $(X, \mathcal{X}, \mu_X)$ et $(Y, \mathcal{Y}, \mu_Y)$ deux espaces mesurés de mesures σ-finies, il existe une et une seule mesure produit de μ_X et μ_Y. Cette mesure produit est en outre σ-finie. On la note $\mu_X \otimes \mu_Y$.*

Démonstration : • On définit $\mu(E) = \int_y \mu(E_y)$ avec E_y la deuxième coupe suivant y de E.

• On montre facilement qu'il s'agit bien d'une mesure, en utilisant le cinquième • de la proposition précédente.
• Il est clair qu'elle vérifie l'hypothèse sur la mesure des rectangles élémentaires.

• Si deux mesures vérifient les propriétés demandées, alors elles coïncident sur le Π-système des rectangles mesurables, en outre les rectangles mesurables engendrent la tribu produit, et cette tribu produit est de mesure σ-finie (facile).
• On peut alors appliquer le lemme 336 (en l'étendant, ce qui est aisé, au cas des mesures σ-finies).□

Corollaire 375 *Soit μ la mesure produit ainsi définie ; alors*

$$\mu(E) = \int_y \mu(E_y) = \int_x \mu(E_x)$$

Démonstration : La première égalité est directement issue de la preuve ci-dessus ; la seconde est due à l'unicité de la solution et à la symétrie du problème.□

Corollaire 376 *Pour qu'un ensemble E de $(X \times Y)$ soit négligeable pour $\mathcal{X} \otimes \mathcal{Y}$, il suffit que presque toutes les coupes premières de E soient négligeables (pareil avec les coupes secondes).*

Démonstration : Un tel ensemble est négligeable si et seulement si sa fonction caractéristique est d'intégrale nulle, c'est à dire si sa coupe est d'intégrale nulle presque partout.□

Corollaire 377 • *La tribu des boréliens sur $\mathbb{R}^{p+q}$ est la tribu produit des deux tribus de boréliens de $\mathbb{R}^p$ et de $\mathbb{R}^q$ (produit au sens des σ-algèbres et pas produit cartésien).*
• *La mesure de Lebesgue sur $\mathbb{R}^{p+q}$ est le produit de la mesure de Lebesgue sur $\mathbb{R}^p$ et de la mesure de Lebesgue sur $\mathbb{R}^q$.*

Démonstration : • On procède par double inclusion.
- tout d'abord soit un pavé ouvert de $\mathbb{R}^{p+q}$; il appartient bien à la tribu produit de $\mathbb{R}^p$ par $\mathbb{R}^q$ car c'est un rectangle mesurable. Or un ouvert de $\mathbb{R}^{p+q}$ est une réunion dénombrable de pavés ouverts (par exemple les pavés ouvert de coordonnées rationnelles inclus dans ce pavé). Donc les ouverts de $\mathbb{R}^{p+q}$ sont bien des mesurables pour la tribu produit, et donc les boréliens étant engendrés par les ouverts, ils sont eux-mêmes inclus dans la tribu produit.
- Soit un rectangle mesurable de $\mathbb{R}^p \times \mathbb{R}^q$; il s'écrit $X \times Y$, et donc $(X \times \mathbb{R}^q) \cap (\mathbb{R}^p \times Y)$, avec X et Y mesurables. X mesurable implique $X \times \mathbb{R}^q$ mesurable, car X appartient à la σ-algèbre engendrée par les ouverts, et donc $X \times \mathbb{R}^q$ appartient à la σ-algèbre engendrée par les ouverts de $\mathbb{R}^{p+q}$.
• Il suffit de considérer l'unicité de la mesure sur $\mathbb{R}^n$ vérifiant le fait que la mesure d'un pavé soit bien le produit des longueurs.□

Cette propriété est valable pour les boréliens MAIS pas pour les lebesguiens.

Le théorème qui suit est un théorème fondamental en théorie de l'intégration.

Théorème 378 (Fubini) *On suppose $(X,\mathcal{X},\mu_X)$ et $(Y,\mathcal{Y},\mu_Y)$ des espaces mesurés de mesures σ-finies. Soit f mesurable de $(X\times Y,\mathcal{X}\otimes\mathcal{Y},\mu_X\otimes\mu_Y)$ dans $\overline{\mathbb{R}}$.*
Alors :
• pour tout $x\in X$ l'application $f_{2,x}:y\mapsto f(x,y)$ est mesurable sur $(X,\mathcal{X})$.
• pour tout $y\in Y$ l'application $f_{1,y}:x\mapsto f(x,y)$ est mesurable sur $(X,\mathcal{X})$.
• si f est positive, alors $y\mapsto\int_X f_{1,y}(x).dx$ est mesurable positive, et

$$\int_Y (\int_X f_{1,y}(x).dx).dy=\int_{X\times Y} f.dz$$

• si f est positive, alors $x\mapsto\int_Y f_{2,x}(y).dy$ est mesurable positive. et

$$\int_X (\int_Y f_{2,x}(y).dy).dx=\int_{X\times Y} f.dz$$

• si f est intégrable, alors pour presque tout x, $f_{2,x}$ est intégrable, et $x\mapsto\int_Y f_{2,x}(y).dy$ est définie presque partout et intégrable, et on a

$$\int_X (\int_Y f_{1,x}(y).dy).dx=\int_{X\times Y} f.dz$$

• si f est intégrable, alors pour presque tout y, $f_{1,y}$ est intégrable, et $y\mapsto\int_X f_{1,y}(x).dx$ est définie presque partout et intégrable, et on a

$$\int_Y (\int_X f_{1,y}(x).dx).dy=\int_{X\times Y} f.dz$$

On remarque bien sûr qu'un • sur deux est équivalent au • précédent.

Démonstration : • (pareil pour le second •) : Soit U un ouvert ; $f_{2,x}^{-1}(U)$ est égal à la seconde section de $f^{-1}(U)$ en y, qui est mesurable d'après l'une des propriétés vues ci-dessus.
• (pareil pour le quatrième •) : on le montre tout d'abord pour une fonction caractéristique d'une partie mesurable de $(X\times Y,\mathcal{X}\otimes\mathcal{Y})$, puis pour une fonction simple, par combinaison linéaire,puis pour une fonction mesurable positive, en utilisant le théorème de convergence monotone et le fait que toute fonction mesurable positive est limite de fonctions simples.
• (pareil pour le sixième •) : Il suffit d'appliquer le cas positif à la partie positive et la partie négative d'une fonction donnée.$\square$

Ce théorème sert dans beaucoup beaucoup de situations. Citons :
- le lemme 1208, utile pour le théorème de Runge.
- de nombreuses choses sur le produit de convolution, voir le théorème 389, dans la partie 7.
- quelques propriétés de la transformée de Fourier, voir par exemple la proposition

440.

- le théorème 424, d'approximation de fonctions L^p par des fonctions C^∞ à support compact.

Corollaire 379 *Si f de $X \times Y$ dans $\mathbb{R}$ est intégrable ou mesurable positive, alors*
$$\int_Y \int_X f(x,y).dx.dy = \int_X \int_Y f(x,y).dy.dx$$

⚠ Bien noter qu'il n'est pas suffisant que l'une de ces deux expressions soit bien définie pour que le résultat soit vrai, ni même que les deux expressions soient bien définies !

Je donne ci-dessous, sans démonstration (voir [3] pour une preuve complète) la formule du changement de variable dans $\mathbb{R}^n$:

Théorème 380 (Changement de variable) *Si K est un compact inclus dans U ouvert de $\mathbb{R}^n$, si ϕ est un C^1-difféomorphisme de U sur U' ouvert de $\mathbb{R}^n$, si f est continue, alors*
$$\int_{\phi(K)} f = \int_K f \circ \phi J\phi$$
avec $J\phi$ le jacobien de ϕ.

Par exemple, cela servira à montrer la commutativité du produit de convolution.

6.9 Mesurabilité et mesurabilité au sens de Lebesgue

On a vu que la tribu des boréliens sur $\mathbb{R}^n$ pouvait être complétée en une autre tribu telle que toute partie comprise (au sens de l'inclusion) entre deux boréliens de même mesure soit mesurable ; cette tribu étant appelée la tribu des lebesguiens. En utilisant cette nouvelle tribu, on a une nouvelle notion de mesurabilité.

Quelques propriétés :

• Une fonction f est mesurable au sens de Lebesgue si et seulement si il existe une fonction g mesurable (au sens des Boréliens) égale à f presque partout

• Alors que la tribu des boréliens sur $\mathbb{R}^{n+p}$ est égale au produit (au sens des tribus) de la tribu des boréliens sur $\mathbb{R}^n$ par celle des boréliens sur $\mathbb{R}^p$, la même propriété n'est plus vraie pour les lebesguiens.

6.10 Fonctions définies par des intégrales

Les fonctions définies par des intégrales seront capitales pour les applications suivantes :

- Le produit de convolution (voir la partie 7) (avec pour conséquence de nombreux résultats de densité)

- L'analyse de Fourier (voir partie 10)

- Les indices de chemins (voir définition 658), et donc toute la construction menant au théorème de Cauchy 661

- Le théorème de Brouwer 311, lorsqu'on le prouve en utilisant la formule de Stokes (ce qui n'est pas le cas dans cet ouvrage), voir par exemple le livre "Calcul différentiel et géométrie", de D. Leborgne, Presses universitaires de France, 1982.

- La fonction gamma $x \mapsto \int_0^\infty t^{x-1}e^{-t}dt$ (permettant d'ailleurs le calcul de l'aire de la sphère unité en dimension n), que l'on trouvera par exemple dans le livre [22].

On peut citer comme autre résultat que ceux qui suivent sur les fonctions de ce type le théorème de Fubini, 378.

6.10.1 Continuité, dérivabilité sous le signe $\int$

Les résultats ci-dessous sont rappelés ici par souci encyclopédique, afin de regrouper les résultats s'appliquant aux fonctions définies par une intégrale.

Théorème 381 *Soit X un espace muni d'une mesure μ sur une tribu $\mathcal{T}$ de X. Soit (E, d) un espace métrique.*
On se donne f une application de $E \times X \to \mathbb{C}$, et on définit $F(t) = \int_X f(t,x)dx$ (intégrale pour la mesure μ, notée souvent aussi $\int_X f(t,x)d\mu(x)$.

Hypothèses	*Conclusion*
Pour tout t l'application $x \mapsto f(t,x)$ est mesurable *Pour presque tout x la fonction $t \mapsto f(t,x)$ est continue en T* *Il existe g L^1 telle que pour tout t et presque tout x $\lvert f(t,x)\rvert \leq g(x)$*	*F est continue en T.*
Pour tout t l'application $x \mapsto f(t,x)$ est mesurable *Pour presque tout x la fonction $t \mapsto f(t,x)$ est continue sur E* *Pour tout compact K de E il existe g L^1 telle que pour tout t dans K et presque tout x $\lvert f(t,x)\rvert \leq g(x)$.*	*F est continue sur E.*
E est un ouvert de $\mathbb{R}$ ou de $\mathbb{C}$ [a] *Pour presque tout t $x \mapsto f(t,x)$ est L^1* *Il existe N négligeable tel que pour tout $x \notin N$ la fonction $t \mapsto f(t,x)$ est dérivable (resp. C^1)* [b]. *Pour tout compact K de E il existe une fonction g L^1 telle que pour tout t dans K et tout $x \notin N$ $\lvert\frac{\delta f}{\delta t}(t,x)\rvert \leq g(x)$.* [a] Hypothèse facile à retenir ; il s'agit de pouvoir définir une dérivée au sens le plus commun, ie dérivée d'une fonction d'une variable réelle ou complexe ! [b] Attention ! Dans le cas d'un ouvert de $\mathbb{C}$ on parle de dérivabilité au sens complexe, et pas de différentiabilité en voyant $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel !	*Pour tout t la fonction $x \mapsto \frac{\delta f}{\delta x}(t,x)$ est L^1* *F est dérivable (resp. C^1), de dérivée $\int_X \frac{\delta f}{\delta t}(t,x)dx$.*
E est un ouvert de $\mathbb{R}$ ou un ouvert de $\mathbb{C}$. *Pour presque tout t $x \mapsto f(t,x)$ est L^1* *Il existe N négligeable tel que pour tout $x \notin N$ la fonction $t \mapsto f(t,x)$ est C^k.* *Pour tout compact K de E et tout $j \in [1,k]$ il existe une fonction g L^1 telle que pour tout t dans K et tout $x \notin N$ $\lvert\frac{\delta^j f}{\delta t^j}(t,x)\rvert \leq g(x)$.*	*Pour tout t la fonction $x \mapsto \frac{\delta^j f}{\delta t^j}(t,x)$ est L^1* *F est C^k, et pour $j \in [0,k]$ $\frac{\delta^j F}{\delta t^j} = \int_X \frac{\delta^j f}{\delta t^j}(t,x)dx$.*

Le théorème 382 est un exemple d'application.

Démonstration : Le premier théorème de continuité a été prouvé plus haut ; voir théorème 369.

Le second découle du fait que pour montrer la continuité en un point, il suffit de montrer la continuité séquentielle en ce point ; on se donne alors une suite tendant vers ce point, et on considère l'ensemble K des points de cette suite, plus la limite en question.

K est séparé parce qu'inclus dans un métrique.

K est compact, car étant donné un recouvrement ouvert de ce compact, on extrait un ouvert contenant la limite, il y a un nombre fini de points en dehors de cet ouvert, et ainsi on extrait un recouvrement fini.

On peut donc appliquer le théorème précédent à K, et on a le résultat souhaité.

Les théorèmes sur la dérivabilité découlent du corollaire 370. $\square$

6.10.2 Fonctions holomorphes sous le signe $\int$

Théorème 382 *Soit Ω un ouvert de $\mathbb{C}$ et f de $\Omega \times X$ dans $\mathbb{C}$, avec X espace muni d'une σ-algèbre $\mathcal{T}$ et μ une mesure sur $(X, \mathcal{T})$.*
On définit la fonction $F(z) = \int_X f(z,x)dx$.
Hypothèses requises :
• Pour tout z $x \mapsto f(z,x)$ est L^1.
• Il existe N négligeable tel que pour tout $x \notin N$ la fonction $z \mapsto f(z,x)$ appartient à $H(\Omega)$.
• Pour tout compact K de Ω, il existe g L^1 sur K telle que pour tout z dans K et pour tout $x \notin N$, $|f(z,x))| \leq g(x)$.
Alors :
• $F(z)$ est définie pour tout z
• F est holomorphe
• $F'(z) = \int_X \frac{\delta f}{\delta z}(z,x)dx$

Démonstration :

Montrons tout d'abord que :

pour tout compact K de Ω, il existe h L^1 sur K telle que pour tout z dans K et pour tout $x \notin N$, $|\frac{\delta f}{\delta z}(z,x))| \leq h(x)$.

Pour montrer ce résultat :

• Soit K un compact de Ω.

• Soit K_δ l'ensemble des z dans $\mathbb{C}$ tels que la distance de z à K soit $\leq \epsilon$ [1].

• L'application qui à z associe la distance de z à K est continue.

• K_δ est borné, et fermé comme image réciproque d'un fermé par une fonction continue ; K_δ est donc compact.

• Pour tout z dans K le cercle $S(z,\delta)$ de centre z et de rayon δ est inclus dans K_δ.

• On choisit δ suffisamment petit pour que K_δ soit inclus dans Ω (c'est possible car sinon on construit une suite dans $K_{1/n} \cap \mathbb{C} \setminus \Omega$, on se plonge dans le compact K_1, on considère une limite de suite extraite, elle n'est pas dans Ω mais elle est dans K, d'où contradiction)

• D'après le théorème 662, pour z' dans le disque de centre z et de rayon δ,

$$f(z',x) = \frac{1}{2i\Pi} \int_0^1 \frac{f(z+\delta e^{2i\Pi u},x)}{\delta e^{2i\Pi u} - z'} du$$

[1] Il ne s'agit pas d'un ϵ voisinage, on a pris $\leq$ et non $<$.

Les hypothèses du théorème de dérivation sous le signe intégral 381 étant vérifiées, on en déduit :

$$\frac{\delta f}{\delta z}(z,x) = \frac{1}{2i\Pi}\int_0^1 \frac{f(z+\delta e^{2i\Pi u},x)}{(\delta e^{2i\Pi u}-z)^2}du$$

Et donc

$$\frac{\delta f}{\delta z}(z,x) \leq \frac{sup_{K_\delta} f(.,x)}{\delta^2}$$

Donc par hypothèse

$$|\frac{\delta f}{\delta z}(z,x)| \leq \frac{1}{\delta}g(x)$$

La fonction h égale à $\frac{1}{\delta^2}g$ convient donc pour le résultat que l'on voulait montrer, à savoir :

pour tout compact K de Ω, il existe h L^1 sur K telle que pour tout z dans K et pour tout $x \notin N$, $|\frac{\delta f}{\delta z}(z,x))| \leq h(x)$.

Alors, on peut conclure en appliquant le théorème de dérivation sous le signe intégral 381 précédent que :

- $F(z)$ est définie pour tout z
- F est holomorphe
- $F'(z) = \int_X \frac{\delta f}{\delta z}(z,x)dx$ $\square$

6.10.3 Primitives

On pourra consulter le théorème 633 pour voir ce qu'il advient de la primitive de f lorsque $\frac{f'}{f}$ converge, et le théorème 593 pour voir ce qu'il advient lorsque $\frac{xf'(x)}{f(x)}$ converge.

6.11 Zoologie de la mesure

6.11.1 Approfondissements sur les mesures complexes

Définition 383 *Etant donnée μ une mesure complexe sur X, on appelle* **variation totale de** μ *ou* **mesure de la variation totale de** μ *et on note $|\mu|$ l'application de l'ensemble des parties mesurables de X dans $[0,+\infty]$ qui à E mesurable associe $sup \sum_i |\mu(E_i)|$, le sup étant pris sur l'ensemble des partitions de X.*

Proposition 384 *$|\mu|$ est une mesure.*

<u>Démonstration :</u>

- Il est clair que $|\mu|(\emptyset) = 0$.
- Soit maintenant $(E_i)_{i\in I}$ avec I dénombrable une partition de E. On va chercher à montrer que $|\mu|(E) = \sum_i |\mu|(E_i)$.

- Montrons tout d'abord que $|\mu|(E) \geq \sum_i |\mu|(E_i)$. Pour cela on se donne $\epsilon > 0$, et on considère, pour tout i dans I, une partition $F_{i,j}$ de E_i avec $\sum_i |\mu(F_{i,j})| \geq (1-\epsilon).|\mu|(E_i)$ (on peut trouver une telle partition, par définition de $|\mu|$). On a alors

$\sum_i (1-\epsilon)|\mu|(E_i) \leq \sum_{i,j} |\mu|(F_{i,j}) \leq \sum_i |\mu|(E_i)$. En faisant tendre ϵ vers 0 on obtient l'inégalité attendue.

- Montrons maintenant que $|\mu|(E) \leq \sum_i |\mu|(E_i)$. Considérons une partition F_j de E. Si on réussit à montrer que la somme des $|\mu(F_j)|$ est inférieur ou égal à la somme des $|\mu|(E_i)$ on aura gagné. On considère alors $G_{i,j} = E_i \cap F_j$. Les $G_{i,j}$ forment une nouvelle partition de E. Par définition de $|\mu|$ on sait que $|\mu|(E_i) \geq \sum_j |\mu(G_{i,j})|$. En sommant sur I on obtient

$$\sum_i |\mu|(E_i) \geq \sum_{i,j} |\mu(G_{i,j})| \geq \sum_j \sum_i |\mu(G_{i,j})| \geq \sum_j |\mu(F_j)|$$

Le résultat est ainsi prouvé.$\square$

Notons que pour prouver ce résultat on utilise allègrement les permutations de sommes infinies lorsque la convergence est absolue.

6.11.2 Presque recouvrement d'un ouvert de $\mathbb{R}^n$ par des petites boules

Proposition 385 *Soit $(x_i, r_i)_{i \in Boules}$, avec les x_i dans $\mathbb{R}^n$, et les r_i tels que $0 < r_i < \Delta < +\infty$.*
Alors il existe $Boules'$ un sous-ensemble <u>dénombrable</u> de $Boules$ tel que

$$\cup_{i \in Boules} \overline{B}(x_i, r_i) \subset \cup_{i \in Boules'} \overline{B}(x_i, 5r_i)$$

ET
$$(i,j) \in Boules'^2 \wedge i \neq j \Rightarrow B(x_i, r_i) \cap B(x_j, r_j) = \emptyset$$

(c'est à dire que les boules de $Boules'$ sont disjointes, et que si on multiplie leurs rayons par 5 on recouvre tout ce que l'on recouvrait avec $Boules$.

Par commodité, on identifiera le couple (x, r) et la boule fermée $\overline{B}(x, r)$.

<u>Démonstration :</u>

• On définit $Boules_k$ pour $k \geq 1$ l'ensemble des $i \in Boules$ tels que $r_i \in [\frac{D}{2^k}, \frac{D}{2^{k-1}}]$. C'est à dire que l'on regroupe les boules par taille, les plus grosses d'abord.

• On définit $BellesBoules_1$ comme un ensemble maximal de boules disjointes dans $Boules_1$. Cela peut se faire grâce au lemme de Zorn (voir lemme 36).

• On définit ensuite $BellesBoules_k$ et $Genantes_k$ par récurrence.

• $Genantes_k$ est le sous-ensemble de $Boules_k$ des boules qui intersectent la réunion des boules de la réunion des $BellesBoules_i$ pour $1 \leq i < n$.

• $BellesBoules_k$ est la réunion de $BellesBoules_{k-1}$ et d'un ensemble maximal de boules disjointes parmi $Boules_k \setminus Genantes_k$.

• Il est clair que $BellesBoules_k$ est un ensemble de boules de diamètre minoré par une constante > 0.

• On montre maintenant que $BellesBoules_k$ est un ensemble dénombrable de boules :

- Le nombre de boules de $BellesBoules_k$ incluses dans $[-i, i]^n$ est fini, puisque le volume d'une boule est minoré par une constante > 0 et qu'il y a un volume fini dans $[-i, i]^n$ (rappelons que les boules sont disjointes).

- Donc $BellesBoules_k$ est dénombrable.

• On définit maintenant $BellesBoules$ (tout court, sans indice !) comme la réunion des $BellesBoules_k$. Il s'agit d'une famille de boules disjointes (facile, on prend deux boules, soit elles appartiennent à un même $BellesBoules_k$ (auquel cas elles sont disjointes), soit l'une appartient à $BellesBoules_k$ et l'autre à).

• $BellesBoules$ est une réunion dénombrable d'ensembles dénombrables, et est donc dénombrable.

• Il reste maintenant à prouver ce que l'on cherchait à prouver, c'est à dire que la famille $BellesBoules$ convient, c'est à dire que si l'on multiplie leurs rayons par 5, les boules de $BellesBoules$ remplissent au moins tout l'espace rempli par $Boules$.

• Soit x appartenant à une boule de $Boules$.

• Soit B une boule de $Boules$ à laquelle appartient x.

• Soit n tel que $B \in Boules_k$.

• Par définition de $BellesBoules_k$, la famille

$$\{B\} \cup BellesBoules_k \cup BellesBoules_{k-1} \cup BellesBoules_{k-2} ... \cup BellesBoules_1$$

n'est pas une famille de boules disjointes.

• Il existe donc une boule B' dans un des $BellesBoules_i$ pour $i \leq n$ qui intersecte B.

• En multipliant le rayon de B' par 5, on recouvre donc B...□

Corollaire 386 *Soit U un ouvert de $\mathbb{R}^n$, et δ un réel > 0.*
Alors il existe une famille dénombrable de boules fermées disjointes de diamètre $< \delta$, toutes incluses dans U, qui recouvrent U à part sur un ensemble de mesure de Lebesgue nulle.

Démonstration :

• On considère les intersections de U avec les couronnes ouvertes de centre 0 et comprises entre les sphères de rayon n et $n + 1$.

• Il est clair que les intersections en question sont ouvertes, bornées.

• Il est donc clair que si l'on résoud la question dans le cas d'un ouvert borné, par réunion dénombrable, on aura résolu la question.

• Soit donc un tel U, ouvert borné.

• On considère l'ensemble des boules fermées de diamètre $< \delta$ incluses dans U.

• On considère maintenant la famille $ChouettesBoules_1$ dénombrable construite suivant la proposition précédente ; c'est à dire qu'en multipliant le rayon des boules par 5 on recouvre tout U.

• Le volume de $ChouettesBoules_1$ est au moins $\frac{1}{5^n}$ fois le volume de U.

• On réitère sur le complémentaire de $ChouettesBoules_1$ dans U ; on construit ainsi $ChouettesBoules_2$.

• On continue sur le complémentaire de $ChouettesBoules_1 \cup ChouettesBoules_2 \cup ... \cup ChouettesBoules_k$ pour construire $ChouettesBoules_{k+1}$.

• On considère maintenant $ChouettesBoules$ la réunions des $ChouettesBoules_k$; cette famille est dénombrable, et son complémentaire dans U est de volume inférieur à $(1 - 1/5^n)^i$ pour tout i, et donc de volume nul...□

Chapitre 7

Produit de convolution

Cette partie sera très enrichie par la lecture de la partie 36.4, consacrée à la convolution en probabilités, et de la partie 9, consacrée à l'approximation de fonctions, et faisant un large usage du produit de convolution.

7.1 Définitions et généralités

Définition 387 *Soient f et g deux applications de $\mathbb{R}^n$ dans $\mathbb{R}$ mesurables. Alors on appelle* **produit de convolution de f et g** *et on note $f * g$ la fonction $x \mapsto \int_{\mathbb{R}^n} f(x-y)g(y)d\mu(y)$.*

La convolution servira beaucoup, beaucoup, beaucoup, pour les résultats d'approximation de la partie 9 (notamment une version utile du lemme d'Urysohn 413) et de la partie 36.4, consacrée à la convolution en probabilités.

Proposition 388

$$f * g = g * f$$

Démonstration : Résulte simplement du changement de variable $u = x - y$.□

Théorème 389 (Domaine de définition de $f * g$) *• Si f et g sont L^1 alors $f * g$ est L^1 et défini presque partout, et $\|fg\|_1 \leq \|f\|_1 \|g\|_1$.*
• Si f est L^∞ et g L^1 alors $f * g$ est L^∞ et défini presque partout.
• Si f est bornée sur tout compact et si g est L^1 à support compact alors $f * g$ est défini partout.

Ce résultat est utilisé par exemple dans la proposition 412.

Démonstration :

• Simple application de Fubini (théorème 378).

• Facile, $\int_{\mathbb{R}^n} |f(x-y)g(y)|d\mu(y) \leq M \int_{\mathbb{R}^n} |g(y)|d\mu(y)$ avec M un majorant essentiel de $|f|$.

• Facile aussi, $f(x-.)$ admet un majorant sur le compact en dehors du quel $g(.)$ est nul.$\square$

Théorème 390 *Si f est L^1 et si g est L^p, pour $p \in [1,\infty]$, alors $f*g$ est définie presque partout et appartient à L^p ; en outre $\|f*g\|_p \leq \|f\|_1 \; \|g\|_\infty$.*

__**Démonstration :**__

Si $p=\infty$, c'est clair, si $p=1$, c'est le théorème précédent.
Considérons donc maitenant $1<p<\infty$.
• $y \mapsto g(y)^p$ est L^1.
• $\int |f(x-y)| \; |g(y)|^p dy$ est fini, par le cas $p=1$.
• Posons q tel que $\frac{1}{p}+\frac{1}{q}=1$.
• $\int (|f(x-y)|^{1/p} \; |g(y)|)^p dy$ est fini
• Puisque $\int (|f(x-y)|^{1/q})^q dy$ est fini aussi, et par l'inégalité de Hölder 398,

$$\int |f(x-y)| \; |g(y)| \; dy$$

est fini aussi et

$$\leq (\int (|f(x-y)|^{1/p} \; |g(y)|)^p dy)^{1/p} (\int |f(x-y)|)^{1/q}$$

$$\leq (\int (|f(x-y)| \; |g(y)|^p) dy)^{1/p} (\int |f(x-y)|)^{1/q}$$

par le cas $p=1$

$$\leq \|f(x-.)\|_1^{1/p} \; \|g(.)^p\|_1^{1/p} \|f(x-.)\|_1^{1/q}$$

$$\leq \|f\|_1 \; \|g\|_p$$

D'où le résultat.$\square$

Définition 391 *Soit f une application définie sur un espace topologique X et à valeur dans $\mathbb{C}$. On appelle* **support de** *f l'adhérence de l'ensemble*

$$\{x \in X ; f(x)=0\}.$$

On dit que f est à **support compact** *si son suppport est compact.*

Théorème 392 (Propriété fondamentale du produit de convolution) • *Si f est C^k et si g est L^1 à support compact, alors $f*g$ est C^k. En outre pour tout ν tel que $|\nu| \leq k$ $D^\nu(f*g)=(D^\nu f)*g$.*
• *On a le même résultat si f est C^k à support compact et g L^1.*

Ce théorème a pour conséquence la propositon 393, ou le résultat d'approximation 412. Il servira aussi pour le théorème 423 : la densité de l'ensemble des fonctions C^∞ à support compact dans $C^k(\mathbb{R}^n)$.

Démonstration :

On démontre simplement le premier $\bullet$, le second étant similaire.

$\bullet$ On montre le résultat sur toute boule $B = B(0, R)$; c'est clairement suffisant pour avoir le résultat désiré.

$\bullet$ On se donne R' tel que le support de G soit inclus dans la boule $B(0, R')$.

$\bullet$ Pour tout x dans B et tout y tel que $g(y)$ est non nul, $x+y$ est dans $B(0, R+R')$.

$\bullet$ Il existe M tel que la somme des $|D^\nu f|$ pour $|\nu| \leq k$ soit inférieure à M sur $B(0, R+R')$.

$\bullet$ On procède alors par récurrence.

$\bullet$ L'initialisation de la récurrence, le cas $k = 0$, est simplement une continuité sous le signe intégral (voir théorème 381).

$\bullet$ Ensuite on suppose le résultat vrai jusqu'au rang k, et on se donne ν tel que $|\nu| = k + 1$; alors pour un certain i $D^\nu = \frac{\delta}{\delta x_i} D^\eta$ pour un certain η.

$\bullet$ Les trois points suivants sont clairement vérifiés :

$$y \mapsto D^\eta f(x-y)g(y)$$

est intégrable

$$x \mapsto D^\eta f(x-y)g(y)$$

est C^1

$$|D^\nu f(x-y)|\, |g(y)| \leq M|g(y)|$$

$\bullet$ Alors :

$$D^\nu \int_{\mathbb{R}^n} f(x-y)g(y)d\mu(y)$$

$$= \frac{\delta}{\delta x_i} D^\eta \int_{\mathbb{R}^n} f(x-y)g(y)d\mu(y)$$

$$= \frac{\delta}{\delta x_i} \int_{\mathbb{R}^n} D^\eta f(x-y)g(y)d\mu(y)$$

(par hypothèse de récurrence)

$$= \int_{\mathbb{R}^n} \frac{\delta}{\delta x_i} D^\eta f(x-y)g(y)d\mu(y)$$

(grâce aux résultats affirmés dans le $\bullet$ précédent et grâce au théorème 381)

$$= \int_{\mathbb{R}^n} D^\nu f(x-y)g(y)d\mu(y)$$

D'où le résultat.$\square$

7.2 Zoologie de la convolution

7.2.1 Convoluée d'un polynôme

Proposition 393 *Si f est un polynôme et si g est L^1 à support compact, alors $f * g$ est un polynôme.*

Démonstration : Il s'agit d'une application directe du théorème 392... Il suffit de se rappeler qu'une application dont une dérivée est nulle est un polynôme.$\square$

7.2.2 Une fonction FONDAMENTALE pour la convolution

Proposition 394 *Il existe une certaine fonction ρ C^∞ de $\mathbb{R}^n$ dans $\mathbb{R}$, positive, d'intégrale 1, de support inclus dans $B(0,1)$.*

<u>Démonstration :</u> On peut par exemple considérer $\rho(x) = K\ exp(-\frac{1}{1-\|x\|^2})$, pour K convenablement choisi. On trouvera au lemme 271 une preuve du fait que cette fonction est convenable.□

Voir lemme 271 pour une liste d'applications.

Corollaire 395 *Pour tout ϵ, il existe une fonction ρ_ϵ C^∞, de support inclus dans $B(0,\epsilon)$, et d'intégrale 1.*

<u>Démonstration :</u> On utilise simplement la fonction ρ définie en 394, avec $\rho_\epsilon(x) = L\rho(x/\epsilon)$, avec L convenablement choisi.□

On a des applications aux résultats d'approximation suivantes :

- Approximation d'ensembles mesurables par des fonctions C^∞, voir proposition 412.

- Approximations de fonctions C^k par des fonctions C^∞ à support compact ; voir le théorème 423.

- Approximations de fonctions L^p par des fonctions C^∞ à support compact : voir le théorème 424.

Chapitre 8

Espaces $\mathcal{L}^p$ et espaces L^p

8.1 Quelques résultats utiles

Définition 396 (Nombres conjugués) *On dit que deux réels p et q sont conjugués si ils sont tous les deux > 0 et si*

$$\frac{1}{p} + \frac{1}{q} = 1$$

Lemme 397 *Pour $\alpha \in]0,1[$ et $(u,v) \in (\mathbb{R}^+)^2$ on a $u^\alpha.v^{1-\alpha} \leq \alpha.u + (1-\alpha).v$*

Démonstration : On passe au ln et le résultat est évident par concavité de ln.□

Théorème 398 (Inégalités de Hölder) *Soient f et g deux fonctions mesurables de X dans $\overline{\mathbb{R}^+}$, et soient p et q deux réels conjugués, alors*

$$\int f.g \leq (\int f^p)^{\frac{1}{p}}.(\int g^q)^{\frac{1}{q}}$$

Démonstration : • Posons $F = (\int f^p)^{\frac{1}{p}}$ et $G = (\int g^q)^{\frac{1}{q}}$.

• On peut supposer sans perte de généralité F et G finis et non nuls.

• Posons $u = (\frac{f(x)}{F})^p$, $v = (\frac{g(x)}{G})^q$, $\alpha = \frac{1}{p}$, on a $1 - \alpha = \frac{1}{q}$, et donc en appliquant le lemme 397 on obtient

$$\frac{f(x).g(x)}{FG} \leq \frac{1}{p}\frac{f(x)^p}{F^p} + \frac{1}{q}\frac{g(x)^q}{G^q}$$

• En intégrant on obtient

$$\frac{1}{FG}\int fg \leq \frac{1}{p} + \frac{1}{q} = 1$$

Ce qui achève la preuve.$\square$

Corollaire 399 (Inégalité de Schwartz) *Soient f et g deux fonctions mesurables de X dans $\overline{\mathbb{R}}$, alors*

$$\int |f.g| \leq \sqrt{\int f^2}.\sqrt{\int g^2}$$

Démonstration : Spécialisation du théorème 398 dans le cas $p = q = 2$.$\square$

Théorème 400 (Inégalité de Minkowski) *Soit $p \in]1, +\infty[$, et soient f et g des fonctions mesurables de X dans $[0, +\infty]$.*
Alors

$$(\int (f+g)^p)^{\frac{1}{p}} \leq (\int f^p)^{\frac{1}{p}} + (\int g^p)^{\frac{1}{p}}$$

Démonstration : • Si $(\int (f+g)^p)^{\frac{1}{p}}$ est infinie, alors par convexité de $x \mapsto x^p$, on peut écrire $(\frac{f+g}{2})^p \leq \frac{f^p+g^p}{2}$, et donc déduire que l'inégalité annoncée est vraie.
• On suppose maintenant que $(\int (f+g)^p)^{\frac{1}{p}}$ est finie.
• On considère q conjugué à p.
• Alors par le théorème 398, on peut écrire les deux inégalités suivantes :

$$\int f.(f+g)^{p-1} \leq (\int f^p)^{\frac{1}{p}}.(\int (f+g)^{(p-1).q})^{\frac{1}{q}}$$

$$\int g.(f+g)^{p-1} \leq (\int g^p)^{\frac{1}{p}}.(\int (f+g)^{(p-1).q})^{\frac{1}{q}}$$

• On additionne et on obtient l'inégalité annoncée.$\square$

8.2 Espaces $\mathcal{L}^p$ et L^p

On se donne $p \in [1, +\infty]$. On note bien que p peut être $+\infty$.

Définition 401 (Normes N_p) *Si $p \in [1, +\infty[$ alors on note N_p l'application qui à une fonction f de X dans $\overline{\mathbb{R}}$ ou de X dans $\mathbb{C}$ associe $(\int |f|^p)^{\frac{1}{p}}$.*
On appelle **majorant essentiel** *d'une fonction f tout M tel que $|f(x)| \leq M$ pour presque tout x.*
Si $p = +\infty$ alors on note N_∞ l'application qui à une fonction f de X dans $\overline{\mathbb{R}}$ ou de X dans $\mathbb{C}$ associe la borne inf des majorants essentiels de f. $N_\infty(f)$ est appelée **borne supérieure essentielle** *de f.*
On dit qu'une fonction est **essentiellement bornée** *si $N_\infty(f)$ est fini.*
On note $\mathcal{L}^p(X, \mu)$ ou $\mathcal{L}^p(X)$ lorsqu'il n'y a pas ambiguité l'ensemble des fonctions f mesurables de (X, μ) dans $\mathbb{R}$ telles que $N^p(f)$ est fini.

Pour la relation $\mathbb{R}$ définie par

$$f \mathbb{R} g \iff f(x) = g(x) \text{ presque partout}$$

on note $L^p(X, \mu)$ ou $L^p(X)$ l'ensemble des classes d'équivalence de l'ensemble des applications de X dans $\mathbb{R}$ contenant au moins une fonction de $\mathcal{L}^p(X, \mu)$.
On définit de même $\mathcal{L}^p_{\mathbb{C}}(X, \mu)$, $L^p_{\mathbb{C}}(X, \mu)$, $\mathcal{L}^p_{\mathbb{C}}(X)$ et $L^p_{\mathbb{C}}(X)$, en considérant des fonctions à valeur dans $\mathbb{C}$.
On note $\lambda^p(X)$ l'espace $L_{\mathbb{C}}(X, \mu) = \mathcal{L}_{\mathbb{C}}(X, \mu)$ avec μ la mesure du dénombrement (il y a identité entre $L(\mathbb{N})$ et $\mathcal{L}(\mathbb{N})$ car pour tout P partie de $\mathbb{N}$ $\mu(P)$ est égal au nombre d'éléments de P, éventuellement $+\infty$, et donc le seul ensemble négligeable est l'ensemble vide).
On note λ^p l'espace $L^p_{\mathbb{C}}(\mathbb{N}, \mu) = \mathcal{L}^p_{\mathbb{C}}(\mathbb{N}, \mu) = L^p_{\mathbb{C}}(\mathbb{N}) = \mathcal{L}^p_{\mathbb{C}}(N)$, avec μ la mesure du dénombrement.
Une famille $(x_i)_{i \in I}$ de nombres complexes est **sommable de somme** *x si pour tout ϵ il existe $J \subset I$ finie telle que pour tout K fini telle que $J \subset K \subset I$ on ait $|x - \sum_{i \in K}| \leq \epsilon$.*

On note que la borne supérieure essentielle d'une fonction est le plus petit majorant essentiel de cette fonction.
N_p est une semi-norme sur $\mathcal{L}^p(X)$ et une norme sur $L^p(X)$.
$l^p(X)$ est l'ensemble des applications des f de X dans $\mathbb{C}$ telles que la famille $(|f(x)|^p)_{x \in X}$ soit sommable.

8.3 Théorèmes sur les L^p

Théorème 402 (Convergence dominée de Lebesgue dans L^p) On suppose ici $p \neq +\infty$.
Soit une suite (f_n) de fonctions mesurables, telle que

$$f_n(x) \to f(x) \text{ pour presque tout } x$$

$$\exists g \in L^p / \forall (x,n) |f_n(x)| \leq g(x)$$

alors la classe de f appartient à L^p et f_n tend vers f pour N^p.

Démonstration : Il suffit d'appliquer le théorème de convergence dominée à f_n^p, avec g^p et f.□

Remarque 403 (Contre-exemple avec $p = +\infty$) *Il suffit de considérer $f_n = \chi_{[}n, +\infty[$, pour avoir toutes les hypothèses vérifiées, sans que la conclusion soit juste.*

Corollaire 404 *Les espaces L^p sont des espaces de Banach, pour $p \in [1, +\infty]$.*

Démonstration : • Tout d'abord on considère le cas $p = \infty$.
On se donne une série f_n normalement convergente pour la norme infinie ; on peut clairement considérer une limite simple f par complétude de $\mathbb{R}$ ou $\mathbb{C}$.
Quitte à remplacer f par une autre fonction de la même classe que f, on peut considérer M fini un majorant de f (et pas seulement un majorant *essentiel*). On peut alors simplement considérer le reste $\sum_{i=n+1..+\infty} N_\infty(x)$ pour avoir la convergence pour la norme infinie.
• On traite maintenant le cas $p \neq \infty$.
• Supposons donnée une série f_n de fonctions mesurables normalement convergente (il est nécéssaire et suffisant pour qu'un espace normé soit complet que toute série normalement convergente soit convergente).
• On note $g(x) = \sum |f_n(x)|$.
• Par le théorème de convergence monotone $\int g = lim \int \sum_{i \in [1,n]} |f_i(x)|$ et cette quantité est finie par hypothèse.
• Par le théorème 402 appliquée à $\sum_{i=1..n} f_i$ majorée par g appartenant à L^p, f_n tend vers f pour N_p.□

Bien noter que le résultat de complétude vaut aussi pour L^∞.

Proposition 405 *Pour $p \in [1, +\infty]$, si X est de mesure finie, alors $L^{p'}(X) \subset L^p(X)$ pour tout $p' \geq p$ (éventuellement p' infini).*

Démonstration : Pas dur, en séparant X en l'ensemble des points où $|f| > 1$ et en l'ensemble des points où $|f| \leq 1$.$\square$

Remarque importante ! Comme le signale la remarque qui suit le théorème 422, l'espace $L^p(\mathbb{R}^n)$ s'exprime en fait comme le complété pour la distance associée à la norme $\|.\|_p$, si $p < \infty$. Ce résultat n'est pas valable pour $p = \infty$; ici l'adhérence serait simplement l'ensemble des applications qui, pour tout $\epsilon > 0$, sont inférieures à ϵ[1] en dehors d'un certain compact K_ϵ.

La démonstration suivante, difficile, est directement extraite (et simplifiée, quitte à renforcer légèrement les hypothèses - le résultat général inclut en fait X union dénombrable de parties de mesure finie) du livre de W. Rudin, Analyse réelle et complexe.

Théorème 406 *Soit $p \in [1, \infty[$, μ mesure positive finie sur X, ϕ forme linéaire bornée (à valeurs dans $\mathbb{C}$) sur $L^p(X)$. Soit q tel que $\frac{1}{p} + \frac{1}{q} = 1$. Alors il existe un unique g (presque partout)* [a] *L^q, tel que*

$$\phi(f) = \int_X fgd\mu \tag{8.1}$$

$$\tag{8.2}$$

et on a alors

$$\|\phi\| = \|g\|_q = \sqrt[q]{\int |g|^q}$$

[a] L'unicité presque partout signifie que si deux fonctions vérifient cette propriété, alors elles sont nécéssairement égales presque partout.

Ce théorème énonce exactement un isomorphisme entre L^q et L^p. Il faut bien noter que L^p signifie $L^p(\mu)$ et L^q signifie $L_q(\mu)$.

Démonstration :

L'unicité découle facilement du fait que si g et g' sont deux fonctions vérifiant la propriété 8.1, alors pour tout E mesurable, l'intégrale sur E de $g - g'$ est nulle ; donc en particulier pour E égal à l'ensemble des x pour lesquels (respectivement) $Re(g(x)) > Re(g'(x))$, $Re(g(x)) < Re(g'(x))$, $Im(g(x)) > Im(g'(x))$, $Im(g(x)) < Im(g'(x))$.

L'existence est beaucoup plus laborieuse à prouver :

• Tout d'abord, l'inégalité de Hölder 398 et l'équation 8.1 impliquent immédiatement $\|\phi\| \leq \|g\|_q$. Il est donc suffisant de montrer l'existence de g et le fait que $\|\phi\| \geq \|g\|_q$.

• Le cas $\phi = 0$ est trivial. Par la suite nous supposerons donc ϕ non nulle.

• Montrons tout d'abord l'existence d'une certaine fonction g telle que $\phi(f) = \int fg$ pour toute $f \in L^\infty(\mu)$. Cela se fait comme suit :

– Définissons $\mathcal{L}(E) = \phi(\chi_E)$ pour χ_E fonction caractéristique de E mesurable de (X, μ).

– $\mathcal{L}$ est additive, au sens où si A et B sont disjoints, $\mathcal{L}(A \cup B) = \mathcal{L}(A) + \mathcal{L}(B)$.

[1] En module !

– Elle est dénombrablement additive. Pour le voir considérons les E_i pour i entier >0 mesurables, disjoints, A_k égal à l'union des A_i pour i égal à $1, 2, ..., k$, et E l'union des E_i. p étant supposé inférieur à ∞, $\|\chi_E - \chi_{A_k}\|_p = (\mu(E \setminus A_k))^{1/p} \to 0$ comme $k \to \infty$. Par continuité de ϕ, ceci implique que $\mathcal{L}(A_k) \to \mathcal{L}(E)$.
– $\mathcal{L}$ est donc une mesure complexe.
– $\mu(E) = 0$ implique que $\mathcal{L}(E) = 0$ car alors $\|\chi_E\|_p = 0$.
– Donc, d'après le théorème de Radon-Nikodyn il existe g L^1 telle que pour tout E mesurable inclus dans X,

$$\phi(\chi_E) = \int_E g d\mu = \int_X \chi_E g d\mu$$

– Ce résultat se généralise par linéarité aux fonctions étagées mesurables.
– On peut ensuite le généraliser aux fonctions dans $L^\infty(\mu)$ car toute fonction f bornée (presque partout) est limite uniforme de fonctions f_i étagées mesurables ; et $\|f_i - f\|_p \to 0$; donc $\phi(f_i) \to \phi(f)$.

• Il reste donc maintenant à montrer que g est L^q et que $\|\phi\| \geq \|g\|_q$; par densité de L^∞ dans L^p on pourra alors conclure que l'équation 8.1 est bien vérifiée.

• On traite tout d'abord le cas $p = 1$. On se donne E mesurable ; alors

$$|\int_E g d\mu| \leq \|\phi\| \|\chi_E\|_1$$

$$= \|\phi\| \mu(E)$$

Donc par le lemme 407 ci-dessous, $\|g\|_\infty \leq \|\phi\|$. On a donc d'une pierre deux résultats : g est L^∞, et $\|\phi\| \geq \|g\|_\infty$.

• Il reste donc à traiter le cas $p > 1$. Cela se fait en considérant α telle que $\alpha g = |g|$ et α de module constant égal à 1. Ceci est un exercice classique et peu difficile. Ensuite :
– On définit E_n l'ensemble des x tels que $|g(x)| \leq n$.
– On définit $f = \chi_{E_n} \times |g^{q-1}| \times \alpha$.
– On constate facilement que $|f|^p = |g|^q$ sur E_n.
– On sait déjà que $\phi(f) = \int fg$; donc

$$\int_{E_n} |g|^q = \int fg = \phi(f) \leq \|\phi\| \times (\int_{E_n} |g|^q)^{1/p}\|$$

donc $\int_{E_n} |g|^q \leq \|\phi\|^q$. Par le théorème de convergence monotone 355, on peut alors dire que $\|g\|_q \leq \|\phi\|$; d'où le résultat. □

Lemme 407 *Supposons X de mesure $\mu(X)$ fini, g appartenant à $L^1(\mu)$, S fermé de $\mathbb{C}$. Alors si pour tout E mesurable de mesure >0 la moyenne sur E de g (ie $\frac{1}{\mu(E)} \int_E g(x)$) appartient à S, alors g appartient à S presque partout.*

Démonstration :

• Considérons $\mathbb{D}$ un disque fermé dans le complémentaire de Δ.

• Il suffit de montrer que $\mu(E) = 0$ avec $E = g^{-1}(\mathbb{D})$; en effet, le complémentaire de S étant (comme tout ouvert) union dénombrable de disques fermés, on aura alors $g^{-1}(\mathbb{D})$ union dénombrables d'ensembles de mesure nulle, donc $g^{-1}(\mathbb{D})$ de mesure nulle.

• Montrons donc $\mu(E) = 0$. Pour cela supposons, pour arriver à une contradiction, $\mu(E) > 0$. Alors, en posant α le centre de $\mathbb{D}$ et r son rayon,

$$|\frac{1}{\mu(E)}\int_E g - \alpha| = |\frac{1}{\mu(E)}\int_E (g-\alpha)|$$
$$\leq \frac{1}{\mu(E)}\int_E (g-\alpha)$$
$$\leq r$$

Or $\frac{1}{\mu(E)}\int_E g$ est censé appartenir à S, d'où contradiction.$\square$

8.4 Zoologie des espaces L^p

8.4.1 Espace l^p

Proposition 408 *On note l^p l'espace $L^p_{\mathbb{C}}(\mathbb{N},\mu) = \mathcal{L}^p_{\mathbb{C}}(\mathbb{N},\mu) = L^p_{\mathbb{C}}(\mathbb{N}) = \mathcal{L}^p_{\mathbb{C}}(N)$, avec μ la mesure du dénombrement. Il s'agit donc de l'espace des suites (x_i) telles que $\sum |x_i|^p$ est fini. La norme N_p, est la norme suivante :*

$$N_p(x) = \|x\| = (\sum |x_i|^p)^{1/p}$$

Pour cette norme,
l^p est complet.

Démonstration :
L'inégalité de Hölder se traduit par

$$N_1(z) \leq N_p(x).N_q(y)$$

si p et q sont conjugués et si $z_i = x_i.y_i$.$\square$

8.4.2 Espace L^2

Généralités

2 est conjugué à lui-même, d'où le cas particulier.

Voici une liste de propriétés, découlant immédiatement des propriétés de L^p :
• Le produit de deux fonctions de L^2 est intégrable (par l'inégalité de Hölder)
•

⊡ Espaces préhilbertiens L^2 et $L^2_{\mathbb{C}}$

Définition 409 *Le* **produit scalaire euclidien usuel** *sur $L^2(X)$ est égal à $(f,g) \mapsto < f|g >= \int f.gd\mu$.*
Ce produit scalaire euclidien fait de $L^2(X)$ un espace hilbertien réel. Le **produit scalaire hermitien usuel** *sur $L^2_{\mathbb{C}}(X)$ est égal à $(f,g) \mapsto < f|g >= \int \overline{f}.gd\mu$.*
Ce produit scalaire hermitien fait de $L^2_{\mathbb{C}}(X)$ un espace hilbertien complexe.

⊡ Espace de Hilbert L^2 et $L^2_{\mathbb{C}}$

L^2 et $L^2_{\mathbb{C}}$ sont préhilbertiens et complets, donc ce sont des espaces de Hilbert. Ceci sera abondamment utilisé dans la partie 10 sur les séries de Fourier.

Chapitre 9

Approximation de fonctions

9.1 Topologie et approximation de fonctions caractéristiques

On trouvera ici des lemmes qui seront des outils utiles pour les démonstrations ultérieures. On peut se référer au livre "analyse réelle et complexe" de Rudin.

9.1.1 Intercalation d'ouverts relativement compacts entre un ouvert et un compact

Lemme 410 *Soit X un espace séparé localement compact, U un ouvert de X, et K un ensemble compact inclus dans U. Alors il existe un ouvert V de X relativement compact* [a] *tel que $K \subset V \subset \overline{V} \subset U$.*

[a] C'est-à-dire d'adhérence dans X compacte.

Le lemme d'Urysohn 411 se prouve facilement en utilisant ce résultat.

Démonstration : • Construisons tout d'abord V ouvert contenant K, avec $\overline{V}$ compact

- Soit K_x un voisinage compact de x, pour $x \in K$
- Soit V_x l'intérieur de K_x
- Les V_x recouvrent K, on peut donc en extraire un recouvrement fini $\cup_{x \in I} V_x$
- la réunion V des V_x, pour x dans I, convient.

• Si $U = X$, V convient.

• Sinon, soit F le complémentaire de U. Bien entendu il est fermé.

• Pour x dans F définissons T_x ouvert contenant K avec $x \notin T_x$. Un tel ouvert existe, par le théorème 176.

• Définissons alors $K_x = \overline{T}_x \cap F \cap \overline{V}$ (voir figure 9.1.

• L'intersection des K_x pour x dans F est vide, car chaque x de F n'appartient pas à T_x, donc pas à $\cap_y T_y$.

• Les K_x sont fermés dans un compact $\overline{V}$. Donc par la propriété 174, on peut en extraire une sous-famille $(K_x)_{x \in J}$ finie telle que l'intersection soit vide.

• Alors $\cap_{x \in J} T_x \cap V$ convient (rappelons qu'un fermé d'un compact est compact, voir corollaire 175).□

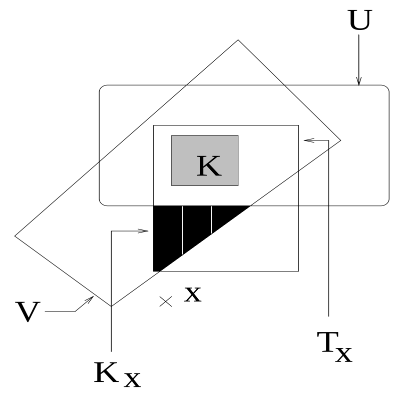


FIG. 9.1 – Construction de K_x pour $x \in F$.

9.1.2 Séparation d'un compact et d'un fermé

Lemme 411 (Lemme d'Urysohn) *Soit X un espace topologique séparé localement compact, U un ouvert de X, K un compact de X inclus dans U. Alors il existe une fonction f continue de X dans $[0,1]$ telle que*

$$x \in K \to f(x) = 1$$
$$x \notin U \to f(x) = 0$$

Ne pas utiliser un théorème difficile dans un cas simple : dans le cas d'un espace métrique, la fonction qui à x associe $\frac{d(x,X\setminus U)}{d(x,X\setminus U)+d(x,K)}$ convient.

Dans le cas de $\mathbb{R}^n$, on trouvera une preuve plus simple avec le théorème 413 (utilisant la convolution). En outre, la fonction construite sera C^∞.

Cela revient à avoir un compact, un fermé disjoint du compact, et à définir une fonction continue égale à 1 sur le compact et à 0 sur le fermé.

Démonstration :

• Par le lemme 410 (appliqué deux fois), construisons V_0 et V_1 deux ouverts relativement compacts [1] tels que

$$K \subset V_0 \subset \overline{V}_0 \subset V_1 \subset \overline{V}_1 \subset U$$

• Soit $q_0, ..., q_n, ...$ une bijection de $\mathbb{N}$ sur $\mathbb{Q} \cap [0,1]$, avec $q_0 = 0$ et $q_1 = 1$.

• Supposons construits (V_{q_i}) pour $i \in \{0, 1, ...n\}$ des ouverts relativement compacts tels que $q_i < q_j$ implique $\overline{V}_{q_i} \subset V_{q_j}$ pour i et j dans $\{0, 1, ...n\}$.

• Alors le lemme 410 permet de construire $V_{q_{n+1}}$.

• On construit ainsi une famille de fermés indexés par les éléments de $Ratio = \mathbb{Q} \cap [0,1]$, avec $q < p \Rightarrow \overline{V}_q \subset V_p$.

[1]D' adhérences compactes.

• Définissons alors, pour $q \in Ratio$, $f_q = (1-q)\chi_{V_q}$, et $g_q = q\chi_{\overline{V}_q} + (1-q)$ (voir figure 9.2).

• Puis $f = sup_{q \in Ratio} f_q$, et $g = inf_{q \in Ratio} g_q$.

• Nous avons donc, par la proposition 103, f semi-continue inférieurement et g semi-continue supérieurement.

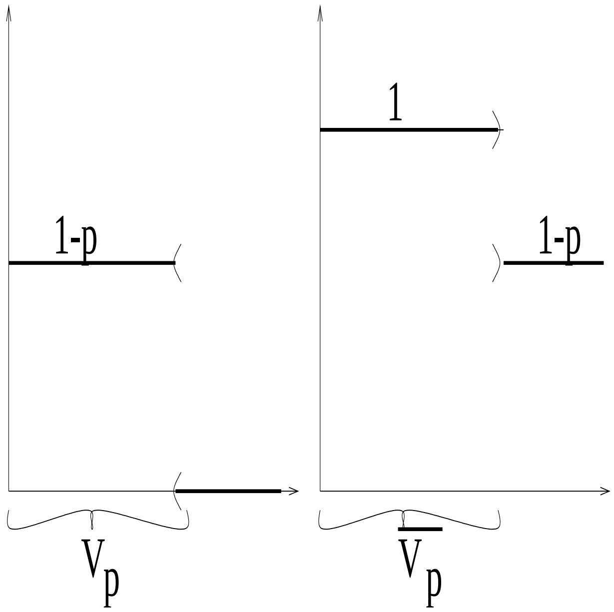


FIG. 9.2 – Graphe de f_q (à gauche) et g_q (à droite).

• Il est clair que $\chi_K \leq g$ et $\chi_U \geq f$.

• Il suffit de montrer (par la proposition 103) que $f = g$; ainsi f, semi-continue à la fois supérieurement et inférieurement, sera continue.

• Supposons que $f(x) > g(x)$.

- Alors $\exists p, q \in Ratio$ tels que

$$f_p(x) > g_q(x)$$

- Donc $x \notin V_q$, $x \in V_p$, et donc $q < p$

- Par contre (voir figure 9.2) $1-p > 1-q$; ce qui est contradictoire avec $q < p$.

• Supposons maintenant que $f(x) < g(x)$.

- Alors on peut trouver $(p,q) \in Ratio^2$ tels que

$$f(x) < 1-p < 1-q < g(x)$$

- Alors par définition du sup et de l'inf :

$$f_p(x) < 1-p < 1-q < g_q(x)$$

- On en déduit alors $x \notin V_p$ et $x \in V_q$, ce qui implique $p < q$, et contredit $1-p < 1-q$.$\square$

On trouvera par exemple une application dans la partie 5.6.9 sur le cube de Hilbert. Les autres versions de lemmes d'Urysohn (voir lemme 413) auront d'autres applications.

9.1.3 Approximation d'un ensemble mesurable par une fonction C^∞

Proposition 412 *Soit E un ensemble mesurable de mesure finie de $\mathbb{R}^n$, et soit $\epsilon > 0$.*
Alors il existe une fonction f C^∞ telle que

$$\chi_E \leq f \leq \chi_{V_{2\epsilon}(E)}$$

avec $V_t(E)$ l'ensemble des éléments à distance $< t$ de E.

Démonstration :

• Soit E un tel ensemble.

• Définissons $f_n = \chi_{V_{1/n}(E)} * \rho_{1/n}$, avec $\rho_{1/n}$ la fonction définie par le corollaire 395 (de support inclus dans $B(0, 1/n)$ et d'intégrale 1, étant en outre de classe C^∞).

• f_n est bien définie et L^1, car $\rho_{1/n}$ et $\chi_{V_{1/n}(E)}$ sont L^1 (voir propriété 389 du produit de convolution).

• f_n est C^∞, par la propriété 392.

• Tout d'abord on remarque que $\chi_E \leq f_n$

En effet, si $x \in E$, alors $f_n(x)$ est l'intégrale de $\rho_{1/n}$ sur une boule de rayon ϵ (sur cette boule en effet $\chi_{V_{1/n}(E)}$ vaut 1 - l'intégrale de f_n y est donc égale à l'intégrale de $\rho_{1/n}$, donc 1).

• Ensuite $f_n \leq \chi_{V_{2\epsilon}(E)}$ pour n assez grand.

- on le montre tout d'abord pour $\chi_{V_{2\epsilon}(E)}(x) = 0$. Pour cela, si $\chi_{V_{2\epsilon}(E)}(x) = 0$, on note que $d(x, E) > 3\epsilon/2$, alors si $1/n \leq \epsilon$,

$$f_n(x) = \int \chi_{V_{1/n}(E)}(x - y)\rho(y)d\mu(y)$$

$$= \int_{\|y\| \leq 1/n} \chi_{V_{1/n}}(x - y)\rho_{1/n}(y)d\mu(y)$$

$$= 0$$

- f est par ailleurs toujours inférieure à 1. D'où le résultat, en choisissant $f = f_n$ pour n assez grand.□

9.1.4 Lemme d'Urysohn

Il ne s'agit ici d'une version dans $\mathbb{R}^n$ du lemme d'Urysohn. On trouvera une version beaucoup plus générale (espace localement compact séparé) avec le lemme 411. Mais pour des applications de la vie de tous les jours, ce théorème suffit, et peut même s'avérer plus puissant, puisqu'il fournit une fonction C^∞ et non simplement une fonction continue.

Théorème 413 (Lemme d'Urysohn, deuxième version) *Soit K un compact de $\mathbb{R}^n$, Ω un ouvert de $\mathbb{R}^n$ contenant K, alors il existe une fonction f C^∞ à support compact telle que $\chi_K \leq f \leq \chi_\Omega$.*

Il est plus élégant de se passer d'invoquer un tel théorème lorsque l'on peut construire manuellement une solution élégante. Notamment on peut construire manuellement une fonction C^∞ de $\mathbb{R}^N$ dans $\mathbb{R}$ comprise entre $\chi_{\overline{B}(0,n)}$ et $\chi_{\overline{B}(0,n+1)}$. En effet, définissons

$$f_{|B(0,1)}(x) = e^{-\frac{1}{1-\|x\|^2}}$$

$$f_{|B(0,1)^c}(x) = 0$$

Cette fonction est C^∞, comme expliqué en 271, et > 0 sur $B(0,1)$.

On définit alors $F_n(x) = \int_{B(x,n)} f(t)d\mu(t)$.

F_n est C^∞, comme on s'en convainc en dérivant sous le signe $\int$ l'expression suivante (équivalente par un simple changement de variable) $F_n(x) = \int_{B(0,n)} f(x + t)d\mu(t)$ (voir théorème 381). Attention, pour appliquer le théorème de dérivation sous le signe somme, il faut bien voir que chaque dérivée $D^\nu f$ est majorée par une fonction L^1, ce qui n'est pas difficile en l'occurence, puisque toutes les dérivées $D^\nu f$ sont continues à support compact.

Il est ensuite évident que F_n est strictement > 0 sur $B(0,n)$, nulle sur $B(0,n+1)^c$, et comprise entre 0 et 1 partout.

On construit de même, pour Ω ouvert contenant la boule unité fermée, en considérant $F_n(x/n)$ pour n assez grand, une fonction égale à 1 sur $\overline{B}(0,1)$, et nulle en dehors de Ω.

Voir par exemple 9.6.3 ou les théorèmes 421, 414 et 422.

Démonstration :

• Il faut tout d'abord se rappeler que la distance entre un compact et un fermé disjoints est toujours > 0 (en effet la distance au fermé est continue, par la proposition 259, et donc son minimum est atteint sur le compact).

• Ensuite on applique le lemme 412.$\square$

9.1.5 Partition C^∞ de l'unité

Théorème 414 (Partition C^∞ de l'unité) *Soit K un compact inclus dans $\mathbb{R}^n$, inclus dans la réunion des Ω_i pour $i \in [1,p]$, avec Ω_i ouvert.*
Alors il existe $f_1, ..., f_p$ des applications C^∞ telles que le support de f_i soit inclus dans Ω_i pour tout $i \in [1,p]$, avec

$$\chi_K \leq \sum_{i\in[1,p]} f_i \leq 1$$

Démonstration :

• On considère la famille $(B_i)_{i\in I}$ des boules B telles qu'il existe $i \in [1,p]$ tel que $B \subset \Omega_i$ [2]. Etant donné i dans I on note $Num(i)$ un entier tel que $B_i \subset \Omega_{Num(i)}$.

• Puisque K est compact, et puisque K est (clairement !) réunion des $(B_i)_{i\in I}$, on se restreint à une réunion finie $(B_i)_{i\in J}$, recouvrant K.

• Commençons par prouver le théorème, sans se préoccuper de la contrainte $\sum_i f_i \leq 1$.

[2] $\exists i \in [1,p]$ tel que $B \subset \Omega_i$ et pas $B \subset \cup_{i\in[1,p]}\Omega_i$!

• Pour cela, définissons Ury_i, pour $i \in J$, une fonction C^∞ égale à 1 sur B_i et nulle en dehors de $\Omega_{Num(i)}$. Cela se fait par le lemme d'Urysohn, version 413, ou éventuellement par la remarque qui suit le dit lemme, qui montre que dans ce cas particulier on peut se passer du résultat général.

• On définit maintenant $MetaUry_i$, pour $i \in [1,p]$, la somme des Ury_j, pour $j \in J$ et $Num(j) = i$.

• Il est clair, comme annoncé plus haut, que la famille $(MetaUry_i)_i$, vérifie le théorème énoncé, à ceci près que la somme n'est pas nécéssairement inférieure ou égale à 1.

• On se donne maintenant une fonction f C^∞ comprise entre χ_K et $\chi_{V_\epsilon(K)}$ [3], avec ϵ inférieur au double de la distance du compact au fermé F défini par son complémentaire :

$$F^c = \{x / \sum_{i\in[1,p]} MetaUry_i(x) > 0\}$$

• On définit alors f_i pour $i \in [1,p]$ par $f_i(x) = \frac{g_i(x)f(x)}{\sum_{j=1}^n g_j(x)}$ si $x \in F^c$, et $f_i = 0$ sinon.

• On vérifie facilement que la famille ainsi construire convient.□

9.2 Approximation de fonctions continues

Théorème 415 (Théorème de Stone) *On se donne K un compact, et A une sous-algèbre unitaire de l'algèbre $C^0(K,\mathbb{R})$ des fonctions continues à valeurs réelles sur $\mathbb{K}$, munie de la norme $f \mapsto \|f\|_\infty = sup_{x\in K}|f(x)|$.*
On suppose que A sépare les points de K, c'est à dire qu'étant donné x et y dans K il existe f dans A tel que $f(x) \neq f(y)$.
Alors A est dense dans $C^0(K,\mathbb{R})$.

Démonstration :

• On montre tout d'abord que $x \mapsto \sqrt{x}$ définie sur $[0,1]$ est dans l'adhérence de l'ensemble des polynômes, pour la norme $\|.\|_\infty$. Pour cela on considère le développement de $x \mapsto \sqrt{1-x}$; qui est bien défini sur $[0,1[$. Le problème est le développement en 1. On observe alors que :

- On a $\sqrt{1-t} = 1 - \sum a_i.t^i$, avec les a_i positifs, pour $t \in [0,1[$

- Et $\sum_{i\in[0,n]} a_i t^i$ est majoré par $1-\sqrt{1-t}$, donc par 1, pour tout t ; en prolongeant par continuité, $\sum_{i\in[0,n]} a_i$ est majoré par 1. En passant à la limite pour $n \to +\infty$, $\sum_{i\in[0,n]} a_i$ est majoré par 1.

Le reste $\sum_{i\geq n} a_i.t^i$ est alors majoré par $\sum_{i\geq n} a_i$, qui tend vers 0 quand $n \to \infty$, indépendamment de t, donc...

Il ne reste plus qu'à composer par $t \mapsto 1-t$ pour avoir le résultat désiré : $x \mapsto \sqrt{x}$ est approchable uniformément par des polynômes sur $[0,1]$.

• Montrons maintenant que si f est dans A, alors $|f|$ est dans A. On se donne $racine_n(t)$ le n-ième polynôme d'une suite de polynômes tendant vers $\sqrt{t}$ sur $[0,1]$. On suppose f comprise entre -1 et 1 (on peut se ramener à ce cas-là en divisant f

[3] $V_\epsilon(K)$, dit ϵ-voisinage de K, est l'ensemble des points situés à une distance $< \epsilon$ de K.

par une constante suffisamment grande - on utilise ici le fait que A est unitaire (et donc contient les constantes)). Alors on constate que

$$racine_n(f.f) \rightarrow |f| \text{uniformément}$$

• On montre maintenant que si n fonctions $f_1, ..., f_n$ sont dans A, alors $max\{f_1, f_2, ..., f_n\}$ (resp. $min\{f_1, f_2, ..., f_n\}$) est dans l'adhérence de A ; en effet on peut toujours exprimer le max (resp. le min) d'un ensemble fini de fonctions f_i par des sommes et différences finies des f_i et de $|f_i|$ (or $|f_i|$ est dans l'adhérence de A par le point ci-dessus).

• On montre maintenant qu'étant donnés deux points x et y distincts de $[0, 1]$ et deux réels X et Y, il existe une application f dans A telle que $f(x) = X$ et $f(y) = Y$... Cela est facile en rappelant que A contient les constantes (puisqu'elle est unitaire) et que A sépare les points.

• On se donne maintenant une fonction f dans $C^0(K, \mathbb{R})$, $\epsilon > 0$ et x dans K. On cherche à montrer qu'il existe g dans A telle que $g(x) = f(x)$ et $g(t) < f(t) + \epsilon$ pour tout t dans $\mathbb{K}$.

Pour cela on considère, en utilisant le • démontré ci-dessus, pour tout t dans K une fonction f_t de A égale à f en x et inférieure à $x + \epsilon/2$ en t. On considère alors pout tout t dans K l'ouvert U_t surlequel f_t est inférieure à $f + \epsilon$; les U_t recouvrent K et on peut donc en extraire un recouvrement fini $U = \cup_{t \in E} U_t$, avec E fini. Il ne reste alors qu'à considérer la fonction min des fonctions f_t pour $t \in E$, et on a bien une fonction comme souhaitée.

• Maintenant on se donne une fonction f dans $C^0(K, \mathbb{R})$, et on cherche à montrer que l'on peut approcher f uniformément par des fonctions de A ; on aura ainsi conclu le théorème.

Pour cela, on se donne ϵ, et on associe à tout t dans K une fonction g_t égale à f en t, inférieure à $f + \epsilon$ (grâce au • ci-dessus). On peut alors associer à tout t un ouvert V_t tel que $g_t > f - \epsilon$ sur V_t. On peut alors prendre pour fonction g le max des g_t pour $t \in F$, avec F fini tel que l'union des V_t pour $t \in F$, et on a bien $f - g < \epsilon$.□

Un corollaire important est la densité de l'ensemble des polynômes trigonométriques dans l'ensembles des fonctions 2Π-périodiques continues. Un autre corollaire est le suivant :

Corollaire 416 (Théorème de Weierstrass) *L'ensemble des polynômes sur un compact K de $\mathbb{R}$ et à coefficients dans $\mathbb{R}$ est dense dans l'ensemble des fonctions continues de K dans $\mathbb{R}$, pour la norme uniforme $\|.\|_\infty$.*

Démonstration : C'est un corollaire immédiat du théorème de Stone ci-dessus.□

Il existe une autre preuve du théorème de Weierstrass, basée sur des arguments de probabilité. En fait précisément, ce corollaire est aussi un corollaire du théorème ci-dessous.

Théorème 417 *Soit f une application continue de $[0, 1]$ dans $\mathbb{R}$.*
On définit $B_n(x) = \sum_{k=0}^{n} C_n^k f(\frac{k}{n}) x^k (1-x)^{n-k}$, n-ième **polynôme de Bernstein** *associé à f.*
Alors la suite B_n converge uniformément vers f.

Démonstration : On remarque que $B_n(x)$ est précisément l'espérance de $f(\frac{X}{n})$, avec X suivant une loi binomiale $B(n, x)$.

On utilise alors le **module de continuité** $w(f,\delta)$ par

$$w(f,\delta) = sup_{|x-y|<\delta}|f(x) - f(y)|$$

il est bien à valeurs dans $\mathbb{R}$ car f est continue sur un compact.

Notons $M = sup|f|$.

Alors $|f(x) - B_n(x)| \leq$

$$w(f,\delta)P(|\frac{X}{n} - x| \leq \delta) + M\ P(\frac{X}{n} - x \geq \delta) + M\ P(\frac{X}{n} - x \leq -\delta)$$

$$\leq w(f,\delta) + 2M\ \frac{1}{n^2\delta^2}Var(X)$$

(grâce à l'inégalité de Tchebytchev 1333)

$$\leq w(f,\delta) + \frac{M}{2n\delta^2}$$

$$\to 0 \text{ quand } n \to \infty$$

D'où le résultat.$\square$

Attention ! Le théorème n'est valable que dans le cas des fonctions continues de $K \subset \mathbb{R}$ dans $\mathbb{R}$; par exemple si l'on considère l'ensemble des fonctions continues de $K \subset \mathbb{C}$ dans $\mathbb{C}$, on constate que l'on ne peut pas approcher $z \mapsto \overline{z}$ du disque unité fermé dans le disque unité fermé, car

$$\int_0^{2\Pi} f(e^{i\theta}).e^{i\theta}d\theta = 2\Pi$$

et donc si on suppose que la suite de polynômes P_n tend uniformément vers $z \mapsto \overline{z}$, la suite d'intégrales ci-dessous tend vers 2Π :

$$\int_0^{2\Pi} P_n(e^{i\theta}).e^{i\theta}d\theta$$

Or toutes les intégrales de cette suite sont nulles (considérer l'intégrale monôme par monôme pour s'en convaincre !).

Pour que tout s'arrange, il faudrait des polynômes en z ET $\overline{z}$.

Théorème 418 (Stone, version complexe) *On se donne A une sous-algèbre unitaire de l'ensemble des fonctions continues de K un compact à valeurs dans $\mathbb{C}$, stable par passage au conjugué et séparant les points de K.*
Alors A est dense dans $C(K,\mathbb{C})$ pour la norme $\|.\|_\infty$.

Démonstration :

• $Re\ f = (f+\overline{f})/2$ et $Im\ f = (f-\overline{f})/2i$; or A est stable par passage au conjugué, donc les parties réelles et imaginaires de fonctions de A sont dans A.

• la sous-algèbre des fonctions réelles de A sépare les points. En effet, soit x et y distincts, il existe une fonction f qui sépare x et y ; soit la partie réelle de f les sépare, soit la partie imaginaire de f les sépare. Dans le premier cas on a bien ce qu'on veut (une fonction réelle de A qui les sépare), dans le deuxième cas on multiplie par i (on considère $i \times Im\ f$) et cette fonction les sépare.

• en appliquant le théorème dans le cas réel, on peut donc conclure en approchant séparément la partie réelle et la partie imaginaire.$\square$

L'hypothèse sur la stabilité de A par passage au conjugué est indispensable ! En effet, on considère l'algèbre $\mathbb{C}[x]$, et l'application du disque unité fermé dans lui-même, définie par $z \mapsto \overline{z}$ n'est pas holomorphe, et donc ne peut être dans l'adhérence de $\mathbb{C}[x]$ (rappel : une limite uniforme de fonctions holomorphes est holomorphe).

FLEMMARD Elebeau n'y croit pas moi je crois que c'est ok

9.3 Approximation de fonctions mesurables

Théorème 419 (Lusin) *Soit f une application mesurable de $\mathbb{R}^n$ dans $\mathbb{C}$, dont le support est inclus dans E de mesure finie. Alors pour tout ϵ il existe g continue de $\mathbb{R}^n$ dans $\mathbb{C}$ telle que :*
• $\mu(\{x/f(x) \neq g(x)\}) < \epsilon$
• $sup|g(x)| \leq sup|f(x)|$

L'hypothèse "support inclus dans E de mesure finie" peut être oubliée. En effet, en résumé, on peut considérer les $g_i = f_{|U_i}$ pour i dans $\mathbb{N}$, avec U_i boule de centre 0 et de rayon i ; alors on peut "approcher" g_i par h_i égale à g_i sauf sur un ensemble de mesure $< \frac{\epsilon}{2^{i+1}}$, et avec $\sup h_i \leq \sup g_i$. En "recollant" les h_i on arrive à une fonction proche de f.

<u>Démonstration :</u>

• On laisse de côté la fonction nulle.

• Premier cas : $0 \leq f \leq 1$ et E compact, $\sup\ f = 1$.

- Donnons nous $\epsilon > 0$

- f est limite croissante d'une série croissante de fonctions simples (s_n), égales à des fonctions caractéristiques d'ensembles mesurables (E_n), comme expliqué pendant la preuve de la proposition 351.

- On considère R suffisamment grand pour que $E \subset B(0, R)$.

- Pour tout n on se donne K_n un compact et Ω_n un ouvert inclus dans $B(0, R)$ avec $K_n \subset E_n \subset \Omega_n \subset B(0, R)$, et $\mu(\Omega_n - K_n) < \epsilon.2^{-n}$ (possible grâce au théorème 327).

- Grâce au lemme d'Urysohn 411 ou 413, on se donne une fonction f_n telle que $\chi_{K_n} \leq f_n \leq \chi_{\Omega_n}$.

- Il ne reste plus qu'à sommer $g(x) = \sum_{n \in \mathbb{N}} \frac{f_n}{2^n}$.

- g est continue, comme somme convergeant normalement d'une série de fonctions continues (proposition 468).

- Sur le complémentaire de la réunion des $\Omega_n \setminus K_n$, $f = g$. Cette réunion est de mesure $\leq \epsilon$

- On a bien $|g| \leq \sup\ f$, puisque chaque f_n est majorée par 1.

D'où le résultat dans le cadre qu'on s'était donné, ie $0 \leq f < 1$ et E compact.

• Passons au cas d'une fonction f telle que $0 \leq f \leq 1$, sans hypothèse de compacité sur E.

- Donnons nous $\epsilon > 0$

- Il existe un compact K tel que la mesure de $E \setminus K$ soit inférieur à ϵ.

- On peut donc travailler sur la restriction de f à K pour construire g, et on a encore le résultat désiré.

• Passons maintenant au cas d'une fonction f bornée.

- En considérant f^+ et f^- [4], et en multipliant par une constante pertinente pour se ramener à des fonctions à valeurs dans $[0, 1]$, on a aussi le résultat désiré.

• On peut maintenant considérer le cas le plus général.

- Considérons $A_m = \{x \in \mathbb{R}^n / |f(x)| > m\}$. A_0 est de mesure finie ; on peut donc appliquer la proposition 328 pour conclure que la mesure des A_m tend vers 0 quand $m \to \infty$.

- On se donne donc m tel que $\mu(A_m) < \epsilon$.

- On peut donc appliquer le résultat partiel ci-dessus à la fonction $f.\chi_{A_m^c}$ (fonction produit de f par la fonction caractéristique du complémentaire de A_m, c'est à dire fonction égale à f sur le complémentaire de A_m et à 0 partout ailleurs).

D'où le résultat.$\square$

9.4 Approximation de fonctions mesurables bornées

Corollaire 420 *Soit f une application mesurable <u>bornée</u> de $\mathbb{R}^n$ dans $\mathbb{C}$, dont le support est inclus dans E de mesure finie. Alors f est limite simple presque partout d'une suite de fonctions g_n continues et bornées (par la même borne).*

FLEMMARD intérêt de l'hypothèse f bornée ?

<u>Démonstration :</u> Corollaire immédiat du théorème 419. Détaillons toutefois un peu :

• On peut se donner une suite (g_m) de fonctions continues telles que g_m soit bornées par la même borne que f et telles que g_m soit égale à f sauf sur un ensemble E_m de mesure au plus 2^{-m}, par le théorème 419.

• $\cap_{m \geq 0} \cup_{p \geq m} E_p$ est de mesure nulle (résultat facile à prouver directement, ou découlant facilement du premier lemme de Borel-Cantelli, voir partie **??**).

• On vient précisément d'écrire le résultat (si, si, regardez bien).$\square$

9.5 Dans les espaces C^k ou L^p

Pour connaitre la topologie usuelle sur $C^k(\mathbb{R}^n)$, on consultera la partie 20.3.

[4] $f^+(x) = max(f(x), 0)$, $f^-(x) = max(-f(x), 0)$

9.5.1 Densité des fonctions C^k à support compact dans $C^k(\mathbb{R}^n)$

Théorème 421 *L'ensemble des fonctions $C^k(\mathbb{R}^n, \mathbb{R})$ à support compact est dense dans $C^k(\mathbb{R}^n, \mathbb{R})$.*

Démonstration : On se donne une fonction $bulbe$ C^∞ telle que $\chi_{\overline{B}(0,1)} \leq \chi_{\overline{B}(0,2)}$, grâce au lemme 413.

On définit $bulbe_n(x) = bulbe(x/n)$

Il est clair que quelle que soit la fonction f dans $C^k(\mathbb{R}^n)$, $f.bulbe_{n+1}$ est égale à f sur la boule de rayon n et de centre 0. Tout compact étant inclus dans une telle boule, la convergence est clairement uniforme sur tout compact de $f.bulbe_{n+1}$ vers f, et pareil avec toutes les dérivées.$\square$

9.5.2 Densité de l'ensemble des fonctions continues à support compact dans $L^p(\mathbb{R}^n)$

Théorème 422 *L'ensemble des fonctions continues à support compact de $L^p(\mathbb{R}^n)$ est dense dans $L^p(\mathbb{R}^n)$, pour $p \neq \infty$.*

Démonstration : • Donnons-nous f dans $L^p(\mathbb{R}^n)$.

• On va chercher une fonction continue g à support compact telle que $\|f - g\|_p < \epsilon$.

• C'est facile, il suffit d'appliquer le théorème 419.$\square$

D'une part L^p est complet, d'autre part l'ensemble des fonctions continues à support compact est dense dans L^p. On en déduit donc que $L^p(\mathbb{R}^n)$ est le complété de l'ensemble des fonctions continues à support compact pour la norme distance associée à la distance L^p.

9.5.3 Densité de l'ensemble des fonctions C^∞ à support compact dans $C^k(\mathbb{R}^n)$

Théorème 423 *Pour tout k dans $\mathbb{N}$, $C_c^\infty(\mathbb{R}^n)$[a] est dense dans $C^k(\mathbb{R}^n)$.*

[a] Ensemble des fonctions C^∞ à support compact de $\mathbb{R}^n$ dans $\mathbb{R}$.

Démonstration :

• Grâce au théorème 421, il suffit de montrer que $C_c^\infty(\mathbb{R}^n)$ est dense dans $C_c^k(\mathbb{R}^n)$[5].

• Pour cela on fixe $n \in \mathbb{N}$, $k \in \mathbb{N}$, f dans $C_c^k(\mathbb{R}^n)$.

• On va utiliser les fonctions ρ et ρ_ϵ définies dans le corollaire 395 et la proposition 394.

• On définit alors (f_m) l'application $f * \rho_{1/n}$. On va montrer que la suite f_m converge uniformément sur tout compact vers f, et qu'il en est de même des dérivées[6].

[5] Ensembles des fonctions C^k à support compact.

[6] Ce qui est caractéristique de la convergence pour la topologie usuelle de C^k, voir 20.3.

• Calculons :

$$(f_m - f)(x) = \int (f(x-y) - f(x))\rho_{1/n}(y)d\mu(y)$$

$$= \epsilon^{-n} \int \rho(\frac{y}{\epsilon})(f(x-y) - f(x))dy$$

$$= \int \rho(z)(f(x-\epsilon.z) - f(x))dz \tag{9.1}$$

(avec le changement de variable $z = y/\epsilon$)

• f étant continue à support compact, elle est uniformément continue (voir théorème 198).

• Donc la quantité 9.1 tend vers 0, l'intégrale de ρ étant bornée, uniformément en x.

• On en déduit que $f_m \to f$ uniformément.

• Pour généraliser au cas des dérivées de f_m et de f, et pour justifier que f_m est bien C^∞, il suffit d'appliquer le théorème 392.□

9.5.4 Densité de l'ensemble des fonctions C^∞ à support compact dans $L^p(\mathbb{R}^n)$

Théorème 424 *L'ensemble des fonctions C^∞ à support compact de $\mathbb{R}^n$ est dense dans $L^p(\mathbb{R}^n)$, pour $p < \infty$.*

Démonstration :

• En vertu du théorème 422, il suffit de montrer que l'ensemble des fonctions C^∞ à support compact de $\mathbb{R}^n$ dans $\mathbb{R}$ est dense dans $L^p_c(\mathbb{R}^n)$[7]$\cap C^0(\mathbb{R}^n, \mathbb{R})$, pour $p < \infty$.

• On se donne donc une fonction f dans $L^p_c(\mathbb{R}^n)$, pour $p < \infty$.

• On va utiliser les fonctions ρ et ρ_ϵ définies dans le corollaire 395 et la proposition 394.

• On définit alors (f_m) l'application $f * \rho_{1/n}$.

• On pose maintenant q tel que $1/p + 1/q = 1$, le conjugué de p.

• Il ne reste plus qu'à calculer :

$$|(f_m - f)(x)|$$

$$\leq \int \rho(z)|f(x - z/n) - f(x)|d\mu(z)$$

$$\leq \int \rho(z)^{1/q}\rho(z)^{1/p}|f(x - z/n) - f(x)|d\mu(z)$$

$$\leq (\int (\rho(z)^{1/q})^q)d\mu(z))^{1/q}(\int (\rho(z)^{1/p})|f(x - z/n) - f(x)|)^p d\mu(z))^{1/p}$$

par l'inégalité de Hölder (théorème 398)

$$\leq (\int \rho(z)d\mu(z))^{1/q}(\int \rho(z)|f(x - z/n) - f(x)|^p d\mu(z))^{1/p}$$

[7] Ensemble des fonctions de L^p à support compact.

• On déduit du calcul précédent, puisque $\int \rho = 1$ par la proposition 394 :

$$N^p(f_m - f)^p \leq \int (\int \rho(z)|f(x - z/n) - f(x)|^p d\mu(z))d\mu(x)$$
$$\leq \int \rho(z) \int |f(x - z/n) - f(x)|^p d\mu(x)d\mu(z)$$

(par le théorème de Fubini 378)

$$\leq \int \rho(z) \underbrace{\|f(. - z/n) - f(.)\|_p^p}_{\to 0 \text{ (justifié ci-dessous)}} d\mu(z)$$

$$\to 0$$

par le théorème de convergence dominée de Lebesgue 402, puisqu'on a convergence simple vers 0, et convergence dominée par $z \mapsto \rho(z)2^p\|f\|_p$ qui est bien une fonction L^1.

• Le fait qu'étant donné z, $\|f(. - z/n) - f(.)\|_p \to 0$ est clair dans le cas où f est continue, puisqu'alors f est uniformément continue, et qu'on intègre $f(. - z/n) - f(.)$ sur un domaine borné (à z donné).

• Sinon f s'exprime de toute façon comme limite dans L^p de fonctions continues à support compact (par le théorème 422) ; donc il suffit d'appliquer l'inégalité triangulaire.□

9.6 Autre approche, dans les espaces L^p

On donne ici des résultats de densité dans les espaces L^p. Comme signalé dans le chapitre Fourier, $L_1 \cap L_2$ est dense dans L_2, ce qui a de grandes applications (prolongement de la transformée de Fourier).

9.6.1 Approximation dans L^1 par des fonctions semi-continues

Théorème 425 (Vitali-Carathéodory) *Soit f appartenant à $L^1(\mathbb{R}^n)$. Alors pour tout ϵ il existe s semi-continue supérieurement et i semi-continue inférieurement, avec s majorée et i minorée, telles que*

$$s \leq f \leq i$$

et

$$\int (i - s)d\mu < \epsilon$$

Démonstration : • On se ramène à f positive en considérant $f = f^+ - f^-$, avec f^+ et f^- des applications positives.
• Ensuite on écrit f comme limite d'une suite s_n de fonctions simples
• On écrit alors f comme limite d'une série de fonctions $f_n = s_n - s_{n-1}$
• f_n étant une combinaison linéaire de fonctions caractéristiques, on peut écrire $f(x) =$

$\sum_{i\in I}\lambda_i\chi_{E_i}(x)$
• f étant dans L^1, on a $\int f d\mu = \sum_{i\in I}\lambda_i\mu(E_i)$.
• De par les propriétés de la mesure de Lebesgue, E_i est compris entre K_i et U_i, avec $\mu(U_i \setminus K_i) \leq \epsilon/2^i$.
• Il ne reste plus qu'à sommer les $\lambda_i.\chi_{U_i}$ pour déterminer i, et un nombre fini suffisamment grand de $\lambda_i.\chi_{K_i}$ pour déterminer s.□

9.6.2 Approximation dans L^p pour $p < \infty$ par des fonctions en escalier à support compact

Définition 426 *Une application à valeurs dans $\mathbb{R}$ est dite* **en escalier** *si c'est une combinaison linéaire (finie) de fonctions caractéristiques d'intervalles de $\mathbb{R}$. Une application à valeurs dans un espace vectoriel F est dite* **en escalier** *si c'est une combinaison linéaire finie d'applications $f.\overrightarrow{x}$ avec f fonction caractéristique et $\overrightarrow{x}$ un élément de F.*

Théorème 427 *Les classes des fonctions en escalier à support compact constituent un sous-espace vectoriel dense de $L^p(\mathbb{R}^n)$ pour $p \in [1, +\infty[$.*

Démonstration : • Il est clair qu'il s'agit bien d'un sous-espace vectoriel .
• L'adhérence de ce sous-espace vectoriel contient les fonctions caractéristiques d'ouverts de mesure finie. En effet soit U un ouvert de mesure finie, $U = \cup_{i\in J} I_i$ avec J au plus dénombrable (par exemple $U = \cup_{i\in\mathbb{N}} U \cap B(0,i)$), avec les I_i des ouverts bornés, et donc $f_n = \sup_{i\in[1,n]}\chi_{I_i}$ est majorée par χ_U, converge simplement vers χ_U, et donc converge vers χ_U pour la norme de L^p par le théorème 402.
• L'adhérence de ce sous-espace vectoriel contient aussi les fonctions caractéristiques d'ensembles mesurables de mesure finie, comme on s'en convaincra aisément en consultant le théorème 327, montrant que la fonction caractéristique d'un ensemble mesurable est limite de fonctions caractéristiques d'ouverts.
• L'adhérence de ce sous-espace vectoriel contient aussi les fonctions simples. En effet ces fonctions sont des combinaisons linéaires des fonctions précédentes.
• L'adhérence de ce sous-espace vectoriel contient enfin toutes les fonctions intégrables positives, puisque celles-ci sont limites de suites de fonctions simples (voir la proposition 351) et que le théorème 402 garantit la convergence dans L^p, et donc toutes les fonctions intégrables négatives, et donc toutes les fonctions intégrables. □

9.6.3 Approximation dans L^p pour $p < \infty$ par des fonctions C^∞ à support compact

Théorème 428 *Pour $1 \leq p < +\infty$ ($p \neq +\infty$), les classes des fonctions indéfiniment dérivables à support compact constituent un sous-espace vectoriel dense de $L^p(\mathbb{R}^n)$.*

Démonstration : Par le théorème précédent, il suffit d'approcher la fonction caractéristique d'un ouvert borné par des fonctions C^∞ à support compact.

Pour cela on utilise une propriété fondamentale de la mesure de Lebesgue, qui est le fait qu'un ensemble mesurable est compris, pour tout $\epsilon > 0$, entre un compact K et un ouvert U tels que $\mu(U \setminus K) < \epsilon$.

Il ne reste plus qu'à trouver une fonction θ C^∞, telle que

$$\chi_K \leq \theta \leq \chi_U$$

Construire une telle fonction est précisément l'objet d'une variante du lemme d'Urysohn, voir lemme 413.□

9.6.4 Approximation de fonctions tendant vers 0 en $\pm\infty$ dans L^∞ par des fonctions C^∞ à support compact

Théorème 429 *L'espace vectoriel des fonctions C^∞ à support compact est dense dans le sous-espace vectoriel de L^∞ des fonctions bornées tendant vers 0 en $\pm\infty$.*

Démonstration : Laissée au lecteur.□

Chapitre 10

Fourier

Il est nécéssaire pour bien faire d'avoir lu la partie29 et la partie8.

10.1 Séries trigonométriques

Définition 430 *Une fonction f est dite T-**périodique** si pour tout x $f(x+T) = f(x)$.*
Pour p fini on note $\mathfrak{L}^p$ l'espace $L^P_{\mathbb{C}}([-\pi,\pi])$ [a] pour la mesure de Lebesgue, mais avec une norme divisée par 2π, c'est à dire que $\|f\| = (\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x)|^p.dx)^{\frac{1}{p}}$.
On note $\mathfrak{L}^\infty$ l'espace $L^\infty_{\mathbb{C}}([-\pi,\pi])$ pour la mesure de Lebesgue.
*On appelle **polynôme trigonométrique** une application de la forme $t \mapsto a_0 + \sum_{i=1}^{N} a_i.cos(i.t) + b_i.sin(i.t)$, pour $N \in \mathbb{N}$, $a_i \in \mathbb{C}$, $b_i \in \mathbb{C}$.*
Le produit scalaire hermitien usuel sur $\mathfrak{L}^2$ est l'application $(f,g) \mapsto \frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{f(t)}.g(t).dt$. Il s'agit bien d'un produit scalaire hermitien.
On note u_n l'application $t \mapsto e^{int}$, pour $n \in \mathbb{Z}$. Nous verrons plus loin qu'il s'agit d'une base hilbertienne de $\mathfrak{L}^2$.

[a] attention, ne pas confondre $\mathfrak{L}^p$ et $\mathcal{L}^p$, ce dernier désignant l'ensemble des fonctions dont la puissance p-ième est intégrable, AVANT de quotienter pour la relation d'égalité presque partout - par passage au quotient on obtient L^p, et en spécialisant aux fonctions définies sur $[-\pi,\pi]$ on obtient $\mathfrak{L}^p$.

Remarques 431 • *On identifiera par la suite (sans préavis !) une fonction définie sur $[-\pi,\pi]$ à une fonction périodique de période 2π. A part dans les cas où la continuité est importante, on se préoccupera peu du problème de définition en π, puisque l'on travaillera généralement sur des propriétés vraies presque partout pour la mesure de Lebesgue.*

• *Il y a une renormalisation (la division par 2π pour la mesure de Lebesgue) pour p fini et pas pour p infini ; cela ne serait pas le cas si l'on raisonnait sur $[0,1]$ au lieu de $[-\pi,\pi]$.*

• *On peut réécrire un polynône trigonométrique sous la forme* $t \mapsto \sum_{i=-N}^{N} c_n.e^{int}$*, avec* $N \in \mathbb{N}$ *et* $c_i \in \mathbb{C}$ *(et réciproquement, une telle fonction est toujours un polynôme trigonométrique).*
• *Un polynône trigonométrique est* 2π*-périodique.*

Théorème 432 *La famille* $(u_n)_{n\in\mathbb{Z}}$ *est une base hilbertienne de* $\mathfrak{L}^2$*.*

Démonstration : • Il est clair qu'il s'agit d'une famille orthonormale.
• Il suffit d'appliquer 9.6.3 et de montrer que l'ensemble des polynômes trigonométriques est dense dans l'ensemble des fonctions continues de $[-\pi, \pi]$ dans $\mathbb{C}$ (on utilise la densité de $C^0([-\pi, \pi], \mathbb{C})$ dans $\mathcal{L}^2$). Pour cela, on procède comme suit :
- on considère f une fonction continue de $[-\pi, \pi]$ dans $\mathbb{C}$
- on définit $P_n(t) = \frac{(\frac{1+cos(t)}{2})^n}{\int_{-\pi}^{\pi}(\frac{1+cos(t)}{2})^n.dt}$

On constate (grâce à la linéarisation) que P_n est un polynôme trigonométrique positif, d'intégrale 1, convergeant uniformément vers 0 sur $[-\pi, -\delta] \cup [\delta, +\pi]$ pour tout $\delta \in]0, \pi[$. Intuitivement, P_n tend vers une fonction comportant une pointe en 0 et nulle partout ailleurs.
- on définit $f_n(x) = \int_{-\pi}^{\pi} f(t)P(x-t).dx$. En remarquant que le translaté d'un polynôme trigonométrique est un polynôme trigonométrique, on montre que f_n est un polynôme trigonométrique.
- on montre alors que la norme infinie de $f_n - f$ tend vers 0, et donc la norme 2 aussi puisque $[-\pi, \pi]$ est de mesure finie (oui, je passe sous silence le calcul, désolé...).□

10.2 Séries de Fourier d'une fonction périodique

Pour étudier une fonction périodique, on se ramène au cas d'une période 2π, et on la considère définie sur $[-\pi, \pi]$.

Définition 433 *Soit f dans $\mathfrak{L}^1$. Alors on définit les* **coefficients de Fourier de** *f pour $n \in \mathbb{Z}$ par*

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i.n.t}.f(t).dt$$

On appelle **noyau de Dirichlet d'ordre** *n l'application $x \mapsto \sum_{i=-n}^{n} u_i(x)$.*
On le note D_n.
On appelle **noyau de Féjer d'ordre** *n l'application*

$$x \mapsto \frac{\sum i = 0^{n-1} D_i}{n}$$

On le note K_n.
On note $s_n(f)$ et on appelle **somme de Fourier d'ordre** *n la somme $\sum_{i=-n}^{n} \hat{f}(n) u_n$.*
On note $\sigma_n(f)$ et on appelle **somme de Féjer d'ordre** *n la somme $\frac{\sum_{i=0}^{n-1} s_i}{n}$.*
On appelle **série de Fourier** *associée à f la série*

$$\sum_{n=-\infty}^{+\infty} \hat{f}(n).e^{i.n.t}$$

$\frac{1}{2\pi} \int_{-\pi}^{\pi} u_n = 1$ si $n = 0$ et 0 sinon, ce qui implique que pour tout $n \geq 0$ on a $\frac{1}{2\pi} D_n = 1$ et pour tout $n \geq 1$ on a $\frac{1}{2\pi} K_n = 1$.

On a $s_n(f) = \frac{1}{2\pi} D_n * f$ et $\sigma_n(f) = \frac{1}{2\pi} K_n * f$ (avec par définition $(g * f)(x) = \int_{-\pi}^{\pi} g(t) f(x-t) dt$, produit de convolution) ce qui justifie le terme de noyau.

Bien noter le signe "-" dans l'exponentielle de la formule.
On remarque que si $f \in \mathfrak{L}^2$, alors $f \in \mathfrak{L}^1$, et $\hat{f}(n) = (u_n, f)$.

Proposition 434 *On a isomorphisme isométrique entre $\mathcal{L}^2$ et $l^2(\mathbb{Z})$, donné par $f \mapsto \hat{f}$.*

<u>Démonstration :</u> C'est simplement une reformulation du théorème 1110. □

Proposition 435 *Toute fonction f dans $\mathcal{L}^2$ est somme de sa série de Fourier pour $\mathcal{L}^2$ (c'est à dire que la série de Fourier de f tend vers f pour la norme 2 et pour f dans $\mathcal{L}^2$).*

<u>Démonstration :</u> C'est simplement une reformulation de la relation de Parseval (théorème 1108).□

Il n'y a pas convergence simple de la série de Fourier vers f, même si f est continue ! $\square$

Un problème majeur va être de montrer des résultats similaires dans $\mathcal{L}^1$.

Proposition 436

$$D_n(x) = \frac{sin((n+1/2)x)}{sin(x/2)}$$

$$K_n(x) = \frac{1}{n}\left(\frac{sin(nx/2)}{sin(x/2)}\right)^2$$

Démonstration :

Si x n'est pas multiple de 2π, alors

$$D_n(x) = u_{-n}(x)\sum_{k=0}^{2n} u_k(x)$$

$$= u_{-n}(x)\frac{(e^{ix})^{2n+1}-1}{e^{ix}-1}$$

$$= \frac{e^{ix(n+1/2)}}{e^{ix/2}}\frac{e^{ix(n+1/2)}-e^{-ixn+1/2}}{e^{ix/2}-e^{-ix/2}}$$

$$= \frac{sin((n+1/2)x)}{sin(x/2)}$$

D'où le résultat sur D_n (si x est multiple de 2π, $D_n(x) = 2n+1$, qui est l'unique prolongement par continuité de $\frac{sin(n+\frac{1}{2}x)}{sin(x/2)}$).

Toujours pour x non multiple de 2π,

$$K_n(x) = \frac{1}{n}\sum_{k=0}^{n-1}\frac{sin((k+1/2)x)}{sin(x/2)}$$

$$= \frac{1}{nsin(x/2)}Im\left(\sum_{k=0}^{n-1}e^{i(k+1/2)x}\right)$$

$$= \frac{1}{nsin(x/2)}Im\left(e^{ix/2}\frac{e^{inx}-1}{e^{ix}-1}\right)$$

$$= \frac{sin(nx/2)^2}{nsin(x/2)^2}$$

D'où le résultat. $\square$

Deux résultats distincts, prouvés de manières similaires :

Théorème 437 (Théorème de Fejer) • *Soit f périodique continue de période 2π. Alors pour tout n $\|\sigma_n(f)\|_\infty \leq \|f\|_\infty$ et $\|\sigma_n(f)-f\|_\infty \to 0$ pour $n \to \infty$.*
• *Soit $f \in \mathfrak{L}^p$, avec $p \in [1,\infty[$[a]. Alors pour tout n $\|\sigma_n(f)\|_p \leq \|f\|_p$ et $\|\sigma_n(f) - f\|_p \to 0$ pour $n \to \infty$.*

[a]Bien noter que ∞ est exclus !

Démonstration : Cette preuve est détaillée dans le livre [22, p81]. Elle utilise à la fois l'inégalité de Hölder et le théorème de Fubini, et le dernier résultat donné sur le noyau de Féjer.□

Un autre résultat, de convergence ponctuelle ce coup-ci :

Théorème 438 (Théorème de Dirichlet) *Si f est L^1 et si f admet une* pseudo-dérivée *à droite et à gauche en x, alors*

$$\sigma_n(f)(x) \to \frac{1}{2}(lim_{t\to x, t<x} f(t) + lim_{t \to x, t>x} f(t))$$

⚠ *Il ne s'agit pas nécéssairement de dérivées à gauche ou à droite, on peut se contenter d'avoir f dans L^1, et admettant en x une limite à gauche et à droite $f_-(x)$ et $f_+(x)$; alors les "pseudo-dérivées" à gauche et à droite sont*

$$f'_g(x) = lim_{t\to x, t<x} \frac{f(t) - f_-(x)}{t-x}$$

$$f'_d(x) = lim_{t\to x, t>x} \frac{f(t) - f_+(x)}{t-x}$$

Démonstration : On renvoie à [15] pour une preuve très claire.□

10.3 Transformation de Fourier

Définition 439 *On se donne f dans $L^1_{\mathbb{C}}(\mathbb{R})$, et on note pour x dans $\mathbb{R}$*

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x).e^{-i.x.t}.dt$$

$\hat{f}$ est appelée **transformée de Fourier** *de f (plus précisément il s'agit de la* **transformée de Fourier** *L^1 de f).*
On note $\mathcal{C}$ l'ensemble des $x \in \mathbb{C}$ tels que $|x| = 1$.

Si f est dans L^1, alors $\hat{f}$ est continue et tend vers 0 en $\pm\infty$.

Proposition 440 (Quelques propriétés de la transformée de Fourier) *Soit $f \in L^1(\mathbb{R})$.*
- *Avec $g : x \mapsto f(x)e^{i\lambda x}$, pour $\lambda \in \mathbb{R}$, $\hat{g}(t) = \hat{f}(t - \lambda)$.*
- *Avec $g : x \mapsto f(x - \lambda)$, pour $\lambda \in \mathbb{R}$, $\hat{g}(t) = \hat{f}(t)e^{-i\lambda t}$.*
- *Si g est L^1 et si $h = f * g$, alors $\hat{h}(t) = \hat{f}(t)\hat{g}(t)$.*

Démonstration : Les deux premiers • sont clairs. Le troisième découle immédiatement du théorème de Fubini 378.□

Les deux théorèmes suivants, fondamentaux, ne seront pas prouvés ici. Ils sont ardus, et prouvés rigoureusement dans [16].

Théorème 441 (Théorème d'inversion) *Si f et $\hat{f}$ appartiennent tous deux à L^1, alors*

$$g : x \mapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(t)e^{ixt}dt$$

est continue, tend vers 0 en $+\infty$ ou $-\infty$ et est égale à f presque partout

Théorème 442 (Théorème de Plancherel) *La transformation de Fourier s'étend en une* **transformation de Fourier** *L^2, définie comme l'unique application $f \mapsto \hat{f}$ de L^2 dans L^2 telle que :*
- *elle coïncide avec la transformée de Fourier L^1 sur $L^1 \cap L^2$*
- *c'est une isométrie de L^2 dans L^2*

Elle vérifie en outre :
- *c'est un isomorphisme d'espaces de Hilbert entre L^2 et L^2*
- *Théorème d'inversion L^2 :*

$$lim_{M\to\infty}\|f_A - \hat{f}\|_2 = 0$$

avec

$$f_M(t) = \frac{1}{\sqrt{2\pi}} \int_{-M}^{M} f(x)e^{-ixt}dx$$

et

$$lim_{M\to\infty}\|\hat{f}_M - f\|_2 = 0$$

avec

$$\hat{f}_M(x) = \frac{1}{\sqrt{2\pi}} \int_{-M}^{M} f(t)e^{-ixt}dt$$

on peut aussi écrire que l'application de $L^1 \cap L^2$ dans $L^1 \cap L^2$ qui à f associe $\tilde{f}$ avec $\tilde{f}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)exp(ixt)dt$ s'étend en une isométrie de L^2 dans L^2, et $f \mapsto \tilde{f}$ est l'inverse de $f \mapsto \hat{f}$ au sens où pour toute f L^2 $\hat{\tilde{f}} = \tilde{\hat{f}} = f$ presque partout.

10.4 Applications des séries de Fourier

10.4.1 Calcul de $\sum_{n=0}^{\infty} \frac{1}{n^2}$

Proposition 443

$$\sum_{n=1}^{\infty} 1/n^2 = \pi^2/6$$

<u>Démonstration :</u>

On applique simplement la formule de Parseval (théorème 1108) à la transformée de Fourier de l'application identité de $[-\pi, \pi[$ dans lui-même.

Pour tout $k \in \mathbb{Z}$, définissons

$$c_k = 2\pi \hat{f}(k)$$

$$c_k = \int_{-\pi}^{\pi} x e^{ikx} dx$$

$$c_0 = 0 \text{c'est à dire} \hat{f}(0) = 0$$

$$k \neq 0 \Rightarrow c_k = [\frac{xe^{ikx}}{ik}]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} e^{ikx} dx$$

$$= \frac{2\pi(-1)^k}{ik}$$

donc pour $k \neq 0$ $\hat{f}(k) = (-1)^k/(ik)$.

Donc par la formule de Parseval :

$$(\sum_{k \in \mathbb{Z}} |\hat{f}(k)|^2 = \sum_{k<0} 1/k^2 + \sum_{k>0} 1/k^2$$

$$= \sum_{k=1}^{+\infty} 2/k^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 dx = \pi^2/3$$

Et $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6.\square$

Exemple Maple
$> Sum(1/k\hat{}\ 2, k = 1..infinity)$
$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$
$> value(\%)$
$$\frac{1}{6}\pi^2$$

10.4.2 Exemple de développement en série de Fourier : fonction créneau, fonction identité par morceaux

Exemple Maple
$> \ restart; fourierc := (f,n)- > (int(f(t) * cos(n * t), t = -Pi..Pi)); fouriers := (f,n)- > (int(f(t) * sin(n * t), t = -Pi..Pi));$ $$fourierc := (f,\, n) \to \int_{-\pi}^{\pi} f(t) \cos(n\, t)\, dt$$ $$fouriers := (f,\, n) \to \int_{-\pi}^{\pi} f(t) \sin(n\, t)\, dt$$ $> \ serie\_fourier := (f,n)- > sum(fouriers(f,k) * sin(k * t), k = 1..n) + sum(fourierc(f,k) * cos(k * t), k = 1..n) + fourierc(f,0)/2;$ $$serie\_fourier :=$$ $$(f,\, n) \to (\sum_{k=1}^{n} \mathrm{fouriers}(f,\, k) \sin(k\, t)) + (\sum_{k=1}^{n} \mathrm{fourierc}(f,\, k) \cos(k\, t))$$ $> \ p := array(0..5); for\ i\ from\ 1\ by\ 2\ to\ 9\ do\ p[(i-1)/2] := serie\_fourier(Heaviside, i) od; p[5] := Heaviside(t);$ $> plot(p, t = -Pi..Pi, title =$ $"Approximations\ d'une\ fonction\ creneau\ par\ seriesde\ Fourier");$ $> \ q := array(0..5); for\ i\ from\ 1\ by\ 2\ to\ 9\ do\ q[(i-1)/2] := serie\_fourier(x- > x, i) od;$ $> \ plot(q, t = -Pi..Pi, title =$ $"Approximations\ d'une\ fonction\ lineaire\ par\ morceaux$ $par\ series\ de\ Fourier");$

Il faut bien noter que les coefficients de Fourier étant calculée simplement en fonction de la période $[-\pi, \pi]$, on ne se préoccupe que de la valeur de la fonction sur ces valeurs, d'où la simple définition $x- > x$ ou $Heaviside$, où je ne me préoccupe pas de périodiciser la fonction.

Chapitre 11

Calcul différentiel

Il est recommandé de bien maîtriser la partie5 avant d'étudier cette partie, et notamment les espaces de Banach.

11.1 Introduction

11.1.1 Généralités

Définition 444 *Soient E et F des espaces vectoriels normés et U un ouvert de E. Soit une application $f : U \to F$, on dit que f est* **différentiable** *(ou* **dérivable***) en $x \in U$ s'il existe une application linéaire continue ϕ de E dans F telle que*

$$lim_{h \to 0} \frac{f(x+h) - f(x) - \phi(h)}{\| h \|} = 0$$

On appelle ϕ la **différentielle** *ou* **dérivée** *de f en x, on la note $Df(x)$.*
f est dite **différentiable** *si elle est différentiable en tout point de U.*

Proposition 445 • *Si f est dérivable en x, alors f est continue en x.*
• *La dérivée de f est unique et*

$$Df(x)(h) = lim_{t \to 0} \frac{f(x+t.h) - f(x)}{t}$$

• *La notion de dérivée ne dépend que des topologies et pas des normes (du moment qu'elles définissent la même topologie) ; si deux normes sont équivalentes, alors une fonction différentiable pour l'une est différentiable pour l'autre, et la différentielle est la même.*
• *$D(\lambda f + \mu g)(x) = \lambda Df(x) + \mu Dg(x)$*
• *Si $E = K$ corps associé aux espaces vectoriels E et F, alors la différentiabilité équivaut à l'existence de la limite pour $t \to 0$ de $\frac{f(x+t)-f(x)}{t}$. On note alors cette limite $f'(x)$, et $Df(x)(t) = t.f'(x)$.*
• *L'application qui à une application différentiable en x_0 associe sa différentielle en X_0 est une application linéaire de l'espace vectoriel des applications de E dans F différentiables en x_0 dans l'espace vectoriel des applications linéaires de E dans F.*
• *Une application linéaire continue f est différentiable en tout point x_0 et $Df(x_0)(h) = f(h)$.*

Démonstration : Un peu laborieux mais rien de bien difficile, en notant $\epsilon(x,h) = f(x+h) - f(x) - Df(x)(h)$, pour h suffisamment petit pour que $x + h$ appartienne à U.□

Définition 446 *Soient E et F des espaces vectoriels normés et U ouvert de E. f de U dans F est* **de classe** C^1 *si elle est différentiable et si l'application qui à x associe la différentielle de f en x est continue (voir* 5.2.2 *pour un rappel de la topologie usuelle sur $\mathcal{L}(E,F)$) .*

Proposition 447 • *Si f est constante sa dérivée est nulle partout, f est C^1.*
• *Si ϕ de E dans F est linéaire continue, alors ϕ est C^1 avec $D\phi(x) = \phi$, pour tout x.*
• *Si f de $E_1 \times E_2 \times ... \times E_n$ dans F est multilinéaire continue, alors f est C^1, et on a*

$$Df(x_1,...,x_n)(h_1,...,h_n) = \sum_{i=1}^{n} f(x_1,x_2,...,x_{i-1},h_i,x_{i+1},...,x_n)$$

Démonstration : pas dur, tout ça !□

Théorème 448 (Différentielle de fonctions composées) *Soit E,F et G des espaces vectoriels normés , et U et V des ouverts de E et F respectivement. Si f de U dans V est différentiable en x et g de V dans G est différentiable en $f(x)$, alors la composée $g \circ f$ est différentiable en x et a pour différentielle*

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$$

Si g et f sont C^1 alors $g \circ f$ est C^1.

Démonstration : on écrit comme pour d'autres preuves $f(x+h) = f(x) + Df(x)(h) + \epsilon(h) \parallel h \parallel$, et de même $g((f(x)+k)$, et on calcule...
Pour voir que la composée est C^1, il suffit de voir que la différentielle est la composée de 3 fonctions continues.□

Définition 449 (Isomorphisme d'espaces normés) *Un* **isomorphisme de l'espace vectoriel normé** E **sur l'espace vectoriel normé** F *est une application $\phi : E \to F$ linéaire continue et bijective d'inverse continue. On note $Isom(E,F)$ le sous-ensemble de $\mathcal{L}(E,F)$ formé des isomorphismes de E dans F.*

Théorème 450 *Soient E et F des espaces de Banach. Le sous-ensemble $Isom(E,F)$ est ouvert dans $\mathcal{L}(E,F)$. L'application $inv : Isom(E,F) \to Isom(F,E)$ qui à u associe u^{-1} est C^1 avec $Dinv(u)(v) = -u^{-1}.v.u^{-1}$.*

Démonstration : Soit u_0 un isomorphisme de E vers F ; alors $u_0 + v = u_0.(Id +$

$u_0^{-1}.v)$. Si $\| v \| < \| u_0^{-1} \|^{-1}$, on a $\| u_0^{-1}.v \| v \| < 1$, et donc la série

$$\sum_{i=0}^{+\infty} (-u_0^{-1}.v)^i$$

est convergente dans $\mathcal{L}(E,F)$ et donne un inverse à $(Id + u_0^{-1}.v)$.
Donc pour $\| v \| < \| u_0^{-1} \|^{-1}$, l'inverse de $u_0 + v$ est $\sum_{i=0}^{+\infty} (-u_0^{-1}.v)^n.u_0^{-1}$.
On a alors $(u_0 + v)^{-1} = u_0^{-1} - u_0^{-1}.v.u_0^{-1} + \sum_{i=2}^{+\infty} (-u_0^{-1}.v)^i.u_0^{-1}$; or la quantité $\frac{(-u_0^{-1}.v)^i.u_0^{-1}}{\|v\|}$ tend vers 0 quand v tend vers 0. L'application $Dinv$ est continue comme composée de fonctions continues, comme on s'en convaincra aisément.□

Proposition 451 (Liens entre différentiabilité Banach sur $\mathbb{R}$ et sur $\mathbb{C}$) *Un $\mathbb{C}$-espace vectoriel peut aussi être considéré comme un $\mathbb{R}$-espace vectoriel ; il suffit de restreindre le produit par un scalaire à un produit par un scalaire réel. En remplaçant E et F en tant que $\mathbb{C}$-espaces vectoriels par E et F en tant que $\mathbb{R}$-espaces vectoriels , une fonction différentiable pour $\mathbb{C}$ est différentiable pour $\mathbb{R}$; par contre la réciproque n'est pas garantie dans le cas général ; il faut que la différentielle sur $\mathbb{R}$ soit définie et que la différentielle sur $\mathbb{R}$ soit linéaire et continue en tant qu'application entre $\mathbb{C}$-espaces vectoriels (c'est à dire appartienne à $\mathcal{L}_{\mathbb{C}}(E,F)$).*

Démonstration : Sans grande difficulté, et laissée au lecteur.□

11.1.2 Applications à valeurs dans un produit d'espaces vectoriels normés

Proposition 452 *Soit f une application de E dans $F_1 \times F_2$, avec E, F_1 et F_2 des espaces vectoriels normés , et soient f_1 et f_2 ses composantes. Alors f est différentiable en x si et seulement si f_1 et f_2 sont différentiables en x, et $Df(x)(h) = (Df_1(x)(h), Df_2(x)(h))$.
En outre, f est C^1 si et seulement si f_1 et f_2 sont C^1.*

Démonstration : Il suffit de voir que les projections canoniques de F sur F_1 et F_2 sont C^1 car linéaires et continues, et que l'injection canonique de F_1 dans F ou de F_2 dans F sont linéaires continues, donc elles aussi C^1.□

Corollaire 453 (Formule de Leibnitz) *Si $f_1 : U \to F_1$ et $f_2 : U \to F_2$ sont différentiables en x et $B : F_1 \times F_2 \to G$ et B de $F_1 \times F_2$ dans G est bilinéaire continue, alors $B(f_1, f_2) : x \mapsto B(f_1(x), f_2(x))$ est différentiable en x et*

$$DB(f_1, f_2)(x)(h) = B(f_1(x), Df_2(x)(h)) + B(Df_1(x)(h), f_2(x))$$

En outre si f_1 et f_2 sont C_1 alors $B(f_1, f_2)$ est C^1.

11.1.3 Applications de plusieurs variables et dérivées partielles

Proposition 454 (Définition des dérivées partielles) *Soit U un ouvert du produit $E_1 \times E_2$ de deux espaces vectoriels normés , soit $f : U \to F$, avec F espace vectoriel normé , et f différentiable en $a = (a_1, a_2)$. Alors les deux applications partielles $x_1 \mapsto f(x_1, a_2)$ et $x_2 \mapsto f(a_1, x_2)$ sont différentiables respectivement en a_1 et a_2. On note les deux différentielles obtenues respectivement $D_1 f(a_1, a_2)$ et $D_2 f(a_1, a_2)$, ou bien $\frac{\delta f}{\delta x_1}$ et $\frac{\delta f}{\delta x_2}$, et on les appelles respectivement* **première dérivée partielle** *et* **deuxième dérivée partielle**. *On a alors*

$$Df(a_1, a_2)(h_1, h_2) = D_1 f(a_1, a_2)(h_1) + D_2 f(a_1, a_2)(h_2)$$

On peut généraliser de même à un produit fini d'espaces vectoriels normés ; si f est différentiable en $(a_1, a_2, ..., a_n)$, alors pour tout i dans $[1, n]$ $x \mapsto f(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n)$ est différentiable en a_i, sa différentielle en a_i est noté $\frac{\delta f}{\delta x_i}(a)$, et

$$Df(a_1, ..., a_n)(h_1, ..., h_n) = \sum_{i=1}^{n} \frac{\delta f}{\delta x_i}(a)(h_i)$$

Démonstration : Facile !□

⚠ Il n'y a pas de réciproque dans le cas général ! Même si toutes les dérivées partielles sont définies la différentielle n'est pas nécéssairement définie. Par contre si les différentielles partielles sont continues alors on peut conclure que f est différentiable et même C^1 (voir partie 11.2.3).

Théorème 455 *Soit E un espace de Banach , U un ouvert de E et $F_1, ..., F_n$ des espaces de Banach . Soit f de U dans $F_1 \times ... \times F_n$.*
On note $p_i(x_1, ..., x_n) = x_i$.
Alors f est différentiable en x si et seulement si chacune des application f_i de E dans F_i $y \mapsto p_i(f(x))$ est différentiable en x et on a alors

$$Df(x)(h) = (Df_1(x)(h_1), Df_2(x)(h_2), ..., Df_n(x)(h_n))$$

Démonstration : • Le sens "seulement si" est clair ; une composée d'applications différentiables est différentiable.
• Le sens "si" et l'égalité annoncée s'obtiennent simplement en considérant

$$f = \sum_{i=1}^{n} u_i \circ f_i$$

avec $u_i(x) = (0, ..., 0, x, 0, , , 0)$.□

Définition 456 *Eventuellement on peut avoir U ouvert de $\mathbb{R}^n$ et $F = \mathbb{R}^m$; on peut alors noter la différentielle sous forme matricielle ; cette matrice est appelée* **matrice jacobienne**. *Elle est de la forme :*

$$\begin{pmatrix} \frac{\delta f_1}{\delta x_1} & \frac{\delta f_1}{\delta x_2} & \dots & \frac{\delta f_1}{\delta x_n} \\ \frac{\delta f_2}{\delta x_1} & \frac{\delta f_2}{\delta x_2} & \dots & \frac{\delta f_2}{\delta x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta f_m}{\delta x_1} & \frac{\delta f_m}{\delta x_2} & \dots & \frac{\delta f_m}{\delta x_n} \end{pmatrix}$$

Si $n = m$, la matrice jacobienne est carrée, on peut donc considérer son déterminant, appelé **jacobien** *de f.*

11.2 Le théorème des accroissements finis

11.2.1 Résultats principaux

Définition 457 *Soient a et b dans $\mathbb{R}$, avec $a < b$, et F un espace vectoriel normé . Une application f de $[a, b]$ dans F est dite* **dérivable** *à droite en x appartenant à $[a, b[$ si la limite à droite $lim_{h\to 0, h>0} \frac{f(x+h)-f(x)}{h}$ existe ; on l'appelle alors* **dérivée à droite** *de f en x.*

Théorème 458 (Théorème des accroissements finis) *Soient a et b dans $\mathbb{R}$ avec $a < b$, et F un espace vectoriel normé . On suppose que les deux fonctions $f : [a, b] \to F$ et $g : [a, b] \to \mathbb{R}$ sont continues sur $[a, b]$ et dérivables à droite sur $[a, b] \setminus D$ avec D au plus dénombrable. Si, pour tout $t \in [a, b] \setminus D$ on a $\| f'_d(t) \| \leq g'_d(t)$, alors $\| f(b) - f(a) \| \leq g(b) - g(a)$.*

<u>Démonstration :</u> • On note $D = \{d_1, d_2, d_3, ...\}$ avec $d_i < d_{i+1}$
• On se donne $\epsilon > 0$.
• On considère E l'ensemble des x tels que

$$\|f(x) - f(a)\| > g(x) - g(a) + \epsilon.(x - a) + \epsilon. \sum_{d_i<x} \frac{1}{2^i} \tag{11.1}$$

• E est ouvert
• Soit x_0 la borne inf de E
• $x_0 > a$, car pour x assez petit, l'inégalité 11.1 est fausse
• $x_0 \notin E$, car E est ouvert ; donc

$$\|f(x_0) - f(a)\| \leq g(x) - g(a) + \epsilon.(x - a) + \epsilon. \sum_{d_i<x} \frac{1}{2^i} \tag{11.2}$$

• $x_0 \neq b$, car E est ouvert et n'est donc pas réduit à un singleton

• On va maintenant distinguer deux cas, selon que c appartienne à D ou non.
- Si $x_0 \in D$, alors $x_0 = d_{i_0}$ pour un certain i_0. Alors par continuité pour $x > x_0$ suffisamment proche de x_0, on a

$$\|f(x) - f(x_0)\| \leq g(x) - g(a) + \epsilon.(x - c) + \epsilon.2^{-i_0} \tag{11.3}$$

(par continuité). Or pour $x > x_0$ on a

$$\sum_{d_i < x} \frac{\epsilon}{2^i} \geq \epsilon.2^{-i_0} + \epsilon \sum_{d_i < x_0} 2^{-i} \tag{11.4}$$

. En sommant 11.2, 11.4 et 11.3, on obtient que x ne vérifie par 11.1 pour x assez proche de x_0 suffisamment proche de x_0 ; ce qui est contradictoire avec E ouvert.
- Si $x_0 \notin D$, par hypothèse $\|f'_d(x_0)\| \leq g'_d(x_0)$. Donc pour x suffisamment proche de x_0 et $x > x_0$ on a

$$\frac{1}{x - x_0}\|f(x) - f(x_0)\| \leq \frac{1}{x - x_0} g(x) - g(x_0) + \epsilon$$

et donc

$$\|f(x) - f(x_0)\| \leq g(x) - g(x_0) + \epsilon.(x - x_0)$$

en additionnant avec 11.2 on obtient alors que $x \notin E$ pour x suffisamment proche de x_0 et $x > x_0$; ce qui est contradictoire puisque E est ouvert.
• On a alors montré que E est vide, et donc il suffit de faire tendre ϵ vers 0 pour avoir le résultat désiré.□

Corollaire 459 *On a le même résultat en remplaçant les dérivées à droite par les dérivées à gauche.*

Démonstration : Facile, en remplaçant x par $-x$!□

Corollaire 460 *Une fonction continue de $\mathbb{R}$ dans un espace vectoriel normé dont la dérivée existe et est nulle sauf sur un ensemble au plus dénombrable est constante.*

Démonstration : Facile !□

on utilise le théorème des accroissements finis pour les théorèmes 482, 471, 478, 370, 581, 724.

On l'utilise aussi pour montrer qu'une fonction dérivable à dérivée bornée est lipschitzienne, ou bien qu'une fonction C^1 est localement lipschitzienne → applications aux équations différentielles.

Proposition 461 *Une fonction f continue de $\mathbb{R}$ dans $\mathbb{R}$ dont la dérivée existe et est positive sauf sur un ensemble au plus dénombrable est croissante.*

Démonstration : L'astuce réside dans le fait que la fonction f dont il est question ici doit jouer le rôle de la fonction g du théorème des accroissements finis ! On utilise pour f une fonction nulle, donc de dérivée nulle ; on considère une fonction g de dérivée positive, et le tour est joué !$\square$

Corollaire 462 **Inégalité des accroissements finis** *f définie de l'ouvert U de l'espace vectoriel normé E et à valeurs dans l'espace vectoriel normé F. Si f est dérivable et si le segment $[x, y]$ est inclus dans U, alors*

$$\| f(y) - f(x) \| \leq \| y - x \| .sup_{z \in [x,y]} \| Df(z) \|$$

Démonstration : Si le sup est infini il n'y a rien à prouver. Sinon on considère la fonction θ qui à un réel t compris entre 0 et 1 associe $\|y-x\|.(sup_{z\in[x,y]}\|Df(z)\|).t$; θ est dérivable, en tout point, de dérivée constante égale à $\|y-x\|.(sup_{z\in[x,y]}\|Df(z)\|)$, que l'on va noter C. L'application ϕ qui à $t \in [0,1]$ associe $f((1-t).x + t.y)$ est dérivable en tout point de $[0,1]$, de dérivée $Df((1-t).x+t.y)(y-x)$. La norme de cette dérivée est majorée par θ', donc par C. On peut donc majorer $\|\phi(1)-\phi(0)\|$ par $\theta(1)-\theta(0)$.$\square$

Corollaire 463 *Une application définie sur un ouvert U de l'espace vectoriel normé E à valeurs dans l'espace vectoriel normé F dérivable et de dérivée nulle est localement constante. Si U est connexe, f est constante.*

Trois corollaires (pour le deuxième il faut un peu y réfléchir, pour le troisième c'est une conséquence du second) :

Corollaire 464 *En définissant la distance entre deux points d'un ouvert connexe comme la longueur inf d'une ligne brisée entre ces deux points (voir la partie topologie pour vérifier qu'un ouvert connexe d'un espace vectoriel normé est connexe par arcs et que toute paire de points dans un tel ensemble peut être reliée par une ligne brisée), et en supposant que f est une application de cet ouvert dans un espace vectoriel normé différentiable telle que pour tout x $\|f'(x)\| \leq k$, alors $\|f(b) - f(a)\|$ est inférieur ou égal à k fois la distance de a à b.*

Définition 465 *Une application* **localement lipschitzienne** *est une application entre espaces métriques telle que pour tout x il existe un voisinage de x sur lequel la restriction de f est lipschitzienne.*

Corollaire 466 *Une application de classe C^1 est localement lipschitzienne.*

11.2.2 Applications : interversion de limite et de dérivation

Définition 467 (Convergence uniforme, rappel) *Une suite d'applications f_n de X dans Y avec X et Y espaces métriques converge uniformément vers f application de X dans Y si*

$$lim_{n \to +\infty} sup_{x \in X} d(f_n(x), f(x)) = 0$$

Proposition 468 *Si les f_n sont continues et convergent uniformément vers f alors f est continue.*

Démonstration :

$$d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y) - f(y)$$

Etant donné ϵ l suffit alors de prendre n assez grand et x et y assez proches pour que $f(x) - f(y) \leq \epsilon$.□

Ce résultat servira par exemple pour le théorème 634, ou le théorème 490.

Théorème 469 *On suppose E et F des espaces vectoriels normés , U ouvert de E, f_n une suite d'applications de U dans F différentiables, f_n convergeant simplement vers f, les Df_n convergeant uniformément vers une certaine application g de U dans $\mathcal{L}(E, F)$,*
alors :
- *f est différentiable et $Df = g$*
- *Pour tout C convexe et borné inclus dans U la convergence de $f_{n|C}$ vers $f_{|C}$ est uniforme*
- *Si les f_n sont C^1 alors f est C^1.*

Démonstration : laborieuse, mais pas vraiment difficile ; il suffit d'écrire $\epsilon_n = sup_{x \in U} \|Df_n(x) - g(x)\|$, avec ϵ_n tendant vers 0 lorsque n tend vers l'infini, et de montrer que $sup_{y \in C} \|f(y) - f_n(y)\| \leq \|f(x) - f_n(x)\| + \epsilon_n.D$ avec D le diamètre de C pour voir la deuxième propriété ; la première propriété se montre facilement à partir de là, et la troisième est un corollaire de la proposition 468.□

Corollaire 470 *On suppose U connexe ouvert de E et f_n de U dans F dérivable ; E et F sont des espaces vectoriels normés , et F est complet (donc F est un Banach). On suppose qu'il existe x_0 tel que $f_n(x_0)$ converge, et que pour tout x il existe V_x voisinage de x tel que la suite des $Df_{n|V_x}$ soit de Cauchy pour la métrique d définie par*

$$d(f,g) = sup_{z \in V_x} \|f(z) - g(z)\|$$

(c'est à dire que la suite des Df_n converge normalement sur un certain voisinage de tout point)
Alors il existe f de U dans F tel que :
• f est dérivable en tout point
• la suite des f_n converge vers f (simplement)
• tout x possède un voisinage V_x tel que les convergences de f_n et Df_n restreints à V_x soient uniformes.
• Si les f_n sont C^1, f l'est aussi.

Démonstration : On définit l'ensemble A des z tels que $f_n(z)$ converge. Etant donné x on considère un voisinage V_x de x convexe et vérifiant l'hypothèse sur le critère de Cauchy (on peut toujours imposer V_x convexe en le restreignant à une boule). On définit alors $\alpha_{n,m} = sup_{z \in V_x} \|Df_n(z) - Df_m(z)\|$; par hypothèse, $\alpha_{n,m}$ tend vers 0 quand n et m tendent vers l'infini. Si $V_x \cap A \neq \emptyset$, alors soit x' dans $V_x \cap A$. Par convexité de V_x, on peut écrire pour tout y dans V_x :

$$\|[f_m(y) - f_n(y)] - [f_m(x) - f_n(x)]\| \leq \alpha_{n,m}.\|y - x'\|$$

or $f_n(x')$ est une suite de Cauchy (comme toute suite convergente dans un métrique), donc $f_n(y)$ est une suite de Cauchy, et donc converge. Donc si $V_x \cap A \neq \emptyset$, $V_x \subset A$. Donc soit $V_x \subset A$, soit $V_x \subset A^c$, donc A et A^c sont ouverts. A étant non vide et U étant connexe, $A = U$. On a donc montré la deuxième assertion.
$\mathcal{L}(E,F)$ est complet pour la norme uniforme puisque F l'est ; donc la suite des dérivées sur V_x converge uniformément. En appliquant le théorème précédent, on voit que f est dérivable de dérivée la limite des dérivées ; en supposant V_x borné on a alors $f_{n|V_n} \to f_{|V_x}$ uniformément, toujours par le théorème précédent.□

11.2.3 Applications : dérivées partielles et dérivées

Proposition 471 *E_1, E_2, ... , E_n et F des espaces vectoriels normés ; U un ouvert de $E = \Pi_i E_i$, f une application de U dans F ; alors si les $\frac{\delta f}{\delta x_i}$ existent sur un voisinage de x et sont continues en x, alors f est différentiable en x.*

Démonstration : • Il est suffisant de montrer que

$$\|f(x_1,...,x_n) - f(a_1,...,a_n) - \sum_{i=1}^{n} \frac{\delta f}{\delta x_i}(a).(x_i - a)\| = o(\|x - a\|)$$

pour tout $(a_1,...,a_n)$ de U.

• Pour cela on décompose $f(x_1,...,x_n) - f(a_1,...,a_n) - \sum_{i=1}^{n} \frac{\delta f}{\delta x_i}(a).(x_i - a)$ en

$$f(x_1,...,x_n) - f(a_1,x_2,...,x_n) - \frac{\delta f}{\delta x_1}(a).(x_1 - a_1)$$

$$+f(a_1,x_2,...,x_n) - f(a_1,a_2,x_3,...,x_n) - \frac{\delta f}{\delta x_2}(a).(x_2 - a_2)$$

$$+...$$

$$+f(a_1,a_2,...,a_i,x_{i+1},...,x_n)-f(a_1,...,a_{i+1},x_{i+2},...,x_n)-\frac{\delta f}{\delta x_{i+1}}(a).(x_{i+1}-a_{i+1})$$

$$+...$$

$$+f(a_1,a_2,...,a_{n-1},x_n) - f(a_1,...,a_n) - \frac{\delta f}{\delta x_n}(a).(x_n - a_n)$$

• Il suffit ensuite de montrer que pour x_i tendant vers a_i, $f(a_1,a_2,...,a_{i-1},x_i,...,x_n) - f(a_1,...,a_i,x_{i+1},...,x_n) - \frac{\delta f}{\delta x_i}(a).(x_i - a_i)$ est un $o(x_i - a_i)$. (ensuite il suffira de sommer)

• Le fait ci-dessus provient des accroissements finis ET de la continuité de la i-ième dérivée partielle (en effet une application directe des accroissements finis donnent un $o(x_i - a_i)$ pour

$$f(a_1,...,a_{i-1},x_i,...,x_n)-f(a_1,...,a_i,x_{i+1},...,x_n)-\frac{\delta f}{\delta x_i}(a_1,...,a_i,x_{i+1},...,x_n).(x_i-a_i)$$

(la différentielle n'est pas prise là où il faudrait qu'elle le soit) $\square$

Théorème 472 *E_1, E_2, ... , E_n et F des espaces vectoriels normés ; U un ouvert de $E = \Pi E_i$, alors f application de U dans F est C^1 si et seulement si les dérivées partielles $\frac{\delta f}{\delta x_i}$ de f existent et sont continues sur U.*

Démonstration : Il est clair que si f est C^1, alors les dérivées partielles existent et sont continues. La réciproque, utilisant la proposition précédente, ne présente pas de difficulté majeure.$\square$

On pourra par exemple trouver une application dans la partie20.3.

11.3 Théorème d'inversion locale et fonctions implicites

11.3.1 Théorème d'inversion globale

Définition 473 *On appelle* **application contractante** *ou* **contraction** *une application lipschitzienne dont le coefficient de Lipschitz est* < 1.

Théorème 474 (Théorème de Banach du point fixe) *Soit X un espace métrique complet et h une contraction de X dans X. Alors :*
• h admet un unique point fixe x_0
• $\forall x\ d(x, x_0) \leq \frac{1}{1-Lip(h)} d(x, h(x))$

Démonstration : Unicité :
• Supposons x_1 et x_2 deux points fixes.
• $d(x_1, x_2) = d(h(x_1), h(x_2)) \leq Lip(h).d(x_1, x_2)$; donc $x_1 = x_2$
Existence :
• Considérons x quelconque dans X, on va travailler sur la suite des $h^n(x)$.
• Supposons $n \leq m$, alors

$$d(h^m(x), h^n(x)) \leq \sum_{i=m}^{n-1} d(h^i(x), h^{i+1}(x))$$

$$\leq \sum_{i=m}^{+\infty} Lip(h)^i.d(x, h(x)) \leq \frac{Lip(h)^m}{1 - Lip(h)}.d(x, h(x))$$

On en déduit facilement les deux résultats annoncés.□

⚠ Il faut que f soit une contraction, c'est à dire une application lischitzienne de constante de Lipschitz < 1 ; avec un rapport 1 cela ne marche pas, ni même avec $d(f(x), f(y)) < d(x, y)$. Par exemple, $x \mapsto x + e^{-x}$ définit une application de $\mathbb{R}^+$ dans $\mathbb{R}^+$, $\mathbb{R}^+$ est bien complet, et on a bien $d(f(x), f(y)) < d(x, y)$, et pourtant f n'admet pas de point fixe.

théorème d'inversion locale 478, théorème de Cauchy-Lipschitz 515, la résolution de l'équation de Volterra (voir [6]).

D'autres théorèmes de points fixes existent : par exemple le théorème du point fixe de Brouwer 311 (avec pour application le corollaire 313), le théorème de Kakutani , le théorème de Schauder , et même pour ceux qui connaissent un peu la calculabilité un théorème de point fixe que l'on trouvera dans le livre "Théorie de la récursion pour la métamathématique", de R. Smullyan (Masson, 1995), avec pour application le théorème de Rice et ses multiples conséquences (attention, il faut connaître un peu le domaine pour pouvoir se lancer dans ce genre d'originalités...).

Lemme 475 *Soient U et V des ouverts des espaces normés E et F. On se donne h de U dans V, h bijective, dérivable en x_0. Alors h^{-1} est dérivable en $h(x_0)$ si et seulement si les deux conditions suivantes sont vérifiées :*
• $Dh(x_0)$ est un isomorphisme de E sur F
• Il existe $K \geq 0$ et un voisinage W de $h(x_0)$ dans F tels que

$$\forall y \in W \|h^{-1}(y) - x_0\| \leq K\|y - h(x_0)\|$$

Démonstration : Tout d'abord montrons que ces deux conditions sont nécéssaires. Pour cela on suppose qu'effectivement h^{-1} est dérivable en $h(x_0)$, et on procède comme suit :
• On dérive les deux expressions

$$h^{-1} \circ h = Id_E$$

et

$$h \circ h^{-1} = Id_E$$

et on montre bien que $Dh(x_0)$ est un isomorphisme.
• Par définition de la dérivée, la quantité ci-dessous tend vers 0 pour $y \to h(x_0)$:

$$\frac{\|h^{-1}(y) - h^{-1}(h(x_0)) - D(h^{-1})(h(x_0))(y - h(x_0))\|}{\|y - h(x_0)\|}$$

Donc pour y dans un certain W cette quantité est plus petite que 1, et donc pour $y \in W$ on a

$$\|h^{-1}(y) - h^{-1}(h(x_0))\| \leq K\|y - h(x_0)\|$$

avec $K = 1 + \|D(h^{-1})(h(x_0))$.

Il reste à prouver la réciproque, c'est à dire que les conditions sont suffisantes.
• Par définition, on a

$$h(x) - h(x_0) = Dh(x_0)(x - x_0) + \|x - x_0\|\epsilon(x)$$

avec $\epsilon(x)$ tendant vers 0 pour $x \to x_0$.
En composant avec $Dh(x_0)^{-1}$ (dont les hypothèses garantissent l'existence), on obtient

$$x - x_0 = Dh(x_0)^{-1}(h(x) - h(x_0)) - \|x - x_0\|.(Dh(x_0)^{-1}(\epsilon(x)))$$

avec $y = h(x)$ (tout y de V peut s'écrire ainsi) et $y_0 = h(x_0)$, on obtient alors

$$h^{-1}(y) - h^{-1}(y_0) - Dh(x_0)^{-1}(y - y_0) = -Dh(x_0)^{-1}.\epsilon(h^{-1}(y)).\|h^{-1}(y) - h^{-1}(y_0)\|$$

On sait par hypothèse que pour y assez proche de y_0 on a $\|h^{-1}(y) - h^{-1}(y_0)\| \leq K.\|y - y_0\|$, donc

$$\frac{Dh(x_0)^{-1}.\epsilon(h^{-1}(y)).\|h^{-1}(y) - h^{-1}(y_0)\|}{\|y - y_0\|} \leq K.\|Dh(x_0)^{-1}\|.\|\epsilon(h^{-1}(y))\|$$

or $h^{-1}(y)$ tend vers x_0 quand $y \to y_0$ donc cette quantité tend vers 0 ; ce qui permet de conclure.□

Théorème 476 (Théorème d'inversion globale) *Soit A une application linéaire continue de E dans F, avec E un espace de Banach et F un espace normé, telle que A^{-1} existe et est continue (A est un homéomorphisme linéaire). Soit ϕ une application lipschitzienne de E dans F telle que $Lip(\phi) < \|A^{-1}\|^{-1}$. Alors :*
• $h = A + \phi$ est inversible
• h^{-1} est lipschitzienne, avec $Lip(h^{-1}) \leq \frac{\|A^{-1}\|}{[1-\|A^{-1}\|.Lip(\phi)]}$
• Si h est C^1 sur U ouvert de E, et si $\forall x \in U\ Dh(x) \in Isom(E, F)$, alors h^{-1} est C^1 sur l'ouvert $h(U)$, et pour tout $x \in U$ la différentielle de h^{-1} est donnée par

$$D(h^{-1})(h(x)) = (Dh(x))^{-1}$$

Démonstration : Raisonnement en plusieurs étapes :
• Etant donné y dans F, et on considère l'équation

$$h(x) = y$$

équivalente à

$$x = A^{-1}.y - A^{-1}(\phi(x))$$

L'application $x \mapsto A^{-1}(\phi(x))$ est Lipschitzienne de rapport < 1, et donc l'application $x \mapsto A^{-1}(Y) - A^{-1}(\phi(x))$ aussi. Donc par le théorème du point fixe de Banach, cette équation a une solution unique x.
h est donc inversible.
• Avec $y = h(x)$ et $y' = h(x')$, on peut écrire

$$x = A^{-1}(y) - A^{-1}(\phi(x))$$

$$x' = A^{-1}(y') - A^{-1}(\phi(x'))$$

et en déduire

$$\|x - x'\| \leq \|A^{-1}\|.\|y - y'\| + \|A^{-1}\|.Lip(\phi).\|x - x'\|$$

en utilisant l'hypothèse sur $Lip(\phi)$, on a alors

$$\|x - x'\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|.Lip(\phi)}\|y - y'\|$$

D'où le résultat sur la constante de Lipschitz de h^{-1}.
• Supposons maintenant h C^1 sur U ouvert de E. $h(U)$ est un ouvert puisque h est un homéomorphisme. Par le lemme précédent, h^{-1} est dérivable sur U. En dérivant

$$h^{-1} \circ h = Id_E$$

et

$$h \circ h^{-1} = Id_E$$

on obtient que $D(h^{-1})(h(x)) = (Dh(x))^{-1}$. En écrivant

$$D(h^{-1})(y) = Dh(h^{-1}(y))^{-1}$$

et en rappelant que $Inv : Isom(E, F) \to Isom(F, E), h \mapsto h^{-1}$ est continue on constate qu'en outre $D(h^{-1})$ est continue.

11.3.2 Théorème d'inversion locale

Définition 477 (Difféomorphisme C^1) *Une application h de U dans V avec U ouvert d'un espace vectoriel normé et V ouvert d'un espace vectoriel normé est un* **difféomorphisme** *C^1 si h est bijective et de classe C^1 et de réciproque de classe C^1. Plus généralement, avec $k \geq 1$, une application h de U dans V avec U ouvert d'un espace vectoriel normé et V ouvert d'un espace vectoriel normé est un* **difféomorphisme** *C^k si h est bijective et de classe C^k et de réciproque de classe C^k.*

Une application bijective et C^1 n'est pas un difféomorphisme C^1 ; il faut aussi que la réciproque soit C^1 !

Théorème 478 (Théorème d'inversion locale) *Soit h de U dans F une application C^1, avec U ouvert de E, et E et F des espaces de Banach. Si la différentielle $Dh(x_0)$ est bijective de E dans F pour un certain x_0 de U, alors il existe U_0 voisinage de x_0 dans E et un voisinage ouvert V_0 de $f(x_0)$ dans F tels que h induit un difféomorphisme C^1 de U_0 dans V_0. On a alors*

$$D(h^{-1})(h(x)) = (Dh(x))^{-1}$$

pour tout x dans U_0.

Voir le théorème 480 et le corollaire 479. Le théorème d'inversion locale permettra aussi de montrer l'équivalence des différentes définitions des variétés de $\mathbb{R}^n$, voir définition 493.

Démonstration : On pourra se référer par exemple à [13].

Dans le cas de la dimension finie, le fait que la différentielle soit bijective implique que les dimensions des espaces soient les mêmes, et que le jacobien (défini puisque les dimensions sont les mêmes) est non nul.

Corollaire 479 *Soit f une application C^1 d'un ouvert U d'un espace de Banach E dans un espace de Banach F ; si f injective et si sa différentielle $f'(x)$ est un isomorphisme en tout $x \in U$, alors f est un C^1-difféomorphisme de U sur $f(U)$, qui est alors un ouvert de F.*

Modulo le résultat selon lequel l'application qui à f dans $Isom(E, F)$ associe

f^{-1} est C^∞ (que l'on trouvera par exemple dans [3]), on montre que si f est un C^1-difféomorphisme de classe C^n, alors f est un C^n difféomorphisme.

11.3.3 Théorème des fonctions implicites

Théorème 480 *f application C^1 de U un ouvert de $E_1 \times E_2$ dans F, avec E_1, E_2 et F des Banach. Si f est différentiable par rapport à la deuxième variable en (a,b) et si cette différentielle est un isomorphisme alors il existe ϕ C^1 définie sur un ouvert U_1 de E_1 contenant a_1 et à valeurs dans un ouvert U_2 de E_2 contenant a_2 telle que pour tout (y_1,y_2) dans $U_1 \times U_2$ on ait*

$$f(y_1,y_2) = f(a,b) \iff y_2 = \phi(y_1)$$

et pour tout x dans U_1 on a

$$D\phi(x) = -[D_2 f(x,\phi(x))]^{-1} \circ D_1 f(x,\phi(x))$$

FLEMMARD

Démonstration : • On va chercher à utiliser le théorème d'inversion locale. Pour cela il faut construire une application différentiable C^1, de différentielle bijective.

• On définit l'application μ de U dans $E_1 \times E_2$ dans $E_1 \times G$, définie par $\mu(y_1,y_2) = (y_1, f(y_1,y_2))$.

• $D\mu(a,b)(h,k) = (h, D_1 f(a,b).h + D_2 f(a,b).k)$

• $D\mu(a,b)$ est bijective (facile), μ est C^1

• On applique le théorème d'inversion locale à μ en (a,b).

• La fonction réciproque associe clairement à un couple (q,r) un couple $(q,\phi(q,r))$, et en fixant r on en déduit ce que l'on veut...□

La figure 11.1 illustre ce théorème.

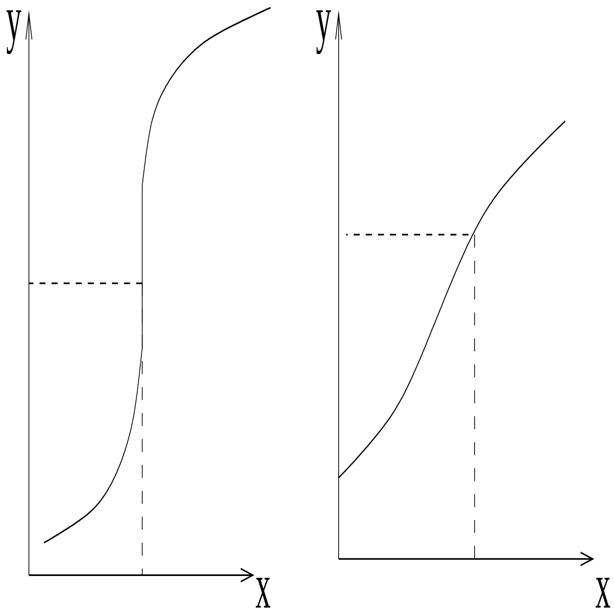


FIG. 11.1 – Illustration du théorème des fonctions implicites dans le cas d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$. La courbe est une courbe de niveau. A gauche la différentielle par rapport à y n'est pas un isomorphisme ; on comprend intuitivement qu'on ne peut pas donner y en fonction de x sur un voisinage. A droite c'est un isomorphisme ; donc on peut.

11.4 Dérivées d'ordre supérieur

11.4.1 Généralités

Définition 481 *Etant donnée une application f d'un ouvert U d'un espace de Banach E dans un espace de Banach F, on dit que f est* **de classe** C^n *(on dit aussi n* **fois continûment différentiable***) si f est différentiable et si sa différentielle est de classe C^{n-1}. L'application est dite C^∞ si elle est C^n pour tout n ; on dit alors qu'elle est* **indéfiniment différentiable**.
On note alors $f^{(1)}(a)$ l'application $Df(a)(x)$, et par récurrence $f^{(n)}(a)$ l'application $Df^{(n-1)}(a)$.
Etant donnée une application f d'un ouvert U d'un espace de Banach E dans un espace de Banach F, on dit que f est n fois différentiable en x appartenant à U si et seulement si f est de classe C^{n-1} sur un voisinage de x et si la n-ième différentielle de f sur ce voisinage est différentiable en a.
Etant donnée f application d'un ouvert U d'un espace de Banach E dans un espace de Banach F, on note $\frac{\delta^2 f}{\delta x_i \delta x_j}$ l'application $\frac{\delta \frac{\delta f}{\delta x_j}}{\delta x_i}$.

11.4.2 Dérivées secondes

Théorème 482 *Soit f une application de U, ouvert d'un espace de Banach , dans F, espace de Banach . Si f est deux fois différentiable en x_0, alors $f''(a)(h)(k) = f''(a)(k)(h)$. Via l'identification du théorème 1014 $f''(a)$ est une forme bilinéaire symétrique.*

Démonstration : • Il s'agit donc de montrer que $f''(a)(h)(k) = f''(a)(k)(h)$.

• On introduit la fonction $\mu(h,k) = f(a+h+k) - f(a+h) - f(a+k) + f(a)$

• $\mu(h,k) = \mu(k,h)$ clairement

• On montre maintenant que $\mu(h,k)$ approxime $f''(a)(h)(k)$ (on pourra plus loin en déduire le résultat souhaité, car $f''(a)(h)(k)$ approximera alors $f''(a)(k)(h)$)

• $\|\mu(h,k) - f''(a)(h)(k)\| \leq \|\mu(h,k) - f'(a+k)(h) + f'(a)(h)\| + \|f'(a+k)(h) - f'(a)(h) - (f''(a).k).h\|$

• Le second terme est petit par la définition de la différentielle $f''(a)$, le premier est petit par définition de la différentielle de $f'(a)$, et par utilisation des accroissements finis.

• On arrive ainsi à montrer que $\mu(h,k) - f''(a)(k)(h)$ est un $o(\|h\|^2 + \|k\|^2)$; par symétrie on a aussi le fait que $\mu(h,k) - f''(a)(k)(h)$ est un $o(\|h\|^2 + \|k\|^2)$; on en déduit $\|f''(a)(k)(h) - f''(a)(h)(k)\| = o(\|h\|^2 + \|k\|^2)$.

• $\|f''(a)(\lambda.k)(\lambda.h) - f''(a)(\lambda.h)(\lambda.k)\| \leq \epsilon.(\|\lambda.h\|^2 + \|\lambda.k\|^2)$ pour tout ϵ, pour λ suffisamment petit

• $\lambda^2 \|f''(a)(k)(h) - f''(a)(h)(k)\| \leq \epsilon.\lambda^2(\|h\|^2 + \|k\|^2)$ pour tout ϵ et pour λ assez petit

• $\|f''(a)(k)(h) - f''(a)(h)(k)\| \leq \epsilon(\|h\|^2 + \|k\|^2)$ pour tout ϵ !

• $f''(a)(k)(h) = f''(a)(h)(k) = 0$ d'où le résultat.□

Théorème 483 *Soit f deux fois différentiables en $a \in U$, avec f définie de U ouvert de $E = E_1 \times E_2 ... \times E_n$ (des espaces de Banach) dans F (un espace de Banach). Alors*

$$(f''(a)(h_1, ..., h_n))(k_1, ..., k_n) = \sum_{i,j} (\frac{\delta^2 f}{\delta x_i \delta x_j}(a).h_i).k_j$$

Démonstration : On utilise simplement deux fois la proposition 454, à f et f'.□

Corollaire 484 *Soit f deux fois différentiables en $a \in U$, avec f définie de U ouvert de $E = E_1 \times E_2 ... \times E_n$ (des espaces de Banach) dans F (un espace de Banach). Alors*

$$\frac{\delta^2 f}{\delta x_i \delta x_j}(a)(h)(k) = \frac{\delta^2 f}{\delta x_j \delta x_i}(a)(k)(h)$$

Démonstration : Il s'agit simplement du théorème 482 après quelques manipulations...□

Corollaire 485 (Théorème de Schwartz) *Soit f application de $\mathbb{R}^n$ dans F (F espace de Banach) deux fois différentiables en x, alors $\frac{\delta^2 f}{\delta x_i \delta x_j}(x) = \frac{d^2 f}{\delta x_j \delta x_i}$*

Démonstration : Il s'agit d'une reformulation dans le cas de $E_i = \mathbb{R}$ du corollaire précédent !□

Proposition 486 (Existence de la dérivée seconde) *Soi f une application de U ouvert de $E = E_1 \times E_2 ... \times E_n$ (des espaces de Banach) dans F (un espace de Banach). Alors si les $\frac{\delta f}{\delta x_i}$ existent et sont continues sur un voisinage de x, et si les $\frac{\delta^2 f}{\delta x_i \delta x_j}$ existent sur un voisinage de x et sont continues en x, alors f est deux fois différentiable en x.*

Démonstration : Il suffit d'appliquer deux fois la proposition 471.□

11.4.3 Généralisations à la dérivée n-ième

On pourra réviser la partie 28.

Théorème 487 (Généralisation du théorème 482) *Si f est une application de E dans F avec E et F des espaces de Banach n fois différentiable en x, alors $f^{(n)}(x)$ appartient à $\mathcal{L}_n(E; F)$ et est une application n-linéaire symétrique.*

Démonstration : On procède par récurrence. Pour $n = 1$, c'est clair. Pour $n = 2$, c'est le théorème 482. Supposons maintenant le résultat prouvé jusqu'au rang $n - 1$, et montrons le pour le rang n, avec $n \geq 3$.
• Il suffit de montrer que si l'on permute deux variables consécutives parmi les h_i on ne change pas la valeur $f^{(n)}(x)(h_1, ..., h_n)$.
• $f^{(n-1)}$ est symétrique, donc on peut permuter sans rien changer h_i et h_{i+1} pour $i > 1$
• Pour $i = 1$ il suffit de rappeler que $f^{(n)} = (f^{(n-2)})''$ et d'utiliser 482.□

11.5 Zoologie du calcul différentiel

11.5.1 Fonctions convexes

Définition 488 *Une fonction f définie sur un convexe U d'un espace vectoriel à valeurs dans $\mathbb{R}$ est dite* **convexe** *(resp.* **strictement convexe***) si*

$$\forall (u,v,t) \in U^2 \times [0,1] f(tu + (1-t)v) \leq tf(u) + (1-t)f(v)$$

(resp.)$\forall (u,v,t) \in U^2 \times]0,1[u \neq v \Rightarrow f(tu + (1-t)v) < tf(u) + (1-t)f(v)$

Dans la suite de cette section, on suppose que U est un convexe d'un espace vectoriel , et que f est une application de Ω dans $\mathbb{R}$, avec Ω ouvert contenant U. Les liens entre dérivabilité et convexité sont les suivants :

Théorème 489 *f C^1 est convexe (resp. strictement convexe) sur U si et seulement si*

$$\forall (u,v) \in U^2 f(v) \geq f(u) + f'(u)(v-u)$$

(resp.)$\forall (u,v) \in U^2 u \neq v \Rightarrow f(v) > f(u) + f'(u)(v-u)$

f C^2 est convexe si et seulement si

$$\forall (u,v) \in U^2 f'(u)(v-u, v-u) \geq 0$$

Si

$$\forall (u,v) \in U^2 u \neq v \Rightarrow f''(u)(v-u, v-u) > 0$$

On pourra voir 12.4 pour les applications de la convexité à la recherche d'extréma, 1325 pour les applications de l'inégalité de Jensen, le lemme 397 (et par suite l'inégalité de Hölder), l'inégalité de Minkovski.

11.5.2 Fonction continue partout dérivable nulle part

Cet exemple, élaboré par Van der Waerden, est extrait du livre [20].

Théorème 490 *Soit T la fonction définie sur $\mathbb{R}$ par $T(x) = min(x - E(x), E(x) + 1 - x)$ (c'est à dire que $T(x)$ est la distance de x à l'entier le plus proche de x).*
La fonction f définie sur $\mathbb{R}$ par $f(x) = \sum_{n=0}^{\infty} \frac{T(10^n x)}{10^n}$ est continue partout dérivable nulle part.

Démonstration :

• Bonne définition, continuité de f : facile, f est limite uniforme d'une suite de fonctions continues (voir proposition 468).

• La non dérivabilité, c'est plus dur.

- T, et donc f, est périodique, de période 1.
- On se limite donc à montrer la non-dérivabilité sur $[0,1[$
- On note les développements décimaux en excluant les développements illimités ne comportant que des 9 [1]
- Soit donc $x \in [0,1[$, on montre la non-dérivabilité de f en x.
- Soit x_n la n-ième décimale de x.
- définissons $h_m = -10^{-m}$ si $x_m = 4$ ou $x_m = 9$, $h_m = 10^{-m}$ sinon.
- Calculons maintenant

$$\frac{f(x+h_m)-f(x)}{h_m}$$

$$= 10^m \sum_{n=1}^{\infty} \frac{\epsilon_n(T(10^n(x+\epsilon'_m 10^{-m})) - T(10^n x))}{10^n}$$

- Raisonnons un petit peu maintenant, sur un cas particulier pour mieux visualiser ($x = 0.33333333...$) :

n	$10^n x$	$10^n(x+h_m)$	$T(10^n(x+h_m)) - T(10^n x)$	$\times 10^{m-n}$
0	$0,333333\ldots$	$0,33\ldots333433\ldots$	$0,000\ldots001$	1
1	$3,33333\ldots$	$3,33\ldots334333\ldots$	$0,000\ldots01$	1
2	$33,3333\ldots$	$33,3\ldots343333\ldots$	$0,000\ldots1$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	$333\ldots3,33\ldots$	$333\ldots4,333\ldots$	1	1
$m+1$	$3333\ldots3,33\ldots$	$333\ldots43,333\ldots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Il faut bien noter que dans le cas général le chiffre de la dernière colonne peut être 1 ou -1 ; quoi qu'il en soit $\frac{f(x+h_m)-f(x)}{h_m}$ est un entier de parité variant avec m et ne peut donc pas converger.$\square$

On en profite pour montrer ce dont est capable Maple. Le dessin se trouve en figure 11.2.

Exemple Maple
$> T := x \to \min(x - \mathrm{floor}(x), \mathrm{floor}(x) + 1 - x)$
$T := x \to \min(x - \mathrm{floor}(x), \mathrm{floor}(x) + 1 - x)$
$> g := x -> sum(T(2^n * x)/2^n, n = 0..17);$
$g := x \to \sum_{n=0}^{17} \frac{\mathrm{T}(2^n\, x)}{2^n}$

[1] Au profit de l'équivalent obtenu en remplaçant ...243999999999... par ...2449999999999....

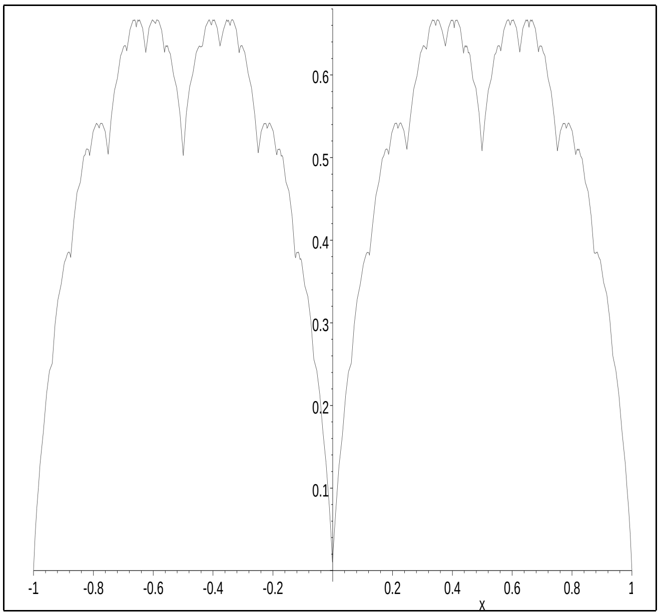


FIG. 11.2 – Tracé d'une courbe continue dérivable nulle part, établie par Van der Waerden

11.5.3 Fonction dérivable dans toutes les directions mais non continue

Définition 491 *Soit f une application d'un ouvert U d'un espace vectoriel normé E dans un espace vectoriel normé F, alors f est dite* **différentiable en** $x \in U$ **suivant la direction** $e \in E$ *si l'application $g : \mathbb{R} \to F$ $t \mapsto f(x + tu)$ est différentiable en 0. La différentielle de g en 0 est alors appelée* **différentielle de** f **en** x **suivant** u.

Proposition 492 *Il existe une application f différentiable dans toutes les directions en x et qui n'est pas continue en x.*

Démonstration : Le livre [20] propose la fonction $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto \frac{x^5}{(y-x^2)^2+x^8}$, avec $f(0, 0) = 0$. On constate la non continuité de f en regardant la limite de $x \mapsto f(x, x^2)$ en 0. La différentiabilité suivant toutes les directions est vite vue (distinguer différents cas, suivant $u = (a, b)$, cas a et b non nuls, cas a nul, ou cas b nul).

On peut aussi regarder la fonction définie par ses coordonnées polaires par

$$f(r\cos(\theta), r\sin(\theta)) = e^{-\frac{(\theta-r)^2}{r^4}}$$

Avec $f(0, 0) = 0$, on a bien une fonction différentiable dans toutes les directions (avec des différentielles nulles !), et on constate sans le moindre calcul que $f(xcos(x), xsin(x))$ est constant égal à 1 pour $x \neq 0$.$\square$

11.5.4 Variétés de $\mathbb{R}^n$, théorème de Jordan

Définition 493 - Proposition*[Variété de $\mathbb{R}^n$] Soit M une partie de $\mathbb{R}^n$, x un point de M. Soit p un entier > 0 et k un entier > 0.*
M est par définition une **variété de dimension p et de classe C^k au voisinage de** *x si l'une des conditions équivalentes suivantes est vérifiée :*
(i) Il existe V voisinage ouvert de x tel qu'il existe un C^k difféomorphisme de V sur $W \subset \mathbb{R}^n$ tel que $g(x) = 0$ et $g(M \cap V) = W \cap P$, avec P l'ensemble des $(y_1, ..., y_n)$ tels que $y_{p+1} = 0$, $y_{p+2} = 0$, ... , $y_n = 0$.
(ii) Après une permutation pertinente des coordonnées $(x_1, ..., x_n)$, il existe V voisinage ouvert de x et ϕ application C^k de $\mathbb{R}^p$ dans $\mathbb{R}^{n-p}$ tel que pour tout y dans V

$$y \in M \iff \phi(y_1, ..., y_p) = (y_{p+1}, ..., y_n)$$

(iii) Il existe un voisinage V ouvert de x, Ω un voisinage ouvert de 0 dans $\mathbb{R}^p$, f une application C^k de Ω dans $\mathbb{R}^n$, tels que f induise un homéomorphisme de Ω sur $M \cap V$, $f(0) = x$ et $f'(0)$ de rang p [a].
(iv) Il existe un voisinage V ouvert de x, Ω un voisinage ouvert de 0 dans $\mathbb{R}^p$, f une application C^k de Ω dans $\mathbb{R}^n$, tels que f induise un homéomorphisme de Ω sur $M \cap V$, $f(0) = x$ et $f'(y)$ de rang p pour tout y dans Ω [b].

[a] Il s'agit d'un élément de $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$.
[b] Il s'agit d'un élément de $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$.

Démonstration :

On notera pendant cette preuve $x_{|I}$, avec $I = \{i_1, ..., i_m\}$ et $i_1 < i_2 < \cdots < i_m$ un sous-ensemble de $[1, n]$ et x un élément de $\mathbb{R}^n$, $(x_{i_1}, \ldots, x_{i_m})$.

• L'équivalence entre (iii) et (iv) est claire ; bien sûr (iv) implique (iii), et réciproquement en supposant (iii) par continuité de la différentielle et continuité du déterminant d'une matrice extraite, on peut trouver un voisinage de 0 dans lequel la différentielle a le même rang. Il suffit alors de se restreindre à ce voisinage.

• Voyons maintenant que (iii) implique (ii).

Supposons (iii). La matrice de la différentielle de f en 0 est de rang p ; modulo une bonne permutation des coordonnées, on peut donc supposer que la matrice extraite de la différentielle pour les indices en ligne et en colonne inférieurs ou égaux à p est inversible.

En se restreignant aux p première coordonnées, f est alors C^k, de différentielle en 0 de rang plein. On peut donc appliquer le théorème d'inversion locale 478, et f ainsi restreint est donc un C^k difféomorphisme au voisinage de 0. En prenant ϕ la composée de f et de l'inverse de la restriction de f aux p-premières coordonnées, on obtient une fonction satisfaisant (ii).

• Voyons maintenant (ii) implique (iii).

Supposons (ii) vérifiée. Définissons alors $f(y) = (y + x_{|[1,p]}, \phi(y + x_{|[1,p]})$[2]. f convient...

• Montrons maintenant que (iii) implique (i).

Supposons (iii) vérifiée.

Définissons alors $g(y) = (y_{|[1,p]} - x_{|[1,p]}, y_{|[p+1,n]} - \phi(y_{|[1,p]})$... g convient pour (i).

[2] Je "recolle" ainsi un élément de $\mathbb{R}^p$ et un élément de $\mathbb{R}^{n-p}$ pour obtenir un élément de $\mathbb{R}^n$

• Il ne reste plus qu'à vérifier que (i) implique (iii).
Supposons donc (i) vérifiée.
Alors soit $f(y) = g^{-1}(y_{|[1,p]}, 0, \ldots, 0)$.
La différentielle de g est injective, donc la restriction à $\mathbb{R}^p$ est injective aussi. Donc f vérifie bien (iii).□

Théorème 494 (Théorème de Jordan) *Toute hypersurface M (i.e. variété de dimension $n-1$) de $\mathbb{R}^n$ C^∞*
• est le noyau d'une application C^∞ dont la différentielle ne s'annule pas sur M.
• est orientable (c'est à dire qu'il existe un champ de vecteur ne s'annulant pas, continu, défini sur M, à valeurs dans $\mathbb{R}^n$)
• partage le plan en deux domaines[a] (l'un borné l'autre non) dont elle est la frontière commune.

[a]Rappelons qu'un domaine est un ouvert connexe.

Démonstration :
Ce théorème, long et loin d'être trivial, nécéssitera différents lemmes. Cette preuve est largement inspiré de la note "Le théorème de Jordan pour les hypersurfaces C^∞", de D. Leborgne, paru dans la Revue de Maths Spé numéro 104, 1993-1994, lui-même inspiré de l'article "Orientability of smooth hypersurfaces and the Jordan Brouwer separation theorem", Expo. Math. 5 (1987), p 283-286.

Lemme 495 *Soit X un espace topologique connexe, et f et g continues de X dans $\mathbb{R}$. Si f et g sont localement égales ou opposées, et si l'intérieur de $f^{-1}(0)$ est réduit à l'ensemble vide, alors f et g sont égales ou opposées.*

Démonstration :
• Définissons :

$$Egales = \{x/f(x) = g(x)\}$$
$$Oppos = \{x/f(x) = -g(x)\}$$

Et notons Ω l'intérieur de $Egales$.
• Supposons tout d'abord Ω non vide, et montrons que $\Omega = X$. Si on a un tel résultat, alors on saura que soit $Egales$ est d'intérieur vide, soit il est égal à tout l'espace. Si $Egales$ est égal à tout l'espace, alors on a bien le résultat souhaité. Si $Egales$ est d'intérieur vide, alors l'ensemble des x tels que $f(x) = g(x) \neq 0$ (inclus dans l'intérieur d'$Egales$) est vide, et donc $Oppos = X$, d'où le résultat souhaité. Donc montrer que si $\Omega \neq \emptyset$ alors $\Omega = X$ est suffisant pour le résultat souhaité.
• Donc, on suppose Ω vide.
• Soit x appartenant à la frontière de Ω.
• Par hypothèse, il existe U ouvert contenant x sur lequel f et g sont égales ou opposées, puisque f et g sont localement égales ou opposées.
• L'intersection de U et Ω est non vide (puisque x est sur la frontière de Ω), et contient un point sur lequel f et g sont non nulles, puisque $f^{-1}(x)$ est d'intérieur vide. Donc f et g sont égales sur U (rappelons que Ω est inclus dans $Egales$).
• On en déduit que x appartient à $Egales$.

• On a donc montré que la frontière de Ω est incluse dans Ω.
• Ω est donc fermé et ouvert.
• Ω est donc égal à X, puisque $\underline{\ X \text{ est connexe}}$. D'où le résultat.□
Nous avons encore besoin d'un lemme de topologie :

> **Lemme 496** *Soit $(\Omega_i)_{i\in I}$ une famille d'ouverts recouvrant $\mathbb{R}^n$, et pour i dans I, f_i application C^∞ de Ω_i dans $\mathbb{R}$, telle que l'intérieur de $f_i^{-1}(0)$ soit d'intérieur vide.*
> *Supposons que lorsque Ω_i et Ω_j s'intersectent, alors f_i et f_j soient égales ou opposées localement sur $\Omega_i \cap \Omega_j$.*
> *Alors il existe une unique fonction f, au signe près, dont la restriction pour tout i à Ω_i soit f_i.*

Bien voir que l'on n'a pas supposé que les Ω_i soient connexes.

Démonstration :

• On considère la famille $(B_j)_{j\in J}$ des boules B telles qu'il existe un certain i tel que $B \subset \Omega_i$. On définit f'_j la restriction de Ω_i à B_j, lorsque B_j est inclus dans Ω_i.

• Supposons qu'on ait construit une fonction f localement égale ou opposé à f'_j sur B_j pour tout j dans J.

• Alors f, pour tout j dans J, est égale à $\epsilon'_j f'_j$ sur B_j, avec $\epsilon'_j \in \{-1, 1\}$, par connexité de B_j et application du lemme précédent 495.

• Alors f, pour tout i dans I, est égale à $\epsilon_i f_i$ sur Ω_i, avec $\epsilon_i \in \{-1, 1\}$. En effet, donnons-nous x et y dans Ω_i, avec $f_i(x)$ et $f_i(y)$ tous deux non nuls, et montrons que nécéssairement f et f_i égales (resp. opposées) en x implique f et f_i égales (resp. opposées) en y.

- le segment $[x, y]$ est recouvert par un nombre fini de boules B_j
- chacune des intersections de ces boules est connexe, et chaque boule est connexe.
- on peut donc appliquer sur chacun de ces connexes le lemme 495, d'où le résultat.

• Il reste donc simplement à construire une fonction f convenable sur les B_j.

• On montre tout d'abord le résultat en dimension 1, sur un segment fermé. Cela se fait en recouvrant le segment en question par un nombre fini de boules ouvertes B_j, en définissant f sur cette réunion finie de proche en proche. Le fait que l'on soit en dimension 1 rend cela facile ; il suffit de choisir des ouverts consécutifs, non inclus les uns dans les autres. Par le lemme 495, on a une solution et une seule, au signe près.

• On procède maintenant par récurrence sur la dimension i, pour montrer l'existence et l'unicité au signe près d'une telle fonction sur un pavé $[-m, m]^i$.

• Pour cela on considère ce pavé comme le produit $[-m, m] \times [-m, m]^{i-1}$.

• On définit une fonction f_t pour t dans $[-m, m]$, définie sur $[-m, m]^{i-1}$, en utilisant l'hypothèse de récurrence, f_t égale à FLEMMARD ça marche pas parce qu'on n'est pas sur que l'intersection d'un ensemble d'intérieur vide avec un ensemble de dimension inférieure est d'intersection vide.

•

11.5.5 Espaces vectoriels normés de dimension finie

◙ Propriétés topologiques

La boule unité fermée d'un espace vectoriel normé est compacte si et seulement si l'espace est de dimension finie (théorème de Riesz 5.2).

Tout espace vectoriel normé de dimension finie est complet. Dans un espace vectoriel normé de dimension finie, toutes les normes sont équivalentes (voir théorème 188). En particulier, les espaces vectoriels normés réels de dimension finie sont tous isomorphes à $\mathbb{R}^n$ pour un certain n, tous les espaces vectoriels normés complexes de dimension finie sont isomorphes à $\mathbb{C}^n$ pour un certain n. Les compacts d'un espace vectoriel de dimension finie sont donc exactement les fermés bornés. Cela implique notamment que tout sous-espace vectoriel de dimension finie d'un espace vectoriel normé est fermé (qu'il s'agisse d'un espace vectoriel normé de dimension finie ou non).

⊡ Propriétés géométriques

Définition 497 *Une partie A d'un espace vectoriel normé est dite* **équilibrée** *si elle contient ax pour tout a de module 1 et tout x de A.*

Théorème 498 *Soit E un espace vectoriel normé de dimension finie. Alors, soit $B \subset E$, B est la boule unité pour une certaine norme si et seulement si B est convexe équilibré compact ayant 0 comme point intérieur. La norme correspondante est alors unique.*

Démonstration : Cette preuve, utilisant la notion de jauge, est détaillée dans [22, p236].□

Définition 499 *On définit ici la notion de* **séparation au sens large (resp. strict)** *: On dit qu'un hyperplan H* sépare au sens large *deux parties A et B (resp.* sépare au sens strict*) si A et B sont inclus dans l'un et l'autre des demi-espaces fermés (resp. ouverts) délimités par H.*

Théorème 500 (Forme géométrique de Hahn-Banach en dim. finie) *Soit E un espace vectoriel normé de dimension finie. Alors :*
• Soit A ouvert convexe non vide, L sous-espace vectoriel de E de dimension finie n'intersectant pas A. Alors il existe un hyperplan H tel que $L \subset H$ et $A \cap H = \emptyset$.
• Soient A et B des convexes non vides et disjoints de E. Alors
- Si A est ouvert, il existe un hyperplan séparant A et B au sens large.
- Si A et B sont ouverts, il existe un hyperplan séparant A et B au sens strict.
- Si A est compact et B fermé, il existe un hyperplan séparant A et B au sens strict.
- Si A et B sont fermés, alors il existe un hyperplan séparant A et B.

Démonstration : Voir [19, p347] pour une preuve complète ; les points sont à démontrer dans cet ordre pour simplifier la preuve.□

Chapitre 12

Extrema

Une référence claire et complète est [7]. En outre on y trouve des algorithmes justifiés rigoureusement.

12.1 Cadre et définitions

Pour ce chapître, on travaillera sur une application f continue d'un ouvert U d'un espace de Banach E dans $\mathbb{R}$.

Définition 501 *On dit que f admet un* **minimum relatif** *ou* **minimum local** *en $x \in U$ si il existe un voisinage V de x tel que pour tout v dans V $f(x) \leq f(v)$.*
On dit que f admet un **minimum relatif strict** *ou* **minimum local strict** *en $x \in U$ si il existe un voisinage V de x tel que pour tout $v \neq x$ dans V $f(x) < f(v)$.*
On dit que f admet un **minimum global** *en $x \in U$ si pour tout v dans U $f(x) \leq f(v)$.*
On dit que f admet un **minimum global strict** *en $x \in U$ si pour tout $v \neq x$ dans U $f(x) < f(v)$.*
On définit de même les notions de **maximum relatif**, **maximum relatif strict**, **maximum global**, **maximum global strict**, *en remplaçant les $\leq$ par des $\geq$ et les $<$ par des $>$.*

12.2 Résultats liés à la compacité

Proposition 502 *Soit E un espace vectoriel normé de dimension finie et f C^0 de E dans $\mathbb{R}$ telle que*

$$lim_{\|x\| \to \infty} f(x) = \infty$$

Alors f est minorée et atteint son minimum.

Démonstration : On se donne $A > 0$ tel que $\|x\| > A$ implique $f(x) > f(0)$. On considère alors $K = \overline{B}(0, A)$. K est fermé car on est en dimension finie (les compacts d'un espace vectoriel normé de dimension finie sont les fermés bornés). f atteint donc sa borne inf (voir corollaire 182).$\square$

Corollaire 503 (Quelques applications) • *La distance d'un point à un fermé non vide est minorée et le minimum est atteint.*
• *Aussi les trois applications suivantes, empruntées à [15] :*
- étant donnée une application f C^0 de $[0,1]$ dans $\mathbb{R}$, il existe un polynôme P minimisant $\|f-P\|_\infty$ parmi les polynômes de degré $\leq n$.
- le **théorème de D'Alembert-Gauss**, *stipulant que tout polynôme à coefficients complexes et de degré ≥ 1 admet une racine, en considérant $z \mapsto |P(z)|$ (corollaire : tout polynôme à coefficients dans $\mathbb{C}$ est scindé dans $\mathbb{C}[X]$).*

Voir le corollaire 1225 sur la trigonalisation de matrices complexes, ou la partie 31.3.2 sur les suites récurrentes linéaires.

12.3 Résultats de calcul différentiel

12.3.1 Résultats au premier ordre

Théorème 504 (Condition nécéssaire du premier ordre) *Si x est un minimum relatif de f et si f est différentiable en x, alors la différentielle de f en x est nulle.*

Démonstration :
$df(x)(h) \geq 0$ car $\frac{f(x+th)-f(x)}{t} \geq 0$ si $t \geq 0$
$df(x)(h) \leq 0$ car $\frac{f(x+th)-f(x)}{t} \leq 0$ si $t \leq 0$$\square$

Pas de réciproque ; par exemple $x \mapsto x^3$ de $\mathbb{R}$ dans $\mathbb{R}$ a une différentielle nulle en 0 et n'a ni maximum ni minimum en zéro.

Définition 505 *Si $df(x) = 0$, on dit que x est un point critique.*

Pour aller plus loin on peut s'intéresser aux extréma liés ; voir pour cela [7].

12.3.2 Résultats du second ordre

Théorème 506 (Condition nécéssaire du second ordre) *Soit f deux fois différentiable en x.*
Alors si f admet un minimum local en x, $f''(x)$ est positive.

Démonstration :
Supposons $f''(x)$ non positive ; alors il existe v tel que $f''(x)(v,v)$ soit < 0. On se donne un voisinage V de x sur lequel $f \leq f(x)$ et sur lequel la formule de Taylor-Young 583 donne

$$f(x+t.v) \leq f(x) + \frac{1}{4} f''(x)(tv, tv)$$

(rappelons que par le résultat précédent $f'(x)$ est nul)

on peut toujours choisir un tel voisinage V car

$$f(x+t.v) = f(x) + \frac{1}{2} f''(x)(tv, tv) + o(t^2)$$

pour t assez petit

et donc $f(x+tv) - f(x)$ est alors négatif pour ces valeurs de t (à part pour $t = 0$).□

Théorème 507 (Condition suffisante du second ordre) *Supposons E de dimension finie. Soit f deux fois différentiables en x, $f'(x) = 0$, et $f''(x)$ définie positive. Alors f admet un minimum relatif strict en x.*

Démonstration : On considère simplement un minimum de $f''(x)(u, u)$ pour $\|u\| = 1$ et la conclusion vient rapidement.□

On peut se passer d'hypothèse de dimension finie à condition d'imposer que $f''(x)$ vérifie $\exists \alpha > 0 / f''(x)(u, u) > \alpha \|u\|^2$.

12.4 La convexité

Pour une introduction à la convexité, voir la partie 11.5.1.

Les résultats liés à la convexité sont très intuitifs, et se justifient rigoureusement sans trop de difficulté : on pourra consulter [7] pour moultes développements.

Théorème 508 *Un minimum local d'une fonction convexe définie sur une partie convexe est en fait un minimum global.*
Une fonction strictement convexe définie sur une partie convexe admet au plus un minimum et si un tel minimum existe il est strict.

12.5 Pour aller plus loin

On trouvera dans [7] des études des cas particuliers des formes quadratiques, et une justification rigoureuse de la méthode de Newton.

Chapitre 13

Equations différentielles

Une référence appréciable pour sa clarté et son souci d'illustration et d'utilisation pratique des résultats (notamment informatique) est le livre "Analyse numérique et équations différentielles" de J.-P. Demailly ([9]). Je traite ici exclusivement le cas de la dimension finie, largement suffisant pour la plupart des problèmes ; pour une analyse plus générale, on pourra consulter [3].

13.1 Lemmes préliminaires

Lemme 509 (Lemme de Gronwall) *Soit ϕ une fonction C^0 de l'intervalle $[a, b]$ dans $\mathbb{R}^+$. Soit $c \in [a, b]$, soient A, B des réels positifs.*
Supposons que pour tout t dans $[a, b]$, on ait

$$\phi(t) \leq A + B \left| \int_c^t \phi(u).du \right|$$

Alors pour tout t dans $[a, b]$

$$\phi(t) \leq A\, e^{B|t-c|}$$

Démonstration :

Pour $t \geq c$, on définit $F(t) = A + B \int_C^t \phi(s).ds$. F est C^1. Calculons la dérivée de $t \mapsto e^{-Bt}.F(t)$; cette dérivée est

$$e^{-Bt}(-B\, F(t) + B\phi(t))$$

donc est ≤ 0 sur $[c, b]$. Le résultat en découle immédiatement pour $t \geq c$.

Pour le cas restant, $t \leq c$, on définit $F(t) = A + B \int_t^c \phi(u)du$; F est C^1. Calculons la dérivée de $t \mapsto e^{Bt}.F(t)$; cette dérivée est

$$e^{Bt}(B\, F(t) - B\phi(t))$$

donc est ≥ 0 sur $[a, c]$. Le résultat en découle immédiatement pour $t \leq c$.$\square$

13.2 Equations différentielles d'ordre 1

Définition 510 *On appelle* **équation différentielle du premier ordre** *une équation de la forme*

$$\frac{dx}{dt} = f(t,x)$$

*avec E un $\mathbb{R}$-espace vectoriel de dimension finie m, U une partie (*non nécessairement ouverte ! *) de $\mathbb{R} \times E$, et f une application de U dans E, avec f continue.*
On appelle **solution** *de cette équation une application ϕ dérivable de I dans E avec I un connexe de $\mathbb{R}$ (i.e. un intervalle) telle que $\{(t,\phi(t))/t \in I\} \subset U$ et $\phi'(t) = f(t,\phi(t))$ pour tout t dans I.*

Proposition 511 *Une solution d'une équation différentielle est nécessairement C^1.*

Démonstration : L'équation exprime notamment le fait que la dérivée de sa solution est C^0.$\square$

Remarque (forme intégrale) : $\phi : I \to E$ est une solution de $\frac{\delta x}{\delta t} = f(t,x)$ de donnée initiale $\phi(c) = x_0$ si et seulement si pour tout t dans I, $\phi(t) = \phi(c) + \int_c^t f(u,\phi(u))du$.

13.2.1 Avec des hypothèses sympathiques sur f

Définition 512 *Une fonction de deux variables $(x,y) \in X \times Y \mapsto f(x,y)$ est dite* **localement lipschitzienne en** *y si pour tout (x,y) il existe un scalaire k et V un voisinage de (x,y) dans $X \times Y$ tel que pour tous x', y_1, y_2 tel que $(x',y_1) \in V$ et $(x',y_2) \in V$ on ait*

$$\|f(x',y_1) - f(x',y_2)\| \leq k.\|y_1 - y_2\|$$

Lemme 513 (Unicité) *Si f est localement lipschitzienne en x, alors si ϕ_1 et ϕ_2 sont deux solutions sur un même intervalle ayant même valeur en un certain t, alors $\phi_1 = \phi_2$.*

Démonstration : Donnons nous ϕ_1 et ϕ_2 deux solutions, égales en t.

Pour simplifier le raisonnement on va supposer qu'il existe $t'' > t$ tel que $\phi_1(t'') \neq \phi_2(t'')$. En cas contraire, on raisonnerait de même en considérant $t'' < t$ vérifiant cette propriété.

Soit t' l'inf de ces t''. Par continuité, $\phi_2(t') = \phi_1(t')$.

Soit V l'intersection de $[t', \infty[$ et d'un voisinage de t' tel que $\{(t,\phi_1(t))/t \in V\}$ et $\{(t,\phi_2(t))/t \in V$ restent dans un voisinage de $(t', \phi'(t'))$ sur lequel f est C-lipschitzienne en y.

Pour v dans V, on a $\phi_i'(v) = f(v, \phi_i(v))$, et donc

$$\phi_1(v) - \phi_2(v) = \int_0^t f(u, \phi_1(u)) - f(u, \phi_2(u))du$$

$$|\phi_1(v) - \phi_2(v)| \leq C \int_0^v |\phi_1(u) - \phi_2(u)|du$$

Donc par le lemme de Gronwall on conclut que ϕ_1 et ϕ_2 sont égales sur V, ce qui est contradictoire avec la définition de t'. □

Lemme 514 (Existence) *On se donne une fonction f C^0 de $[t_0 - a, t_0 + a] \times \overline{B}(x_0, b) \times \overline{B}(\lambda_0, c)$ dans $\mathbb{R}^n$, la boule $\overline{B}(x_0, b)$ étant une boule de $\mathbb{R}^n$, et la boule $\overline{B}(\lambda_0, c)$ étant une boule compacte ou un seul point d'un espace métrique.*
On suppose qu'il existe C tel que $\|f(t, x_1, \lambda) - f(t, x_2, \lambda)\| \leq C\|x_1 - x_2\|$, et on se donne M tel que $|f(t, x)|$ est borné par M sur $[t_0 - a, t_0 + a] \times \overline{B}(x_0, b) \times \overline{B}(\lambda_0, c)$ (un tel M existe nécessairement par continuité de f sur $[t_0 - a, t_0 + a] \times \overline{B}(x_0, b) \times \overline{B}(\lambda_0, c)$ qui est compact).
Alors il existe une fonction $\phi(., \lambda)$ telle que $\frac{\delta\phi(t,\lambda)}{\delta t} = f(t, \phi(t, \lambda), \lambda)$, définie sur $[t_0 - T, t_0 + T] \times \overline{B}(\lambda_0, c)$, avec T le min de a et b/M, et telle que pour tout λ $\phi(t_0, \lambda) = x_0$; en outre cette fonction est continue par rappport à t et λ.

Le paramètre t de la fonction $f(t, x, \lambda)$ peut être vu comme le paramètre temporel ; c'est une fonction de t que l'on cherche comme solution. ϕ est la fonction solution, dépendant de t. Quant à λ, c'est un paramètre désignant les conditions initiales. Le lemme est directement donné sous-une forme très générale ; mais le cas d'une boule $\overline{B}(\lambda_0, c)$ compacte réduite à un point n'est pas à négliger ; il s'agit en fait du cas le plus courant, l'intérêt d'introduire une boule étant simplement de montrer la continuité par rapport aux conditions initiales.

L'hypothèse de l'existence de C sera notamment vérifiée si les dérivées partielles de f par rapport aux n composantes de ϕ existent et sont continues.

Démonstration :

- On pose $x_0(t, \lambda) = x_0$
- On définit par récurrence $x_{k+1}(t, \lambda) = x_0 + \int_{t_0}^t f(u, x_k(u, \lambda))du$
- Il est clair par récurrence que $|x_k(t, \lambda) - x_0| \leq b$
- x_k est C^0 en t clairement
- x_k est C^0 en λ par continuité sous le signe intégral (théorème 381)
- On montre maintenant par récurrence que pour tout k

$$|x_{k+1}(t, \lambda) - x_k(t, \lambda)| \leq \frac{MC^k|t - t_0|^{k+1}}{(k+1)!}$$

- Le cas $k = 0$ est immédiat

- $|(x_{k+1} - x_k)(t, \lambda)| = |\int_{t_0}^{t} f(u, x_k(u, \lambda), \lambda) - f(u, x_{k-1}(u, \lambda), \lambda).du| \leq C|t - t_0| \int_{t_0}^{t} MC^{k-1} \frac{|u-t_0|^k}{k!} du \leq \frac{MC^k |t-t_0|^{k+1}}{(k+1)!}$

- $t - t_0$ étant borné sur l'ensemble qu'on s'est donné, la suite des x_k converge uniformément. Du coup la limite est continue par rapport à (t, d). Le fait que la limite vérifie l'équation est conséquence du passage à la limite.□

Théorème 515 (Cauchy-Lipschitz) *Si f est localement lipschitzienne en x, étant donné (t_0, x_0), il existe une et une seule solution maximale (i.e. sur un intervalle maximal) de l'équation différentielle.*
En outre, cette fonction maximale n'admet pas de limite au bord de l'intervalle où elle est définie, si ce bord est fini et n'est pas le bord de l'intervalle de définition de f.

Démonstration : L'existence et l'unicité découlent des lemmes ci-dessus. Lorsque la solution ne tend pas vers l'infini au bord du domaine ($\neq \infty$ et $\neq$ du bord de l'intervalle de définition de f), on peut prolonger par le lemme d'existence.□

Théorème 516 (Existence de solutions globales) *On suppose désormais U de la forme $I \times E$, avec I intervalle ouvert de $\mathbb{R}$. On suppose en outre qu'il existe une fonction k continue de I dans $\mathbb{R}^+$ telle que $f(t, .)$ est k-lipschitzienne de rapport de Lipschitz $<\ k(t)$ sur E. Alors, toute solution maximale est une solution globale (sur I).*

Démonstration : Donnons-nous un intervalle compact K inclus dans J. Il est suffisant de montrer qu'il existe une solution définie sur K. Il suffit donc d'appliquer le lemme 514, k étant majorée sur K par une certaine constante C.□

Théorème 517 (Equadif dépendante d'un paramètre) *On remplace $f(t, x)$ par $f(t, x, \lambda)$; on suppose que f est continue lipschitzienne en x, de constante de Lipschitz indépendante de t et λ, avec λ appartenant à un espace topologique L, t dans un intervalle compact de $\mathbb{R}$ et $x \in B(x_0, r) \subset E$ avec E un Banach. En outre, f est bornée par M.*
Alors étant donné $t_0 \in \mathbb{R}$ on peut à λ associer une solution ϕ_λ sur $J = I \cap [t_0 - r/M, t_0 + r/M]$, et $(t, \lambda) \mapsto \phi_\lambda(t)$ est continue.

Démonstration : La démonstration utilise le théorème du point fixe de Banach ; pour plus de précisions, on consultera [9].□

13.2.2 Sans hypothèse sympathique sur f

f sera ici simplement supposée C^0.

Définition 518 *On dira que ϕ est une* **solution ϵ-approchée** *de l'équation différentielle $\frac{\delta x}{\delta t} = f(t,x)$ si ϕ est définie continue C^1 par morceaux sur un intervalle J, si $\phi(t_0) = x_0$ et si pour tout t dans J $(t, \phi(t))$ est bien dans U, avec $\|\phi'(t) - f(t, \phi(t))\| \leq \epsilon$, à part aux points de discontinuité, auxquels on doit avoir $\|\phi'_d(t) - f(t, \phi(t))\|$ et $\|\phi'_g(t) - f(t, \phi(t))\|$ tous deux $\leq \epsilon$, avec ϕ'_d et ϕ'_g les dérivées à droites et à gauche.*

Théorème 519 *Supposons f C^0, définie sur $I \times B(x_0, r)$, avec I intervalle de $\mathbb{R}$, à valeurs dans E. Supposons $|f| \leq M$. Alors avec $J = I \cap [t_0 - r/M, t_0 + r/M]$, $\forall \epsilon > 0$, l'équation $\frac{\delta x}{\delta t} = f(x,t)$ a une solution ϵ-approchée affine par morceaux telle que $\phi(t_0) = x_0$.*

__**Démonstration :**__

• On montre qu'on peut construire une telle solution sur $J^+ = J \cap [t_0, \infty[$, le résultat s'obtenant de même sur $J^- = J \cap] - \infty, t_0]$.

• Définissons tout d'abord ϕ_0 affine, définie par $\phi_0(t_0) = x_0$, et $\phi'_0(t_0) = f(t_0, x_0)$.

• On note que ϕ_0 est bien telle que $(t, \phi(t))$ soit dans le domaine de définition de f pour t dans J^+.

• Par continuité de f, ϕ_0 est une solution ϵ-approchée sur $[t_0, t]$, pour t suffisamment petit. En considérant t_1 le sup de ces t, on obtient t_1 [1], et par continuité ϕ_0 est une solution ϵ-approchée sur $[t_0, t_1]$.

• Si t_1 est différent de $sup\ J$, alors on recommence le même processus, en remplaçant x_0 par $x_1 = f(t_1)$, et t_0 par t_1, et r par $r - \|x_1 - x_0\|$ (> 0 par définition de J) ; on nomme ϕ_1 la nouvelle application obtenue. Puis t_2, puis t_3, et ainsi de suite, jusqu'à ce que t_i soit le sup de J, auquel cas la preuve est terminée.

• Supposons maintenant que la suite des t_i croît sans jamais atteindre la borne sup de J ; notons T le sup des t_i.

• Remarquons que $\|x_{n+1} - x_n\| \leq M|t_{n+1} - t_n|$ et que donc $\| \sum_{m=n}^{n+p} x_{n+1} - x_n\| \leq M(T - t_n)$; donc par le critère de Cauchy (rappelons que E est un Banach), la suite (x_n) tend vers une certaine limite x.

• Considérons maintenant

$$\|f(t_n, x_n) - f(t, \phi_n(t))\|$$

pour $T \geq t \geq t_n$, et examinons ce qu'il se passe pour $n \to \infty$.

Cette quantité est inférieure ou égale à

$$\|f(\underbrace{t_n}_{\to T}, \underbrace{x_n}_{\to x}) - f(T,x)\| + \|f(T,x) - f(\underbrace{t}_{\to T}, \underbrace{\phi_n(t)}_{\to x \text{ car } \|\phi_n(t) - x_n\| \leq (t-t_n)M})\|$$

[1] Je passe sous silence le cas $t_1 = \infty$, qui termine la preuve immédiatement.

et donc tend vers 0 en l'infini, et donc finit par être inférieure à ϵ à un certain rang n ; donc $t_{n+1} \geq T$, d'où la contradiction.$\square$

Corollaire 520 *Supposons E de dimension finie. f étant toujours bornée dans un voisinage de (t_0, x_0), on peut toujours trouver un voisinage de t_0 sur lequel l'équation admet des solutions ϵ-approchées pour tout ϵ.*

On utilise le fait que dans le théorème précédent, le voisinage obtenu est indépendant de ϵ.

13.3 Equation différentielle d'ordre n

Etant donnée une équation de la forme

$$\frac{d^n x}{dt} = f(t, x, \frac{dx}{dt}, ..., \frac{d^{n-1}x}{dt^{n-1}})$$

où f est supposée localement lipschitzienne en x, $\frac{dx}{dt}, ..., \frac{d^{n-1}x}{dt^{n-1}}$, avec $U \subset \mathbb{R} \times E^n$, f de U dans E continue à valeurs dans $\mathbb{R}^m$, on se ramène à

$$\frac{dx}{dt} = x_1$$

$$\frac{dx_1}{dt} = x_2$$

$$\cdots$$

$$\frac{dx_i}{dt} = x_{i+1}$$

$$\frac{dx_{n-1}}{dt} = f(t, x, x_1, ..., x_{n-1})$$

Ces équations différentielles correspondent donc à une équation d'ordre 1 dans l'espace E^n.

Il reste à reformuler les différents résultats sur les équations différentielles d'ordre 1 au cas de l'ordre n :

Etant donnés t_0 et x_0, x_1, ... , x_{n-1}, il existe (au moins) une solution maximale x définie sur un intervalle ouvert contenant t_0 telle que $x(t_0) = x_0$, $x'(t_0) = x_1$, ... $x^{(n-1)}(t_0) = x_{n-1}$ (rappelons que l'on a supposé f continue). f étant localement lipschitzienne en x, $\frac{dx}{dt}, ..., \frac{d^{n-1}x}{dt^{n-1}}$, alors il y a unicité. Pour l'existence sans l'unicité, le théorème de Cauchy-Péano permet de se passer du caractère localement lipschitzien.

Si U est de la forme $I \times E^n$, et s'il existe une fonction continue dépendant seulement de t majorant le coefficient de lipschitz, alors les solutions maximales sont définies sur I tout entier.

Ces résultats découlent immédiatement des résultats à l'ordre 1, grâce à la transformation décrite ci-dessus.

13.4 Zoologie des équations différentielles

13.4.1 Equation différentielle linéaire du premier ordre

Définition 521 *Une* **équation différentielle linéaire du premier ordre** *est une équation différentielle du premier ordre avec*

$$\frac{dx}{dt} = A(t).x + B(t)$$

avec $A(t)$ une application linéaire continue de E dans E et $B(t) \in E$, pour tout t dans I intervalle de $\mathbb{R}$.
E est toujours un $\mathbb{R}$-espace vectoriel , A est continue, B est continue.
*L'***équation différentielle homogène associée** *est*

$$\frac{dx}{dt} = A(t).x$$

Théorème 522 (Théorème de Cauchy) *Si A est continue de I dans $\mathcal{L}(E)$[a], alors l'équation différentielle linéaire du premier ordre admet une solution ϕ telle que $\phi(t_0) = x_0$ définie sur tout I.*
Il existe une unique solution définie sur tout I.

[a] Ensemble des applications linéaires continues de E dans E.

Démonstration : Il s'agit directement d'une application du théorème 516. □

Théorème 523 *En notant cette solution ϕ_{x_0} (à t_0 fixé) la solution de l'équation linéaire homogène associée, l'application qui à x_0 associe ϕ_{x_0} est linéaire bijective de E dans l'ensemble des solutions de l'équation homogène.*

Démonstration :
• L'injectivité est claire, si deux fonctions sont différentes en t_0 alors elles sont différentes tout court.
• La surjectivité est non moins claire, par définition de ϕ_{x_0}.
• La linéarité, enfin est immédiate ; il suffit de voir que si x et y sont solutions, alors $\lambda.x + \mu.y$ est aussi solution.□

Corollaire 524 (Cas de la dimension finie) *On en déduit au passage que la dimension de l'espace des solutions d'une équation diférentielle est finie et égale à la dimension de E, lorsque la dimension de E est finie.*

⊡ **Cas général, dimension non nécessairement finie**

Dans cette partie, et seulement celle-ci, on traitera un cadre plus général. L'intérêt est seulement de donner un exemple d'utilisation en dimension non finie. On admettra le fait que le théorème de Cauchy-Lipschitz est aussi valable dans le cas d'un espace de Banach, même s'il n'est pas de dimension finie (la démonstration est d'ailleurs la même).

⋄ A **non constant**

Définition 525 *On appelle* **équation résolvante** *de l'équation différentielle linéaire de la définition 521 l'équation à paramètre dans l'ensemble des applications C^1 de I dans $\mathcal{L}(E)$:*

$$U'(t) = A(t) \circ U(t)$$

Afin de pouvoir travailler sur cette équation, nous aurons besoin du théorème de Cauchy ; aussi devons nous bien voir :

- que $\mathcal{L}(E)$ est un Banach (de manière générale l'ensemble des applications linéaires continues d'un espace vectoriel normé dans un Banach est un Banach)

- que l'application $t \mapsto (\phi \mapsto A(t) \circ \phi)$ est continue de I (pour la topologie usuelle) dans $\mathcal{L}(E)^{\mathcal{L}(E)}$ (pour la topologie produit)

Le deuxième point découle facilement de la continuité de A.

On peut donc appliquer le théorème de Cauchy, et exhiber pour tout t_0 dans I une solution unique $Resolvant_{t_0}$ sur tout I de l'équation résolvante vérifiant

$$Resolvant_{t_0}(t_0) = Id_E.$$

Définition 526 *La solution $Resolvant_{t_0}(t_0)$ est appelée* **résolvante d'origine** t_0.

Voyons maintenant les propriétés sympathiques de la résolvante, qui découlent des résultats ci-dessus.

Proposition 527 • *Pour tous a, b et c dans I, $Resolvant_a(b).Resolvant_b(c) = Resolvant_a(c)$.*
• *Pour tout a et tout b $Resolvant_a(b)$ est dans $GL(E)$.*

__**Démonstration :**__

• Il suffit de dériver tout ça, et d'utiliser l'unicité donnée par le théorème de Cauchy.
• C'est une conséquence évidente du fait que $Resolvant_a(b).Resolvant_b(a) =$

$Resolvant_b(a).Resolvant_a(b) = Resolvant_a(a) = Resolvant_b(b) = Id_E$.□

Théorème 528 • *La solution x de l'équation homogène associée vérifiant $x(t_0) = x_0$ est l'application $t \mapsto Resolvant_{t_0}(t).x_0$.*
• *Une solution particulière x de l'équation générale de la définition 521 vérifiant $x(t_0) = x_0$ est donnée par*

$$x(t) = Resolvant_{t_0}(x_0) + \int_{t_0}^{t} Resolvant_u(t)B(u)du$$

Démonstration :

• La résolvante est construite pour ça. Il suffit d'écrire la dérivée de

$$t \mapsto Resolvant_{t_0}(t).y_0$$

pour avoir la résultat souhaité.

• Il suffit d'écrire que x peut s'exprimer sous la forme $x(t) = Resolvant_{t_0}(t)(C(t))$; ensuite, par une méthode bien similaire à la méthode de variation des constantes, on écrit

$$x'(t) = \frac{(\delta((x,y) \mapsto Resolvant_{t_0}(x)(y)) \circ (t \mapsto (t, C(t))))}{\delta t}$$

$$= (((\frac{\delta Resolvant_{t_0}(x)}{\delta x})(y)dx + (Resolvant_{t_0}(x)dy)) \circ (t, C(t))).(1, C'(t))$$

$$= A(t)Resolvant_{t_0}(t)C(t) + Resolvant_{t_0}(t)C'(t)$$

Donc x sera solution si $B(t) = Resolvant_{t_0}(t)C'(t)$, c'est à dire si $C(t) = Resolvant^{-1}_{t_0,t}(x_0) + \int_{t_0}^{t} Resolvant^{-1}_{t_0}(u)B(u)du$, donc si $x(t) = Resolvant_{t_0}(x_0) + Resolvant_{t_0}(t).\int_{t_0}^{t} Resolvant_u(t_0)B(u)du$.

Donc on a bien

$$x(t) = Resolvant_{t_0}(t)x_0 + \int_{t_0}^{t} Resolvant_u(t)B(u)du$$

D'où le résultat. □

◇ **A constant**

Théorème 529 *Si $A(t) = A$ est constant, alors l'équation différentielle linéaire définie en 521 admet pour unique solution x vérifiant $x(t_0) = x_0$ l'application*

$$x : t \mapsto exp((t - t_0)A)(x_0)$$

Pour se rappeler de ce qu'est l'exponentielle d'un endormophisme continu d'un Banach on pourra consulter 15.12.4.

Démonstration : Le théorème de Cauchy nous donne l'unicité, et il est immédiat que cette fonction convient, par le théorème donnant la dérivée de l'application $t \mapsto exp(t\ f)$ (voir partie 15.12.4).□

On note que $t \mapsto exp((t-t_0)A)(x_0)$ est le résolvant de l'équation différentielle.

Théorème 530 *Une solution particulière x de l'équation différentielle générale définie en 521 dans le cas où $A(t) = A$ est constant et vérifiant $x(t_0) = x_0$ est donnée par*

$$x(t) = exp((t-t_0)A).x_0 + \int_{t_0}^{t} exp((t-u)A)B(u)du$$

Démonstration : On peut simplement argumenter en utilisant le fait signalé ci-dessus, ie que $t \mapsto exp((t-t_0)A)$ est le résolvant (d'origine t_0) de l'équation différentielle, mais on peut aussi faire le calcul directement en cherchant des solutions de la forme $t \mapsto exp(tA)C(t)$.□

Remarque : tout comme lorsque A est constant on a $Resolvant_{t_0}(t) = exp((t-t_0)A)$, on a $Resolvant_{t_0}(t) = exp(\int_{t_0}^{t} A(u)du$ LORSQUE pour tous t et s dans I $A(t)$ et $A(s)$ commutent ($A(t)A(s) = A(s)A(t)$).

⊡ Cas de la dimension finie

⋄ $A(t)$ non constant

Si E est de dimension finie n, alors l'espace des solutions de l'équation homogène associée est de dimension finie n, comme on le souligne dans le corollaire 524. Pour simplifier les notations, on identifie E et $\mathbb{R}^n$, sans perte de généralité.

Proposition 531 *On se donne une famille $x_1, ..., x_m$ de solutions de l'équation différentielle homogène. Alors ces solutions sont libres si et seulement si l'ensemble des $x_i(t)$ est libre pour un certain t, si et seulement si l'ensemble des $x_i(t)$ est libre pour tout t.*

Démonstration :

Il est clair que si les $x_i(t)$ forment une famille libre pour un certain t, alors les x_i forment une famille libre.

Il est clair que si pour tout t, les $x_i(t)$ forment une famille libre, il en est de même.

Il reste donc juste à voir que si les x_i forment une famille libre, alors les $x_i(t)$ forment une famille libre, quel que soit t. Cela découle simplement du théorème 523.□

⚠ Il n'est par contre pas vrai que dans le cas général, (x_i) famille libre $\rightarrow (x_i(t))$ famille libre (par exemple n'importe quelle famille libre de fonctions de $\mathbb{R}$ dans $\mathbb{R}$).

Ainsi lorsque l'on aura obtenu une famille libre de n solutions de l'équation différentielle homogène et une solution de l'équation générale, alors on pourra en déduire toutes les solutions de l'équation différentielle, en considérant la somme de la solution de l'équation générale plus une combinaison linéaire quelconque des n solutions libres de l'équation homogène associée.

Généralement, le problème ne sera pas d'obtenir les solutions de l'équation homogène, mais plutôt d'obtenir les solutions de l'équation générale.

Pour cela on utilisera notamment la méthode de la **variation des constantes**.

On suppose que $x_1, ..., x_n$ sont des solutions libres de l'équation homogène associée.

On cherche alors x solution particulière de l'équation générale, avec x de la forme

$$x(t) = \lambda_1(t)x_1(t) + \lambda_2(t)x_2(t) + \cdots + \lambda_n(t)x_n(t)$$

On note bien que toute fonction peut s'exprimer de la sorte, puisque pour tout t la famille des $x_i(t)$ est libre.

Faisons maintenant « varier les constantes » :

$$\begin{aligned}\frac{dx}{dt}(t) = & \underbrace{\lambda_1\frac{dx_1}{dt} + \cdots + \lambda_n\frac{dx_n}{dt}} \\ & = \lambda_1 Ax_1 + \lambda_2 Ax_2 + ... + \lambda_n Ax_n \text{ par définition des } x_i. \\ & + \frac{\delta\lambda_1}{\delta t}(t)x_1(t) + \ldots \frac{\delta\lambda_n}{\delta t}(t)x_n(t)\end{aligned}$$

Donc x vérifie l'équation générale si et seulement si

$$\sum_i \lambda_i'(t)x_i(t) = B(t)$$

En écrivant $\Lambda = (\lambda_1, ..., \lambda_n)$ et $M(t) = \begin{pmatrix} (x_1)_1 & (x_1)_2 & \ldots & (x_1)_n \\ (x_2)_1 & (x_2)_2 & \ldots & (x_2)_n \\ \vdots & \vdots & \vdots & \vdots \\ (x_n)_1 & (x_n)_2 & \ldots & (x_n)_n \end{pmatrix}$[2]

On obtient l'équation $\Lambda' = M(t)^{-1}B(t)$ (notez que $M(t)$ est inversible, de par la proposition 531).

On peut donc en déduire les Λ, à une constante près. Les constantes en question ne changent de toute façon rien, puisque cela revient à ajouter une combinaison linéaire des solutions de l'équation homogène.

◇ **Cas $A(t) = A$ constant**

Tout d'abord on peut donner la forme générale des solutions, de manière simple lorsque l'endomorphisme est diagonalisable en dimension finie.

[2] On vérifiera facilement qu'il s'agit du résolvant (voir partie 525).

Théorème 532 *On va supposer ici que $A(t) = A$ est constant, et qu'il s'agit d'un endomorphisme diagonalisable, en dimension finie n. Alors avec (e_i) une base de E dans laquelle A s'identifie à une matrice diagonale, avec $A(e_i) = \lambda_i.e_i$, les solutions de l'équation homogène associée à l'équation différentielle linéaire définie en 521 sont les combinaisons linéaires de fonctions de la forme*

$$f_i : t \mapsto exp(\lambda_i.t)e_i$$

pour $i \in [1, n]$

Démonstration : Il est facile de voir que ces fonctions sont bien dans l'ensemble des solutions de l'équation différentielle homogène associée. En vertu du corollaire diffé524 il suffit donc de vérifier que les solutions en question forment bien une famille libre ; ce fait se déduit immédiatement du fait que la famille des $f_i(0)$ est libre.

13.4.2 Equations différentielles autonomes

Définition 533 *Une équation différentielle est dite* **autonome** *si f ne dépend pas de t.*
On appelle **point d'équilibre** *d'une équation différentielle autonome un point x_0 tel que $f(x_0) = 0$.*
On appelle **point stable** *d'une équation différentielle autonome un point d'équilibre x_0 tel que*
$\forall \epsilon > 0 \exists \eta$ tel que pour tout x solution de l'équation différentielle et tout t_0 tel que $\|x(t_0) - x_0\| \leq \eta$
• *x est définie sur $[t_0, \infty[$*
• *$\|x(t) - x_0\| \leq \epsilon$ pour tout $t \geq t_0$*
On appelle **point asymptotiquement stable** *d'une équation différentielle autonome un point d'équilibre x_0 tel que pour un certain η, pour tout x solution de l'équation différentielle et tout t_0 tel que $\|x(t_0) - x_0\| \leq \eta$,*
• *x est définie sur $[t_0, \infty[$*
• *$lim_{t \to \infty} x(t) = x_0$*

C'est à dire qu'une équation différentielle autonome est de la forme $\frac{\delta x}{\delta t} = f(x)$.
Les résultats d'unicité permettent de dire que si

$$\frac{dx}{dt} = f(x) \text{ et } x(u) = x_0$$

et

$$\frac{dy}{dt} = f(y) \text{ et } y(v) = x_0$$

avec I intervalle maximal de définition de x et J intervalle maximal de définition de y, alors $J + u = I + v$ et $y(t + v) = x(t + u)$ pour tout t tel que $t + v \in J$.

Quelques exemples :

• Equation $x' = x$: 0 est l'unique point d'équilibre ; il n'est ni stable ni asymptotiquement stable.

• Equation $x' = -x$: 0 est l'unique point d'équilibre ; il est stable et asymptotiquement stable.

• Equation $x' = M.x$ avec M antisymétrique : 0 est point d'équilibre ; il est stable, mais pas asymptotiquement stable.

• Equation $x' = u$, avec u vecteur non nul : pas de point d'équilibre.

13.4.3 Equation de la chaleur

⊡ Le problème

Définition 534 *On définit* $\Omega =]0,1[\times\mathbb{R}^{+*}$ *;* $\overline{\Omega}$ *désigne l'adhérence de* Ω *dans* $\mathbb{R}^2$*, c'est à dire* $[0,1]\times\mathbb{R}^+$*.*
On cherche u *continue de* $\overline{\Omega}$ *dans* $\mathbb{R}$*, telle que* $u_{|\Omega}$ *soit* C^∞ *et vérifie*

$$\frac{\delta u}{\delta t} - \frac{\delta^2 u}{\delta x^2} = 0$$

(équation de la chaleur)
avec les conditions aux limites (CLs) :

$$u(0,t) = u(1,t) = 0$$

et les conditions initiales (CIs) :

$$u(x,0) = h(x)$$

avec h *une certaine fonction continue de* $[0,1]$ *dans* $\mathbb{R}$*,* C^1 *sur* $]0,1[$*, telle que* $h(0) = h(1) = 0$ *(indispensable pour que les CLs puissent être vérifiées).*

⊡ La méthode

Pour attaquer cette équation, comme d'autres équations vérifiant une équation de la même forme (dérivée première en fonction du temps égale à une dérivée seconde (un laplacien) en coordonnées d'espace), on cherche en fait une solution $u(x,t) = f(x)g(t)$ s'exprimant comme produit d'un terme d'espace par un terme de temps. On ne se préoccupera pas pour le moment de la CI.

Une fois des solutions trouvées, on remarquera que les solutions forment un espace vectoriel. On cherchera alors une solution combinaison linéaire vérifiant la CI. Pour cela, puisqu'on aura remarqué que nos solutions en x sont des sinusoïdes (ayant toutes pour fréquence un multiple d'une certaine fréquence fondamentale) on considérera l'antisymétrisée de la fonction des CI, pour considérer un développement en série de Fourier qui ne comporte que des sinusoïdes (il n'y aura pas de cosinusoïdes puisque l'on considèrera une fonction impaire !).

On obtiendra ainsi une solution. Il existe une preuve d'unicité, qui ne sera pas exposée ici : on la trouvera par exemple dans [22, p103].

⊡ **Les calculs**

Ecrivons
$$u(x,t) = f(x)g(t)$$

Alors l'équation de la chaleur s'écrit

$$f(x)g'(t) = f''(x)g(t)$$
$$\frac{g'(t)}{gt} = \frac{f''(x)}{f(x)} = \lambda$$

Ce terme est indépendant de x (à cause du terme de gauche) et de t (à cause du terme de droite). λ est donc une constante.

Le cas $\lambda > 0$ et le cas $\lambda = 0$ nous amènent, via les CLs, au cas $u = 0$, peu intéressant. Il reste donc seulement le cas $\lambda < 0$.

On a une équation de degré 2, qu'on peut réécrire comme une équation de degré 1,

$$f''(x) = \lambda f(x) \text{ équivaut à } \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ \lambda & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

Cette équation est linéaire, et admet donc des solutions sur tout $]0,1[$; il s'agit d'une équation homogène (pas de second membre) elle est dans un espace de dimension 2, donc l'espace des solutions est de dimension 2.

On a deux solutions évidentes : $x \mapsto cos(\omega x)$ et $x \mapsto sin(\omega x)$ (avec $\omega^2 = -\lambda$). La solution cosinusoïde ne satisfait pas les CLs, donc on garde les sinusoïdes. On déduit des CLs que ω doit être de la forme $\omega_n = n\Pi$ pour $n \in \mathbb{N}$.

Pour g, l'espace des solutions est de dimension 1, il s'agit d'une exponentielle décroissante, $x \mapsto Ce^{-\omega^2 t}$ (on applique l'équation de la chaleur pour trouver le ω^2, où on remarque et on utilise l'équation $\frac{g'}{g} = \lambda$ de la page précédente).

On a donc des solutions en u : $(x,t) \mapsto e^{-\omega_n^2 t} sin(\omega_n x)$. On note que les solutions sont un espace vectoriel. On a donc pour solution de l'équation de la chaleur et des CLs <u>au moins</u> (on n'a pas prouvé que c'étaient là les seules solutions) les combinaisons linéaires de solutions de cette forme. On va en fait considérer aussi les combinaisons linéaires infinies $\sum_{n\geq 0} a_n e^{-\omega_n^2 t} sin(\omega_n x)$, pourvu que la série $\sum a_n$ soit absolument convergente ; ainsi les théorèmes de dérivation sous le signe intégrale s'appliquent et la combinaison linéaire obtenue est bien une solution de l'équation.

On se préoccupe maintenant des CIs, en cherchant une solution combinaison linéaire des solutions trouvées ci-dessus.

On voudrait donc $h(x) = \sum_{n\geq 0} a_n \sin(n\Pi x)$. On va donc décomposer h en série de Fourier. Pour cela on va choisir h impaire, pour n'avoir que des sinusoïdes.

Donc on définit h sur $[-1,0]$ par $h(-x) = -h(x)$. Ensuite on prolonge h par 2 périodicité. On a alors $h(x) = \sum_{n\geq 0} a_n \sin(n\Pi x)$, avec convergence normale de la somme des a_n puisque h est C^1 (pas de problème en 0 ou en 1 car $h(0) = h(1) = 0$).

L'unicité de la solution ainsi obtenue ne sera pas détaillée ici ; voir [22, p103].

13.4.4 Equations à variables séparées

Définition 535 *On appelle* **équation à variables séparée** *une équation différentielle que l'on peut réécrire sous la forme*

$$x' = f(t)g(x)$$

• Si g s'annule en une valeur particulière x_p, alors la fonction constante $x = x_p$ est clairement solution particulière.

• L'équation peut se réécrire $\frac{dx}{g(x)} = f(t)dt$, et donc on obtient en intégrant chaque membre une expression de $\int 1/g$ en fonction de $\int f$. Il est clair que $\int 1/g$ est continue strictement monotone sur les intervalles sur lesquels $g(x)$ ne s'annule pas, et que donc on en déduit x en fonction de t en considérant l'inverse de $\int 1/g$. Par les résultats d'unicité si la condition initiale (t_0, x_0) est telle que $g(x_0) \neq 0$ alors par le lemme 513 (si les fonctions en jeu vérifient bien les conditions énoncées !) on a $\forall t, g(x(t)) \neq 0$ si x est une solution maximale (en effet en cas contraire x serait la fonction constante égale à x_p avec $g(x_p) = 0$), et donc on obtient bien ainsi des solutions maximales.

On trouvera dans [9] un exemple bien détaillé.

13.4.5 Equation de Bernoulli

Définition 536 *On appelle* **équation de Bernoulli** *une équation de la forme*

$$x' = p(t)x + q(t)x^\alpha$$

avec $\alpha \in \mathbb{R} \setminus \{1\}$ et p et q continues.

On se place sur les intervalles où x ne s'annule pas. On peut alors diviser par x^α, et poser $z = x^{1-\alpha}$, on obtient alors

$$\frac{dz}{dt} = (1-\alpha)(p(t)z + q(t))$$

équation linéaire en z, que l'on sait donc résoudre.

13.4.6 Equation de Ricatti (polynôme à coefficients dépendant de t de degré 2 en x)

Définition 537 *On appelle* **équation de Riccati** *une équation de la forme*

$$x' = a(t)x^2 + b(t)x + c(t)$$

avec a, b et c trois fonctions continues.

N'ayant pas de solution magique, on choisit de supposer que l'on est capable d'exhiber une solution particulière x_p.

On pose alors $z = x - x_p$, et on obtient miraculeusement

$$z'(t) = [2a(t)x_p(t) + b(t)] \times z(t) + a(t)z^2(t)$$

On se ramène donc à un cas particulier d'équation de Bernoulli, que l'on résoud comme expliqué en 13.4.5.

13.4.7 Equations homogènes

Définition 538 *On appelle équation homogène une équation de la forme $x' = f(x/t)$, avec f C^0 d'un intervalle de $\mathbb{R}$ dans $\mathbb{R}$.*

Une telle équation se résoud classiquement en posant $y = x/t$; on se ramène alors à l'équation à variables séparées $y' = (f(y) - y)/t$, qui se résoud elle-même par la méthode exposée en partie 13.4.4. On a des solutions constantes de la forme $z = a$, soit $y = ax$, pour a vérifiant $f(a) = a$.

On montre facilement que l'image d'une solution par une homothétie est encore une solution (seul comptant le rapport x/t).

13.4.8 Equation de Lagrange

Définition 539 *L'équation de Lagrange est $x(t) = a(x'(t))t + b(x'(t))$, avec a et b des fonctions C^1.*

On la résoud de la manière suivante :

• chercher les solutions à x' constant sur un intervalle ; ce sont les $x' = c$, avec $a(c) = c$. Les solutions sont alors les $x = ct + b(c)$.

• poser $y = x'$, $t = g(x)$, pour chercher d'autres solutions. On obtient

$$\frac{dt}{dy} = \frac{a'(y)t + b'(y)}{y - a(y)}$$

qui est une équation différentielle linéaire en t.

Chapitre 14

Formes différentielles

Ce chapître a seulement pour but de fournir un cadre utile dans la vie de tous les jours, pour maîtriser les outils utiles pour s'attaquer à de nombreux théorèmes. L'intérêt n'étant pas d'épuiser les richesses des formes différentielles, de nombreuses définitions et de nombreux théorèmes seront donnés sans justification, notamment dans les fondements des formes différentielles, au niveau des propriétés d'algèbre multilinéaire.

14.1 Généralités, rappels sur les applications multilinéaires

14.1.1 Définition d'une forme différentielle

Tout de go, définissons tout d'abord ce qu'est une application différentielle :

Définition 540 *Soit U un ouvert de E un $\mathbb{R}$-espace de Banach, soit F un $\mathbb{R}$-espace de Banach.*
On appelle **forme différentielle de degré p sur U à valeurs dans F** *une application de U dans $\mathcal{A}_p(E;F)$* [a]*.*
La forme différentielle est dite de classe C^n si l'application est C_n (pour $n \in \mathbb{N} \cup \{\infty\}$).
On note $\Omega_p^{(n)}(U,F)$ l'ensemble des p formes différentielles de U dans F de classe C^n.

[a] Espace des applications p-linéaires alternées continues de E dans F. Cet espace est un Banach, car c'est un sous-espace vectoriel fermé de $\mathcal{L}_p(E,F)$ (qui est un Banach comme chacun sait).

Exemples :
- Une application C^n de E dans F est une 0-forme différentielle de E dans F, de classe C^n.
- Si $n > 0$, sa différentielle est une forme 1-différentielle de classe C^{n-1}.

Nous allons définir plus loin de nombreuses opérations sur cet outil, mais tout d'abord nous devons rappeler certaines propriétés des applications multilinéaires.

14.1.2 Propriétés des applications multilinéaires

Définition 541 *Etant donnée une application ϕ bilinéaire de $F \times G$ dans H, on définit une* **multiplication d'applications p-linéaires alternées** *par :*

$$\mathcal{A}_p(E,F) \times \mathcal{A}_q(E,G) \to \mathcal{A}_{p+q}(E,H)$$

$$(f,g) \mapsto f \wedge_\phi g$$

définie par

$$(f \wedge_\phi g)(x_1, \dots, x_{p+q}) =$$

$$\sum_\sigma \epsilon(\sigma)\phi(f(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(p)}), g(x_{\sigma(p+1)}, x_{\sigma(p+2)}, \dots, x_{\sigma(p+q)}))$$

La sommation étant étendue à l'ensemble des permutations σ de $[1,n]$ telles que $\sigma(1) < \sigma(2) < \dots < \sigma(p)$ et $\sigma(p+1) < \sigma(p+2) < \dots < \sigma(p+q)$.

Il conviendrait de montrer que $f \wedge_\phi g$ est bien $p+q$ linéaire, continue et alternée.

souvent on s'abstiendra de noter $\wedge_\phi$; on se contentera de $\wedge$. ϕ sera souvent implicitement l'application la plus intuitive ; par exemple si H et G sont égaux et si F est $\mathbb{R}$, on utilisera le produit d'un élément d'un Banach par un réel.

Proposition 542 (Propriété du produit d'applications multilinéaires) $\wedge_\phi$ *est bilinéaire.*

Proposition 543 (Propriétés du produit de formes multilinéaires) • *Si f appartient à $\mathcal{A}_p(E,\mathbb{R})$ et g appartient à $\mathcal{A}_q(E,\mathbb{R})$, alors $f \wedge g = (-1)^{pq} g \wedge f$*
• *Si f ppartient à $\mathcal{A}_p(E,\mathbb{R})$, g appartient à $\mathcal{A}_q(E,\mathbb{R})$ et h appartient à $\mathcal{A}_r(E,\mathbb{R})$, alors $(f \wedge g) \wedge h = f \wedge (g \wedge h)$.*
• *Si les f_i sont des formes linéaires continues sur E (dans $\mathbb{R}$), pour $i \in [1,n]$, alors*

$$f_1 \wedge \dots \wedge f_n(x_1, \dots, x_n) = \sum_{\sigma \in \sigma^n} \epsilon(\sigma) f_i(x_{\sigma(i)}) = det\ (f_i(x_j))_{i,j}$$

Proposition 544 (Propriétés des application p-linéaires avec $\dim E = n$) *E est ici supposé isomorphe à $\mathbb{R}^n$.*
Toute application p-linéaire de E dans F s'écrit de manière unique

$$x \mapsto \sum_{1 \le i_1 < i_2 < \cdots < i_p \le n} \underbrace{c_{i_1,\ldots,i_p}}_{\in F} e^*_{i_1} \wedge e^*_{i_2} \wedge \ldots e^*_{i_p}$$

avec la famille des e^*_i la base duale de la base des e_i (base canonique de $\mathbb{R}^n$), c'est à dire que les e^*_i sont les formes qui donnent les coordonnées d'un point. En particulier, si $p = n$, l'application s'écrit $x \mapsto (e^*_1 \wedge e^*_2 \wedge \cdots \wedge e^*_n)(x)c$, avec c un élément de F, et si $p > n$, l'application est nécéssairement nulle.

14.1.3 Application de tout ça aux formes différentielles

Définition 545 (Produit extérieur de formes différentielles) *U désigne un ouvert d'un espace de Banach E. F, G et H sont des espaces de Banach .*
On se donne ϕ une application bilinéaire de $F \times G$ dans H. On suppose que $f \in \Omega^{(n)}_p(U, F)$ et que $g \in \Omega^{(n)}_q(U, G)$.
On définit alors le **produit extérieur des formes différentielles** *f* **et** *g $f \wedge_\phi g \in \Omega^{(n)}_{p+q}(U, H)$ par*

$$f \wedge_\phi g : x \mapsto f(x) \wedge_\phi g(x)$$

La notation est abusive du fait que l'on garde la même notation que pour le produit d'applications multilinéaires. Là aussi on négligera souvent de préciser $\wedge_\phi$, et on gardera $\wedge$.

Exemple :

Si $F = G = H = \mathbb{R}$, ϕ est alors généralement implicitement le produit usuel. Alors le produit de formes différentielle est anticommutatif et associatif, et

$$(\omega_1 \wedge \omega_2 \wedge \cdots \wedge w_n)(x)(h_1, \ldots, h_n) =$$

$$\sum_{\sigma \in \sigma_n} \epsilon(\sigma) w_1(x)(h_{\sigma(1)}) w_2(x)(h_{\sigma(2)}) \ldots w_n(x)(h_{\sigma(n)}) = det\ (\omega_i(e_j)_{i,j})$$

Chapitre 15

Quelques rappels et compléments d'analyse

15.1 Rappels sur le corps des réels

On considère un corps $\mathbb{K}$, commutatif, muni d'une relation d'ordre total $\leq$.

Définition 546 (Définitions de base) $\mathbb{K}$ *est* **totalement ordonné** *si* $\forall (x, y) \in \mathbb{K}^2\ 0 \leq x \wedge 0 \leq y \rightarrow 0 \leq x + y \wedge 0 \leq x.y$
$\mathbb{K}$ *est* **archimédien** *si* $\forall (x, y) \in \mathbb{K}^2\ 0 \leq x \wedge 0 < y \rightarrow \exists n \in \mathbb{N}/x < y + ... + y$*(*$n$ *fois)*
$\mathbb{K}$ *a la* **propriété de la borne supérieure** *si toute partie non vide majorée de* K *admet une borne supérieure.*
On appelle **corps réel** *un corps commutatif totalement ordonné possédant la propriété de la borne supérieure. On le note* $\mathbb{R}$*. On admettra l'existence et l'unicité à isomorphisme près d'un tel corps. Il possède en outre la* **propriété de la borne inférieure** *et il est archimédien (comme on peut le prouver facilement).*

Propriétés :
• $\mathbb{R}$ est de caractéristique nulle ; $1 + 1 + ... + 1 + 1$ est différent de 0.
• $\mathbb{R}$ est infini.

Définition 547 (Quelques définitions supplémentaires) • *On appelle* **valeur absolue** *de $x \in \mathbb{R}$ et on note $|x|$ le réel $sup(\{x, -x\})$; c'est une norme, dite* **norme usuelle**, *de $\mathbb{R}$ en tant que $\mathbb{R}$-espace vectoriel ; la métrique associée est dite* **distance usuelle** *de $\mathbb{R}$.*
• *On note $x^+ = max(x, 0)$, et $x^- = max(-x, 0)$, si x est un réel.*
• *On note $f^+(x) = (f(x))^+$, si f est une fonction à valeurs réelles. On définit de même $f^-(x) = (f(x))^-$.*
• *On peut définir x^+ à partir de x et $|x|$, $|x|$ à partir de x^+ et x^-, $sup(x, y)$ à partir de $|x - y|$, x et y, en utilisant simplement des additions et des soustractions ; je ne donne pas le détail des formules, que l'on retrouve facilement, et que personne ne se fatigue à apprendre par coeur.*
• *On appelle* **partie entière** *d'un réel x et on note $E(x)$ ou $\lfloor x \rfloor$ le plus grand entier qui lui est inférieur ou égal. Il est caractérisé par $E(x) \in \mathbb{N} \wedge E(x) \leq x < E(x) + 1$. $x - E(x)$ est appelé* **partie décimale** *ou* **partie fractionnaire** *de x, et est noté $[x]$.*
• *On appelle* **intervalle** *de $\mathbb{R}$ toute partie de $\mathbb{R}$ contenant le segment d'extrémités x et y pour tous x et y dans cette partie.*
• *On appelle* **longueur d'un intervalle** *I non vide le réel $sup_{(x,y) \in I^2} |x - y|$.*
• *Une partie A de $\mathbb{R}$ est dite* **bornée** *si $\{|x - y| / (x, y) \in A^2\}$ est majoré ; si A est non vide le sup de cet ensemble est alors le diamètre de la partie A, noté $\delta(A)$.*

Propriétés :
• Le diamètre d'une boule ou d'une sphère de rayon r est $\leq 2.r$, si cette partie est non vide (ce qui est toujours le cas si l'espace n'est pas réduit à zéro à moins qu'il ne s'agisse d'une boule ouverte de rayon nul)
• Un intervalle peut-être de la forme $[a, b]$, $]a, b[$, $[a, b[$, $]a, b]$, avec éventuellement $a = -\infty$ si l'intervalle est ouvert à gauche, et/ou $b = +\infty$ si l'intervalle est ouvert à droite.
• Un segment est un intervalle fermé, de diamètre sa longueur.

Définition 548 *On appelle* **droite numérique achevée** *et on note $\overline{\mathbb{R}}$ l'ensemble totalement ordonné $\mathbb{R} \cup \{-\infty, +\infty\}$, avec $+\infty$ plus grand élément et $-\infty$ plus petit élément (le reste de l'ordre étant l'ordre usuel).*
On étend les définitions de segments et d'intervalles à $\overline{\mathbb{R}}$.

15.2 Les nombres complexes

On suppose le corps des nombres complexes déjà connu ; on rappelle $i^2 = (-i)^2 = -1$.

Définition 549 *Le corps des complexes peut par exemple être construit comme le produit de $\mathbb{R}$ par $\mathbb{R}$, muni des opérations suivantes :*

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \times (c, d) = (a \times c - b \times d, a \times d + b \times c)$$

On note alors $i = (0, 1)$.

On note $Im(z)$ la partie imaginaire d'un nombre complexe $z = (a, b)$, c'est à dire $Im(a, b) = b$, et $Re(z)$ sa partie réelle, c'est-à-dire $Re(a, b) = a$.

15.3 Définition de l'intégration au sens de Riemann

Cette partie est un bref rappel sur l'intégrale de Riemann. Pour une définition complète, on consultera par exemple [11].

L'intégrale de Riemann est tout d'abord définie pour les fonctions en **escalier**.

Définition 550 *Etant donnée f une telle fonction et σ une* **subdivision** *de l'intervalle $[a, b]$, ie une famille $(\sigma_0, ..., \sigma_n)$ vérifiant $a = \sigma_0 < \sigma_1 < \sigma_2 < ... < \sigma_n = b$, l'intégrale sur I de f, pour σ* **adaptée** *à f, c'est à dire telle que sur chaque $]\sigma_i, \sigma_{i+1}[$ f soit constante, est par définition $\int_I f = \sum_{k=0}^{n-1} \lambda_i(\sigma_{i+1} - \sigma_i)$, où $f_{|]\sigma_i,\sigma_{i+1}[} = \lambda_i$.*

Quelques propriétés :

• Définition de l'intégrale indépendante de la subdivision adaptée choisie.

• linéarité de l'intégrale (i.e $\int_I \lambda f + \mu g = \lambda \int_I f + \mu \int_I g$).

• Loi de Chasles : $\int_{[a,b]} f + \int_{[b,c]} f = \int_{[a,c]} f$.

• Valeur de l'intégrale inchangée si on modifie la valeur de f en un nombre fini de points.

• $f \leq g$ implique $\int_I f \leq \int_I g$ (en particulier l'intégrale d'une fonction positive est positive, l'intégrale d'une fonction $\leq M$ sur un intervalle de longueur l est inférieure à lM)

Définition 551 *Une application de $\mathbb{R}$ dans E ($\mathbb{K}$-espace vectoriel normé de dimension finie, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) est dite continue par morceaux sur $I = [a, b]$ s'il existe une subdivision σ de I telle que sur chaque $]\sigma_i, \sigma_{i+1}[$ elle est continue et admet des limites aux bords.*

Remarque :
Notons $F_b([a,b],\mathbb{R})$ [1] l'ensemble des fonctions bornées de $[a,b]$ dans $\mathbb{R}$. Il s'agit d'un espace vectoriel normé complet pour $\|f\|_\infty = sup_{x\in[a,b]}|f(x)|$[2] Notons $\mathcal{E}$ l'ensemble des fonctions en escalier de $[a,b]$ dans $\mathbb{R}$. On voit alors que $\mathcal{E}$ est inclus dans $F_b([a,b],\mathbb{R})$ et que l'application intégration $f \mapsto \int_[a,b]f$ est linéaire, de norme $\leq |b-a|$; l' intégration est ainsi uniforméméent continue. $\overline{\mathcal{E}}$ est appelé l'ensemble des **fonctions réglées** sur $[a,b]$. D'après un théorème de prolongement, et comme (1) $\mathcal{E}$ est dense dans $\overline{\mathcal{E}}$ (par définition de l'adhérence) (2) $\overline{\mathcal{E}}$ est complet puisque fermé de F_b qui est lui-même complet, on peut prolonger $\int_[a,b]$ de manière unique en une application linéaire continue de $\overline{\mathcal{E}}$ dans $\mathbb{R}$ (encore de norme $\leq |b-a|$). Cette méthode permet ainsi de définir l'intégrale sur $[a,b]$ sans avoir à montrer que la limite est indépendante des subdivisions choisies (il suffit juste de voir que les fonctions continues par morceaux appartiennent à F_b).

Toute application f continue par morceaux sur I est limite uniforme d'application en escalier. Soit $(f_n)_{n\in\mathbb{N}}$ une telle suite ; $\int_I f_n$ converge vers une limite, indépendante du choix de f_n. Par définition, cette limite commune est l'intégrale de f.

Quelques propriétés :

• linéarité de l'intégrale (i.e $\int_I \lambda f + \mu g = \lambda \int_I f + \mu \int_I g$).

• Loi de Chasles : $\int_{[a,b]} f + \int_{[b,c]} f = \int_{[a,c]} f$

• Valeur de l'intégrale inchangée si on modifie la valeur de f en un nombre fini de points.

• $f \leq g$ implique $\int_I f \leq \int_I g$ (en particulier l'intégrale d'une fonction positive est positive, l'intégrale d'une fonction $< M$ sur un intervalle de longueur l est inférieure à lM)

• Une fonction continue positive d'intégrale nulle sur I est nulle sur I.

L'**inégalité de Schwartz** et l'**inégalité de Minkowski**, données dans la partie 8, sont valables dans le cadre de l'intégrale de Riemann.

On note aussi la formule de la moyenne :

[1] $\mathbb{R}$ pouvant d'ailleurs éventuellement être remplacé par E Banach, ce qui sera utile pour la résolution d'équations linéaires à coefficients non constants en dimension infinie.
[2] $|f(x)|$ remplacé par $\|f(x)\|$ si E espace de Banach au lieu de $\mathbb{R}$.

Proposition 552 *Formule de la moyenne : Si f et g continues par morceaux sur $[a,b]$, f à valeurs dans $\mathbb{R}$, g à valeurs dans $\mathbb{R}^+$, $m = inf_{x \in I}\ f(x)$, $M = sup_{x \in I}\ f(x)$, on a*

$$m \int_I g \leq \int_I fg \leq M \int_I g$$

Si de plus f est continue, alors

$$\exists c \in I / \int_I fg = f(c) \int_I g$$

Un corollaire important est le cas d'une fonction g constante égale à 1 : $m|b-a| \leq \int f \leq M|b-a|$ et si f est continue, $\int_I f = (b-a)f(c)$ pour un certain $c \in [a,b]$.

Maintenant quelques propriétés des sommes de Riemann ;

Définition 553 *On appelle* **subdivision pointée** *de $I = [a,b]$ un couple (σ, ξ) avec σ une subdivision $(\sigma_0, ..., \sigma_n)$ et ξ une famille $(\xi_0, ..., \xi_{n-1})$ avec $\xi_i \in [\sigma_i, \sigma_{i+1}]$.*

Définition 554 *On appelle* **somme de Riemann** *associée à la subdivision pointée (σ, ξ) et on note $Riemann(f, \sigma, \xi)$ la valeur $\sum_{i \in [0,n-1]} f(\xi_i)|\sigma_{i+1} - \sigma_i|$. On appelle* **pas d'une subdivision ou d'une subdivision pointée** *la quantité $sup_{i \in [0,n-1]} |\sigma_{i+1} - \sigma_i|$. On le note $pas(\sigma)$.*

Alors, si f est continue,

$$lim_{pas(\sigma) \to 0} Riemann(f, \sigma, \xi) = \int_I f$$

Si f est continue sur $[a,b]$, alors $F : x \mapsto \int_a^x f$ est C^1 et est une primitive de f, i.e. $F' = f$.

Enfin, en intégrale de Riemann l'intégration par parties est valable :

Théorème 555 (Intégration par parties) *Soient f et g des primitives de fonctions réglées sur $I = [a,b]$, à valeurs dans $\mathbb{R}$. Alors*

$$\int_I fg' = [fg]_a^b - \int_a^b f'g$$

avec par définition $[fg]_a^b = f(b)g(b) - f(a)g(a)$.

Cette formule sert à peu près partout, par exemple la proposition 443 (calcul de la somme des $1/n^2$), le théorème 593, pour les intégrales de Wallis (partie 15.10.1).

On trouvera dans le formulaire (partie 38) une preuve de la formule de Stirling utilisant l'intégration par parties.

Le changement de variable est aussi valable :

Théorème 556 (Changement de variable) *Si f est continue sur $[c, d]$ et si θ est C^1 de $[a, b]$ dans $[c, d]$, alors on a*

$$\int_{\theta(a)}^{\theta(b)} f = \int_a^b (f \circ \theta)\theta'$$

15.4 Lien entre intégrale de Riemann et intégrale de Lebesgue

Théorème 557 *Soit λ la mesure de Lebesgue sur $\mathbb{R}$.*
(i) Soit $I = [a, b]$ intervalle compact ($-\infty < a < b < \infty$), f de I dans $\mathbb{R}$ ou $\mathbb{C}$ continue (ou simplement réglée). Alors f est mesurable, L^1 sur $[a, b]$ et

$$\int_I f d\lambda \text{(au sens de Lebesgue)} = \int_a^b f(x)dx \text{(au sens de Riemann)}$$

(ii) Soit $I =]a, b[$ intervalle de $\mathbb{R}$ ($\infty \leq a < b \leq \infty$), f continue de $]a, b[$ dans $\mathbb{C}$ (ou simplement réglée sur tout intervalle compact de I). f est alors mesurable, et f est L^1 si et seulement si son intégrale au sens de Riemann est absolument convergente (ie si $\lim_{X \to a^+, Y \to b^-} \int_X^Y |f(x)|dx < \infty$) et alors

$$\int_a^b f(x)dx = \int f d\lambda$$

Le terme de gauche désignant l'intégrale généralisée au sens de Riemann (ie $\lim_{X \to a^+, Y \to b^-} int_X^Y |f(x)|dx$) et le terme de droite l'intégrale au sens de Lebesgue.

Démonstration : □

15.5 Suites et séries

15.5.1 Suites

⊡ Définitions

Définition 558 *La suite* $(y_n)_{n\in\mathbb{N}}$ *est une* **suite extraite** *de la suite* $(x_n)_{n\in\mathbb{N}}$ *si il existe* ϕ *de* $\mathbb{N}$ *dans* $\mathbb{N}$ *strictement croissante telle que* $\forall k\ y_k = x_{\phi(k)}$.
x *est un* **point d'accumulation** *ou une* **valeur d'adhérence** *de la suite* x_n *si* $x \in \cap_n \overline{\{x_k / k \geq n\}}$.
x_n **converge** *vers* x *si pour tout voisinage* V *de* x *il existe* N *tel que pour tout* $n > N\ x_n \in V$.
La suite x_n *est dite* **convergente** *si elle converge vers un certain* x.

⊡ Propriétés

• Cas général

Une suite finie admet une suite extraite constante.
Une suite infinie admet une suite extraite monotone.
Une suite bornée admet une suite extraite convergente.
x est point d'accumulation des x_n si pour tout V voisinage de x il existe une infinité de n tels que x_n est dans V.
La notion de convergence d'une suite est équivalente à la notion de limite en l'infini, avec la topologie que l'on s'est donnée en 5 sur $\mathbb{N} \cup \{+\infty\}$.
Dans un espace séparé, la limite est unique.
Si une suite converge vers x, alors toute suite extraite de cette suite converge vers x.
Une limite de suite extraite est une valeur d'adhérence (pas de réciproque dans le cas général).

• Cas métrique

Dans le cas d'un espace métrique, x est donc point d'accumulation si pour tout ϵ il existe une infinité de n tels que $d(x_n, x) < \epsilon$.
Dans un espace métrique, la limite est unique.
Dans le cas d'un espace métrique, une valeur d'adhérence est une limite de suite extraite.
Dans un espace métrique, une suite de Cauchy a au plus une valeur d'adhérence, et si elle en a une, elle converge et cette valeur est sa limite.
Dans un espace métrique compact (qui est donc aussi complet), une suite n'admettant qu'une seule valeur d'adhérence est une suite convergente.
Dans un espace métrique il sera souvent utile d'introduire la définition $X_n = \{x_p / p \geq n\}$; l'ensemble des valeurs d'adhérence est alors l'intersection des $\overline{X_n}$, c'est donc un fermé, et la suite est une suite de Cauchy si le diamètre des X_n tend vers 0.

• Cas d'un espace vectoriel normé

L'ensemble des suites d'un espace vectoriel normé est un espace vectoriel , avec les opérations de multiplication par un scalaire terme à terme, et d'addition terme à terme (valable même si l'espace n'est pas normé).
L'ensemble des suites bornées en est un sous-espace vectoriel. On peut le normer par

la fonction qui à une suite associe le sup des normes de ses éléments. On notera cette norme $\| \cdot \|$.
L'ensemble des suites de Cauchy est un sous-espace vectoriel de l'espace vectoriel des suites bornées.
L'ensemble des suites convergentes est un sous-espace vectoriel de l'espace vectoriel des suites de Cauchy. L'application qui à une suite convergente associe sa limite (qui est unique) est linéaire et continue de norme 1.
Si notre espace vectoriel normé est muni de deux normes différentes $\| \cdot \|_1$ et $\| \cdot \|_2$ et si il existe α tel que $\| \cdot \|_1 \leq \alpha \| \cdot \|_2$ alors les suites bornées pour $\| \cdot \|_2$ sont des suites bornées pour $\| \cdot \|_1$, idem pour les suites de Cauchy, et idem pour les suites convergentes. Dans le cas de normes équivalentes, on a donc les mêmes suites de Cauchy, les mêmes suites bornées, et les mêmes suites convergentes. Ainsi, il est de Banach pour $\|.\|_1$ si et seulement si il est de Banach pour $\|.\|_2$.

• Cas d'un corps $\mathbb{K} = \mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$
L'ensemble des suites est une $\mathbb{K}$-algèbre commutative et unitaire ; il est muni pour cela du produit par un scalaire, de l'addition terme à terme, du produit terme à terme (l'unité étant la suite constante égale à 1).
L'ensemble des suites bornées est une sous-algèbre unitaire de cette algèbre.
L'ensemble des suites de Cauchy est une sous-algèbre unitaire de cette sous-algèbre.
L'ensemble des suits convergentes est une sous-algèbre unitaire de cette sous-algèbre.
L'application qui à une suite associe sa limite est un morphisme d'algèbres. Ce morphisme est continu. Le noyau de ce morphisme (l'ensemble des suites qui convergent vers 0) est un idéal de l'ensemble des suites bornées.

• Cas du corps $\mathbb{R}$
Toute suite monotone bornée converge.
Si deux suites sont adjacentes, alors elles convergent et ont même limite.

• Cas du corps $\mathbb{C}$
Dans le cas d'une suite complexe on peut se ramener à l'étude de suites réelles ; la suite converge si les parties réelles et imaginaires convergent.

15.5.2 Suites réelles

⊡ Généralités

L'ensemble des suites réelles est un anneau commutatif, un espace vectoriel sur $\mathbb{R}$, ou même une $\mathbb{R}$-algèbre.

Définition 559 *Une suite réelle est dite* **majorée**, **minorée**, **bornée**, **finie**, *si son image est majorée, minorée, bornée, finie.*

L'ensemble des suites réelles bornées est un sous-anneau, un sous-espace vectoriel, une sous-algèbre de l'ensemble des suites réelles.

Définition 560 *Pour la définition de la limite d'une suite, on pourra consulter 5, et notamment la topologie de* $\mathbb{N} \cup \{+\infty\}$.
Une suite **converge** *vers un réel* x *si sa limite en* $+\infty$ *est* x. *Une suite* **diverge** *si et seulement si elle ne converge pas.* $\mathbb{R}$ *étant séparé, la limite d'une suite est unique.*
Une suite **tend vers** $+\infty$ **(resp.** $-\infty$**)** *si sa limite en* $+\infty$ *est* $+\infty$ *(resp.* $-\infty$*) dans* $\overline{\mathbb{R}}$.

L'ensemble des suites réelles convergentes est un sous-anneau, un sous-espace vectoriel, une sous-algèbre de l'ensemble des suites réelles bornées.

⊡ Suites monotones. Applications

Théorème 561 *Toute suite croissante majorée de* $\mathbb{R}$ *est convergente ; sa limite est le* sup *de son image.*
Toute suite décroissante minorée de $\mathbb{R}$ *est convergente ; sa limite est l'*inf *de son image.*
Si deux suites sont l'une croissante, l'autre décroissante, et si leur différence tend vers 0*, alors elles sont dites* **adjacentes** *; deux suites adjacentes convergent vers la même limite.*

Démonstration : Ces résultats, très simples, constituent un (petit !) entraînement à l'utilisation des ϵ et des δ.□

Une illustration des suites adjacentes est le résultat 603.

Théorème 562 $\mathbb{R}$ *est complet.*

Démonstration : Il est évident qu'une suite convergente est de Cauchy ; et réciproquement étant donnée une suite de Cauchy x_n on peut considérer les deux suites $n \mapsto sup_{k \geq n}\ x_k$ et $n \mapsto inf_{k \geq n}\ x_k$; ces deux suites sont adjacentes.

Théorème 563 (Théorème des segments emboîtés) *L'intersection d'une suite décroissante de segments dont la longueur tend vers* 0 *est un singleton.*

Démonstration : Il suffit de considérer la suite des sup et la suite des inf ; ces suites sont adjacentes.□

15.5.3 Séries

Dans cette partie on travaillera avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Définition 564 *On se donne une suite $(u_n)_{n\in\mathbb{N}}$, à valeurs dans $\mathbb{K}$. On définit $U_n = \sum_{k=0}^{n} u_k$. On se pose la question de la convergence de la suite U_n. Si la limite existe, on l'appellera* **somme** *de la série et on la notera*

$$\sum_{k=0}^{+\infty} u_k = lim_{k\to+\infty} U_k$$

On dit que (U_n) est la **série** *de terme général (u_n).*

Définition 565 *On dit que la série de terme général (u_n) converge* **absolument** *si la série de terme général $(|u_n|)$ converge. Une série est dite* **semi-convergente** *si elle converge sans converger absolument.*

Définition 566 *On dit que deux séries sont* **de même nature** *si et seulement si elles sont simultanément convergentes ou simultanément divergentes.*

Définition 567 *On appelle* **somme partielle d'ordre** p *de la série de terme général $(u_n)_{n\geq 0}$ la somme $\sum_{n=0}^{p} u_n$.*
On appelle **reste d'ordre** p *de la série de terme général $(u_n)_{n\geq 0}$ la somme de la série de terme général $(v_n)_{n\geq 0}$ définie par $v_n = u_{n+p}$, lorsque cette série converge. On notera par la suite R_p le reste à l'ordre p de la série de terme général (U_n). On a $R_p = \sum_{n\geq p} u_n$.*

Quelques remarques simples :

• La série de terme générale (u_n) converge si et seulement si la suite $n \mapsto \sum_{i=1}^{n} u_i$ converge.

• s'il y a convergence de la série, alors $u_n \to 0$.

• la série de terme général (x^n) converge si et seulement si $|x| < 1$. Elle converge d'ailleurs si et seulement si elle converge absolument.

• une série converge (resp. converge absolument) si et seulement si son reste à un certain ordre p converge (resp. converge absolument) ; et en ce cas son reste à tout ordre converge (resp. converge absolument).

• Si une série converge, alors la suite (R_n) tend vers 0 en $+\infty$.

• $\mathbb{C}$ et $\mathbb{R}$ étant complets, on peut directement appliquer le critère de Cauchy aux séries dans ces espaces :

Proposition 568 *La série de terme général* (u_n) *converge (resp. converge absolument) si et seulement si*

$$\forall \epsilon \exists N / \forall n > 0 \, |\sum_{i=N}^{N+n} u_i| < \epsilon$$

resp. si et seulement si

$$\forall \epsilon \exists N / \forall n > 0 \sum_{i=N}^{N+n} |u_i| < \epsilon$$

Voir pour application la partie15.12.1.

• si une série est semi-convergente, alors la série de terme général $u_n^+ = max(u_n, 0)$ et la série de terme général $u_n^- = max(-u_n, 0)$ sont divergentes (tendent vers $+\infty$).

• si une série converge absolument alors elle converge (preuve en considérant les parties réelles et imaginaires, et en les décomposant en partie positive et en partie négative - preuve possible aussi en utilisant le critère de Cauchy).

⊡ Séries à termes positifs, absolue convergence

Quelques remarques simples :

Proposition 569 • *Une série à termes positifs converge si et seulement si elle est absolument convergente.*

• *Toute série de terme général extrait du terme général d'une série à termes positifs convergente est convergente, de somme inférieure à la somme de la série initiale.*

• *Toute série déduite d'une série convergente à termes positifs par permutation (éventuellement infinie) des termes est une série convergente de même somme.*

• *Si deux séries à termes positifs sont ordonnées par $u_n \leq v_n$, alors si $\sum v_n$ converge on peut affirmer que $\sum u_n$ converge, et si $\sum u_n$ diverge on peut affirmer que $\sum u_n$ diverge.*

• *si $u_n = O(v_n)$ avec v_n terme générale d'une série absolument convergente (par exemple une série convergente à termes positifs), alors u_n est le terme général d'une série absolument convergente (nb : il n'est pas requis que u_n soit une suite ≥ 0).*

• *Si $u_n \simeq v_n$ et si l'une de ces deux séries est à termes positifs (au moins au voisinage de l'infini) alors les deux séries sont de même nature.*

On déduit notamment de ceci le fait que si u_n est le terme général d'une série à termes positifs ou nuls, alors les séries de termes généraux $ln(1 + u_n)$ et $ln(1 - u_n)$ sont de même nature (et de même nature que la série de terme général u_n, dans le cas où $u_n \to 0$).

Si u_n est une série à terme général positif ou nul, alors $\Pi(1 + u_n)$ converge si et seulement si $\sum u_n$ converge.

◇ Séries de Riemann, séries de Bertrand

Proposition 570 • *Série de Riemann : $x \in \mathbb{R}$, $\sum 1/n^x$ converge si et seulement si $x > 1$ (démonstration par la comparaison avec une intégrale, voir plus loin)*
• *Série de Bertrand : $(x, y) \in \mathbb{R}^2$, $\sum 1/(n^x.ln(n)^y)$ converge si et seulement si $x > 1$ ou ($x = 1$ et $y > 1$) (démonstration par la comparaison avec une intégrale, voir plus loin)*

On trouvera une preuve amusante de ces deux résultats en partie15.12.3 (plus originale que la preuve par comparaison avec une intégrale).

◇ **Utilisation de u_{n+1}/u_n**

Proposition 571 *On se donne (u_n) et (v_n) 2 suites à termes > 0. Alors si $u_{n+1}/u_n < v_{n+1}/v_n$ (au moins à partir d'un certain rang), alors*
• la convergence de la série de terme général v_n entraine la convergence de la série de terme général u_n
• la divergence de la série de terme général u_n entraine la divergence de la série de terme général v_n

Démonstration : Les deux affirmations étant contraposées, il suffit de prouver la première, et elle n'est pas bien dure en se ramenant à la comparaison de deux séries de termes généraux équivalents, l'une étant positive...□

Corollaire 572 (Critère de D'Alembert) *Si (u_n) est une série à termes positifs telle que u_{n+1}/u_n tend vers k, alors si $k < 1$ la série converge, et si $k > 1$, alors la série diverge.*

Démonstration :
• Cas où $k < 1$: on se donne k' avec $k < k' < 1$; la proposition ci-dessous, jointe au fait que la série de terme général k'^n converge, permet de conclure que la série de terme général u_n converge.
• Cas où $k > 1$: en prenant k' tel que $1 < k' < k$, on constate que la suite u_n ne tend pas même vers zéro, ce qui serait la moindre des choses pour le terme général d'une série convergente.□

◇ **Règle de Cauchy**

Théorème 573 (Règle de Cauchy) *$u_n > 0$, alors $limsup\ \sqrt[n]{u_n} = l < 1 \quad \rightarrow \sum u_n$ converge*
$limsup\ \sqrt[n]{u_n} = l > 1 \quad \rightarrow \quad \sum u_n$ diverge

Démonstration : Facile, je ne daigne pas détailler, il suffit dans le premier cas de choisir x compris entre l et 1 ; et dans le deuxième cas, la suite n'a pas même le bon goût de tendre vers 0.□

◇ **Critère de Raabe-Duhamel (alias $\frac{u_{n+1}}{u_n}$, suite)**

Proposition 574 (Règle de Raabe-Duhamel) *Soit $u_n > 0$, avec $\frac{u_{n+1}}{u_n} = 1 - \frac{a}{n} + r_n$, avec r_n terme général d'une série absolument convergente ; alors $u_n \simeq K/n^a$ pour un certain $K > 0$ (et donc la série de terme général u_n converge $\iff a > 1$).*

Démonstration : On définit $v_n = ln(n^a.u_n)$; il est clair que si cette suite converge on a bien le résultat souhaité.

Or $v_{n+1} - v_n = ln(u_{n+1}/u_n) + a.ln(\frac{n+1}{n}) = -a/n + r_n + O(-a/n + r_n)^2 + O(1/n^2)$. En utilisant le fait que $|(a/n).w_n| \leq \frac{1}{2}(a^2/n^2 + w_n^2)$, on en déduit que $v_{n+1} - v_n = O(1/n^2) + O(r_n^2) + r_n = O(r_n + \frac{1}{n^2})$. Or r_n et $\frac{1}{n^2}$ sont tous deux termes généraux de séries absolument convergentes, donc $v_{n+1} - v_n$ est terme général d'une série absolument convergente, donc convergente, et donc (v_n) est une suite convergente.

15.6 Du théorème de Rolle aux formules de Taylor

Théorème 575 (Théorème de Rolle) *Soit f une fonction continue définie sur le segment $[a, b]$, dérivable sur $]a, b[$, avec $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.*

Démonstration :
• Si f est constante sur $[a, b]$, le résultat est clair.

• Si f n'est pas constante, alors f atteint soit son maximum soit son minimum dans $]a, b[$ (il y a là une application de la compacité ; l'image d'un compact par une application continue est un compact, théorème 171) ; la dérivée de f en ce point (ou l'un de ces points) est nécessairement nulle.$\square$

NB : comme le signale judicieusement le livre [15], on peut aussi avoir le même théorème sur $[a, +\infty[$, avec $f(a) = lim_{x \to +\infty} f(x)$.

du théorème de Rolle :

• le polynôme dérivé d'un polynôme scindé de $\mathbb{R}$ dans $\mathbb{R}$ est scindé.

• le théorème 579.

• les résultats 589 et 590.

• le théorème qui suit :

Théorème 576 (Théorème des accroissements finis pour une application de $\mathbb{R}$ dans $\mathbb{R}$)
On se donne f continue de $[a, b]$ dans $\mathbb{R}$, dérivable sur $]a, b[$.
Alors il existe $c \in [a, b]$ tel que $f(b) - f(a) = f'(c).(b - a)$.

Démonstration : Il suffit de soustraire $x \mapsto f(a) + \frac{f(b)-f(a)}{b-a}.(x - a)$ à la fonction f pour se ramener au théorème de Rolle.$\square$

Ne pas confondre avec le théorème 458.

cela sert pour le théorème 577.

Théorème 577 (Théorème des valeurs intermédiaires pour la dérivée) *Ce théorème est aussi dit théorème de Darboux.*

Soit f dérivable d'un intervalle I dans $\mathbb{R}$; alors $f'(I)$ est un intervalle.

__**Démonstration :**__

• On suppose tout d'abord I ouvert ; le cas général s'en déduit clairement.

• On se donne X et Y dans $f'(I)$, puis Z compris entre X et Y.

• On suppose $f'(x) = X$ et $f'(y) = Y$; on cherche z tel que $f'(z) = Z$.

• On se donne $h > 0$ tel que $\frac{f(x+h)-f(x)}{h} < Z < \frac{f(y+h)-f(y)}{h}$ (on se rend compte qu'un tel h existe en considérant les limites des membres de droite et de gauche pour $h \to 0$)

• On définit $g(u) = \frac{f(u+h)-f(u)}{h}$; c'est à dire que $g(u)$ est la pente de f "en regardant sur une largeur h" (voir figure 15.1).

• On applique le théorème des valeurs intermédiaires, en tant que théorème appliqué à une fonction continue, à g (voir 209). On en déduit qu'il existe un certain u tel que $g(u) = Z$. On applique alors le théorème des accroissements finis 576, pour voir qu'il existe un certain z entre u et $u + h$ tel que $f'(z) = Z$.$\square$

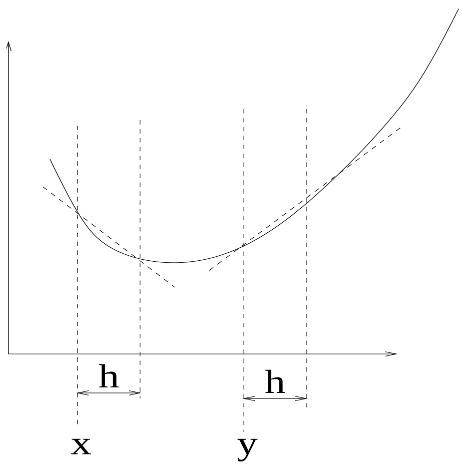


FIG. 15.1 – Théorème de Darboux : on considère les pentes "sur une largeur h"...

Définition 578 (Polynôme de Taylor) *Etant donnée une fonction f de $\mathbb{R}$ dans un espace vectoriel normé au moins n fois dérivable en a, on définit le* **développement de Taylor de f à l'ordre n en a** *qui est une fonction de $\mathbb{R}$ dans notre espace vectoriel normé définie par $P_{f,a,n}(x) = f(a)+\sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$.*

Théorème 579 (Formule de Taylor-Lagrange) *Soit $[a,b]$ un segment de $\mathbb{R}$, avec $a \neq b$, f de classe C^n de $[a,b]$ dans $\mathbb{R}$ $n+1$ fois dérivable sur $]a,b[$; alors il existe $c \in]a,b[$ tel que*

$$f(b) = P_{f,a,n}(b) + \frac{f^{n+1}(c)}{(n+1)!}(b-a)^{n+1}$$

Démonstration :
• On définit
$$g(x) = f(b) - P_{f,x,n}(b) - K.(\frac{(b-x)^{n+1}}{(n+1)!})$$
avec K choisi de manière à avoir $g(a) = 0$ (toujours possible car $b-a$ est non nul par hypothèse).

• On dérive g, on a pile poil les bonnes hypothèses pour appliquer Rolle 575 à g, on en déduit qu'il existe c tel que $g'(c) = 0$ (puisque $g(b) = 0 = g(a)$). Or le calcul de g' permet alors d'écrire que $g'(c) = 0$ implique que $f^{n+1}(c) = K$.

• Il ne reste plus qu'à écrire que $g(a) = 0$.□

Corollaire 580 (Formule de Mac Laurin) *Il s'agit simplement de la même formule, dans le cas $a = 0$ (l'usage veut que ce cas particulier soit nommé "formule de Mac Laurin").*

$$f(b) = P_{f,0,n}(b) + \frac{f^{n+1}(c)}{(n+1)!}(b)^{n+1}$$

Démonstration : Conséquence immédiate du résultat précédent.□

Quelques exemples d'applications, extraites de [15] :

• $x \mapsto \exp(x)$ est limite de la suite des $P_{\exp,0,n}$

• au voisinage de $+\infty$ et pour tout polynôme P, $P(x) = o(e^x)$.

• Si f est C^2 de $\mathbb{R}$ dans $\mathbb{R}$, et avec M_i pour $i \in \{0,1,2\}$ défini comme le sup de $|f^{(i)}|$, on a $M_1 \leq 2\sqrt{M_0.M_2}$ (preuve en écrivant (par Taylor-Lagrange) que $f(a+t) = f(a) + f'(a)t + \frac{1}{2}f''(c)t^2$ pour un certain $c \in]a, a+t[$, en divisant par t pour obtenir $f'(a) = \frac{f(a+t)-f(a)}{t} - \frac{1}{2}f''(c)t^2$ et donc $|f'(a)| \leq \frac{2M_0}{t} + \frac{M_2 t}{2}$ minimal pour $t = 2\sqrt{M_0}M_2$).

• Si f est C^{n+1} de $\mathbb{R}$ dans $\mathbb{R}$, et si f et $f^{(n+1)}$ sont bornées, alors les f_i pour $i \in \{1, 2, ..., n\}$ sont bornées (la preuve utilise le fait que l'ensemble des polynômes de degré $\leq n$ est de dimension finie, bornée, et le fait que toutes les normes sont équivalentes en dimension finie).

Théorème 581 (Inégalité de Taylor-Lagrange) *Soit $[a, b]$ un segment de $\mathbb{R}$, avec $a \neq b$, f de classe C^n sur $[a, b]$ et $n+1$ fois dérivable sur $]a, b[$; on suppose en outre que $f^{(n+1)}$ est bornée par M sur $]a, b[$.*
Alors $\|f(b) - P_{f,a,n}(b)\| \leq M.\frac{(b-a)^{n+1}}{(n+1)!}$.

Démonstration :
• On définit

$$g(x) = f(b) - P_{f,x,n}(b)$$

• On considère l'application qui à x associe $-M.\frac{(b-x)^{n+1}}{(n+1)!}$.
Pour tout x dans $[a, b]$ on a $\|g'(x)\| \leq h'(x)$; le théorème des accroissements finis 458 permet alors d'écrire que $\|g(b) - g(a)\| \leq h(b) - h(a)$ - ce qui est précisément ce qu'on voulait prouver puisque $g(b) = 0$.$\square$

Il y a plus simple dans le cas de f application à image dans $\mathbb{R}$: il s'agit alors d'une conséquence rapide de la formule de Taylor-Lagrange.

Théorème 582 (Formule de Taylor avec reste intégral) *Cette fois-ci f est supposée $\underline{C^{n+1}}$ sur $[a, b]$. Alors :*

$$f(b) = f(a) + P_{f,a,n}(b) + \frac{1}{n!}\int_a^b (b-t)^n.f^{n+1}(t)dt$$

Démonstration : • Comme d'habitude pour ce genre de formule, on définit $g(x) = f(b) - P_{f,x,n}(b)$.

• g est C^1 ; donc $g(b) - g(a)$ est égale à l'intégrale de g' entre a et b ; il s'agit précisément de la formule de Taylor avec reste intégral...$\square$

La formule étant passablement infecte à apprendre (je trouve...) le plus simple est sans doute de faire le calcul soi-même, il n'est pas difficile, et peut être retrouvé très rapidement...

Le théorème de Bernstein, stipulant que toute fonction f C^∞ de $[-a,a]$ dans $\mathbb{R}$, avec $a > 0$, et dont toutes les dérivées d'ordre pair $f^{(2n)}$ sont ≥ 0 sur $[-a,a]$, est développable en série entière sur $]-a,a[$, est démontré dans [15] en application de la formule de Taylor avec reste intégral.
FLEMMARD applications en analyse numérique

Théorème 583 (Formule de Taylor-Young) *On se donne une fonction f de $\mathbb{R}$ dans un espace vectoriel normé E, n fois dérivable en a. Alors*

$$f(x) - P_{f,a,n}(x) = o((x-a)^n)$$

Démonstration : Par récurrence, avec le théorème des accroissements finis 458.□

Ceci fournit bien évidemment une méthode de détermination de développements limités. Mais c'est aussi un outil pour calculer des dérivées de fonctions, comme on va le voir avec les corollaires ci-dessous, permettant de généraliser la formule bien connue $f'(a) = lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$. Il faut noter que la formule se généralise à tout ordre, mais la formule est plus compliquée... Enfin on peut en déduire des résultats sur l'ensemble des zéros d'une fonction suffisamment dérivable, comme illustré par le corollaire 587 (les informations étant bien faibles par rapport aux informations fournies dans le cas complexe...).

Exemple Maple

$> \quad p := array(1..5); \; for \; i \; from \; 2 \; by \; 2 \; to \; 10 \; do \; p[i/2] := convert(series(sin(x), x, i), polynom) \; od;$

$$p := \mathrm{array}(1..5, [])$$

$$p_1 := x$$

$$p_2 := x - \frac{1}{6}x^3$$

$$p_3 := x - \frac{1}{6}x^3 + \frac{1}{120}x^5$$

$$p_4 := x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7)$$

$$p_5 := x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \frac{1}{362880}x^9$$

$> \; plot(p, x = -5..5);$
Le résultat est montré sur la figure 15.2.

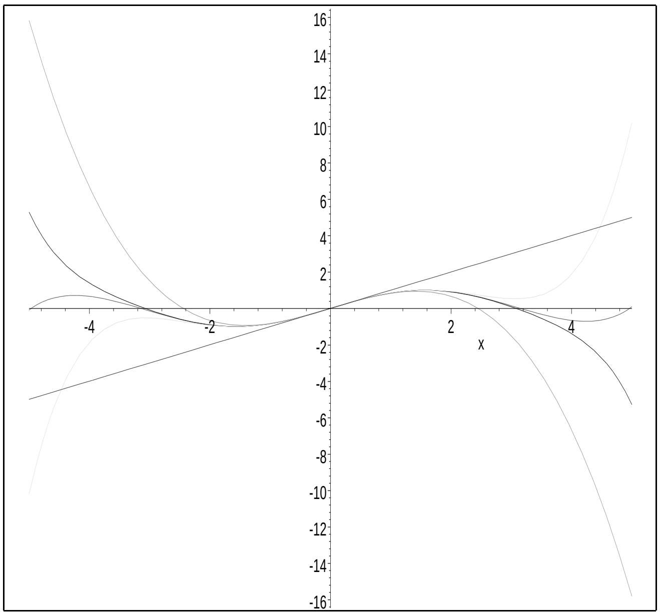


FIG. 15.2 – Approximations successives de $\sin(x)$ par la formule de Taylor-Young

Corollaire 584 (Calcul de la dérivée seconde d'une fonction) *Soit f une fonction 2 fois dérivable en a. Alors*

$$f''(a) = lim_{h\to 0}\frac{f(a+h)+f(a-h)-2.f(a)}{h^2}$$

Démonstration : Par la formule de Taylor-Young, on sait que

$$f(a+h) = f(a) + f'(a).h + \frac{f''(a)}{2}.h^2 + o(h^2)$$

et

$$f(a-h) = f(a) - f'(a).h + \frac{f''(a)}{2}.h^2 + o(h^2)$$

On en déduit bien l'égalité annoncée...□

Corollaire 585 (Calcul de la dérivée n-ième d'une fonction) *Si f est n fois dérivable en 0, alors*

$$f^{(n)}(0) = lim_{h\to 0}\sum_{k=0}^{n} C_n^k(-1)^{n-k}f(kh)$$

Démonstration : Notons $A(h) = \sum_{k=0}^{n} C_n^k(-1)^{n-k}f(kh)$, et utilisons la formule de Taylor-Young à l'ordre n pour chaque $f(kh)$, on obtient

$$A(h) = \sum_{(k,i)\in[0,n]^2} C_n^k(-1)^{n-k}\frac{f^{(i)}(0)}{i!}k^i h^i + o(k^n h^n)$$

Le terme en h^p est alors $\frac{f^{(p)}(0)}{p!}\sum_{k=0}^{n} C_n^k(-1)^{n-k}k^p$.

Il ne reste plus qu'à voir que $\sum_{k=0}^{n} C_n^k(-1)^{n-k}k^n = n!$ et$\sum_{k=0}^{n} C_n^k(-1)^{n-k}k^p = 0$ pour $0 \leq p < n$; voir proposition 1296. $\square$

Définition 586 *On appelle* **ordre** *d'un zéro d'une fonction f C^∞ l'entier p minimal tel que $f^{(p)}(a) \neq 0$. Si un tel entier n'existe pas, le zéro est dit d'***ordre infini**.

Corollaire 587 *Tout zéro d'ordre fini n d'une fonction dérivable au moins n fois est isolé.*

Démonstration : Il suffit d'écrire Taylor-Young, et de se placer suffisamment près du zéro...$\square$

Une fonction C^∞ qui possède une infinité de zéros sur un compact donné en compte au moins un qui est d'ordre infini.

15.7 La trigonométrie

Je ne donnerai ici que quelques définitions, les preuves et formules étant données dans le formulaire (partie 38.9).

Les fonctions trigonométriques se définissent généralement à partir de l'exponentielle complexe : avec $exp(t)$ lui-même défini par une somme de série (voir partie 18.9.1).

On définit $cos(x) = Re(e^{ix})$ et $sin(x) = Im(e^{ix})$, pour x réel[3]. On dérive certaines propriétés (voir partie 38.9) de ces fonctions. La périodicité de $x \mapsto e^{ix}$ donne notamment de nombreux résultats.

Exemple Maple
$> combine(cos(x)\hat{}\,5, trig)$ $$\frac{1}{16}cos(5x) + \frac{5}{16}cos(3x) + \frac{5}{8}cos(x)$$

[3]Pour z complexe, $cos(z) = \frac{e^{iz}-e^{iz}}{2}$ et $sin(z) = \frac{e^{iz}+e^{-iz}}{2i}$.

15.8 Pratique du calcul de primitives

Il est nécessaire avant tout de bien connaître les primitives signalées dans le chapitre 38. Dans TOUS les cas, on aura bien soin de regarder sur quels intervalles on peut définir la primitive. Pour cela on déterminera les intervalles sur lesquels la fonction est continue, puis on regarde aux bornes de chaque intervalle si la fonction peut-être prolongée par continuité.

15.8.1 Primitives de fractions rationnelles

On trouvera en FLEMMARD une méthode de décomposition en éléments simples d'une fraction rationnelle. Une fois une fraction rationnelle décomposée en éléments simples, il suffit d'appliquer les recettes de cuisine évoquées à la partie "primitives" du chapître 38. Les discontinuités sont les pôles de la fraction rationnelle.

15.8.2 Primitives de $P(cos(x), sin(x))$

P désigne un polynôme à deux indéterminées.

Cette situation se présente couramment dans la vie de tous les jours et il est indispensable d'être bien préparé pour y faire face.

Il est évidemment suffisant de savoir intégrer un monôme, c'est-à-dire un élément de la forme $f_{n,m} = x \mapsto \cos(x)^n \sin(x)^m$

• Si n est impair, il suffit de remplacer $\cos(x)^n$ par $(1 - \sin(x)^2)^{\frac{n-1}{2}} cos(x)$, et le changement de variable $u = sin(x)$ nous ramène au calcul de la primitive d'un polynôme.

• Si m est impair, il suffit de remplacer $\sin(x)^n$ par $(1 - \cos(x)^2)^{\frac{n-1}{2}} sin(x)$, et le changement de variable $u = cos(x)$ nous ramène au calcul de la primitive d'un polynôme.

• Si n et m sont pairs, une méthode générale est de linéariser. Voir pour cela la partie15.7. On peut aussi procéder par une intégration par parties, pour se ramener à $I_{n+m,0}$ ou $I_{0,n+m}$

- c'est-à-dire que dans l'intégration par parties on intègre $cos(x)^n sin(x)$ et on dérive $sin(x)^{m-1}$ si $n > m > 0$, pour se ramener à une primitive de $cos(x)^{n+2} sin(x)^{m-2}$

- et si $m > n > 0$ on intègre $sin(x)^m cos(x)$ et on dérive $cos(x)^{n-1}$ pour se ramener à une primitive de $cos(x)^{n-2} sin(x)^{m+2}$.

15.8.3 Primitives de $G(x) = F(cos(x), sin(x))$

F désigne une fraction rationnellle à deux indéterminées.

Les discontinuités seront généralement en nombre infini, périodiques ; il convient d'étudier précisément ce qu'il se passe au niveau de chaque discontinuité.

F désigne une fraction rationnelle à deux indéterminées.

On aura alors recours à la règle de Bioche (que l'on peut justifier rigoureusement, voir par exemple le livre d'analyse de Arnaudiès et Fraysse).

Proposition 588 *La règle de Bioche nous dit que :*

• *Si G est paire, ie $G(-x) = G(x)$, on fait le changement de variable $u = sin(x)$*

• *Si G est impaire, ie $G(-x) = -G(x)$, on fait le changement de variable $u = cos(x)$*

• *Si G est ω-périodique, c'est à dire si $G(x) = G(x + \omega)$, on fait le changement de variable $u = tan(\frac{\pi x}{\omega})$.*

• *Si G n'est rien de tout ca, on n'aura plus d'autre choix que le changement de variable $u = tan(x/2)$. Rappelons que dans ce cas, $cos(x) = \frac{1-u^2}{1+u^2}$, $sin(x) = \frac{2u}{1+u^2}$, $dx = \frac{2}{1+u^2}du$*

15.8.4 Primitives de $H(x) = F(ch(x), sh(x))$

On pensera bien à vérifier en quels points s'annulent le dénominateur.

F désigne une fraction rationnelle à deux indéterminées.

On considère alors la fonction $G(x) = F(cos(x), sin(x))$.

• Si G est paire, ie $G(-x) = G(x)$, on fait le changement de variable $u = sh(x)$

• Si G est impaire, ie $G(-x) = -G(x)$, on fait le changement de variable $u = ch(x)$

• Si G est π-périodique, c'est à dire si $G(x) = G(x + \pi)$, on fait le changement de variable $u = th(x)$.

• Si G n'est rien de tout ca, on peut faire le changement de variable $u = th(x/2)$. La partie 38.10 fournit les formules nécessaires. Dans beaucoup de cas pratiques, il sera en fait préférable de faire $u = e^x$.

15.8.5 Primitives abéliennes

Une étude plus complète (notamment incluant des justifications géométriques) se trouve dans le livre d'analyse de Arnaudiès et Fraysse.

⊡ $\int R(x, \sqrt{ax+b})$

R est supposée être une fraction rationnelle à deux indéterminées, a et b des réels. Il suffit alors de faire le changement de variable $u = \sqrt{ax+b}$ pour que magiquement tout s'arrange et que l'on n'ait plus qu'une fraction rationnelle à intégrer.

⊡ $\int R(x, \sqrt{ax^2+bx+c})$ **avec** $a > 0$

R est supposée être une fraction rationnelle à deux indéterminées.

On fait alors le changement de variable $u = \sqrt{ax^2+bx+c} - \sqrt{a}x$. En fait cela revient simplement à considérer un repère dans lequel l'équation de l'hyperbole est plus sympathique, c'est-à-dire un repère dont les axes sont les asymptotes de l'hyperbole.

On trouvera dans FLEMMARD d'autres variantes sympathiques.

⊡ $\int R(x, \sqrt{ax^2+bx+c})$ **avec** $a < 0$

Géométriquement, on constate simplement que la courbe $x \mapsto (x, \sqrt{ax^2+bx+c})$ est dans le cas $a < 0$ un morceau d'ellipse, et on se donne une paramétrisation sympathique, qui nous donne x comme un cosinus de u (à divers facteurs près) et $\sqrt{ax^2+bx+c}$ comme un sinus de u (là encore à divers facteurs près).

On fait le changement de variable $cos(u) = \sqrt{\frac{a^2}{b^2/4-ac}}(x+b/a)$. Cela nous ramène à $\sqrt{ax^2+bx+c} = \sqrt{\frac{ac-b^2/4}{a}}sin(u)$, et nous simplifie tout ça de manière à en faire une fraction rationnelle.

⊡ $\int R(x, \sqrt[n]{\frac{ax+b}{cx+d}})$

R est supposée être une fraction rationnelle à deux indéterminées.

Si $n = 2$, il suffit de multiplier par $\sqrt{\frac{cx+d}{cx+d}}$ pour se ramener au cas précédent ; sinon on constate que la fonction réciproque de cette fonction est rationnelle, et donc on pose simplement $u = \sqrt[n]{\frac{ax+b}{cx+d}}$ pour se ramener à une fraction rationnelle.

15.9 Zoologie de la dérivation

15.9.1 Une application du théorème de Rolle

Proposition 589 *Soient f et g deux applications continues de $[a, b]$ dans $\mathbb{R}$, dérivables sur $]a, b[$. Si $g(b) \neq g(a)$, alors il existe $c \in]a, b[$ tel que le déterminant*

$$\begin{vmatrix} f(b) - f(a) & f'(c) \\ g(b) - g(a) & g'(c) \end{vmatrix}$$

s'annule.

Démonstration : Il suffit d'appliquer le théorème de Rolle 575 à la fonction h définie par $h(t) = f(t) - f(a) - \frac{g(t)-g(a)}{g(b)-g(a)}.(f(b) - f(a))$.$\square$

Alors que le théorème des accroissements finis exprime le fait qu'une application de $\mathbb{R}$ dans $\mathbb{R}$ a une tangente parallèle à n'importe laquelle de ses cordes, le théorème 589 montre que c'est en fait valable pour n'importe quelle courbe dérivable dans $\mathbb{R}^2$.

Corollaire 590 (Règle de l'Hôpital) *On se donne deux fonctions f et g de $\mathbb{R}$ dans $\mathbb{R}$ continues en $a \in \mathbb{R}$, et dérivables sur un voisinage de a privé de a, avec $f(a) = g(a) = 0$.*
On suppose que g et g' sont non nulles sur un voisinage de a privé de a.
Alors si $\frac{f'}{g'}$ tend vers l en a, alors $\frac{f}{g}$ tend vers l en a.

Démonstration : Il s'agit exactement d'une application de la proposition 589.$\square$

15.10 Zoologie de l'intégration

15.10.1 Intégrales de Wallis

Définition 591 (Intégrale de Wallis) *On définit la nième intégrale de Wallis par $I_n = \int_0^{\Pi/2} sin(x)^n.dx$.*

Supposons $n \geq 2$. On peut écrire $I_n = \int_0^{\Pi/2} sin^{n-2}(x).(1 - cos^2(x)).dx$ car $cos^2 + sin^2 = 1$.

Ensuite on peut écrire $I_n = I_{n-2} - \int_0^{\Pi/2} sin^{n-2}(x).cos^2(x).dx$.

Par une intégration par partie (théorème 555) (on intègre $sin^{n-2}(x).cos(x)$ et on dérive $cos(x)$) on obtient

$I_n = \frac{n-1}{n} I_{n-2}$

Le produit $n.I_n.I_{n-1}$ est donc constant, égal à $\Pi/2$.

Il reste à remarquer que $I_n \leq I_{n-1} \leq I_{n-2} = \frac{n}{n-1} I_n$ pour pouvoir dire que $I_n \simeq I_{n-1}$.

On peut alors conclure que :□

Proposition 592 $I_n \simeq \sqrt{\frac{\Pi}{2n}}$.

15.10.2 Primitives de f, $\frac{f'}{f} = a/x + o(1/x)$

Théorème 593 *On se donne f une fonction > 0 et C^1 au voisinage de $+\infty$, et on suppose que $\frac{f'}{f} = a/x + o(1/x)$, avec $a \neq -1$.*
On définit $F(x) = \int_b^x f(t)dt$.
Si l'intégrale $\int_b^{+\infty} f$ converge, on définit $R_f(x) = \int_x^{\infty} f(t)dt$.
Alors :

$$\begin{cases} a > -1 \Rightarrow \begin{cases} \int_b^{+\infty} f \text{ diverge} \\ F(x) \simeq \frac{xf(x)}{a+1} \end{cases} \\ a < -1 \Rightarrow \begin{cases} \int_b^{+\infty} f \text{ converge } R_f(x) \simeq -\frac{xf(x)}{a+1} \end{cases} \end{cases}$$

Démonstration :

• Si $a = 0$, l'intégrale $\int_b^{\infty} f(t)dt$ diverge car

$$f'(x)/f(x) = o(1/x)$$

$$-\epsilon ln(x/b) \leq ln(f(x)/f(b))$$

$$f(x)/f(b) \geq (x/b)^{-\epsilon}$$

Donc si $\epsilon \leq 1$, on minore bien notre fonction par quelque chose de trop gros pour être intégrable.

Il suffit ensuite de faire une intégration par partie (théorème 555) pour justifier que $F(x) = xf(x) - bf(b) - \int_b^x tf'(t)dt$.

L'intégrale $\int_b^x tf'(t)dt$ est un $o(\int_b^x f(t)dt)$ puisque $tf'(t)$ est un $o(f(t))$, f étant positive et d'intégrale divergente.

Donc $F(x) \simeq xf(x)$.

• Si $a \neq 0$, alors l'intégrale de a/x est de signe constant et est divergente en $+\infty$, donc on peut écrire $ln(f(x)) \simeq a\ ln(x)$.

- Si $a > -1$, avec $\epsilon = (a+1)/2$ et η tel que pour $x > \eta$, $ln(f(x)) \geq (a-\epsilon)ln(x)$, on a pour $x > \eta$ $f(x) \geq x^{a-\epsilon}$ avec $a - \epsilon > -1$, et donc l'intégrale diverge. On peut alors écrire, grâce à une intégration par parties (théorème 555) :

$$F(x) = xf(x) - bf(b) - \int_b^x f(t)dt \tag{15.1}$$

Puisque l'intégrale diverge et puisque $a\ f(t) \simeq tf'(t)$

$$aF(x) \simeq \int_b^x tf'(t)dt$$

En replaçant cette expression dans l'équation 15.1, on obtient

$$F(x) = xf(x) - aF(x) + o(F(x))$$

et donc

$$F(x) \simeq \frac{xf(x)}{1+a}$$

- Si $a < -1$, avec $\epsilon = (a+1)/2$ et η tel que pour $x > \eta\ ln(f(x)) \leq (a+\epsilon)ln(x)$, donc pour $x > \eta\ f(x) \leq x^{a+\epsilon}$ avec $a + \epsilon < -1$, et donc l'intégrale converge.

Par définition,

$$R_f(x) = lim_{Y\to\infty} \int_x^Y f(t)dt$$

On peut alors écrire, grâce à une intégration par parties (théorème 555) :

$$R_f(x) = lim_{Y\to\infty}(Yf(Y) - xf(x) - \int_x^Y tf'(t)dt) \tag{15.2}$$

Or $\int_x^Y tf'(t)dt$ converge pour $Y \to \infty$, et $\int_x^\infty tf'(t)dt = aR_f(x) + o(R_f(x))$. donc

$$R_f(x) = -\frac{1}{1+a}xf(x) + lim_{Y\to\infty}Y\ F(Y) + o(F(x))$$

Par ailleurs $Y\ f(Y)$ a une limite en $+\infty$ (au vu de l'équation 15.2) ; cette limite ne peut être que 0, vue la convergence de l'intégrale de f.

D'où

$$R_f(x) = -\frac{1}{1+a}xf(x) + o(R_F(x))$$

ce qui est bien le résultat souhaité.$\square$

15.10.3 Méthode de Laplace

Théorème 594 (Méthode de Laplace) *Soit f une fonction C^2 sur $]a, b[$, avec $(a, b) \in \overline{\mathbb{R}}^2$, telle que f admette un maximum unique en $c \in]a, b[$, avec $f''(c) < 0$, f n'ayant pas $f(c)$ pour valeur d'adhérence pour $x \to a$ ni pour $x \to b$, et soit g une fonction continue sur $]a, b[$ avec $g(c) \neq 0$.*
On suppose en outre que pour tout t l'intégrale

$$\int_a^b |g(x)|e^{tf(x)}dx$$

est convergente.
Alors

$$\int_a^b g(x)e^{tf(x)}dx \simeq_{t\to\infty} \frac{\sqrt{2\Pi}g(c)e^{tf(c)}}{\sqrt{-tf''(c)}}$$

Démonstration :

• En remplaçant $f(x)$ par $f(x-c)$ et $g(x)$ par $g(x-c)$ on se ramène au cas où $c = 0$, quitte à changer a et b (s'ils ne sont pas infinis).

• Quitte à remplacer g par $-g$, on suppose $g(0) > 0$.

• On se donne $\epsilon > 0$.

• On se donne $\delta > 0$, suffisamment petit pour que $|x| \leq \delta$ implique

$$f(0) + \frac{f''(0)}{2}x^2(1+\epsilon) \leq f(x) \leq f(0) + \frac{f''(0)}{2}x^2(1-\epsilon)$$

et

$$(1-\epsilon)g(0) \leq g(x) \leq (1+\epsilon)g(0)$$

• On précise alors $I(\delta, t) = \int_0^\delta g(x)e^{tf(x)}dx$;

$$\int_0^\delta g(x)e^{tf(0)+t\frac{f''(0)}{2}(1+\epsilon)x^2} \leq I(\delta) \leq \int_0^\delta g(x)e^{tf(0)+t\frac{f''(0)}{2}(1-\epsilon)x^2}$$

• On cherche alors à préciser $U(\delta, t, \pm) = \int_0^\delta g(x)e^{tf(0)+t\frac{f''(0)}{2}(1\pm\epsilon)x^2}$.

$$\int_0^d g(0)(1-\epsilon)e^{tf(0)+t\frac{f''(0)}{2}(1\pm\epsilon)x^2} \leq U(\delta, t, \pm) \leq \int_0^d g(0)(1+\epsilon)e^{tf(0)+t\frac{f''(0)}{2}(1\pm\epsilon)x^2}dx$$

- On effectue un changement de variable $y = x\sqrt{-tf''(0)(1-\epsilon)/2}$.

- En ne traitant que l'inégalité de droite (l'autre étant similaire) :

$$U(\delta, t, \pm) \leq \frac{\int_0^{\delta\sqrt{-tf''(0)(1\pm\epsilon)/2}} g(0)(1+\epsilon)e^{tf(0)-y^2}dy}{\sqrt{-tf''(0)(1\pm\epsilon)/2}}$$

- On a alors

$$U(\delta,t,\pm) \leq \frac{g(0)e^{tf(0)}\int_0^{\delta\sqrt{-tf''(0)(1-\epsilon)/2}}(1+\epsilon)e^{-y^2}dy}{\sqrt{-tf''(0)(1\pm\epsilon/2)}} \leq e^{tf(0)}g(0)\frac{2\Pi}{2\sqrt{-tf''(0)(1\pm\epsilon)}}$$

• Majorons maintenant $U(\delta,t,-)$, en prenant garde au fait que δ dépend de ϵ :

$$U(\delta,t,-) \leq \int_0^{\delta} g(0)(1+\epsilon)e^{tf(0)+t\frac{f''(0)}{2}(1-\epsilon)x^2}dx$$

avec $y = \sqrt{-tf''(0)(1-\epsilon)}x$,

$$\leq \int_0^{\delta\sqrt{-tf''(0)(1-\epsilon)x}} \frac{g(0)(1+\epsilon)e^{tf(0)-y^2}dy}{\sqrt{-tf''(0)(1-\epsilon)x}}$$

$$\leq g(0)(1+\epsilon)e^{tf(0)}\frac{\int_0^{\infty} e^{-y^2}dy}{\sqrt{-tf''(0)(1-\epsilon)x}}$$

Or $\int_0^{\infty} e^{-y^2}dy = \frac{\sqrt{\pi}}{2}$, donc

$$\leq \frac{\sqrt{2\pi}g(0)e^{tf(0)}(1+\epsilon)}{2\sqrt{-tf''(0)}\sqrt{1-\epsilon}}$$

De même on minore $U(\delta,t,+)$ par

$$\int_0^{\delta} g(0)(1-\epsilon)e^{tf(0)+t\frac{f''(0)}{2}(1+\epsilon)x^2}dx$$

avec $y = \sqrt{-tf''(0)\frac{1+\epsilon}{2}}$ on arrive à

$$\geq \frac{\sqrt{2\pi}g(0)e^{tf(0)}(1-\epsilon)}{2\sqrt{-tf''(0)}(1+\epsilon)}$$

pour $t \geq T_\epsilon$.

• On conclut alors que $I(\delta,t)$, qui vérifie

$$U(\delta,t,+) \leq I(\delta,t) \leq U(\delta,t,-)$$

est compris entre $\frac{1-\epsilon}{1+\epsilon}F(t)$ et $\frac{1+\epsilon}{\sqrt{1-\epsilon}}F(t)$, avec $F(t) = \frac{\sqrt{2\pi}g(0)e^{tf(0)}}{2\sqrt{-tf''(0)}}$.

• $I(\delta,t)$ est équivalent à $e^{tf(0)}g(0)\frac{2\Pi}{2\sqrt{-tf''(0)(1-\epsilon)}}$

• On considère par ailleurs $V(\delta,t) = \int_\delta^b g(x)e^{tf(x)}$.

- $|V(\delta,t)| \leq \int_\delta^b |g(x)|e^{(t-1)f(x)}e^{f(x)}dx$

- Pour $x \in [\delta,b]$, $f(x) \leq f(\delta)$, si du moins on prend la peine d'imposer δ suffisamment petit.

- Donc pour un tel δ, $|V(\delta,t)| \leq e^{(t-1)f(\delta)}\int_\delta^b |g(x)|e^{f(x)}dx$

- Or $\int_\delta^b |g(x)|e^{f(x)}dx \leq \int_a^b |g(a)|e^{f(x)} < \infty$ (< par hypothèse)

• $V(\delta,t) = o(I(\delta,t)$ pour $t \to \infty$ (car $f(\delta) < f(0)$)

• On en déduit

$$\int_0^b g(x)e^{tf(x)} \simeq e^{tf(0)}g(0)\frac{\sqrt{2\pi}}{2\sqrt{-tf''(0)(1-\epsilon)}}$$

• En effectuant la même manœuvre sur $]a,c]$ et en sommant (les équivalents étant tous deux positifs) on obtient le résultat désiré...□

15.11 Zoologie des suites

15.11.1 Moyenne de Césaro

Théorème 595 (Moyenne de Césaro) *On se donne une suite u_n à valeurs dans un $\mathbb{R}$-espace vectoriel normé E. Soit λ_n une suite de réels > 0, telle que $\sum \lambda_n$ diverge.*
Alors si $u_n \to l$, alors $v_n \to l$, avec $v_n = \frac{\sum_{k=0}^n \lambda_k.u_k}{\sum_{k=0}^n \lambda_k}$.

__**Démonstration :**__

• On se ramène tout d'abord au cas où $l = 0$, simplement en considérant la suite de terme général $u_n - l$ (il est clair que v_n est ainsi diminué lui aussi de l).

• On se donne $\epsilon > 0$; il existe un certain N tel que pour tout $n > N$, $|u_n| < \epsilon$.

• Pour $n > N$, on a alors

$$v_n = \frac{\sum_{k=0}^N \lambda_k.u_k}{\sum_{k=0}^n \lambda_k}$$

$$+\frac{\sum_{k=N+1}^n \lambda_k.u_k}{\sum_{k=0}^n \lambda_k}$$

• Le premier terme de la somme ci-dessus tend clairement vers 0, le second est plus petit qu'ϵ (en valeur absolue), d'où le résultat.□

15.11.2 Constante Γ d'Euler

Définissons $u_n = \sum_{k=1}^n 1/k - ln(n)$.

$$\frac{1}{n+1} < \int_n^{n+1} \frac{1}{t}dt < \frac{1}{n}$$

car $t \mapsto 1/t$ est décroissante sur $\mathbb{R}_+^*$.

En écrivant $u_{n+1} - u_n$, on obtient donc que u_n décroît.

Par ailleurs, $u_{n+1} - u_n > \frac{1}{n+1} - \frac{1}{n}$, donc par sommation $u_n - u_1 > 1/n - 1$, donc $u_n > 0$. Donc u_n, décroissante positive, tend vers une limite finie, notée Γ et appelée **constante d'Euler**.

Définition 596 *La limite de la suite* $u_n = \sum_{k=1}^{n} 1/k - ln(n)$ *est appelée* **constante d'Euler** *et est notée* Γ.

On obtient facilement en Maple (par la commande "evalf(gamma,500)") :

$$\Gamma \simeq .57721566490153286060651209008240243104215933593992359880576723488486772677766467093694706329174674951463144724980708248096050401448654283622417399764492353625350033374293733773767394279259525824709491600873520394816567085323315177661152862119950150798479374508570574002992135478614669402960432542151905877553526733139925401296742051375413954911168510280798423487758720503843109399736137255306088933126760017247953783675927135157722610273492913940798430103417771778088154957066107501016191663340152278935867965497252036212879226555953669628176388792726801324310104765059637039473949576389065729679296010090151251959509222435014093498712282479497471956469763185066761290638110518241974448678363808617494551698927923018773910729457815543160050021828440960537724342032854783670151773943987003023703395183286900015581939880427074115422278197165230110735658339673487176504919418123000406546931429992977795693031005030863034185698032310836916400258929708909854868257773642882539549258736295961332985747393$$

15.12 Zoologie des séries

15.12.1 Construction de séries divergentes positives toujours plus petites

Le comportement d'une série étant asymptotique, on pourrait se demander s'il ne serait pas possible de construire une suite de terme général (u_n) qui soit "à la limite" de la convergence, fournissant un critère par simple comparaison, permettant de décider si une série positive diverge ou non.

Il n'en est rien, car étant donnée une suite $u_n > 0$ telle que la série $\sum_{n=0}^{+\infty} u_n$ diverge, on peut construire une suite $v_n > 0$ telle que $v_n = o(u_n)$ et $\sum_{n=0}^{\infty} v_n$ diverge.

Pour cela, soit $v_n = \frac{u_n}{U_n}$. Il est clair que $v_n = o(u_n)$. Il reste à voir que la série de terme général (v_n) diverge. Pour cela on utilise le critère de Cauchy ;

$$\begin{aligned}
\forall N \in \mathbb{N}, \forall n > 0, V_{N+n} - V_N &= \sum_{i=N+1}^{N+n} \frac{u_i}{U_i} \\
&\geq \frac{u_{N+1}}{U_{N+n}} + ... + \frac{u_{N+n}}{U_{N+n}} \\
&\geq \frac{U_{N+n} - U_{N-1}}{U_{N+n}} \to 1
\end{aligned}$$

comme $n \to \infty$, donc le critère de Cauchy n'est pas vérifié.$\square$

15.12.2 Somme des inverses des nombres premiers

On définit p_n comme le n-ième nombre premier, dans l'ordre croissant. On se préoccupe de la nature de la série $\sum 1/p_n$ de terme général $1/p_n$.

Définissons $u_n = \Pi_{i=1}^n \frac{1}{1-\frac{1}{p_i}}$, et considérons $ln(u_n) = \sum_{i=1}^n ln(1 - 1/p_n)$, il apparaît que la série est de même nature que la *suite* (et non pas série !) (u_n).

On remarque alors que

$$\frac{1}{1-\frac{1}{p_n}} = \sum_{i=0}^{+\infty} 1/p_n^i$$

et on peut écrire tout nombre $1/n$ comme produit d'un certain nombre d'inverses de nombres premiers ; donc le produit u_n est supérieur à la somme des inverses des entiers plus petits que n (un nombre entier étant produit de nombres premiers inférieurs ou égaux à lui-même...) ; donc la suite (u_n) diverge, car $\sum 1/n$ diverge.

D'où

$$\sum_{n\geq 0} 1/p_n = +\infty$$

15.12.3 Groupement de termes

On a vu que dans le cas de séries à termes positifs, on pouvait permuter les termes autant qu'on le souhaitait sans changer la convergence de la série ; dans le cas général ce n'est pas vrai (considérer $u_n = (-1)^n$ par exemple, ou bien la série de terme général $u_n = (-1)^n/n$, que l'on peut faire tendre vers n'importe quelle limite réelle en changeant l'ordre des termes).

On a toutefois des résultats partiels :

Théorème 597 *On considère une suite u_n à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.*
Soit ϕ une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ avec $\phi(0) = 0$.
On définit $Tranche_n = u_{\phi(n)} + u_{\phi(n)+1} + ... + u_{\phi(n+1)-1}$ (ceux qui n'ont pas compris pourquoi on appelait cette série "Tranche" sont priés de relire soigneusement cette ligne).
On définit $GrosseTranche_n = |u_{\phi(n)}| + |u_{\phi(n)+1}| + ... + |u_{\phi(n+1)-1}|$.
On suppose que $GrosseTranche_n$ tend vers 0.
Alors la série de terme général (u_n) converge si et seulement si la série de terme général $Tranche_n$ converge, et si elles convergent elles ont même somme.

⚠ Peu importe le caractère convergent de la série de terme général $(GrosseTranche_n)$!

Démonstration : Pas dur... On associe à tout n la somme partielle d'ordre ϕ_n, on montre que cette suite converge (c'est en fait l'hypothèse), et on montre en utilisant la seconde hypothèse que la différence entre U_n (la somme partielle d'ordre n) et $U_{\phi(p)}$, avec p maximal tel que $\phi(p) \leq n$, est petite pour n grand...□

⚠ Dans le cas d'une série à termes positifs, on n'a pas besoin de l'hypothèse "$GrosseTranche_n$ tend vers 0".

⇒ Application possible : règle de la loupe.

Proposition 598 (Règle de la loupe) *Soit u_n une suite décroissante de réels > 0 ; alors les séries $\sum u_n$ et $\sum 2^n.u_{2^n}$ sont de même nature.*

Démonstration : Facile avec le groupement de termes ;

$$\frac{1}{2}2^{n+1}u_{2^{n+1}} \leq u_{2^n+1} + ... + u_{2^n-1} \leq 2^n.u_{2^n}$$

D'où le résultat, en constatant qu'on peut grouper les termes par $\phi(n) = 2^n$, puisque l'on est dans le cas d'une série à termes positifs.□

L'hypothèse de décroissance est nécéssaire, comme on peut s'en convaincre en considérant $u_{2^n} = \frac{1}{2^{2^n}}$ et $u_p = 1$ si $p \notin \{2^n/n \in \mathbb{N}\}$.

⇒ Application possible : série de Riemann, série de Bertrand.

Si $u_n = 1/n^x$, $2^n.u_{2^n} = \frac{2^n}{2^{nx}} = 2^{n.(1-x)}$ donc $\sum u_n$ converge si et seulement si $x > 1$

Si $u_n = 1/(n^x.log(n)^y)$, le résultat ci-dessus donne la convergence de la série $\sum u_n$ si $x > 1$, et la divergence si $x < 1$; il reste le cas limite $u_n = 1/(n.log(n)^y)$, dans ce cas $2^n.u_{2^n} = 1/(n.ln(2))^y$, dont la série converge si et seulement si $y > 1$.

15.12.4 Exponentielle d'un endomorphisme

Théorème 599 *Etant donné $f \in \mathcal{L}(E)$ (i.e. f un endomorphisme continu de E), avec E un espace de Banach, on appelle* **exponentielle de l'endomorphisme** *f et on note $exp(f)$ l'endomorphisme qui à x associe $\sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{f \circ \cdots \circ f}_{n \text{ fois}}(x)$*

Bien entendu, la convergence est normale sur tout borné de $\mathcal{L}(E)$, donc uniforme sur tout borné.

Théorème 600 (Dérivation de l'exponentielle) *Soit E un Banach et soit $f \in \mathcal{L}(E)$, alors l'application $t \mapsto exp(tf)$ est dérivable, de dérivée $exp(tf) \circ f = f \circ exp(tf)$.*

Démonstration : Il suffit de voir que l'on a convergence uniforme de la série dérivée, et d'appliquer le théorème de dérivation sous le signe intégrale 381.□

Attention à ne pas généraliser en croyant que $d(f \mapsto exp(f)) = (g \mapsto exp(f)g)$, ce n'est pas vrai en général, et du ici au fait que l'homothétie commute avec tout endomorphisme !

pour la pratique du calcul de l'exponentielle d'un endomorphisme en dimension finie, voir la partie 31.3.3.

15.12.5 Transformation d'Abel

Théorème 601 (Transformée d'Abel) *Soit E un espace vectoriel normé , (r_n) une suite de réels, (e_n) une suite de E. On note $E_n = \sum_{k=0}^{n} e_k$. Alors pour tout $M < N$*

$$\sum_{n=M}^{N} r_n e_n = [rE]_M^N - \sum_{n=M}^{N-1} (r_k - r_{k+1}) E_k$$

avec $[rE]_M^N = r(N)E(N) - r(M)E(M-1)$ par définition (attention au -1 !).

Démonstration :

$$\begin{aligned}
\sum_{n=M}^{N} r_n e_n &= \sum_{n=M}^{N} r_n (E_n - E_{n-1}) \\
&= \sum_{n=M}^{N} r_n E_n - \sum_{k=M-1}^{N-1} r_{n+1} E_n \\
&= [rE]_M^N - \sum_{n=M}^{N-1} (r_k - r_{k+1}) E_k
\end{aligned}$$

D'où le résultat ; soulignons l'analogie avec l'intégration par parties (théorème 555). □

les théorèmes 602 et 605.

⊡ **Théorème d'Abel**

Théorème 602 (Théorème d'Abel) *Soit (r_n) une suite de réels tendant <u>en décroissant</u> vers 0 (au moins au voisinage de l'infini), soit (e_n) le terme général d'une série dans un Banach dont les sommes partielles sont bornées, alors la série de terme général (u_n) avec $u_n = r_n.e_n$ converge, ie*

$$\sum_{n=0}^{\infty} r_n.e_n \text{ converge}$$

<u>Démonstration :</u> Application directe de la transformation d'Abel ! □

Un cas particulier classique est le cas $e_n = (-1)^n$:

Corollaire 603 (Série alternée) *Si $u_n = (-1)^n \epsilon_n$ et si ϵ_n décroît vers 0, alors la série de terme général u_n converge.*

⚠ La preuve utilisant Abel est un peu forte pour ce résultat, qui s'obtient facilement en considérant les suites U_{2n} et U_{2n+1} ; elles sont clairement adjacentes. On voit alors en outre que la limite U de U_n est toujours comprise entre U_{2n+1} et U_{2n}.

⊡ **Théorème de Dirichlet**

Définition 604 *On dit qu'une suite (x_n) est à* **variation bornée** *si la série de terme général (y_n) est absolument convergente, avec $y_n = x_{n+1} - x_n$,*

Grâce au critère de Cauchy, on peut constater qu'une suite à variation bornée est convergente.

Théorème 605 (Théorème de Dirichlet) *Soit ϵ_n une suite à variation bornée tendant vers 0, et soit (v_n) une suite telle que la série de terme général (v_n) ait ses sommes partielles bornées, alors la série de terme général (u_n) converge, avec $u_n = \epsilon_n.v_n$.*

<u>Démonstration :</u> Encore grâce à la transformation d'Abel, avec un zeste de critère de Cauchy. □

L'hypothèse $\epsilon_n \to 0$ est nécéssaire ; pour avoir un contre-exemple, considérer $\epsilon_n = 2 + \frac{1}{n^2}$ et $x_n = (-1)^n$.

15.12.6 Produit de convolution de deux séries

⊡ **Pour commencer**

Définition 606 *On se donne deux séries $\sum u_n$ et $\sum v_n$. Par définition, le* **produit de convolution** *de ces deux séries est la série de terme général w_n défini par*

$$w_n = \sum_{k=0}^{n} u_k.v_{n-k}$$

Proposition 607 *Si deux séries sont absolument convergentes, alors le produit de leurs sommes est égal à la somme de leur produit de convolution (qui est une série absolument convergente).*

<u>Démonstration :</u> • On suppose tout d'abord les séries à termes positifs.

En notant U_n, V_n et W_n les sommes partielles des séries de termes généraux u_n, v_n, w_n, on constate facilement que

$$W_n \leq U_n.V_n \leq W_{2n}$$

(on peut voir ça facilement en cochant sur le plan quadrillé les coordonées (a, b) telles que $u_a.v_b$ intervienne dans les différentes sommes ci-dessus, on obtient un carré et deux triangles imbriqués les uns dans les autres :voir figure 15.3...)

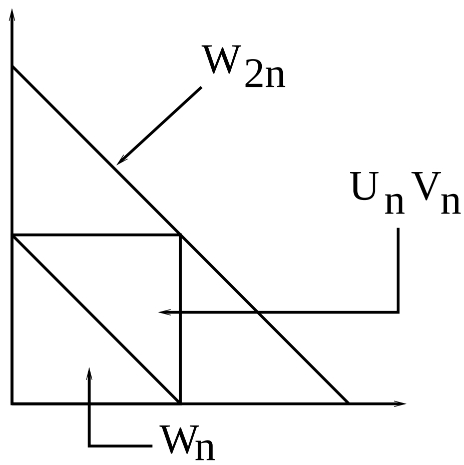


FIG. 15.3 – Produit de convolution

• Le cas général est un peu plus prise de tête.

On reprend les mêmes notations, et on ajoute les notations $|U|_n$, $|V|_n$ et $|W|_n$ pour les sommes partielles des séries de termes généraux $|u_n|$, $|v_n|$ et $|w_n|$.

La convergence absolue du produit de convolution découle trivialement du cas positif. La convergence de W_n vers $(lim\ U_n.lim\ V_n)$, elle, vient du fait que

$$|U_n.V_n - W_n| \leq |U|_n.|V|_n - |W|_n$$

⊡ **Pour les pros : Cauchy-Mertens**

On peut en fait affaiblir les hypothèses (le résultat étant à peine affaibli) :

Théorème 608 (Cauchy-Mertens) *On se donne deux séries de termes généraux (u_n) et (v_n), la série de terme général (u_n) étant supposé absolument convergente, et la série de terme général v_n étant convergente. Alors le produit de convolution de ces deux séries est convergent, et la somme du produit de convolution est égale au produit des sommes de ces deux séries.*

Démonstration : On utilise les notations usuelles pour U_n, V_n, W_n ; on introduit aussi les quantités U, V et W, égales respectivement à la somme des u_n, à la somme des v_n et à la somme des w_n.

• On remplace v_0 par $v_0 - V$. En changeant cela on ne change pas la nature de la série de terme général v_n. La somme U n'est pas changée, et la somme W est diminuée de $V.U$, et V est remplacé par 0 ; donc il n'y a pas de perte de généralité.

• On suppose donc que $V = 0$. Par un raisonnement analogue à celui illustré sur la figure 15.3, on constate que

$$W_n = \sum_{k=0}^{n} u_k.V_{n-k}$$

On se donne M un majorant de $|V_n|$.

On se donne alors $\epsilon > 0$, et N tel que la somme des $|u_i|$ pour $i > N$ soit inférieure à ϵ.

On décompose alors W_n en deux sommes :

$$|W_n| \leq \sum_{k=0}^{N} |u_k.V_{n-k}| + \sum_{k=N+1}^{n} |u_k.V_{n-k}|$$

or :

- $\sum_{k=0}^{N} |u_k.V_{n-k}| \leq (N+1).(max_{k\in[0,N]}|u_k|).(max_{k\in[0,N]}|V_{n-k}|)$

et par ailleurs
- $\sum_{k=N+1}^{n} |u_k.V_{n-k}| \leq M \sum_{k=N+1}^{n} |u_k| \leq \epsilon M$

d'où

$$\leq (N+1).(max_{k\in[0,N]}|u_k|).(max_{k\in[0,N]}|V_{n-k})| + \epsilon.M$$
$$\leq \epsilon.M + \epsilon \text{ pour } n \text{ assez grand}$$

D'où le résultat.□

Et si c'est le produit qui converge ?

On se donne deux série de terme général u_n et v_n, et on définit w_n le terme général de la série produit de convolution de ces deux séries.

On suppose que $\sum w_n$ converge. Alors les séries de termes généraux (u_n) et (v_n) convergent. Ce résultat est cité dans [15], avec un embryon de démonstration (il s'agit d'utiliser Césaro).

15.12.7 Transformation de Toeplitz

Définition 609 *Etant donnée une famille $(c_{i,j})_{(i,j)\in\mathbb{N}^2}$ de coefficients complexes, on définit la* **transformation de Toeplitz** *associée à cette famille comme étant l'application qui à une suite complexe $(u_n)_{n\in\mathbb{N}}$ associe la suite $(Tu_n)_{n\in\mathbb{N}}$ définie par $Tu_n = \sum_{i=0}^{\infty} c_{n,i}.u_i$.*
On dit que T est régulière si et seulement si pour toute suite u_n convergente, Tu_n est définie pour tout n et la suite (Tu_n) converge vers la même limite que (u_n).

Il s'agit de convergence de suites et non de séries ; seuls les termes des Tu_n sont définis par des séries.

Proposition 610 *La transformation de Toeplitz T associée à $c_{n,m}$ est régulière si et seulement si les trois conditions suivantes sont vérifiées :*
• Pour tout j, $c_{i,j} \to 0$ comme i tend vers $+\infty$.

• $\sum_{j=0}^{\infty} c_{i,j}$ tend vers 1 comme $i \to \infty$

• $\sum_{j=0}^{\infty} |c_{i,j}|$ est défini pour tout i et est borné par une constante M indépendante de i

Démonstration :
• Supposons que la transformation T soit régulière.

- le premier • exprime simplement le fait que T est régulière appliqué à la suite $u_n = \delta_{n,j}$.

- Le second • exprime simplement le fait que T est régulière appliqué à la suite constante égale à 1.

- Le troisième • se montre facilement en utilisant le théorème de Banach-Steinhaus. On définit :

$$T_i(x) = \sum_{j=0}^{\infty} c_{i,j} x_j$$

définie pour x une suite convergente ; T_i est bornée indépendamment de i.

Chaque T_i est une application linéaire continue de l'espace de Banach des suites convergentes de $\mathbb{C}$ (pour la norme $\|.\|_\infty$) dans $\mathbb{C}$.

Par le théorème de Banach-Steinhaus (théorème 252), on peut donc trouver M tel que pour toute suite x,

$$|T_i(x)| \leq M.\|x\|_\infty$$

On considère alors, pour i donné et m quelconque, la suite $x^{(m)}$ définie par :

$$x_j = \overline{c}_{i,j}/|c_{i,j}| \text{ si } j < m \text{ et } c_{i,j} \neq 0$$

$$x_j = 0 \text{ sinon}$$

$x^{(m)}$ est convergente, de limite 0, bornée par 1.

En calculant $T_i(x^{(m)})$, on constate que la somme des $|c_{i,j}(x)|$ est bornée par M. D'où le point • .

• Il reste à voir la réciproque, c'est à dire que l'on suppose les trois points réalisés, et on cherche à montrer que T est régulière.

- On se donne (u_n) une suite convergente.

- La suite $\sum_j c_{i,j} u_j$ est évidemment convergente car $c_{i,j} u_j = O(c_{i,j})$ avec $\sum_j c_{i,j}$ absolument convergente. Donc Tu_n est bien défini pour tout n.

- Il reste à montrer que pour toute suite (u_n) la suite de terme général $Tu_n = \sum_{i=0}^{\infty} c_{n,j} u_j$ converge vers la même limite que (u_n)

- dans le cas général, on se ramène facilement au cas d'une limite nulle 0, en remplaçant u_n de limite l par $u_n - l$ (rappelons que par hypothèse $\sum_j c_{i,j} \to 1$ quand $i \to \infty$).

- On se donne donc une suite u_n de limite nulle.

- On se donne N tel que $|u_n| < \frac{\epsilon}{M}$ pour $n \geq N$.

$$|Tu_i| = \sum_{j=0}^{N} c_{i,j}.u_j + \sum_{j=N+1}^{\infty} c_{i,j}.u_j$$

- Le terme $\sum_{j=0}^{N} c_{i,j}.u_j$ tend vers 0 pour i tendant vers $+\infty$, puisque par hypothèse $c_{i,j} \to 0$ pour tout j quand $i \to +\infty$.

- Le terme $\sum_{j=N+1}^{\infty} c_{i,j}.u_j$ est borné par ϵ, par définition de M et N.

- On a donc bien le résultat souhaité.□

Corollaire 611 *Ce résultat montre facilement que la moyenne de Césaro est une transformation de Toeplitz régulière ;* $c_{n,m} = 1/(n+1)$ *si* $m \leq n$ *et* 0 *sinon.*
Si u_n*, suite à valeurs complexes, converge vers* l*, alors* $Tu_n = \frac{1}{n+1} \sum_{i=0}^{n} u_i$ *converge aussi vers* l*.*

Chapitre 16

Développements limités - comparaison de fonctions

16.1 Définitions de base

Définition 612 *On se donne a un point d'un espace topologique X, et on note $\mathcal{V}'(A)$ l'ensemble des voisinages de a privés de a, et $U \in \mathcal{V}'(A)$.*
Etant données deux applications f et g, f de U dans un espace vectoriel normé et g de U dans $\mathbb{R}$, on dit que f est **dominée** *par g, noté $f = O(g)$, si il existe une constante M telle que $\|f\| \leq M.|g|$ sur un certain $V \in \mathcal{V}'(A)$.*
Etant données deux applications f et g, f de U dans un espace vectoriel normé et g de U dans $\mathbb{R}$, on dit que f est **négligeable** *devant g, noté $f = o(g)$, si pour tout ϵ il existe $V \in \mathcal{V}'$ tel que $\|f\| \leq \epsilon.|g|$ sur un certain $V \in \mathcal{V}'(A)$.*
Etant données deux applications f et g de U dans $\mathbb{R}$, on dit que f est **équivalente** *à g et on note $f \simeq g$ si $f - g = o(g)$.*

Ces notions sont locales ; $o(f)$ est lié, implicitement, à a ; la notation, dépourvue de la mention de a, est légèrement abusive car dépendant du contexte.

La notation est doublement abusive ; $f = o(g)$ signifie en fait que f appartient à l'ensemble des fonctions négligeables devant g – il ne s'agit pas d'égalité mais d'appartenance.

Proposition 613 • *$f = o(1)$ si et seulement si f tend vers 0 en a.*

• *$f = O(1)$ si et seulement si f est bornée au voisinage de a.*

• *$\simeq$ est une relation d'équivalence.*

• *f admet une limite non nulle en a et $f \equiv g$ implique que g admet la même limite en a.*

• *$f = o(g)$ équivaut à l'existence de V dans $\mathcal{V}'$ et d'une fonction ϵ de V dans $\mathbb{R}$ tendant vers 0 en a telle que*

$$\forall x \in V \|f(x)\| \leq \epsilon(x).|g(x)|$$

• *$f \simeq g$ équivaut à l'existence de V dans $\mathcal{V}'$ et d'une fonction ϵ de V dans $\mathbb{R}$ tendant vers 0 en a telle que*

$$\forall x \in V \ f(x) = (1 + \epsilon(x)).g(x)$$

• *$f \simeq g$ implique que f et g ont même signe au voisinage de a*

• *$f \simeq 0$ implique que f est nulle au voisinage de a (et pas seulement que f tend vers 0 !)*

• *En $+\infty$ $x \mapsto ln(x^a)$ est négligeable devant $x \mapsto x^b$ qui est négligeable devant $x \mapsto e^{cx}$ pour a, b et $c > 0$.*

Ces notations sont appelées notations de Landau.

Voir par exemple 947.

Définition 614 *On dit que f de $\mathbb{R}$ dans $\mathbb{C}$ admet un* **développement limité** *en $a \in \mathbb{R}$ à l'ordre n s'il existe une fonction polynôme P de $\mathbb{R}$ dans $\mathbb{C}$ telle que au voisinage de a $f(x) = P(x) + o((x-a)^n)$.*

Bien que peu utilisé dans la pratique, il existe une autre définition, plus générale, utilisant les polynômes à coefficients dans un espace vectoriel normé :

Définition 615 (Généralisation) *On dit que f de $\mathbb{R}$ dans un espace vectoriel normé E admet un* **développement limité** *en $a \in \mathbb{R}$ à l'ordre n s'il existe un polynôme P à coefficients dans E telle que au voisinage de a $f(x) = P(x) + o((x-a)^n)$.*

Je n'étudierai pas ce cas, pour lequel on se réfèrera au livre [3].

On définit parfois la notion de développement limité au sens fort :

Définition 616 *On dit que f de $\mathbb{R}$ dans un espace vectoriel normé E admet un* **développement limité au sens fort** *en $a \in \mathbb{R}$ à l'ordre n s'il existe un polynôme P à coefficients dans E telle que au voisinage de a $f(x) = P(x) + O((x-a)^{n+1})$.*

Proposition 617 • *S'il existe un polynôme P à coefficients dans E telle que au voisinage de a, $f(x) = P(x) + O((x-a)^{n+1})$, on peut toujours imposer que P soit de degré $\leq n$. En effet il suffit de prendre le reste dans la division euclidienne de P par $(x-a)^{n+1}$; le quotient multiplié par $(x-a)^n$ sera un négligeable devant $(x-a)^{n+1}$. Ceci a pour conséquence l'unicité du développement limité.*

• *Si f est continue en a, alors elle admet un développement limité en a d'ordre 0.*

• *Si f est dérivable en a, alors elle admet un développement limité en a d'ordre 1.*

• *Si f admet un développement limité à l'ordre 0 en a, alors elle est continue en a ou admet un prolongement par continuité en a.*

• *Si f admet un développement limité à l'ordre 1 en a, alors elle est dérivable en a ou est prolongeable en une fonction dérivable en a.*

• *si f est n fois dérivable en a, alors elle admet un développement limité en a d'ordre n (voir le théorème 583).*

⚠ Il n'y a pas de réciproque au dernier • ! Une fonction peut admettre un développement limité d'ordre n sans être dérivable plus d'une fois ! On donnera un exemple dans la partie 16.2

16.2 Opérations sur les équivalents et les développements limités

Proposition 618 *On suppose f, g, h et $k > 0$ au voisinage de a ; x est un réel.*

• $f \simeq g$ *et* $h \simeq k \Rightarrow fh \simeq gk$

• $f \simeq g \Rightarrow f^x \simeq g^x$

⚠ Par contre on n'a PAS la possibilité d'additionner des équivalents, même positifs.

Exemples : (fonction admettant un développement limité à l'ordre 2 sans être 2 fois dérivable :

Considérons la fonction f :

$$t \mapsto t^3.sin(1/t)$$

prolongée par continuité en 0 ($f(0) = 0$).

On considère son développement limité en 0.

Il est clair que $f(t) = O(t^3)$ car sin est bornée.

Pourtant, en faisant le calcul, on constatera que f n'est pas deux fois dérivable en 0.

Proposition 619 (Unicité du développement limité) *Si f admet un développement limité en a à l'ordre n, ce développement limité est unique.*

Démonstration : Ecrire la différence entre deux polynômes de degré n comme un $o((x-a)^n)$; puis considérer le premier coefficient sur lequel ils diffèrent.□

NB : On peut définir les développements limités sur des intervalles de $\mathbb{R}$; dans ce cas on note que le développement limité en a ou en b pour un développement limité sur $]a, b[$ est unique.

Proposition 620 (Troncature des développements limités) *Si* $f(t) = P(t) + o((t-a)^n)$ *alors a fortiori* $f(t) = P(t) + o((t-a)^p)$ *pour* $p < n$.

Proposition 621 *Si* $f(t) = P(t)(t-a)^k + o((t-a)^n)$ *avec* $k < n$ *alors* g *définie par* $g(t) = f(t)/(t-a)^k$ *est prolongeable par continuité en* a *et* $g(t) = P(t) + o((t-a)^{n-k})$.

Proposition 622 (Intégration d'un développement limité) *Supposons f dérivable au voisinage de a, et*

$$f'(t) = a_0 + a_1(t-a) + a_2(t-a)^2 + ... + a_n(t-a)^n + o((t-a)^n)$$

Alors

$$f(t) =$$

$$f(a)+a_0(t-a)+\frac{a_1}{2}(t-a)^2+\frac{a_2}{3}(t-a)^3+...+\frac{a_n}{n+1}(t-a)^{n+1}+o((t-a)^{n+1})$$

Démonstration : On suppose donc

$$f(t) = P(t) + \epsilon(t)(t-a)^n$$

On se ramène par translation à $a = 0$ et on écrit

$$f(t) = f(0) + \int_0^t P(u)du + \int_0^t \epsilon(u)du$$

$$\left|f(t) - f(0) - \int_0^t P(u)du\right| = \left|\int_0^t \epsilon(u)du\right|$$

$$\leq sup_{u\in[0,t]}|\epsilon(u)| \int_0^t |u^n|.du$$

D'où le résultat.□

On a un résultat similaire avec des développements limités au sens fort.

Proposition 623 (Intégration d'un développement limité au sens fort) *Supposons f dérivable au voisinage de a, et*

$$f'(t) = a_0 + a_1(t-a) + a_2(t-a)^2 + ... + a_n(t-a)^n + O((t-a)^{n+1})$$

Alors

$$f(t) =$$

$$f(t_0)+a_0(t-a)+\frac{a_1}{2}(t-a)^2+\frac{a_2}{3}(t-a)^3+...+\frac{a_n}{n+1}(t-a)^{n+1}+O((t-a)^{n+2})$$

Démonstration : Tout à fait similaire à la preuve précédente.□

⚠ On NE PEUT PAS, dans le cas général, dériver terme à terme un développement limité ! Toutefois, c'est possible pour une fonction de classe C^∞ par exemple.

Par simplicité, on va maintenant supposer que $a = 0$. On en déduit bien évidemment le cas général.

Définition 624 *On définit la relation suivante sur l'ensemble des fonctions continues en* 0 *:*

$$f \equiv^n g \text{ si } f - g = o(x^n)$$

On dit que f *est* **tangente** *à* g *à l'ordre* n*.*

Proposition 625 • *Tout d'abord une évidence :*

$$f \equiv^n g$$

si et seulement si

$$\lim_{t\to 0} t^{-n}(f(t) - g(t)) = 0$$

• $f \equiv^n g$ *et* $h \equiv^n k$ *impliquent* $f + h \equiv^n g + k$
• $f \equiv^n g$ *et* $h \equiv^n k$ *impliquent* $f.h \equiv^n g.k$
• $f \equiv^n g$ *et* $f(0) \neq 0$ *implique* $1/f \equiv^n 1/g$
• $f \equiv^n g$ *et* P *polynôme implique* $P \circ f \equiv^n P \circ g$
• $f \equiv^n g$ *et* $h(x) = O(x^p)$ *impliquent* $f \circ h \equiv^{np} g \circ h$

Démonstration : Seuls les trois derniers points méritent notre attention, les autres étant clairs.

• Supposons donc $f \equiv^n g$ et $f(0) \neq 0$.

$$x^{-n}(1/f(x) - 1/g(x)) = \frac{g(x) - f(x)}{x^n f(x) g(x)}$$

Or f et g ont une limite non nulle en 0. Donc $\frac{g(x)-f(x)}{x^n f(x)g(x)}$ tend vers 0 en 0 ; d'où le résultat.

• Il est suffisant de le montrer avec P un monôme ; et ce résultat découle de

$$f^p(x) - g^p(x) = (f(x) - g(x)) \sum_{k=0}^{p-1} f(x)^k g(x)^{p-1-k}$$

(le terme en $\sum$ étant borné)

• On procède comme suit :

- Ecrivons $f(x) - g(x) = \epsilon(x)x^n$, avec $\epsilon(x) \to 0$ quand $x \to 0$

- Ecrivons $h(x) = M(x)x^p$, avec M bornée au voisinage de 0

- Alors $f \circ h(x) - g \circ h(x) = \epsilon(M(x)x^p)M(x)^n x^{np}$

Le résultat est alors acquis ; on note l'efficacité de la méthode consistant à écrire explicitement les $o(.)$ avec des $\epsilon(.)$ et des $M(.)$. $\square$

Ces résultats vont permettre de combiner des développements limités de la manière expliquée ci-dessous.

Proposition 626 (Somme, produit, quotient, composé) *Soient f et g admettant des développements limités en 0 à l'ordre n.*
Alors $f+g$ et fg admettent des développements limités en 0 à l'ordre n, et f/g aussi si $g(0) \neq 0$.

• *Le développement limité de $f+g$ est la somme des développements limités de f et g.*

• *Le développement limité de fg est le produit des développements limités de f et g (on peut tronquer les termes d'ordre $> n$).*

• *Le développement limité du composé $g \circ f$, si $f(0) = 0$, si $f \equiv^n P$ et si $g \equiv^n Q$, est $Q \circ P$, valable à l'ordre n (on peut tronquer les termes d'ordre supérieur à n).*

• *Le développement limité de f/g est égal au développement limité du quotient des développements limités de f et g.*

Démonstration : • Le cas de l'addition, du produit découlent immédiatement des résultats précédents.

• En reprenant les notations de l'énoncé, on écrit

$$g \circ f(x) = Q(f(x)) + \epsilon(f(x))f(x)^n$$

On sait ensuite que $Q(f(x))$ est équivalent à $Q(P(x))$ à l'ordre n (résultat de la proposition précédente). Et on sait que $f(x)^n$ est équivalent à $Q(x)^n$ à l'ordre n, toujours par la proposition précédente. D'où le résultat.

• Pour le quotient, il suffit, en vertu du résultat sur le produit, d'étudier le développement limité de l'inverse d'une fonction g donnée.

Or si P est le développement limité de g à l'ordre n, alors $1/g \equiv^n 1/P$ (résultat prouvé un peu plus haut).

Il suffit donc de calculer le développement limité de l'inverse. D'où le résultat (on verra un peu plus loin comment déterminer le développement limité en question). $\square$

Proposition 627 *Les résultats précédents restent vrais si l'on travaille avec des développements limités au sens fort.*

Démonstration : Les preuves sont les mêmes à peu de choses près. $\square$

Dans la pratique du calcul des développements limités , bien penser à ne PAS calculer les termes de degré trop élevés, qui devront de toute façon être oubliés.

Proposition 628 *Le développement limité du quotient de P par Q est donné par la division suivant les puissances croissantes (voir théorème 943).*
C'est-à-dire :

Avec $Q(0) \neq 0$, P/Q est équivalent en 0 à T, avec

$$P = TQ + X^{n+1}R$$

Démonstration : Le résultat découle immédiatement de la formule $P = TQ + X^{n+1}R$, en calculant $X^{-n-1}.(P/Q - T)$.$\square$

16.3 Développements asymptotiques

La plupart des résultats de cette partie étant relativement faciles à établir à partir des résultats précédents ou par des méthodes similaires, les théorèmes seront généralement énoncés sans preuve.

Les développements limités sont limités au cas de fonctions admettant une limite en a ; on aimerait pouvoir manier des fonctions tendant vers l'infini, en les approchant par des fractions rationnelles au lieu de les approcher par des polynômes, par exemple.

Pour cela on étend les notations o, O et $\simeq$ à des fonctions non nécéssairement définies continues en a, et a peut désormais être $+\infty$ ou $-\infty$.

On dira que $f = O(g)$, si f est inférieure en module à Cg pour C une certaine constante et sur un certain voisinage de a.
On dira que $f = o(g)$ si pour tout ϵ positif, f est inférieure en module à $\epsilon.g$ sur un certain voisinage V_ϵ (dépendant de ϵ bien sûr) de a.

On dira que $f \simeq g$ si $f - g = o(f)$.

De nombreuses propriétés restent valables ou s'étendent :

- $f \simeq g$ et $h \simeq k$ impliquent $fh \simeq gk$
- $f \simeq g$ et $1/f$ non nul au voisinage impliquent $1/f \simeq 1/g$
- si $h(b) = a$ et si h est continue en b, alors $f \simeq g$ implique $f \circ h \simeq g \circ h$.

Définition 629 *On appelle* **échelle de comparaison au voisinage de** a[a] *une famille d'applications définies au voisinage de* a*, non équivalentes à* 0*, et totalement ordonné par* $. = o(.)$*.*

[a] Qui peut être fini ou non !

Exemples : • Les monômes x^k pour $k \in \mathbb{Z}$ constituent une échelle de comparaison au voisinage de l'infini.

• Les monômes x^k pour $k \in \mathbb{Z}$ constituent une échelle de comparaison au voisinage de zéro.

Mais on peut faire plus fin :

• Les monômes x^r pour $r \in \mathbb{R}$ constituent une échelle de comparaison au voisinage de l'infini.

• Les monômes x^r pour $k \in \mathbb{R}$ constituent une échelle de comparaison au voisinage de zéro.

• Les $x \mapsto x^\alpha ln(x)^\beta$ constituent une échelle de comparaison (ordonnée par l'ordre lexicographique sur (α, β)) au voisinage de $+\infty$

• Les $x \mapsto x^\alpha ln(x^\beta)$ ne constituent PAS une échelle de comparaison, car par exemple $ln(x^2)$ ($= 2ln(x)$!) et $ln(x)$ ne sont pas ordonnée pour $. = o(.)$!

• Les $x \mapsto x^\alpha e^{xP(x)}$ pour $P \in \mathbb{R}[X]$ constituent une échelle de comparaison, ordonnée par l'ordre lexicographique sur les coefficients de P par ordre décroissant puis le coefficient α (attention, il s'agit d'un ordre lexicographique sur des suites pouvant avoir un nombre de termes arbitrairement grand...) au voisinage de $+\infty$.

• Avec $ln_n(x) = \underbrace{ln \circ ln \circ ln \circ ... \circ ln}_{n \text{ fois}}(x)$, les $x \mapsto x^\alpha \Pi_{i=1}^n ln_i^{\beta_i}(x)$[1] forment une échelle de comparaison au voisinage de $+\infty$.

• Toujours au voisinage de $+\infty$, la famille des $\pi_{i=1}^n ln_i^{\beta_i}(x) x^\alpha \pi_{j=1}^p e^{\lambda_j} x^{\mu_j}$ forme aussi une échelle de comparaison. L'exponentielle l'emporte sur les puissances qui l'emportent sur les logarithmes.

[1] pour α dans $\mathbb{R}$, n dans $\mathbb{N}$ et les β_i dans $\mathbb{R}$

Définition 630 *On se donne f une application d'une partie X de $\mathbb{R}$ dans $\mathbb{R}$, et a appartenant à l'adhérence de X dans $\overline{\mathbb{R}}$, et on se donne $\mathcal{E}$ une échelle de comparaison au voisinage de a.*

On dit que f admet un développement asymptotique pour l'échelle $\mathcal{E}$ à la précision ϕ avec ϕ un élément de $\mathcal{E}$ s'il existe une famille de réels presque tous nuls[a] $(\lambda_\psi)_{\psi\in\mathcal{E}}$ *tels que*

$$f(x) = \sum_{\psi\in\mathcal{E}} \lambda_\psi \psi(x) + o(\phi)$$

[a] C'est à dire que seul un nombre fini de ces réels sont non nuls.

16.4 Zoologie des comparaisons de séries, de fonctions

16.4.1 Equivalent de la suite des sommes partielles $U_n = \sum_{k=0}^n u_k$

Citons un résultat facile :

Théorème 631 *Etant donnée une suite réelle $(u_n)_{n\in\mathbb{N}}$, on définit la suite des sommes partielles associée $U_n = \sum_{k=0}^n u_k$; lorsque (U_n) est une suite convergente, on définit aussi la somme $U = lim_{n\to\infty} U_n$, et le reste $R_n = \sum_{k=n+1}^\infty u_k = U - U_n$. On définit de même, avec (v_n) une suite réelle, la série $V_n = \sum_{k=0}^n v_k$, et si $V_n \to V$, $R'_n = V - V_n$. Alors :*

Hypothèses	*Conclusions*
$u_n > 0$ *pour* $n \geq n_0$ $u_n \simeq v_n$ U_n *converge*	V_n *converge* $R_n \simeq R'_n$
$v_n > 0, u_n > 0$ $v_n = o(u_n)$ V_n *diverge*	U_n *diverge* $U_n \simeq V_n$

16.4.2 Equivalent de $F(x) = \int_a^x f(t)dt$

Théorème 632 *On se donne désormais deux fonctions f et g définies sur $]a, +\infty[$ intégrables sur $]a, x[$ pour tout $x > a$.*

$$\boxed{\textit{ON SUPPOSE QUE } f \textit{ EST POSITIVE sur }]b, \infty[\textit{ avec } b \geq a}$$

On définit $F(x) = \int_a^x f(t)dt$ et $G(x) = \int_a^x g(t)dt$. Si F a une limite en $+\infty$ on définit $R_f(x) = \int_x^\infty f(t)dt$, et si G a une limite en $+\infty$ on définit $R_g(x) = \int_x^\infty f(t)dt$.
Alors on a les résultats suivant au voisinage de $+\infty$:

$$\begin{cases} \int_a^x f(t)dt = +\infty & \begin{cases} g = o(f) \Rightarrow G = o(F) \\ g = O(f) \Rightarrow G = O(F) \\ g \simeq f \Rightarrow F \simeq G \end{cases} \\ \int_a^x f(t)dt < +\infty & \begin{cases} g = o(f) \Rightarrow & \begin{cases} \int_a^\infty g(t)dt < +\infty \\ R_g = o(R_f) \end{cases} \\ g = O(f) \Rightarrow & \begin{cases} \int_a^\infty g(t)dt < +\infty \\ R_g = O(R_f) \end{cases} \\ g \simeq f \Rightarrow & \begin{cases} \int_a^\infty g(t)dt < +\infty \\ R_g = \simeq R_f) \\ g \geq 0 \text{ sur }]c, \infty[\text{ avec } c \geq a \end{cases} \end{cases} \end{cases}$$

Démonstration : Certaines preuves se trouvent un peu plus haut ; les autres sont faciles.□

16.4.3 Comparaison séries-intégrales, cas $\frac{f'}{f}$ convergent

Théorème 633 *On se donne une fonction f strictement positive et de classe C^1 au voisinage de $+\infty$.*
On considère d'une part la série $\sum f(n)$, et d'autre part l'intégrale $\int_a^\infty f$.
On définit $\mathcal{S}_n = \sum_{i=0}^n f(i)$, $F(x) = \int_a^x f(t)dt$ (avec a suffisamment grand).
Soit $K(r) = \frac{r}{1-e^{-r}}$, prolongé par continuité en 0 par $K(0) = 1$.
Si la somme converge, on définit $R_n = \sum_{i=n+1}^\infty f(i)$.
Si l'intégrale converge, on définit $R_f(x) = \int_x^\infty f(t)dt$ (pour x assez grand).
On suppose $\frac{f'}{f} = r + o(1)$ alors
l'intégrale converge $\iff$ la série converge, on a alors $R_n \simeq K(r)F(n)$.
l'intégrale diverge $\iff$ la série diverge, on a alors $\mathcal{S}_n \simeq K(r)F(n)$.

Voir le théorème 593 pour plus d'informations sur le cas $a = 0$.

Démonstration : On définit $g(t) = e^{-rt}f(t)$.

Alors on a $\frac{g'}{g} = o(1)$.

Donc on peut trouver un certain η tel que sur $]\eta, \infty[$ $|\frac{g'}{g}| \leq \epsilon$. Donc pour x appartenant à $[n-1, n]$ et pour $x > \eta$,

$$|ln(\frac{g(x)}{g(n)})| \leq \epsilon$$

et donc

$$\frac{g(x)}{g(n)} \leq e^{\epsilon} \text{ et } \frac{g(n)}{g(x)} \leq e^{\epsilon}$$

$$(e^{-\epsilon} - 1)\ f(n) \leq f(x) - f(n) \leq (e^{\epsilon} - 1)\ f(n)$$

Alors

$$\int_{n-1}^{n} f(t)dt = \int_{n-1}^{n} e^{rt}(f(t) - f(n))dt + \int_{n-1}^{n} e^{rt} f(n)dt$$

Si $r = 0$ l'intégrale tout à droite est simplement $f(n)$, et le résultat en découle ; si $r \neq 0$, on réécrit cette expression sous la forme

$$\left| \int_{n-1}^{n} f(t)dt - K(r)f(n) \right| \leq e^{rn}(e^{\epsilon} - 1)g(n) = (e^{\epsilon} - 1)f(n)$$

Et on a le résultat souhaité.□

16.5 Zoologie des développements limités

C'est en développant qu'on devient un bon développeur... On trouve la liste des résultats classiques dans le formulaire 38.

Le développement asymptotique de $x^{\frac{1}{x}}$ est :

$$x^{\frac{1}{x}}$$
$$= exp\left(\frac{ln(x)}{x}\right)$$
$$= 1 + \frac{ln(x)}{x} + \frac{1}{2}\left(\frac{ln(x)}{x}\right)^2 + O\left(\left(\frac{ln(x)}{x}\right)^3\right)$$
$$= 1 + \frac{ln(x)}{x} + \frac{1}{2}\left(\frac{ln(x)}{x}\right)^2 + o\left(\frac{1}{x^2}\right)$$

On trouvera de nombreux exemples (méthode de Laplace, méthode du col...) dans le livre "Calcul infinitésimal", de Dieudonné.

Chapitre 17

Interversions

La hantise classique du mathématicien est de justifier les permutations entre sommes, intégrales, limites, dérivation. L'objectif de cette partie est de fournir tous les outils pour affronter cette tâche...

17.1 Interversion de limites et de dérivation

Pour cela, on consultera la partie11.2.2.

17.2 Interversion de limites et de limites

Il s'agit bien sûr d'intervertir une limite au sens d'un espace de fonctions avec une limite au sens de l'espace sur lequel sont définies ces fonctions... Formellement ça donne ça :

Théorème 634 *Soit X un espace topologique, A une partie de X, a un point de l'adhérence de A.*

Soit (E, d) un espace métrique, et enfin soit f_n une suite d'applications de A dans E, et f une application de A de E.

Supposons que :

• $lim_{x \to a} f_n(x) = e_n$

• f_n *converge* *uniformément* *vers f sur A*

• E *est COMPLET*

Alors il existe un certain e limite de e_n et $lim_{x \to a} f(x) = e$

Démonstration :

• On considère $\tilde{f}_n$ la fonction égale à f_n sur A et prolongée par continuité en a en posant $\tilde{f}_n(a) = e_n$; $\tilde{f}_n$ est donc continue en a.

• On se donne $\epsilon > 0$.

• La convergence des f_n étant uniforme, on peut trouver N tel que

$$n, m > N \Rightarrow \forall x, d(f_n(x), f_m(x)) < \epsilon$$

• On passe à la limite pour $x \to a$ et on obtient

$$n, m > N \Rightarrow d(e_n, e_m) \leq \epsilon$$

(ceci implique que la suite (e_n) est de Cauchy, donc par complétude de E qu'elle a une certaine limite e)

• La convergence de $\tilde{f}_n$ est donc uniforme (par le critère de Cauchy pour la norme uniforme) (voir 242)

• Les $\tilde{f}_n$ étant continues en a, leur limite (uniforme) est donc continue en a (voir proposition 468) ; d'où le résultat.□

Corollaire 635 *Soit f_n des fonctions continues admettant une limite e_n en $a \in \overline{A}$, à valeurs dans E espace de Banach, telles que la série des f_n converge normalement (pour la norme infinie $\|.\|_\infty$) vers f.*

Alors la série de terme général e_n converge vers un certain e, et f tend vers e en a.

Démonstration : C'est exactement la même propriété, dans le cas des séries...□

17.3 Interversion d'une limite et d'une intégrale

Définition 636 (Application réglée) *Une application de $\mathbb{R}$ dans un espace topologique est dite* **réglée** *si et seulement si elle admet une limite à droite et une limite à gauche en tout point.*

On ne demande pas du tout que la limite à droite soit égale à la limite à gauche, ni qu'aucune de ces deux limites soit égale à la valeur de l'application en ce point.

Les applications réglées ont été définies en parties 15.3 (sur l'intégrale de Riemann) comme les éléments de l'adhérence de l'ensemble des fonction en escalier (adhérence pour la norme $\|.\|_\infty$). Ces deux définitions sont équivalentes pour peu que

X soit métrique.

Théorème 637 *On se donne E un espace de Banach , et f_n une suite d'applications réglées du segment $[a,b]$ de $\mathbb{R}$ dans E convergeant uniformément vers f.*
Alors f est réglée et $\int_{[a,b]} f = lim \int_{[a,b]} f_n$.

Démonstration : Le fait que f soit réglée est un corollaire immédiat du théorème 634 (interversion des limites de suites et de fonctions sous certaines hypothèses).

Pour conclure il suffit d'observer que $|\int_{[a,b]} f - \int_{[a,b]} f_n| \leq \int_{[a,b]} |f - f_n| \leq (b-a)\|f - f_n\|_\infty$.$\square$

Corollaire 638 *Si $\sum f_n$ converge normalement (pour la norme $\|.\|_\infty$), de $[a,b]$ segment de $\mathbb{R}$ dans un espace de Banach E, et si les f_n sont réglées, alors la somme des f_n est une fonction f réglée, d'intégrale la somme des intégrales des f_n.*

Démonstration : C'est exactement la même propriété, dans le cas des séries.$\square$

Chapitre 18

Séries entières

18.1 Définitions

Définition 639 *On appelle* **série entière** *une série de fonctions de terme général $z \mapsto (a_n).z^n$, avec la suite (a_n) une suite de nombres complexes. On la note $\sum_n a_n.z^n$.*
La suite $(a_n)_{n\in\mathbb{N}}$ est appelée **suite des coefficients de la série entière***.*
On appelle **domaine de convergence** *d'une série entière l'ensemble des z tels que la somme $\sum a_n.z^n$ est bien définie.*

18.2 L'indispensable : le lemme d'Abel

Le lemme d'Abel doit être présent en mémoire en toute circonstance et doit pouvoir être exhibé sans la moindre hésitation si un jour quelqu'un vous aborde, pointe un révolver sur vous et vous dit "Le lemme d'Abel ou la vie !".

Lemme 640 (Lemme d'Abel) *Soit z un nombre complexe tel que la suite $a_n.z^n$ soit bornée. Alors pour tout z' tel que $|z'| \leq |z|$, la série $\sum a_n.z'^n$ est absolument convergente.*

Démonstration : $|a_n.z'^n| = |a_n.z^n|.|z'/z|^n \leq M(\frac{|z'|}{|z|})^n$, et $\frac{|z'|}{|z|} < 1$... Le résultat en découle.□

Maintenant que me voilà rassuré quant à votre avenir au cas où vous fassiez agresser par un mathématicien amateur désireux (à juste titre) d'apprendre ce résultat fondamental, le lemme d'Abel permet d'introduire une notion fondamentale en matière de séries entières : le rayon de convergence.

Définition 641 *Le* **rayon de convergence** *de la série entière* $\sum a_n.z^n$ *est le sup des réels* $r > 0$ *tel que* $a_n.z^n$ *soit bornée. Le rayon de convergence peut éventuellement être infini.*

Théorème 642 (conséquence du lemme d'Abel) *Si la série* $\sum a_n.z^n$ *est de rayon de convergence* R*, alors :*

• *pour tout* z *de module* $< R$ *la série de terme général* $\sum a_n.z^n$ *est absolument convergente.*

• *pour tout* z *de module* $> R$ *la série de terme général* $\sum a_n.z^n$ *diverge, et en fait la suite* $a_n.z^n$ *n'est pas même bornée.*

Démonstration : Il suffit de regarder le lemme d'Abel dans le blanc des yeux.□

Le disque ouvert de centre 0 et de rayon R s'appelle le **disque de convergence**. Lorsque $R = +\infty$, on appelle $\mathbb{C}$ tout entier le **disque de convergence**.

Le théorème que l'on vient de voir permet donc de classer les nombres complexes en deux catégories, l'une où la somme est convergente, avec plein de belles propriétés que l'on va voir, et un ensemble où il n'y a pas convergente ; toutefois, il reste une zone limite, sur laquelle nous n'avons pas de résultat, et sur laquelle nous en auront bien peu ; le cercle de rayon le rayon de convergence, lorsque celui-ci est fini.

18.3 A l'intérieur du disque de convergence

Ah, joie du matheux : on va avoir des choses faciles et élégantes à dire.

Théorème 643 (Fondamental) *Si* $\sum a_n.z^n$ *a pour rayon de convergence* R*, la série de terme général* $\sum a_n.z^n$ *converge normalement, donc uniformément,* <u>*sur tout compact*</u> *contenu dans le disque de centre* 0 *et de rayon* R*.*

Démonstration : Soit K un tel compact, qu'on va supposer non vide. Alors K est nécéssairement inclus dans un disque fermé de rayon r inférieur à R (en cas contraire considérer une suite x_n dans K tels que $R - |x_n| < 1/n$, et considérer sa limite, qui doit appartenir à K - tout compact étant fermé).

$|a_n.z^n|$ est donc majoré par $a_n.r^n$, lequel est absolument convergent. D'où la convergence normale.□

Ce théorème va nous permettre de déduire pas mal de petites choses fondamentales :

• La somme d'une série entière est continue sur le disque ouvert de convergence (preuve facile : la limite uniforme d'une suite de fonctions continues est continue, donc pour prouver la continuité en x on considère un compact contenant x et inclus dans le disque de convergence ; il y a convergence normale, donc uniforme, sur ce compact)

⚠ Même s'il y a convergence sur un disque FERMÉ on ne peut pas en déduire que la somme est continue sur le disque FERMÉ !

• la somme f de la série entière $\sum a_n.z^n$ admet pour développement limité à l'ordre n en 0 $f(z) = \sum_{k=0}^{n} a^k.z^k + o(z^n)$ (en fait, plus précisément, $O(z^{n+1})$) (preuve en écrivant $f(z) - \sum_{k=0}^{n} a^k.z^k$, et en factorisant par z^{n+1} ; le quotient est une série entière convergente, donc continue en 0).

• la somme f de la série entière $\sum a_n.z^n$ peut s'approximer uniformément par une suite de polynôme sur le disque fermé de centre 0 et de rayon r pour tout r STRICTEMENT inférieur au rayon de convergence. ⚠ ATTENTION il ne s'agit pas d'un corollaire du théorème de Weierstrass d'approximation de fonctions continues uniformément par des polynômes car ici les fonctions vont de $\mathbb{C}$ dans $\mathbb{C}$! (la preuve est en fait claire par les résultats précédents ; un disque fermé est compact, donc on applique le dernier théorème).

On a un résultat parfois utile à souligner aussi :

> **Théorème 644** *Soit f la somme d'une série entière de rayon de convergence R. Alors f admet un prolongement continu sur le disque FERMÉ de rayon R si et seulement si f est limite uniforme de polynômes sur le disque ouvert.*

<u>Démonstration :</u> Il est clair que si f est limite uniforme d'une suite de polynômes sur le disque ouvert, alors

Réciproquement, supposons f prolongée par continuité sur le disque fermé.

• Soit $\epsilon > 0$.

• f est uniformément continue sur ce disque, par le théorème 198.

• On peut donc déterminer $\alpha > 0$ tel que $|z - z'| \leq R.\alpha \Rightarrow |f(z) - f(z')| \leq \epsilon$.

• On peut développer en série entière $z \mapsto f(\frac{z}{1+\alpha})$ sur le disque ouvert de rayon $R.(1 + \alpha)$. On peut donc approximer cette fonction sur le disque fermé de rayon R.

• Soit donc un polynôme P tel que $|P(z) - f(\frac{z}{1+\alpha})| \leq \epsilon$ pour tout z de module $\leq R$.

• On constate alors que z et $z/(1+\alpha)$ sont à une distance $\leq R.\alpha$; donc par définition de α trois lignes plus haut, on peut écrire que $|f(z) - f(\frac{z}{1+\alpha})| \leq \epsilon$.

• On a donc bien approximé f par P à 2ϵ près sur le disque fermé.$\square$

Bien noter que l'on a écrit que f était limite uniforme d'une suite de polynômes, mais pas que cette suite de polynômes était la série $\sum_{k=0}^{n} a_n.z^n$! Il se peut que ce ne soit pas le cas...

18.4 A la limite du disque de convergence

On va juste présenter quelques cas, pour montrer que tout est possible ; soit convergence sur tout le cercle limite, soit divergence sur tout le cercle limite, soit tantôt convergence tantôt divergence...

• $\sum z^n/n$ a clairement pour rayon de convergence 1. En 1, il est clair que la série diverge. Pour z de module 1 et différent de 1, alors la somme des z^k entre 0 et n est bornée (ça fait $\frac{1-z^{n+1}}{1-z}$).

• $\sum z^n$ a clairement pour rayon de convergence 1. Et il est non moins clair que sur tout le cercle de rayon 1 la somme diverge.

• $\sum z^n/n^2$ a toujours aussi clairement un rayon de convergence 1, et il est non moins clair que sur tout le cercle de rayon 1 la somme converge.

18.5 Comment déterminer un rayon de convergence ?

18.5.1 Formule d'Hadamard

Théorème 645 (Formule d'Hadamard) *On se donne $\sum a_n.z^n$ une série entière. Soit $L = limsup|a_n|^{1/n}$, alors $R = 1/L$ (avec $1/0 = +\infty$ et $1/+\infty = 0$).*

Démonstration :
• Cas $0 < L < +\infty$: soit $R = 1/L$, montrons que R est rayon de convergence.

- si $|z| < R$, $a_n.z^n = O((L.|z|)^n)$, or $L.|z|$ est plus petit que 1 ; donc notre série est un $O()$ d'une série absolument convergente, donc elle converge absolument.

- si $|z| > R$, $a_n.z^n$ ne tend pas vers 0, car son $limsup$ ne tend pas vers 0.

• Les autres cas se déduisent facilement de celui-ci... □

Théorème 646 (Règle de D'Alembert) *On se donne $\sum a_n.z^n$ une série entière. On suppose que les a_n sont non nuls, au moins à partir d'un certain rang, et on suppose que a_{n+1}/a_n tend vers L. Alors $R = 1/L$.*

Démonstration :
Ce théorème est une conséquence immédiate du critère de D'Alembert 572 dans le détermination de la convergence de séries.□

18.6 Dérivation des séries entières

Définition 647 *On se donne $\sum_{n\geq 0} a_n.z^n$ une série entière ; on appelle* **série dérivée d'ordre** p *la série $\sum_{n\geq 0} \frac{(n+p)!}{n!} a_{n+p}.z^n$.*

On appelle **série dérivée** *(tout court !) d'une série entière la série dérivée d'ordre* 1.

Théorème 648 *La série dérivée $\sum_{n\geq 0} \frac{(n+p)!}{n!} a_{n+p}.z^n$ de la série $\sum a_n.z^n$ une série entière de rayon de convergence R a le même rayon de convergence R.*

Démonstration :
• On procède par récurrence, la série dérivée d'ordre p étant la dérivée d'ordre 1 de la dérivée d'ordre $p-1$; il suffit de montrer le résultat pour $p = 1$.

• On suppose donc $p = 1$.

• La série dérivée est donc une série de terme général $(n/z).a_n.z^n$. Donc en valeur absolue pour $n \geq z$, le terme général est supérieur en module à celui de la série $\sum a_n.z_n$. Donc le rayon de convergence est inférieur ou égal à R.

• Montrons maintenant qu'il est supérieur ou égal à R. On se donne z de module $< R$, et un réel $r < R$ tel que $|z| < r < R$.

• $(n+1).|a_{n+1}|.|z|^n \leq ((\frac{|z|}{r})^{n+1}.(n+1)/z).|a_{n+1}|.|r|^{n+1}$

• Le terme ci-dessus est plus petit que $|a_{n+1}|.|r|^{n+1}$ pour n assez grand.□

Un corollaire immédiat :

Corollaire 649 *Si f est une somme de série entière, alors $f^{(n)}(0)$ est égal à $n!.a_n$ avec $f = \sum a_n.z_n$.*

18.7 Produit de séries entières

Définition 650 *Etant donnée deux séries entières $\sum_{n\geq 0} a_n z^n$ et $\sum_{n\geq 0} b_n z^n$, on définit la* **série entière produit** *par $\sum_{n\geq 0} c_n z^n$, avec $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, et la* **séries entière somme** *par $\sum_{n\geq 0} d_n z^n$ avec $d_n = a_n + b_n$.*

Il est bien évident que le rayon de convergence de la série entière somme de deux séries entières est au moins le $\min R$ des deux rayons de convergence, et que la fonction somme est dans le disque $D(0, R)$ la somme des deux fonctions sommes des deux autres séries.

En appliquant les résultats de la partie 15.12.6 on montre facilement que deux séries entières $\sum a_n z^n$ et $\sum b_n z_n$ de rayon de convergence $\geq R$ on un produit convergeant sur le disque $D(0, R)$ (R toujours le $\min$ des deux rayons de convergence) et que la série produit sur $D(0, R)$ a une somme égale au produit des deux fonctions sommes obtenues pour les deux séries entières.

18.8 Développement en série entière

Définition 651 (Développement en série entière) *On suppose $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.*

Une application f d'un ouvert U de $\mathbb{K}$ dans $\mathbb{K}$ est dite **développable en série entière au voisinage de** *$a \in U$ s'il existe $\sum a_n.z_n$ de rayon de convergence $r > 0$ telle que $D(a, r) \subset U$ et $\forall z \in D(a, r)$ on ait $f(z) = \sum a_n.(z - a)^n$.*

f est dite **analytique** *sur $V \subset U$ avec V un ouvert de $\mathbb{K}$ si elle est développable en série entière au voisinage de chaque point de V.*

Théorème 652 *Soit $\sum a_n.z^n$ une série entière de rayon de convergence R. La somme f de cette série entière est analytique sur son disque ouvert de convergence.*
Le développement en série entière de f en a est donné sur le disque centré sur a et de rayon $R - |a|$ par

$$f(z) = \sum_n \frac{f^{(n)}(a)}{n!}(z-a)^n$$

Démonstration :

$$f^{(p)}(z) = \sum_n \frac{(n+p)!}{n!} a_{n+p} z^n$$

et donc

$$|f^{(p)}(z)| \leq \sum_n \frac{(n+p)!}{n!} |a_{n+p}| z^n$$

Pour $z \leq r < R$, on peut donc écrire

$$\sum_n |\frac{f^{(n)(z)}}{n!}|(r-|z|)^n \leq \sum_{n,p\geq 0} \frac{(n+p)!}{n!p!} |a_{n+p}|.|z|^p (r-|z|)^n$$

Montrons que la série de droite converge ; pour cela on limite la somme à n et p inférieurs à N et on fera tendre ensuite N vers $+\infty$

$$\sum_{0\leq n,p \leq N} C_{n+p}^n |a_{n+p}|z|^p (r-|z|)^n$$

$$\leq \sum_n |a_n| (\sum_{p=0}^n C_n^p |z|^p (r-|z|)^{n-p})$$

$$\leq \sum_n |a_n| r^n < \infty$$

La série étant absolument convergente, on peut permuter les termes comme on le souhaite, et donc

$$\sum_n \frac{f^n(z)}{n!}(z'-z)^n$$

$$= \sum_{n,p\geq 0} C_{n+p}^p a_{n+p} z^p (z'-z)^n$$

$$= \sum_n a_n (\sum_{p=0}^n C_n^p z^p (z'-z)^n)$$

$$= \sum_n a_n z^n$$

D'où le résultat.□

Corollaire 653 *L'ensemble des points d'analycité d'une application est ouvert.*

Démonstration : Evident au vu du résultat ci-dessus.

18.9 Zoologie des séries entières

18.9.1 L'exponentielle complexe

Définition 654 *On appelle* **exponentielle** *l'application qui à* $z \in \mathbb{C}$ *associe* $\sum_{n\in\mathbb{N}} \frac{z^n}{n!}$. *On la note* $x \mapsto e^x$ *ou* $x \mapsto exp(x)$.

NB : Il est à noter que cette série entière est bien définie partout, par le critère de D'Alembert. La convergence de la série est donc uniforme sur tout compact, et exp est donc holomorphe et entière [1]. La dérivation terme à terme, légitimée par la convergence uniforme des dérivées (voir théorème 381), montre que la dérivée de exp est exp.

D'autres propriétés de l'exponentielle sont fondamentales ; on les trouvera dans le formulaire, avec les schémas de preuve, en partie 38.9.

[1] Une fonction entière est une fonction holomorphe sur tout le plan.

18.10 Séries formelles et série génératrice

Définition 655 *Etant donné A un anneau, on note $A[[X]]$ et on appelle* **ensemble des séries formelles sur** *A l'ensemble des suites à valeurs dans A. Une telle suite $(a_n)_{n\in\mathbb{N}}$ sera notée*

$$\sum_{n\in\mathbb{N}} a_n X^n$$

On dit parfois aussi que $\sum a_n X^n$ est la **série génératrice** *associée à la suite $(a_n)_{n\in\mathbb{N}}$.*
On munit $A[[X]]$ d'une structure d'anneau en définissant un produit et une somme par

$$(\sum_{n=0}^{\infty} a_n X^n) + (\sum_{n=0}^{\infty} b_n X^n) = (\sum_{n=0}^{\infty} (a_n + b_n) X^n)$$

$$(\sum_{n=0}^{\infty} a_n X^n) \times (\sum_{n=0}^{\infty} b_n X^n) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} (a_k b_{n-k}) X^n$$

Etant donnée V une variable aléatoire à valeurs dans $\mathbb{N}$, on appelle aussi **série génératrice** *associée à V la série $\sum P(V=n)X^n$.*

On trouvera une application des séries génératrices aux probabilités dans la partie 36.12.9. Le livre [10] donne aussi une application aux nombres de Catalan (ie c'est un dénombrement basé sur un produit de séries formelles). La proposition 1296 est un autre exemple.

Chapitre 19

Fonctions holomorphes

Un ouvrage de référence est [16], dont nous suivons ici la démarche.

19.1 Cadre

On va ici se préoccuper de fonction de Ω dans $\mathbb{C}$, avec Ω un ouvert de $\mathbb{C}$.

19.2 Généralités

Définition 656 *Une fonction est dite* **dérivable au sens complexe en** $a \in \Omega$ **si** $lim_{x \to a} \frac{f(x)-f(a)}{x-a}$ *existe et est finie. On note alors cette limite* $f'(a)$.
Une fonction est dite **holomorphe sur** Ω *si elle est dérivable au sens complexe en tout point de* Ω. *On note* $H(\Omega)$ *l'ensemble des fonctions holomorphes sur* Ω.
On notera $D(a,r)$ *(resp.* $D'(a,r)$*) avec* $r > 0$ *l'ensemble des* x *de* Ω *tels que* $|x-a| < r$ *(resp.* $0 < |x-a| < r$*).*
Un domaine est un ouvert connexe non vide.

On remarque immédiatement que :

- $H(\Omega)$ est un anneau pour l'addition et la multiplication usuelles.

- la composée de deux fonctions holomorphes est holomorphe.

- tout polynôme est holomorphe sur $\mathbb{C}$

- l'inverse d'une fonction holomorphe ne s'annulant pas est holomorphe

- l'exponentielle complexe est holomorphe sur $\mathbb{C}$

- toute série entière est holomorphe ; si $f(z) = \sum_{n=0}^{\infty} c_n.(z-a)^n$, alors $f'(z) = \sum_{n=1}^{\infty} n.c_n.(z-a)^{n-1}$

19.3 Vers le théorème de Cauchy

Proposition 657 *Soit μ une mesure complexe sur un espace mesurable X, ϕ une fonction complexe mesurable, Ω un ouvert du plan qui ne rencontre pas $\phi(X)$. Alors avec*

$$f(z) = \int_X \frac{d\mu(t)}{\phi(t) - z}$$

définie pour $z \in \Omega$ On a $f \in H(\Omega)$.

Démonstration : • On développe en série entière $1/(\phi(t) - z)$ sur un disque suffisamment petit pour être inclus dans Ω et pour que la convergence soit uniforme.

• On permute alors l'intégrale et la somme (merci la convergence uniforme), et on obtient bien le résultat désiré.$\square$

Définition 658 *On appelle* **courbe** *une application continue d'un intervalle $[a, b]$ de $\mathbb{R}$ dans $\mathbb{C}$.*

On appelle **chemin** *une application continue C^1 par morceaux d'un intervalle $[a, b]$ de $\mathbb{R}$ dans $\mathbb{C}$.*

Une courbe ou un chemin γ est dit **fermé** *si $\gamma(a) = \gamma(b)$.*

Etant donné γ une courbe, on note γ^* l'image de $[a, b]$ par γ.

Etant donné γ un chemin et f une fonction continue sur γ^*, on note $\int_\gamma f(z).dz$ l'intégrale $\int_{[a,b]} f(t).\gamma'(t).dt$, on appelle cette intégrale l'intégrale de f le long de γ.

Deux chemins γ et γ' sont dits équivalents, si pour toute fonction f continue sur γ^* et γ'^* l'intégrale de f le long de γ est égale à l'intégrale de f le long de γ'.

La **longueur** *d'un chemin γ est l'intégrale de la fonction constante égale à 1 le long de cet arc.*

On appelle **indice** *de z pour $z \in \Omega$ par rapport à γ, avec Ω le complémentaire de γ^*, le complexe*

$$Ind_\gamma(z) = \frac{1}{2.i.\Pi} \int_\gamma \frac{dt}{t - z}$$

si ϕ est une bijection C^1 de $[a, b]$ dans $[c, d]$, si γ est un chemin d'intervalle de définition $[a, b]$ et γ' est un chemin d'intervalle de définition $[c, d]$, alors si $\gamma = \gamma' \circ \phi$ alors l'intégrale le long de γ est la même que l'intégrale le long de γ' ; c'est à dire qu'un reparamétrage C^1 transforme un chemin en un chemin équivalent.

Théorème 659 (Indice) *L'indice de z par rapport à γ est entier, constant sur chaque composante connexe de Ω (le complémentaire de γ^*), et nul sur la seule composante connexe de Ω qui ne soit pas bornée.*

Démonstration : • Pour voir qu'il y a une seule composante connexe non bornée, c'est facile, il suffit de voir que γ^* est inclus dans un disque ; le complémentaire de ce disque est connexe et donc inclus dans une composante connexe.

• Pour voir que l'indice est un entier, on suppose l'arc défini sur $[0, 1]$, on considère la fonction qui à t associe $exp(\int_0^t \frac{\gamma'(u)}{\gamma(u)-z}.du)$. En dérivant cette fonction, on obtient qu'elle est proportionnelle à $\gamma(t) - z$.

• On déduit de ça que notre fonction prend la valeur 1 en 1, ce qui est pile poil ce qu'il nous fallait pour que notre fonction soit un multiple de $2i\Pi$.

• L'indice est constant sur chaque composante connexe, puisqu'il est continu et que chaque composante connexe a donc une image connexe.

• $|\frac{1}{2i\Pi} \int_0^1 \frac{\gamma'(s).ds}{\gamma(s)-z}| \leq |\frac{M}{2\Pi} \int_0^1 \gamma'(s).ds|$, avec M un majorant de $|\frac{1}{\gamma(s)-z}|$. M tendant vers 0 pour z tendant vers l'infini, l'indice est de module inférieur à 1 pour z assez grand, et donc il est nul sur la composante connexe non bornée.$\square$

Quelques remarques :

- On montre facilement que l'indice d'un point z par rapport au chemin $[0, 1] \to \mathbb{C}$, $t \mapsto e^{2i\Pi.t}$, est 1 si $|z| < 1$ et 0 sinon.

- On montre facilement que l'intégrale de la dérivée d'une fonction holomorphe le long d'une chemin fermé est nulle. Par suite, l'intégrale d'un polynôme le long d'un chemin fermé est nulle.

Lemme 660 (Théorème de Cauchy dans le cas d'un triangle dans un convexe)
Soit un triangle de sommets a, b et c. L'intégrale le long de ce triangle est en fait l'intégrale suivant $[a, b]$, plus l'intégrale suivant $[b, c]$, plus l'intégrale suivant $[c, a]$.

On suppose Ω convexe.

Alors soit f une fonction continue sur Ω, et holomorphe sur $\Omega \setminus \{p\}$, avec $p \in \Omega$.

Alors l'intégrale de f le long du triangle est nulle.

(On montrerait facilement le même résultat pour un carré où n'importe quel autre polygone, en le triangulant)

Démonstration :
Il faut distinguer trois cas

• p n'est sur aucun des trois côtés du triangle. Alors on coupe le triangle en quatre plus petits triangles, comme sur la figure 19.1 (schéma de gauche), et on constate que l'intégrale sur au moins l'un des triangles doit être de valeur absolue le quart de la valeur absolue de l'intégrale le long du gros triangle ; or la longueur du petit triangle est la moitié de la longueur du gros. On construit ainsi par récurrence une suite de triangles $\mathbb{D}_n$ de longueur $L.2^{-n}$. On considère le point x, intersection des triangles.

Soit ϵ un réel > 0. f est dérivable en x. Il existe donc un triangle $\mathbb{D}_n$ tel que pour z dans $\mathbb{D}_n$, $f(z) - f(x) - f'(x).(z - x)$ est de module inférieur à $\epsilon.|z - x|$. Or l'intégrale d'une fonction polynôme sur un chemin fermé est nulle, puisqu'un polynôme est la dérivée d'un autre polynôme.

On obtient ainsi que l'intégrale le long du petit triangle est majorée par $\epsilon.(2^{-n}.L)^2$; et puisque l'intégrale sur le grand triangle initial est majorée par 4^n fois le module de l'intégrale sur le triangle $\mathbb{D}^n$, on en déduit que cette intégrale est nulle.

• On suppose maintenant que p soit égal à a (b ou c se traitent bien sûr de même). Alors on place x et y comme sur la figure 19.1 (schéma de droite) ; l'intégrale de f sur les triangles xyb et ybc est nulle ; et celle sur axy peut être rendue aussi petite qu'on le souhaite, puisqu'il suffit de faire tendre x et y vers a (rappelons que f est continue, sur un compact, donc bornée).

• Supposons maintenant que p soit un point de $]a, b[$; il suffit alors de raisonner sur abp, bcp et cap pour conclure.□

Ceux qui ont un peu d'avance auront constaté que l'hypothèse implique en fait que f soit holomorphe sur tout Ω ; mais nous avons besoin de notre lemme avec ces hypothèses-ci afin d'arriver à prouver les résultats qui impliqueront ces résultats.

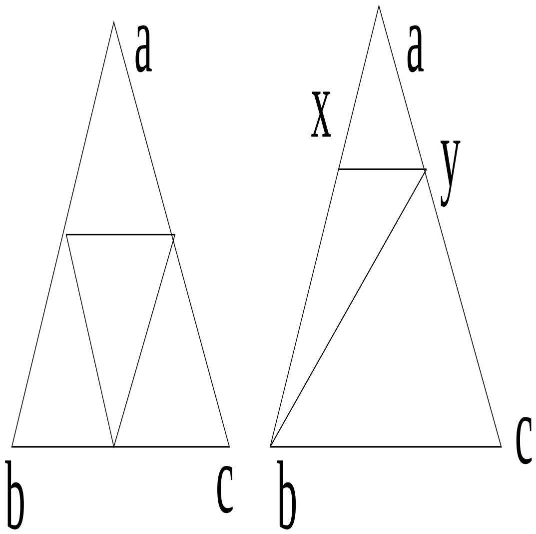


FIG. 19.1 – Découpage en petits triangles ; on utilise les milieux des côtés.

On passe maintenant à une version un peu plus forte :

Théorème 661 (Théorème de Cauchy dans un ensemble convexe) *On suppose Ω ouvert et convexe, p dans Ω, f continue sur Ω et $f \in H(\Omega \setminus \{p\}$. Alors l'intégrale de f le long de γ est nulle pour tout chemin γ fermé tel que $\gamma^* \subset \Omega$.*

Démonstration : On fixe un point a de Ω, et on définit $F(z)$ pour z dans $Omega$ comme l'intégrale sur $[a, z]$ de f.

On raisonne alors sur des triangles a, z, x pour considérer la limite de $\frac{F(z)-F(x)}{z-x}$ pour x tendant vers z. On montre facilement que cette limite est f, et donc que f est une dérivée et est continue, et donc le résultat est clair.$\square$

Théorème 662 (Formule de Cauchy dans un ensemble convexe) *On se donne γ un chemin fermé dans un ouvert convexe Ω, et f holomorphe sur Ω. Si $z \in \Omega$ et $z \notin \gamma^*$ alors*

$$f(z).Ind_\gamma(z) = \frac{1}{2i\Pi} \int_\gamma \frac{f(u)}{u-z}.du$$

Démonstration :
On se donne z vérifiant les hypothèses. On définit alors g par $g(u) = \frac{f(u)-f(z)}{u-z}$ si

$u \in \Omega$ et $u \neq z$, et $g(z) = f'(z)$.

La fonction g est continue, et holomorphe en tout point de $\Omega \setminus \{z\}$, donc d'après le théorème 661, on a $\int_\gamma g(u).du = 0$. En coupant g en ses deux termes $\frac{f(u)}{u-z}$ et $\frac{f(z)}{u-z}$, on obtient le résultat désiré.$\square$

On arrive maintenant à un résultat fondamental d'analyse complexe, facilement démontrable grâce aux résultats qui précèdent.

Théorème 663 (Développement en série entière des fonctions holomorphes) *Toute fonction holomorphe est développable en série entière.*

Démonstration : On se donne a dans Ω, et un disque suffisamment réduit $D(a, r)$ centré en a pour être inclus dans Ω.

Alors on applique la formule de Cauchy (théorème 662) à f sur le convexe $D(a, r)$, avec pour γ l'application de $[0, 1]$ dans $\mathbb{C}$ définie par $t \mapsto a + re^{2i\Pi.t}$.

On obtient une expression de $f(z)$ qui permet d'appliquer la proposition 657, et on a fini...$\square$

Remarquons qu'une fonction holomorphe est développable en série entière, donc sa dérivée est développable en série entière, donc sa dérivée est holomorphe. La dérivée d'une fonction holomorphe est donc une fonction holomorphe. En fait une fonction $\mathbb{C}$-dérivable (i.e. dérivable au sens complexe) est C^∞.

Enfin un théorème qui peut servir et qui ne coûte pas cher à montrer maintenant qu'on en est là :

Théorème 664 (Théorème de Morera) *Soit f une fonction continue complexe dans un ouvert Ω dont l'intégrale sur tout triangle est nulle. Alors f est holomorphe sur Ω.*

Démonstration : • On considère un disque ouvert D inclus dans Ω centré sur a.

• On construit une fonction F sur D dont f est la dérivée, par $F(z) = \int_{[a,z]} f(u).du$ (un disque, c'est convexe...).

• F est holomorphe, donc sa dérivée f est holomorphe.

• Puisque tout cela est valable pour n'importe quel disque inclus dans Ω, f est holomorphe sur Ω.$\square$

Voir le théorème 674.

Maintenant on va voir plein de conséquences de ces fort jolis théorèmes.

Théorème 665 *Soit f une fonction holomorphe sur Ω un ouvert connexe. Soit Z l'ensemble des z tels que $f(z) = 0$. Alors soit Z est égal à Ω, soit Z n'a pas de point d'accumulation dans $\mathbb{C}$.*
Si Z n'est pas $\mathbb{C}$ alors peut pour tout a dans Z trouver un entier positif unique m tel que $f(z) = (z-a)^m.g(z)$, avec g holomorphe non nulle en a. L'ensemble des zéros de f est dans ce cas au plus dénombrable.

On verra une jolie application avec le théorème de Runge 1210.

Démonstration : Soit Z' l'ensemble des points d'accumulation de Z. $Z' \subset Z$, car f étant continue, Z est fermé.

On considère le développement en série entière de f sur un disque D centré sur a quelconque dans $Z(f)$;

$$f(z) = \sum_{n \in \mathbb{N}} c_n.(z-a)^n$$

pour tout $z \in D$.

Si les c_n ne sont pas tous nuls, on considère le plus petit entier m tel que $c_m \neq 0$. On sait alors que g, définie par $g(z) = \frac{f(z)}{(z-a)^m}$ si $z \neq a$ et $g(a) = c_m$, vérifie les conditions demandées. Par continuité de g, on peut alors déduire que a est un point isolé de Z, puisque g est non nul sur un voisinage de a.

Le fait que Z ne contienne aucun point d'accumulation implique que Z contient un nombre fini de points sur toute boule de rayon n, et donc que Z est au plus dénombrable.

De tout ça on déduit que si a est dans Z', alors il y a un disque autour de a qui est aussi dans Z'. Donc Z' est ouvert, puisqu'il contient un disque centré sur a pour tout a dans Z'. Mais il est aussi fermé, puisqu'il est un ensemble de points d'accumulations. Donc s'il n'est pas vide et que l'on travaille dans un connexe, Z' est égal à Ω.$\square$

On remarque au passage que deux fonctions holomorphes égales sur un ensemble ayant un point d'accumulation sont donc nécéssairement égales (leur différence est holomorphe et nulle sur un ensemble ayant un point d'accumulation). Ce résultat est connu sous le nom de **principe de prolongement analytique**.

Proposition 666 (Principe du prolongement analytique) *Si deux fonctions holomorphes sont égales sur un ensemble ayant un point d'accumulation alors elles sont égales partout où elles sont définies.*

Définition 667 *On appelle m l'***ordre** *du zéro de f en a.*

Si f est holomorphe sur un ouvert Ω privé d'un point a, et n'est pas holomorphe en a, on dit que f admet une **singularité isolée** *en a.*

La singularité est dite **artificielle** *si en changeant $f(a)$ on peut rendre f holomorphe en a.*

Théorème 668 *Si f admet une singularité isolée en a et est bornée sur un voisinage de a, alors la singularité est artificielle.*

Démonstration : • On définit h par $h = (z \mapsto (z-a)^2.f(z))$, et $h(a) = 0$.

• h est holomorphe, on la développe en série entière, $h(z) = \sum_{n \geq 2} c_n.(z-a)^n$ (h est nulle et de dérivée nulle en a, puisque f est bornée sur un certain voisinage de a).

• Il ne reste alors plus qu'à poser $f(a) = c_2$.$\square$

On peut faire encore plus fort :

Théorème 669 *Soit $f \in H(\Omega \setminus \{a\})$, alors l'un des cas suivants se produit :*

- f admet une singularité artificielle en a ou pas de singularité du tout

- il existe des c_i en nombre fini tels que $z \mapsto f(z) - \sum \frac{c_i}{(z-a)^i}$ admette une singularité artificielle en a.

- L'image de tout voisinage de a par f est dense dans $\mathbb{C}$

Démonstration :
• Supposons qu'on ne soit pas dans le troisième cas, et choisissons z tel que z n'appartienne pas à l'adhérence de $f(D'(a,r))$ avec $r > 0$.

• Définissons $g(z) = \frac{1}{f(z)-w}$

• g est holomorphe sur $D'(a,r)$, et est bornée dans un voisinage de a ; donc g est prolongeable en une fonction holomorphe sur $D(a,r)$.

• Si $g(a) \neq 0$, alors f est prolongeable en une fonction holomorphe, et on n'en parle plus, c'est le premier cas.

• Sinon, alors on considère m l'ordre du zéro de g en a, et on développe en série entière $z \mapsto \frac{(z-a)^m}{g(z)}$.

• La flemme de finir, mais c'est pas très dur à partir de là...□

Définition 670 *Dans le deuxième cas,* $\sum_{i=1}^{m} \frac{c_i}{(z-a)^i}$ *est appelé* **partie principale du pôle de** f **en** a *;* m *est appelé l'ordre du pôle en* a*.*

c_1 *est appelé* **résidu** *de* f *en* a *; on le note* $Res(f; a)$*.*

Dans le troisième cas, on dit que f *a une* **singularité essentielle** *en* a*.*

Dans le premier cas, on dit que f *a une* **singularité artificielle** *en* a*.*

Théorème 671 *On se donne* f *une série entière,* $f(z) = \sum_{n\in\mathbb{N}} c_n.(z-a)^n$*, pour* $|z| < R$*. Alors pour tout* r *tel que* $0 < r < R$ *on a*

$$\sum_{\mathbb{N}} |c_n|^2.r^{2n} = \frac{1}{2\Pi}\int_0^{2\Pi} |f(a+re^{i\theta})|^2.d\theta$$

<u>Démonstration :</u> Considérer la formule de Parseval (voir théorème 1108), avec la base des $\theta \mapsto e^{-in\theta}$.□

Quelques corollaires pas trop difficiles :

Corollaire 672 (Théorème de Liouville) *Une fonction holomorphe sur* $\mathbb{C}$ *tout entier (On dit que cette fonction est* **entière** *) est soit constante soit non bornée.*

- si f holomorphe n'est pas constante sur un domaine (i.e. ouvert connexe) Ω, alors tout voisinage de a contient un point b tel que $|f(b)| > |f(a)|$.

Autre corollaire :

Théorème 673 (Estimations de Cauchy) f *holomorphe sur un disque ouvert* D *de rayon* R*,* $|f|$ *bornée par* M *sur* D*, alors* $|f^{(n)}(a)| \leq \frac{n!\,M}{R^n}$ *pour tout* $n \geq 0$*.*

Ceci servira pour le théorème 674 et pour le théorème 724.

Passons maintenant à des propriétés de passage à la limite :

Théorème 674 *Soit f_n une suite de fonctions holomorphes sur Ω tendant vers f uniformément sur tout compact de Ω. Alors f est holomorphe, et les f'_n convergent uniformément sur tout compact vers f'.*

Démonstration :
• f est continue comme limite uniforme sur tout compact d'une suite de fonctions continues.

• Pour le caractère holomorphe de f, on regarde ce qu'il se passe sur des diques ouverts (Ω étant réunion de tels disques) et il suffit ensuite de considérer l'intégrale de f sur le contour d'un triangle inclus dans un disque (un tel disque étant convexe) ; l'intégrale d'une limite uniforme étant la limite de l'intégrale, on déduit que l'intégrale de f sur tout triangle est nulle. Le théorème de Morera (voir théorème 664) permet de conclure.

• On utilise ensuite le théorème 673 pour voir que $|f'(z) - f'_n(z)| \leq \frac{1}{r}\|f - f_n\|_K$, avec K un compact ; d'où la convergence uniforme des dérivées, et le résultat désiré.□

Théorème 675 *On suppose Ω convexe, $a_1, ..., a_n$ des points distincts de Ω, et f holomorphe sur $\Omega \setminus \{a_1, ..., a_n\}$. On suppose que f admet en pôle en chaque a_i, et on se donne un chemin fermé γ ne passant pas par les a_i. Alors*

$$\frac{1}{2i\Pi}\int_\gamma f(z).dz = \sum_{k=1}^{n} Res(f; a_k).Ind_\gamma(a_k)$$

Démonstration : On applique le théorème de Cauchy à la fonction f moins ses parties principales en les a_i ; l'intégrale de cette fonction est donc nulle. Il ne reste alors qu'à considérer l'intégrale des parties principales, ce qui est facile au vu de résultats antérieurs (voir le théorème 659, et le fait que x^n pour $n \neq -1$ a une primitive holomorphe sur $\mathbb{C} \setminus \{0\}$).

Théorème 676 *• Si f est holomorphe et admet un zéro d'ordre m en a, alors le résidu de f'/f en a est m.*

• Si f est holomorphe sur $\Omega \setminus \{a\}$, alors le résidu de f'/f en a est égal à $-m$.

Démonstration : Pas dur... Il suffit de réécrire la fonction soit en divisant par $(z-a)^m$ (premier •), soit en soustrayant la partie principale du pôle (second •)...□

Théorème 677 *Soit f une fonction holomorphe, et γ un chemin $\theta \mapsto a + re^{i\theta}$, avec $\overline{D}(a, r)$ inclus dans Ω.*

On définit $\Gamma = f \circ \gamma$. Soit w n'appartenant pas à Γ^*.

Alors le nombre de zéros de $f - w$ dans $D(a, r)$, comptés avec leurs ordres de multiplicité, est égal à l'indice de w par rapport à Γ.

Démonstration : Le nombre de zéros de $f - w$ dans $D(a, r)$ est égal à la somme des résidus de $f'/(f - w)$ dans $D(a, r)$, et cette somme est bien l'indice de w par rapport à Γ.$\square$

Théorème 678 (Théorème de l'image ouverte) *On se donne Ω un ouvert connexe, i.e. un domaine, et f holomorphe sur Ω. Alors si f n'est pas constante, et pour tout z_0 dans Ω, f induit sur un voisinage ouvert V de z_0 une application surjective de V sur un ouvert W, telle que pour tout w dans $W \setminus \{w_0 = f(z_0)\}$, il y ait exactement m points distincts $z \in V$ dont l'image par f est w, avec m l'ordre du zéro de $f - w_0$ en z_0.*

Démonstration :

• on considère un cercle orienté suffisamment petit autour de w_0 pour que le disque D de même centre et de même rayon ne comporte pas de zéro ni de $f - w_0$ ni de f' dedans, à part z_0 lui-même.

• on considère le contour de ce cercle suffisamment petit

• on considère l'image par f de ce contour, et la composante connexe W de w_0 dans le complémentaire de cette image (W est ouvert comme composante connexe d'un ouvert, le complémentaire de l'image d'un compact étant évidemment fermé puisque complémentaire d'un compact (rappelons que l'image d'un compact par une application continue est un compact)).

• on prend alors pour V l'intersection du disque ouvert D et de l'image réciproque de W.

• L'indice de w_0 par rapport à $\Gamma = f \circ \gamma$ est m, ainsi donc que l'indice de tout w dans V. D'où le résultat...$\square$

Remarquons un corollaire intéressant, qui donne son nom à ce théorème ; l'image de tout ouvert par une fonction holomorphe est un ouvert.

Il est clair au vu du théorème précédent que si l'on a $f'(z)$ non nul, avec f holomorphe, alors on a localement une bijection autour de z. On peut améliorer ce résultat ; la réciproque locale, est elle aussi holomorphe ; cela fait l'objet du théorème suivant.

Théorème 679 *Soit f holomorphe, f de dérivée non nulle en a alors on peut trouver un ouvert V contenant a tel que f induise une bijection de V sur $f(V)$; la réciproque de f est holomorphe sur $f(V)$.*

Démonstration : Tout ce qui reste à prouver est le caractère holomorphe de la réciproque g de f sur $f(V)$.

Pour cela on considère $\frac{g(z)-g(a)}{z-a}$, on utilise la continuité de g (qui découle du fait que f est une application ouverte, i.e. que l'image de tout ouvert par une fonction holomorphe est une fonction holomorphe), et le fait que $f'(a)$ est non nul, et tout ça coule de source...$\square$

On va maintenant montrer que l'on a le droit de modifier "un peu" une courbe sans changer l'indice d'un point par rapport à cette courbe.

Théorème 680 *Si γ_1 et γ_2 sont deux chemins d'intervalle de paramétrage $[0, 1]$ (ou autre chose...) et si pour tout $t \in [0, 1]$ on a $|\gamma_1(t) - \gamma_2(t)| \leq |\gamma_1(t)|$, alors $Ind_{\gamma_1}(0) = Ind_{\gamma_2}(0)$.*

Démonstration :

• On pose $\gamma = \gamma_2/\gamma_1$.

• On a alors $\frac{\gamma'}{\gamma} = \frac{\gamma_2'}{\gamma_2} - \frac{\gamma_1'}{\gamma_1}$, donc en intégrant sur $[0, 1]$ on déduit que la différence entre l'indice de 0 par rapport à γ_2 et l'indice de 0 par rapport à γ_1 est l'indice de 0 par rapport à γ.

• $|1 - \gamma(t)| < 1$; donc l'indice de 0 par rapport à γ est 0.$\square$

Corollaire 681 (Théorème de Rouché) *f et g holomorphes sur Ω, le disque fermé de centre a et de rayon r étant inclus dans Ω, et $|f(z) - g(z)| < |f(z)|$ sur le cercle de centre a et de rayon r. Alors f et g ont le même nombre de zéros sur le disque ouvert de centre a et de rayon r (en comptant leurs multiplicités).*

Démonstration : On considère $\gamma(t) = e^{2i\Pi t}$, et $\gamma_1 = f \circ \gamma$ et $\gamma_2 = g \circ \gamma$. On applique alors le théorème précédent...$\square$

Cela servira notamment pour montrer le théorème 682 (preuve d'ailleurs fort sympathique). Ainsi que le rappelle Rudin dans [16], on peut aussi utiliser ce résultat pour montrer que tout polynôme de degré n a n racines dans $\mathbb{C}$ (en montrant que tout polynôme de degré n a le même nombre de zéros que z^n, dans un disque de rayon suffisamment grand.

19.4 Topologie de $H(\Omega)$

Pour cela on consultera20.2.2.

19.5 Zoologie des applications holomorphes

19.5.1 Théorème de Montel

Théorème 682 (Théorème de Montel) *Soit Ω un ouvert connexe de $\mathbb{C}$, K compact de Ω, $M \geq 0$, $m > 0$, $k \in K$. Alors il existe un certain $NombreMaxDeZeros$ tels que le nombre de zéros de f dans K, pour f bornée [a] par M sur K et telle que $|f'(k)| \geq m$, est majoré par $NombreMaxDeZeros$.*

[a]En module.

Démonstration :

• On considère l'ensemble U_n des applications f holomorphes sur Ω telles que le nombre de zéros de f sur K soit $< n$ ($n \in \mathbb{N} \cup \{+\infty\}$; on montre que U_n est ouvert.

- Pour cela donnons-nous f dans U_n

- Soit $NombreZeros$ le nombre de zéros de f dans K. $NombreZeros$, par définition de U_n, est fini.

- Considérons les disques fermés $\overline{D}(z_i, \epsilon_i)$ inclus dans Ω, et tels que f ne s'annule pas sur le disque ouvert, sauf peut-être en z_i.

- Il est clair que les disques ouverts $D(z_i, \epsilon_i)$ recouvrent K.

- On en extrait un nombre fini. Les $D(z_i, \epsilon_i)$ pour $i \in I$, avec I fini, recouvrent donc K.

- On considère alors $LesCercles$ le compact constitué des cercles de rayon ϵ_i et de centre les z_i pour i dans I.

- On considère alors η l'inf de f sur $LesCercles$.

- On considère alors l'ensemble V des fonctions g telles que $sup_{LesCercles}|g - f|$ est inférieur strictement à η. Par le théorème de Rouché 681, les fonctions g dans V ont un nombre de 0 égal au nombre de zéros de f.

- Le résultat est ainsi prouvé.

• On en déduit donc que l'application qui à f associe son nombre de zéros sur K est semi-continue supérieurement.

• L'ensemble $\mathcal{F}$ des applications f bornées en module par M sur K et telles que $|f'(k)| \geq m$ étant compact, le nombre de zéros est borné, le maximum est atteint (car une application semi-continue supérieurement sur un compact atteint son maximum, voir proposition 172), et il n'est pas infini par la condition $|f'(k)| \geq m$.$\square$

19.5.2 Fonctions holomorphes majorées par un polynôme

Théorème 683 *Soit f une fonction holomorphe sur $\mathbb{C}$, et P un polynôme, avec $|f| \leq |P|$. Alors f est un polynôme, de degré $\leq$ au degré de P.*

Démonstration :
Il suffit d'utiliser le théorème 673, qui nous dit que

$$|f^{(n)}(0)| \leq n!M/R^n$$

avec M un majorant de $|f|$ sur le disque de centre 0 et de rayon R. Puisque $M \leq L + K \times R^p$ (par hypothèse), on en déduit :

$$|f^{(n)}(0)| \leq n!(L + K \times R^p)/R^n$$

et donc $f^{(n)} = 0$ pour $n > p$, d'où le résultat.$\square$

19.5.3 Fonctions holomorphes tendant vers l'infini en l'infini

Théorème 684 *Soit f une fonction holomorphe sur $\mathbb{C}$. On suppose*

$$lim_{|z| \to \infty} |f(z)| = \infty$$

Alors f est une fonction polynôme.

Démonstration :
• Soit Z l'ensemble des zéros de f.

• Au vu de l'hypothèse, pour z assez grand en module, $|f(z)| \geq 1$. Donc Z est inclus dans un compact K.

• Si Z est infini, alors Z possède un point d'accumulation dans K, et donc d'après le théorème 665 f est nulle. Ce cas étant résolu, on peut supposer Z fini.

• On considère alors $1/f$. C'est une fonction holomorphe sur le complémentaire de Z. Sur Z, en utilisant le théorème 665, on constate que $1/f$ admet des pôles, et non pas des singularités essentielles ; on peut donc lui soustraire une fraction rationnelle P/Q, afin que $1/f - P/Q$ soit holomorphe.

• $1/f - P/Q$ est majoré par un polynôme, donc d'après le théorème 683 c'est un polynôme.

• $1/f - P/Q = R$, avec R un polynôme, donc $f = (Q-R)/P$, fraction rationnelle.

• f étant holomorphe, f n'a pas de pôle, et donc se simplifie en un polynôme.□

19.5.4 Sphère de Riemann $\widehat{\mathbb{C}}$ - Fonctions holomorphes sur $\widehat{\mathbb{C}}$

Définition 685 (Sphère de Riemann) *Soit $\widehat{\mathbb{C}}$ l'espace topologique obtenu en rajoutant un point, noté ∞, à $\mathbb{C}$, une base de voisinages de ce point étant constituée des ensembles $U \cup \{\infty\}$, où U parcourt les complémentaires dans $\mathbb{C}$ des parties compactes (en somme, c'est le compactifié d'Alexandrov de $\mathbb{C}$). On note S^2 la sphère unité de l'espace $\mathbb{R}^3$.*

La projection stéréographique depuis le pôle nord N

$$p_N : \begin{array}{ccl} \mathcal{S}^2 & \to & \widehat{\mathbb{C}} \\ (x,y,t) & \mapsto & (x+iy)/(1-t) \text{ si } t \neq 1 \\ N & \mapsto & \infty \end{array} ,$$

est un homéomorphisme de la sphère $\mathcal{S}^2$ sur $\widehat{\mathcal{C}}$.

La droite projective complexe $\mathbb{P}^1(\mathbb{C})$ est le quotient de $\mathbb{C}^2 \setminus_{\{(0,0)\}}$ par l'action diagonale par multiplication du groupe multiplicatif des nombres complexes non nuls $\mathbb{C}^*$, munie de la topologie quotient. La classe d'un couple (z_1, z_2) de nombres complexes non tous les deux nuls est notée $(z_1 : z_2)$.

L'application

$$r : \begin{array}{ccl} \mathbb{P}^1(\mathbb{C}) & \to & \widehat{\mathbb{C}} \\ (z_1, z_2) & \mapsto & z_1/z_2 \text{ si } z_2 \neq 0 \\ (1:0) & \mapsto & \infty \end{array}$$

est un homéomorphisme de la droite projective $\mathbb{P}^1(\mathbb{C})$ sur $\widehat{\mathbb{C}}$.

Proposition 686 *Soit a, b, c, d quatre nombres complexes tels que $ad - bc \neq 0$. Alors l'endomorphisme linéaire de $\mathbb{C}^2$ défini par la matrice*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

passe au quotient en un homéomorphisme de $\mathbb{P}^1(\mathbb{C})$ qui, lu sur $\widehat{\mathbb{C}}$, est l'homographie

$$h_A : z \mapsto \frac{az + b}{cz + d},$$

prolongée par $h_A(-d/c) = \infty$ et $h_A(\infty) = a/c$ si $c \neq 0$, ces deux conditions étant remplacées par $h_A(\infty) = \infty$ si $c = 0$. On a ainsi défini un morphisme injectif de $PSL(2, \mathbb{C})$ dans Homéo($\widehat{\mathbb{C}}$).

La droite projective complexe hérite d'une structure de variété complexe de dimension (complexe) 1. La sphère de Riemann est $\widehat{\mathbb{C}}$ munie de la structure complexe héritée de celle de $\mathbb{P}^1(\mathbb{C})$ par le biais de r. Celle-ci peut être décrite par un jeu de deux cartes. Soit P_1 l'ouvert de $\mathbb{P}^1(\mathbb{C})$ correspondant aux couples ayant la première coordonnée non nulle et P_2 l'ouvert correspondant aux couples ayant la deuxième coordonnée non nulle. Les ouverts correspondants dans $\widehat{\mathbb{C}}$ sont respectivement $U_1 = \mathbb{C}$ et $U_2 = \widehat{\mathbb{C}} \setminus \{0\}$, dont l'intersection est $\mathbb{C}^*$.

La matrice

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

qui permute les coordonnées dans $\mathbb{P}^1(\mathbb{C})$, réalise un homéomorphisme (de changement de carte) entre P_1 et P_2, qui est holomorphe, car il se lit $z \mapsto 1/z$ dans $U_1 \cap U_2 = \mathbb{C}^*$.

Une application $f : \widehat{\mathbb{C}} \to \mathbb{C}$ est alors dite holomorphe (respectivement méromorphe, où f n'est alors supposée définie qu'en dehors de l'ensemble de ses pôles) si elle est holomorphe (respectivement méromorphe) dans les cartes, c'est-à-dire, d'une part, dans $\mathbb{C}$ au sens usuel et, d'autre part, dans U_2 au sens que $z \mapsto f(1/z)$ est holomorphe (respectivement méromorphe) dans $\mathbb{C}$ au sens usuel.

Une application $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ est alors holomorphe si et seulement si elle est méromorphe en dehors de l'image réciproque de ∞.

Proposition 687 *Les applications holomorphes de $\widehat{\mathbb{C}}$ dans $\mathbb{C}$ sont les fonctions constantes.*

<u>Démonstration :</u> Soit f une application holomorphe de $\widehat{\mathbb{C}}$ dans $\mathbb{C}$. Par continuité en ∞, la fonction f est bornée sur le complémentaire d'une partie compacte de $\mathbb{C}$.

Comme f est également bornée sur ce compact, f est bornée donc constante. □

Théorème 688 *Les applications holomorphes de $\widehat{\mathbb{C}}$ dans $\widehat{\mathbb{C}}$ sont les fractions rationnelles.*

Démonstration : Soit f une telle application. On suppose f non constante. Comme les fonctions holomorphes non constantes, f a ses zéros isolés, si bien que f n'admet qu'un nombre fini de pôles et de racines (car $\widehat{\mathbb{C}}$ est compact). Soit P une fraction rationnelle ayant pour racines (comptées avec multiplicité) les pôles de f et pour pôles les racines de f, de sorte que fP n'ait ni pôle, ni racine. La fonction fP définit alors une application holomorphe de $\widehat{\mathbb{C}}$ dans $\mathbb{C}$, qui est constante d'après la proposition précédente, si bien que f est une fraction rationnelle. □

Théorème 689 *Les automorphismes de la sphère de Riemann sont les homographies.*

Démonstration : Si $P = A/B$ est une fraction rationnelle et w un élément de $\widehat{\mathbb{C}}$, le nombre de solutions dans $\widehat{\mathbb{C}}$, comptées avec multiplicité, de l'équation $P(z) = w$ est le maximum des degrés de A et de B. Par conséquent P est bijective si et seulement si A ou B est de degré 1, l'autre étant de degré 0 ou 1. □

Proposition 690 *Les applications holomorphes de $\widehat{\mathbb{C}}$ dans $\widehat{\mathbb{C}}$ qui sont holomorphes dans $\mathbb{C}$ au sens usuel sont les polnômes.*

Démonstration : Soit f une telle application. Par le même raisonnement que précédemment, f est constante sauf si $f(\infty) = \infty$. Pla cons-nous donc dans ce dernier cas de figure. La fonction f admet alors un nombre fini (éventuellement nul) de zéros, car une infinité de zéros aurait un point d'accumulation dans $\widehat{\mathbb{C}}$, donc dans $\mathbb{C}$ puisque $f(\infty) = \infty$ et f serait alors nulle.

Soit P un polynôme ayant pour racines (comptées avec multiplicité) les zéros de f, de sorte que la fonction $g = f/P$ n'ait pas de racine dans $\mathbb{C}$. Si $g(\infty) = \infty$, alors la fonction $1/g$ définit une application holomorphe de $\widehat{\mathbb{C}}$ dans $\mathbb{C}$ qui s'annule uniquement en ∞, ce qu'interdit la proposition précédente. On a donc $g(\infty) = \lambda$ pour un certain nombre complexe λ. La proposition précdente montre alors que $g = \lambda$, si bien que $f = \lambda P$. □

Chapitre 20

Analyse fonctionnelle

En analyse fonctionnelle, une référence classique, détaillée et complète, est le livre [2].

20.1 Résultats fondamentaux

20.1.1 Hahn-Banach

⊡ Le théorème

Théorème 691 (Théorème de Hahn-Banach des $\mathbb{R}$-espace vectoriel) *Soit E un $\mathbb{R}$-espace vectoriel , et p une application de E dans $\mathbb{R}$ telle que :*
• $\forall (x, y) \in E, p(x + y) \leq p(x) + p(y)$
• $\forall x \in E, \forall \lambda \in \mathbb{R}^+, p(\lambda.x) = \lambda.p(x)$
Alors toute forme linéaire l sur F sous-espace vectoriel de E telle que $l(x) \leq p(x)$ peut être prolongée en une forme linéaire L sur E telle que $\forall x,\ L(x) \leq p(x)$.

NB : noter que p norme ou semi-norme convient.

voir la partie "applications" juste un peu plus bas.

Démonstration : Cette preuve fait intervenir le lemme de Zorn (voir lemme 36).

On considère l'ensemble I des formes linéaires f prolongeant l sur un certain sous-espace vectoriel $D(f)$ de E contenant F, et telle que $f \leq p$ pour tout x de $D(f)$.

On munit I de la relation d'ordre définie par

$$f_1 \leq f_2 \iff D(f_1) \subset D(f_2) \wedge \forall x \in D(f_1) f_1(x) = f_2(x)$$

I est inductif. En effet, si J est une partie de I totalement ordonnée, alors la fonction f définie par $D(f) = \cup_{g \in J} D(g)$ et $f(x) = g(x)$ si $g \in J$ et $x \in D(g)$ est un majorant de J.

Par le lemme de Zorn (lemme 36), on en déduit que I possède un élément maximal f.

On suppose maintenant que $D(f) \neq E$ (on va chercher à montrer que cette hypothèse est contradictoire). Alors on considère y n'appartenant pas à $D(f)$. On définit f' sur $D(f) + \mathbb{R}.y$ par $f'(x + t.y) = f(x) + \alpha.t$, α étant choisi tel que pour tout x dans $D(f)$ on ait $\alpha \leq p(x + y) - f(x)$ et $\alpha \geq f(x) - p(x - y)$; ce qui est possible car $f(x_1) + f(x_2) \leq p(x_1 + x_2) \leq p(x_1 + y) + p(x_2 - y)$.

D'où le résultat.□

⊡ Des applications

◇ Sur les formes linéaires

Corollaire 692 *Soit g une forme linéaire continue sur un sous-espace vectoriel d'un $\mathbb{R}$-espace vectoriel normé E. Alors il existe une forme linéaire continue f sur E prolongeant g et telle que $\|f\| = \|g\|$.*

Démonstration : Application directe du théorème de Hahn-Banach.□

(la norme sur l'espace dual, évoquée ici, est la norme usuelle, ici la norme de f, forme linéaire continue, est le sup des $\|f(x)\|$ pour x de norme 1)

Corollaire 693 *Soit x dans un $\mathbb{R}$-espace vectoriel normé E, alors il existe une forme linéaire continue f sur E telle que $\|f\| = \|x\|$ et $f(x) = \|x\|^2$.*

Démonstration : Il suffit de prolonger une application linéaire adéquate définie sur $\mathbb{R}.x$.□

Corollaire 694 *Pour tout x d'un $\mathbb{R}$-espace vectoriel normé E, on a $\|x\| = sup_{f \in E'/\|f\|_\infty = 1} \|f(x)\| = max_{f \in E'/\|f\|_\infty = 1} \|f(x)\|$.*

Démonstration : L'inégalité $\|x\| \geq \sup_{f \in E'/\|f\|_\infty \leq 1} \|f(x)\|$ est évidente. Choisissons alors f_0 donné par le corollaire précédent ($\|f_0\| = \|x\|$ et $f_0(x) = \|x\|^2$). On a

$$\|f\|_\infty = \frac{\|f_0\|_\infty}{\|x\|} = 1$$

$$\|f(x)\| = \frac{\|f_0(x)\|}{\|x\|} = \frac{\|x\|^2}{\|x\|} = \|x\|$$

D'où le résultat annoncé.□

Ce corollaire servira pour la partie 20.8.1.

◇ **En géométrie**

Théorème 695 (>Séparation des convexes 1) *Soient A et B des convexes non vides disjoints d'un $\mathbb{R}$-espace vectoriel normé E ; si A est ouvert, alors il existe une forme linéaire continue f sur E et un réel α tels que $f(x) \leq \alpha$ pour $x \in A$ et $f(x) \geq \alpha$ si $x \in B$.*

On notera que cela signifie précisément qu'il existe un hyperplan affine fermé (rappelons que l'image inverse d'un singleton par une forme linéaire non nulle est un hyperplan fermé si et seulement si cette forme linéaire est continue) séparant (au sens large) A et B.

On peut en fait étendre le résultat à $f(x) < \alpha$ (et non simplement $\leq$).

Démonstration : On va avoir besoin de deux lemmes.

Lemme 696 *On se donne U un ouvert convexe contenant 0, et on définit μ_U* **la jauge** *associée à U, c'est-à-dire que $\mu_U(x)$ est l'inf des réels $t > 0$ tels que $t^{-1}.x \in U$.*

Alors il existe un certain réel M tel que

$$\forall x \ \mu_U(x) \leq M.\|x\|$$

$$U = \{x/\mu_U(x) < 1\}$$

$$\forall (x, \lambda) \in E \times \mathbb{R}^+ \ \mu_U(\lambda.x) = \lambda.\mu_U(x)$$

$$\forall (x, y) \in E \ \mu_U(x + y) \leq \mu_U(x) + \mu_U(y)$$

(les deux dernières conditions permettent d'utiliser le théorème de Hahn-Banach)

Démonstration : Pas très très dur (en le faisant dans cet ordre)...□

Lemme 697 *Soit U un convexe non vide et y n'appartenant pas à U. Alors il existe une forme linéaire continue f sur E avec $f(x) < f(y)$ pour tout x dans E.*

Démonstration :

• On montre le résultat dans le cas où 0 appartient à U, et on généralise par une simple translation

• En supposant donc que $0 \in U$, on considère la jauge μ_U (définie comme précédemment).

• On définit la forme linéaire g sur $\mathbb{R}.y$ par $g(t.y) = t$.

• Il est clair que $g(x) \leq \mu_U(x)$

• On peut donc prolonger g à E tout entier ; appelons f la forme linéaire obtenue avec $f \leq \mu_U$.

• f est continue de par le lemme 696 (1ère propriété), et f vérifie les hypothèses demandées (deuxième propriété du lemme 696).$\square$

On peut maintenant en revenir à la démonstration du théorème, toujours non démontré.

• On note $U = \{x - y/(x, y) \in A \times B\}$.

• U est convexe

• U est ouvert (car A l'est)

• U ne contient pas 0

• On considère la fonction f donnée par le lemme 697 avec $y = 0$, c'est-à-dire négative sur tout U.

• Le fait que f soit négative sur tout U se traduit exactement par le fait que pour tout $(x, y) \in A \times B$ on ait $f(x) < f(y)$. On considère alors α le sup des $f(x)$ pour x dans A, et le résultat est démontré.

L'extension ($<$ au lieu de $\leq$) se montre comme suit :

- supposons qu'il existe x_0 tel que $f(x_0) = \alpha$.

- A ouvert implique qu'il existe ϵ tel que $B(x_0, \epsilon)$ soit incluse dans A. (1)

- f non nulle implique qu'il existe g dans E de norme 1 tel que $f(g) > 0$. (2)

- (1) implique que $x'_0 = x_0 + \frac{\epsilon}{2} g \in A$.

- (2) implique que $f(x'_0) > \alpha$ ce qui est absurde !$\square$

Théorème 698 (Séparation des convexes 2) *Soient A et B deux convexes disjoints de E (toujours un espace vectoriel normé), non vides. On suppose A fermé et B compact ; alors il existe une forme linéaire continue $f \neq 0$ avec $f(A) \leq c_1$ et $f(B) \geq c_2$ avec $c_1 < c_2$.*

Cela signifie exactement que A et B sont séparés par un hyperplan fermé (puisque image inverse d'un singleton par une forme linéaire continue non nulle) au sens strict.

On montrera en utilisant ce théorème que la topologie faible est séparée ; voir le corollaire 732.

Démonstration : • On se donne ϵ positif.

• On note A_ϵ le ϵ-voisinage de A (ie la réunion des boules ouvertes de rayon ϵ de centre dans A), et B_ϵ le ϵ-voisinage de B.

• On remarque que A_ϵ et B_ϵ sont ouverts (comme tous les ϵ voisinages)

• Pour ϵ assez petit, A_ϵ et B_ϵ sont disjoints

• D'après le théorème précédent, on peut séparer A_ϵ et B_ϵ au sens large par un hyperplan fermé.

• On a $f \leq \alpha$ sur A et $f > \alpha$ sur B compact donc $f \geq \beta > \alpha$ sur B.□

On retiendra donc que l'on peut séparer dans un espace vectoriel normé par un hyperplan fermé :

• au sens large, deux convexes disjoints dont l'un (au moins) est ouvert

• au sens strict, deux convexes disjoints dont l'un est fermé et l'autre compact.

◇ **En topologie**

Corollaire 699 *Soit F un sous-espace vectoriel de E (qui est toujours un espace vectoriel normé), qui n'est pas dense dans E. Alors il existe une forme linéaire continue sur E, non nulle, qui est nulle sur F.*

Démonstration : • On se donne x qui n'est pas dans l'adhérence de F.

• $\{x\}$ est compact.

• On peut séparer F et $\{x\}$ au sens strict ; soit f la forme linéaire correspondante. On suppose que $f(F) < K < f(x)$

• $f(F) < K$ implique $f(F) = 0$, puisque F est un espace vectoriel .□

Corollaire 700 *Si F est un sous-espace vectoriel de E et si toute forme linéaire continue sur F est nulle sur E, alors F est dense dans E.*

Le théorème de Runge 1210 sera démontré grâce à ce corollaire.

Démonstration : C'est une reformulation du corollaire précédent.□

20.1.2 Le théorème de Baire et ses conséquences

Théorème 701 (Théorème de Baire) *Soit X un espace topologique. Si X est localement compact, ou s'il est métrique complet, alors*
• Toute intersection dénombrable d'ouverts denses est dense
• Une réunion dénombrable de fermés recouvrant X comporte un fermé d'intérieur non vide

Comme le signale le livre [2], on peut en fait énoncer plus précisément que l'intérieur de la réunion d'une suite de fermés d'intérieurs vides est vide.

Démonstration : Ce théorème ayant été prouvé (voir théorème 249) je ne fais que le rappeler ici. Rappelons juste que les deux • sont équivalents (considérer les complémentaires des fermés du deuxième •)□

Notons que le théorème de Baire est en particulier valable pour les espaces de Banach.

Théorème 702 (Théorème de Banach-Steinhaus) *Ce théorème est dit aussi théorème de la* **borne uniforme**.

On se donne E et F des espaces de Banach, et $(T_i)_{i\in I}$ une famille d'applications linéaires continues de E dans F.

Si pour tout x, $\exists M / \forall i \in I / \|T_i(x)\| < M\|x\|$.

Alors $\exists M / \forall i \in I / \|T_i\| < M$.

Ce théorème est plus intuitif sous son petit nom de "théorème de la borne uniforme". L'hypothèse est que l'on a une famille d'applications bornées sur chaque point ; la conclusion est que l'on peut les borner uniformément (bien vérifier que l'on a des Banach).

Notez bien que la famille des T_i n'est pas nécessairement dénombrable !

Démonstration : La aussi je ne donne pas de preuve, puisqu'elle se trouve au théorème 252.$\square$

On verra une application à la transformation de Toeplitz (proposition 610), qui fournit une preuve élégante de la moyenne de Césaro (corollaire 611).

Corollaire 703 *Soient E et F deux Banach, et T_n une suite d'applications linéaires continues de E dans F, avec $T_n(x)$ convergeant pour tout x - on note par la suite $T(x)$ sa limite.*

Alors $\|T_n\|$ est borné, T est linéaire continue, et $\|T\| \leq liminf \|T_n\|$.

Démonstration : Application directe du théorème de Banach-Steinhaus.$\square$

Corollaire 704 *Soit E un espace vectoriel normé et X un sous-ensemble de E.*

On suppose que pour tout f appartenant à E' l'ensemble $f(X)$ est borné.

Alors X est borné.

Démonstration : On applique Banach-Steinhaus dans E', avec pour famille d'applications linéaires les applications qui à $f \in E'$ associe $f(x)$, pour $x \in X$.

Il faut bien noter que le dual d'un espace vectoriel normé est un Banach, et que ce résultat est nécessaire à cette preuve (voir corollaire 241).

Noter aussi que ce résultat exprime que "faiblement borné" implique "fortement borné".

Cette façon de voir est d'ailleurs une belle illustration de la notion de "borne uniforme". Si une partie est bornée suivant "toutes les directions" (traduire : suivant toute forme linéaire), alors elle est bornée " tout court "...$\square$

Théorème 705 (Théorème de l'application ouverte) *Soient E et F des espaces de Banach, et T une application linéaire continue surjective de E dans F. Alors T est ouverte (c'est à dire que l'image de tout ouvert par T est un ouvert.*

Démonstration : voir le théorème 255.□

Bien entendu, dans le cas où T est bijective, on en déduit le théorème d'isomorphisme de Banach 256, qui stipule qu'une bijection linéaire continue est de réciproque continue (et donc est un homéomorphisme).

Il faut noter un corollaire important : si un espace vectoriel E muni de la norme N_1 est un espace de Banach, et si E muni de la norme N_2 est aussi un espace de Banach, alors si N_1 est plus fine que N_2, alors en fait N_2 est équivalente à N_1.

Théorème 706 (Théorème du graphe fermé) *Soit $T : E \to F$, linéaire entre les Banach E et F. L'application T est continue si et seulement si le graphe de T est fermé dans $E \times F$.*

Démonstration : Voir le théorème 258.□

Voir le théorème 736.

20.1.3 Autres définitions et propriétés indispensables

Il est indispensable de connaître la topologie faible, la topologie quotient, la topologie produit, la topologie forte, pour la suite. On travaillera exclusivement sur un espace de Banach E, son dual sera un espace de Banach noté E' (comme tout dual d'espace vectoriel normé). On notera S la sphère unité de E, c'est à dire l'ensemble des vecteurs de norme 1.

En résumé (on se reportera à la partie topologie 5 pour toute les preuves) :

• Dans un espace vectoriel normé les opérations algébriques (multiplications par un scalaire et somme) sont continues. La norme est continue elle aussi.

• La topologie associée à la norme sur E est parfois appelée topologie forte.

• La topologie faible sur E est la topologie engendrée par la famille des applications linéaires continues ; c'est a dire que c'est la topologie la moins fine qui rendre toutes ces applications linéaires continues continues (non c'est pas une erreur s'il y a deux fois le mot continu !), c'est à dire qu'une base d'ouverts est constituée par les intersections <u>FINIES</u> de "bandes" de la forme $\{x/|f_i(x - x_0)| < \epsilon_i\}$, pour certains f_i dans E', certains $\epsilon_i > 0$, et un certain x_0 dans E. La boule unité fermée de E' (déterminée par la norme $\|.\|_\infty$) ci-dessous rappelée) est compacte POUR LA TOPOLOGIE FAIBLE * (théorème de Banach Alaoglu).

• la topologie forte sur le dual est la topologie engendrée par la norme $\|.\|_\infty$ qui à $f \in E'$ associe $sup_{x \in S}\|f(x)\|$. La topologie forte est plus fine que la topologie faible, elle-même plus fine que la topologie faible *.

• Etant donné X un espace topologique, $\mathcal{R}$ une relation d'équivalence sur X, la topologie quotient est l'ensemble des parties Y de $X/\mathcal{R}$ telles que $p^{-1}(Y)$ soit un ouvert de X, avec p la projection canonique de X sur $X/\mathcal{R}$. Il faut savoir que p est continue et ouverte.

• La topologie induite par une famille applications de X dans d'autres espaces topologiques, est la topologie la moins fine qui rende toutes ces applications continues. Une application f à valeurs dans X muni de la topologie engendrée par la famille des f_i est continue si et seulement si sa composée avec chaque f_i est continue. Il faut noter que la topologie faible est la topologie engendrée par les applications linéaires continues.

• La topologie produit, définie sur un produit d'espaces topologiques, est la topologie engendrée par les projections canoniques sur chacun des espaces topologiques du produit. Une applications à valeurs dans le produit est alors continue si et seulement si chacune de ses projections canoniques est continue. Un produit est séparé si et seulement si chacun des facteurs l'est. Le théorème de Tykhonov affirme qu'un produit de compacts est compact.

20.1.4 Quelques convergences dans les espaces de fonctions

⊡ Quelques rappels de topologie

Les résultats sont parfois donnés sans preuve ; on se réfèrera à la partie 5.

◇ Convergence simple

Définition 707 (convergence simple) *On dit qu'une suite f_n d'applications de X dans Y avec Y un espace topologique* **converge simplement** *vers f si pour tout x dans X $f_n(x)$ tend vers $f(x)$ pour n tendant vers $+\infty$.*

Proposition 708 (La convergence simple correspond-elle à une topologie ?) *Soit l'espace Y^X des applications de Y dans X, avec Y un espace topologique. La topologie produit sur Y^X a pour suites convergentes les suites simplement convergentes. C'est pourquoi on appelle cette topologie la* **topologie de la convergence simple**.

Démonstration :

• Soit f_n une suite d'éléments de Y^X, convergeant simplement vers une certaine fonction f. Montrons qu'elle converge aussi vers f pour la topologie produit.

Soit U un ouvert pour la topologie produit, contenant f . Alors (par définition) il existe $x_1, ..., x_n$ dans X et V_i voisinage de $f(x_i)$ tel que $\{g \in Y^X / \forall i \in [1, n], g(x_i) \in$

$V_i\} \subset U$. Il est alors clair qu'à partir d'un certain rang les f_n sont dans U.

• Supposons maintenant que f_n est une suite d'éléments de Y^X, convergeant vers une certaine fonction f pour la topologie produit. Donnons nous alors x dans X ; et U un voisinage de $f(x)$. Alors $V = \{g \in Y^X / g(x) \in U\}$ est un voisinage de f dans Y^X, donc f_n est dans V à partir d'un certain rang, donc $f_n(x) \in V$ à partir de ce même rang. Ceci montre que $f_n(x)$ tend vers $f(x)$.□

Grâce à ce résultat on obtient facilement quelques propriétés, dues à la stabilité de certaines propriétés topologiques par passage au produit :

Corollaire 709 (Caractéristiques de la topologie de la convergence simple) *On considère la topologie de la convergence simple sur Y^X.*

• *si Y est séparé la topologie de la convergence simple est séparée*

• *si Y est compact, alors la topologie de la convergence simple est compacte*

• *si Y est connexe (resp. par arcs), alors la topologie de la convergence simple est connexe (resp. par arcs).*

Démonstration : Un produit de séparés est séparé, un produit de compacts est compact, un produit de connexes est connexe, un produit de connexes par arcs est connexe par arcs.□

◇ **Convergence uniforme, convergence uniforme sur des parties**

Définition 710 *On dit qu'une suite f_n d'applications de X dans Y avec Y un espace métrique* **converge uniformément** *vers f si pour tout ϵ positif il existe N tel que pour tout $n \geq N$ et tout x dans X $d(f(x), f_n(x)) < \epsilon$.*

Etant donnée $\mathcal{S}$ une partie de $P(X)$, on dit que la suite (f_n) de fonctions de X dans Y (avec Y un espace métrique) est **uniformément convergente sur les éléments de** *$\mathcal{S}$ si pour tout $L \in \mathcal{S}$ la suite $(f_{n|L})$ est uniformément convergente sur L.*

Souvent, X sera un espace topologique localement compact et $\mathcal{S}$ sera l'ensemble des compacts de X.

Définition 711 (Topologie de la convergence uniforme) *Soit K un compact et F un espace métrique. L'espace des applications continues de K dans F, noté $C^0(K, F)$ est métrique avec la distance*

$$d(f, g) = sup_{x \in K} d(f(x), g(x))$$

La topologie associée est dite **topologie de la convergence uniforme***.*

Définition 712 (Topologie de la convergence uniforme sur tout compact) *Dans ce cas on peut définir la famille d'écarts (N_K), pour K compact non vide de X, par :*

$$N_K(f, g) = \sup_{x \in K} d(f(x), g(x)) \in [0, \infty]$$

Et la topologie engendrée par ces écarts a pour suites convergentes les suites uniformément convergentes sur les compacts de X. C'est pourquoi on appelle la topologie engendrée par ces applications **topologie de la convergence uniforme sur tout compact***.*

Si la famille $(K_i)_{i \in I}$ (I non nécessairement dénombrable !) est telle que tout compact K de X est inclus dans un certain K_i, alors la famille des N_{K_i} suffit.

La topologie de la convergence uniforme sur tout compact a donc pour base d'ouverts les $N_K^{-1}(f, [0, \epsilon[)$ pour $\epsilon > 0$, K compact non vide et f application de X dans Y.

Proposition 713 (Métrisabilité : topologie de convergence uniforme) *Si X est en fait un espace topologique compact, et si on se limite à l'ensemble $C^0(X, Y)$ des applications continues de X dans Y alors l'application $d(f, g) = sup_X d(f(x), g(x))$ est une distance et définit une topologie (sur $C^0(X, Y)$) pour laquelle les suites convergentes sont les suites uniformément convergentes au sens de la définition 710.*

Proposition 714 (La topologie de la convergence uniforme sur tout compact est-elle métrisable ?) *On suppose X localement compact, réunion dénombrable de compacts K_n, $K_m \subset K_{m+1}$, Y métrique ; alors la topologie engendrée par la distance*

$$d(f, g) = \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{N_{K_k}(f, g)}{1 + N_{K_k}(f, g)}$$

admet pour suites convergentes les suites uniformément convergentes sur tout compact au sens de la définition 710.

Exemple : Soit $x \in K$. Montrer que la fonction qui à $f \in C^0(K, F)$ associe $f(x)$ est continue pour la topologie de la convergence uniforme (resp. de la convergence uniforme sur tout compact).

$\diamond$ **Comparatif entre toutes ces notions de convergence**

Proposition 715 *Supposons que X est un espace topologique localement compact, et Y un espace métrique.*

Convergence pour la topologie de la convergence uniforme

$$\Downarrow$$

Convergence pour la topologie de la convergence uniforme sur tout compact

$$\Downarrow$$

Convergence pour la topologie de la convergence simple

Topologies dues aux mesures

(X, μ) étant un espace mesuré, les espaces de fonctions $\mathcal{L}^p(X)$ et $L^p(X)$ ont été définis et étudiés en partie 8.

Rappelons juste que $L^p(X)$ désigne l'ensemble des classes d'équivalences de l'ensemble des applications de X dans $\mathbb{R}$ pour la relation d'équivalence "être égales presque partout" qui contiennent au moins un élément dans $\mathcal{L}^p(X)$.

Rappelons aussi que $L^p(X)$, si X est réunion d'une suite croissante (pour l'inclusion) de compacts de mesure finie, pour $1 \leq p < \infty$, est le complété pour la norme $\|.\|_p = f \mapsto \sqrt[p]{\int_X |f|^p}$ de l'ensemble des fonctions continues à support compact (le résultat n'est pas valable pour $p = \infty$; ici l'adhérence serait simplement l'ensemble des applications continues qui, pour tout $\epsilon > 0$, sont inférieures à ϵ[1] en dehors d'un certain compact K_ϵ).

On définit en outre deux autres notions de convergence, liées à la notion de mesure : la convergence en mesure et la convergence presque partout.

[1]En module !

Définition 716 *Soit f_n une suite de fonctions de X dans Y, avec X un espace mesuré, et Y un espace topologique.*

On dit que f_n **converge presque partout** *vers f s'il existe N négligeable inclus dans X tel que f_n converge simplement vers f sur le complémentaire de N.*

Soit f_n une suite de fonctions de X dans $\mathbb{C}$ avec X un espace mesuré.

On dit que f_n **converge en mesure vers** *f si pour tout ϵ la limite pour $n \to \infty$ de la mesure de $\{x/|f_n(x) - f(x)| > \epsilon\}$ est nulle.*

On a alors les résultats suivants entre nos différentes notions de convergence des f_n vers f (lorsque toutes sont définies) :

Notez bien que $p < \infty$.

Convergence uniforme

$\Downarrow$

Convergence uniforme sur tout compact

$\Downarrow$

Convergence simple

$\Downarrow$

Convergence presque partout

$\Downarrow$ si les f_n sont majorées en module par une fonction g appartenant à L^p

Convergence dans L^p

$\Downarrow$

Convergence en mesure

$\Downarrow$

Convergence presque partout d'une suite extraite
et
Convergence presque partout et X de mesure finie $\Rightarrow$ convergence en mesure

Ci-dessous une liste de contre-exemples, pour bien se mettre en tête qu'il ne faut pas confondre convergences et convergences :

• convergence uniforme sur tout compact n'implique par convergence uniforme

En effet, sur $[0, +\infty[$ la suite f_n définie par

$$f_n(x) = 1 \text{ si } x < n$$
$$f_n(x) = 0 \text{ sinon}$$

converge uniformément sur tout compact vers la fonction constante égale à 1, mais ne converge pas uniformément vers cette fonction.

• converge simple n'implique pas convergence uniforme sur tout compact

Il suffit de prendre $f_n(x) = max(1 - nx, 0)$ sur $[0, 1]$. $f_n(x)$ converge clairement vers 0 pour $x > 0$ et vers 1 pour $x = 0$. La convergence n'est pas uniforme car le sup de $|f_n - f|$ reste égal à 1 ; elle n'est pas non plus uniforme sur tout compact car $[0, 1]$ étant compact on aurait alors convergence uniforme.

• convergence presque partout n'implique pas convergence simple.

Evident : $f_n(x) = 1$ pour tout x de $[0, 1]$, $f(x) = 1$ pour tout x de $[0, 1[$ et $f(1) = 0$.

• Convergence presque partout et même convergence simple n'impliquent pas convergence dans L^p si les f_n ne sont pas majorées en module par une fonction de L^p.

Par exemple, sur $\mathbb{R}$, $f_n(x) = n$ si $x \in]0, 1/n[$, $f_n(x) = 0$ sinon (on pourrait aussi avoir ce résultat avec des fonctions continues, en considérant des fonctions affines par morceaux...).

• Convergence dans L^p n'implique pas convergence presque partout.

On considère $f_n(x) = 1$ si il existe $u \in \mathbb{N}$ tel que $x + u$ est compris au sens large entre $\sum_{k=0}^{n} 1/k$ et $\sum_{k=0}^{n+1} 1/k$, 0 sinon.

• Convergence en mesure n'implique pas convergence dans L^p

Même contre-exemple que pour "convergence presque partout et même convergence simple n'impliquent pas convergence dans L^p si les f_n ne sont pas majorées en module par une fonction de L^p".

• Convergence presque partout n'implique pas convergence en mesure si la mesure de X n'est pas finie

Facile ; sur $\mathbb{R}$, l'application f_n qui à x associe $sin(x/n)$.

20.2 Théorèmes d'Ascoli et conséquences

20.2.1 Théorie

Définition 717 (Equicontinuité) *Soit $\mathcal{F}$ une famille d'applications $X \to Y$ où X est un espace topologique et Y un espace métrique. On dit que $\mathcal{F}$ est* **équicontinue** *si, pour tout $\epsilon > 0$ et tout $x \in X$ il existe un voisinage V_x de x dans X tel que $d(f(x), f(y)) < \epsilon$ pour tout $f \in F$ et tout $y \in V_x$.*
Si X est aussi métrique, $\mathcal{F}$ est dite **uniformément équicontinue** *si pour tout $\epsilon > 0$ il existe $\alpha > 0$ tel que pour tous x, y vérifiant $d(x, y) < \alpha$ et tout $f \in \mathcal{F}$, on ait $d(f(x), f(y)) < \epsilon$.*

Exemples : Une famille finie est toujours équicontinue.

On a un équivalent du théorème de Heine pour les familles équicontinues sur un espace compact.

Théorème 718 *Si X est métrique compact et si Y est métrique, si $\mathcal{F}$ est une famille d'applications équicontinues de X dans Y, alors la famille $\mathcal{F}$ est uniformément équicontinue.*

Démonstration : On considère α_x le rayon d'une boule inclus dans le V_x correspondant à un ϵ donné ; on recouvre l'espace avec ces boules, on en extrait un recouvrement fini, puis on prend le min des α_x, et on a le résultat.□

Théorème 719 (Théorème d'Ascoli) • *Soit F un espace métrique*[a], *et E un espace topologique ; soit $\mathcal{F}$ une famille équicontinue en $e \in E$ de fonctions de E dans F.*

Alors $\overline{\mathcal{F}}$[b] *est équicontinue en e.*

• *Si $\mathcal{F}$ est équicontinue en tout point, alors $\overline{\mathcal{F}}$ est équicontinue en tout point.*

• *Avec $\mathcal{E}$ une partie dense de E, la topologie de la convergence simple, la topologie de la convergence uniforme sur tout compact, la topologie induite par la convergence simple sur $\mathcal{E}$*[c], *induisent la même topologie sur $\mathcal{F}$.*

[a] Hypothèse facile à retenir ; on ne pourrait pas définir la notion de famille équicontinue si F n'était pas métrique.

[b] Adhérence prise pour la topologie de la convergence simple, c'est à dire pour la topologie produit dans F^E.

[c] C'est-à-dire la topologie induite par les projections canoniques de F^E sur les $(F_{x_i})_i$

Démonstration : • • (on prouve les deux premiers • en un seul coup) On se donne $\epsilon > 0$. On a donc un certain U voisinage de e tel que pour tout x dans U et tout $f \in \mathcal{F}$ $d(f(x), f(e)) < \epsilon$. On cherche à montrer que cela est en fait vrai pour tout $f \in \overline{\mathcal{F}}$. On se donne une telle fonction f, et un certain x dans U.

On définit alors V_x l'ensemble des applications g de E dans F telles que

$$d(g(x), f(x)) < \epsilon \text{ et } d(g(a), f(a)) < \epsilon.$$

V_x est un voisinage de f pour la topologie simple, donc il doit intersecter $\mathcal{F}$; soit g dans l'intersection obtenue. Il suffit alors d'écrire

$$d(g(x), g(a)) \leq d(g(x), f(x)) + d(f(x), f(a)) + d(f(a), g(a)) \leq 3\epsilon$$

• Il est clair que la topologie de la convergence simple sur $\mathcal{E}$ est moins fine que le topologie de la convergence simple elle-même moins fine que la topologie de la convergence uniforme sur tout compact (rappelons qu'un singleton, comme tout ensemble fini séparé, est compact). Le seul problème est la réciproque. On se donne donc U un ouvert pour la topologie de la convergence uniforme sur tout compact, f dans U, et on cherche à montrer que U contient un voisinage pour la topologie de la convergence simple sur $\mathcal{E}$ de f.

U étant ouvert en f pour la topologie forte, il existe un compact K et un réel > 0 ϵ tels que

$$\{g \in \mathcal{F} / \forall y \in K d(f(y), g(y)) < \epsilon\}$$

soit inclus dans U.

$$\forall x \exists U_x \text{ ouvert en } x \ / \forall g \in \mathcal{F} \forall y \in U_x d(g(x), g(y)) < \epsilon/5$$

Alors par la propriété de Borel-Lebesgue, il existe un sous-ensemble I fini de K tel que $K \subset \cup_{x \in I} U_x$. Les U_x étant ouverts non vides et $\mathcal{E}$ étant dense dans E, on choisit pour $x \in I$ un point $y_x \in \mathcal{E}$.

Considérons alors $W = \{g \in \mathcal{F} / \forall x \in I \ d(g(y_x), f(y_x)) < \epsilon/5\}$

$$\forall z \in K \ \exists x \in I / z \in U_x$$

et

$$d(g(z), f(z)) \leq$$

$$d(g(z), g(x)) + d(g(x), g(y_x)) + d(g(y_x), f(y_x)) + d(f(y_x), f(x)) + d(f(x), f(z)) \leq \epsilon$$

Donc $g \in U$, donc $W \subset U$ et donc U est un voisinage de f pour la topologie de la convergence simple sur $\mathcal{E}$. D'où le résultat.$\square$

Théorème 720 (Théorème d'Arzéla-Ascoli) *Une partie $\mathcal{F}$ incluse dans $C^0(K, F)$ est relativement compacte pour la topologie de la convergence uniforme si et seulement si on a les deux conditions suivantes :*
- *La famille $\mathcal{F}$ est équicontinue*
- *Pour tout $x \in K$ l'ensemble des $f(x)$ pour $f \in \mathcal{F}$ est relativement compact*

Voir simplement la partie applications, ci-dessous ; mais aussi le théorème 747.

Démonstration : Tout d'abord supposons que notre famille $\mathcal{F}$ est relativement compacte dans $C^0(K, F)$. Pour tout x l'évaluation $\hat{x} : C^0(K, F) \to F$ est continue ; donc l'image $\hat{x}(\overline{\mathcal{F}})$ est compacte, or il contient $\{f(x) | f \in \mathcal{F}\}$; donc l'adhérence de ce dernier ensemble est un fermé d'un compact, et est donc compacte, d'où le second point. Par ailleurs, comme $\mathcal{F}$ est relativement compacte, avec $\epsilon > 0$, on peut trouver $f_1, ..., f_n \in \mathcal{F}$ tels que pour tout $f \in \mathcal{F}$, $d(f_i, f) < \epsilon$; la famille des f_i étant équicontinue (comme toute famille finie), pour $x \in K$ donné, il existe un voisinage V_x de x tel que $d(f_i(x), f_i(y)) < \epsilon$ pour tout $y \in V_x$. Comme $d(f(x), f(y)) \leq d(f_i(x), f_i(y)) + 2.d(f_i, f)$, on voit que, pour tout $y \in V_x$ on a $d(f(x), f(y)) \leq 3.\epsilon$.
Réciproquement (voir figure 20.1), supposons les deux conditions données remplies, et montrons que la famille $\mathcal{F}$ est relativement compacte. Pour cela on considère $C_0(K, F)$

comme un sous-ensemble de F^K muni de la topologie produit, cette inclusion induisant sur $C_0(K,F)$ la topologie de la convergence simple. Posons $C_x = \overline{\{f(x)|f \in \mathcal{F}\}}$. Par la seconde condition (qui n'intervient qu'ici), C_x est compact. Comme $\mathcal{F}$ est inclus dans le produit des C_x, $\overline{\mathcal{F}}$ est compact dans F^K. Il faut alors montrer que $\overline{\mathcal{F}}$ est inclus dans $C^0(K,F)$, et que la topologie produit sur $\overline{\mathcal{F}}$ et la topologie de la distance sont les mêmes, ce qui finira la preuve.

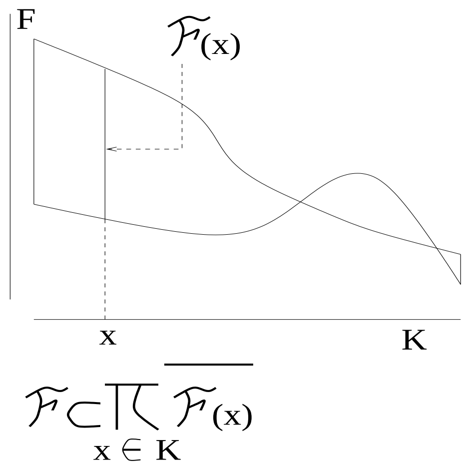


FIG. 20.1 – Illustration de la preuve du théorème d'Arzéla-Ascoli. L'adhérence de la famille considérée est un fermé d'un produit de compacts, donc est un compact ; il reste à vérifier que l'adhérence est bien incluse dans $C^0(K,F)$, et que la topologie produit induit bien la topologie de la distance uniforme.

Lemme 721 *Si $\mathcal{F}$ est équicontinue alors $\overline{\mathcal{F}} \subset C^0(K,F)$ (adhérence pour la topologie produit).*

Démonstration : Le théorème d'Ascoli (2ème point) implique que la famille $\overline{F}$ est équicontinue, ce qui implique clairement le résultat.$\square$

Lemme 722 *La topologie induite par la topologie produit sur $\overline{\mathcal{F}}$ et la topologie de la distance (= topologie de la convergence uniforme) sont les mêmes.*

Démonstration : Chaque fonction $\hat{x}$ (évaluation en x) étant continue, tout ouvert de $\overline{\mathcal{F}}$ est un ouvert pour la topologie de la convergence uniforme.
Il reste à voir que tout voisinage de f_0 appartenant à $\overline{\mathcal{F}}$ dans $\overline{\mathcal{F}}$ pour la métrique, contient un voisinage de f_0 dans $\overline{\mathcal{F}}$ muni de la topologie produit. Soit $\epsilon > 0$, et considérons $\{g \in \overline{\mathcal{F}} | max\ d(g(x), f_0(x)) \leq \epsilon\}$ (qui décrit une base de voisinages de $\overline{\mathcal{F}}$ pour la métrique). Pour tout x, on obtient par la condition 1 un voisinage ouvert de x dans K tel que si $y \in V_x$ on ait $d(h(x), h(y)) < \epsilon/3$ pour tout $h \in \overline{\mathcal{F}}$. Par compacité de K on peut trouver x_1,...x_n tels que $K = \cup_{i=1}^n V_{x_i}$. Considérons alors

$\mathcal{V} = \{g \in \overline{\mathcal{F}} | d(g(x_i), f_0(x_i)) < \epsilon/3\}$; c'est un voisinage de f_0 pour la topologie produit. Si $g \in \mathcal{V}$ et $x \in K$ soit x_{i_0} tel que $x \in V_{x_{i_0}}$, on a alors

$$d(g(x), f_0(x)) \leq d(g(x), g(x_{i_0})) + d(g(x_{i_0}), f_0(x_{i_0})) + d(f_0(x_{i_0}), f_0(x)) < \epsilon$$

donc $d(g(x), f_0(x)) \leq \epsilon$ pour tout $x \in K$. Par conséquent le voisinage

$$\{g \in \overline{\mathcal{F}} | \max d(g(x), f_0(x)) \leq \epsilon\}$$

de f_0 pour la topologie de la convergence uniforme contient $\mathcal{V}$ qui est un voisinage de f_0 pour la topologie produit. En résumé pour cette preuve, un sens est trivial, et l'autre sens se prouve en utilisant une boule pour la distance, et en appliquant à la fois l'équicontinuité de $\mathcal{F}$ et la compacité de K.□
Ces deux preuves achèvent donc le théorème d'Arzéla-Ascoli. En résumé il faut donc, pour le sens difficile :
• Utiliser la condition sur les parties relativement compactes de F pour conclure à la relative compacité de $\mathcal{F}$ dans l'espace produit
• Utiliser l'équicontinuité pour montrer que $\overline{F} \subset C^0(K, F)$
• Utiliser l'équicontinuité de $\mathcal{F}$ et la compacité de K pour montrer que les deux topologies sont égales.□

20.2.2 Applications

⊡ **Topologie de $H(\Omega)$**

On travaille sur $H(\Omega)$, avec Ω un ouvert de $\mathbb{C}$, munie de la topologie de la convergence uniforme sur tout compact.

Définition 723 *On dit d'une partie $\mathcal{F}$ de $H(\Omega)$ qu'elle est* **bornée** *si pour tout compact K de Ω il existe une certaine constante C_K telle que pour toute f dans $\mathcal{F}$ et tout k dans K, $|f(k)| \leq C_K$.*

Théorème 724 *Les parties compactes de l'ensemble des donctions holomorphes sur Ω $H(\Omega)$ sont les parties fermées et bornées.*

Démonstration : • Montrons tout d'abord (partie facile) que les parties compactes sont fermées et bornées.

- Les parties compactes sont fermées, par le lemme 177 (tout compact d'un espace séparé est fermé)

- Les parties compactes sont bornées ; c'est évident.

• Supposons que K soit une partie fermée bornée de $H(\Omega)$.

- Montrons tout d'abord que K est équicontinue en tout point x de Ω. Soit donc un tel x.

- x est centre d'un certain disque compact inclus dans Ω

- toute f de K est bornée par un certain M sur ce disque compact de rayon R

- donc la dérivée de f en tout point du disque de centre x et de rayon $R/2$ est majorée par $2M/R$, grâce à l'estimateur de Cauchy (théorème 673).

- donc $\mathcal{F}$ est équicontinue en x, par le théorème des accroissements finis 458.

- Etant donné x dans Ω, l'ensemble des $f(x)$ pour f dans $\mathcal{F}$ est borné, donc relativement compact.

- Par le théorème d'Arzéla-Ascoli 720, K est donc relativement compact, or il est fermé, donc il est compact.$\square$

Corollaire 725 *L'ensemble des fonctions holomorphes sur Ω $H(\Omega)$ muni de la topologie de la convergence uniforme est métrisable, mais pas normable.*

Démonstration : Pour voir que $H(\Omega)$ est métrisable, il suffit de consulter le lemme 261 et le théorème 714.

D'après le théorème de Riesz (192), si $H(\Omega)$ était normable, alors la boule unité fermé serait compacte si et seulement si l'espace était de dimension finie. Or $H(\Omega)$ n'est pas de dimension finie.$\square$

20.3 La hiérarchie des $C^k(\Omega)$, avec Ω ouvert de $\mathbb{R}^n$

Définition 726 *Etant donné Ω un ouvert de $\mathbb{R}^n$, on note $C^k(\Omega)$ l'ensemble des fonctions k fois continument dérivables de Ω dans $\mathbb{C}$.*

$C^k(\Omega)$ est stable par produit, et si f est dans $C^k(\Omega)$ et ne s'annule pas alors $1/f$ est dans $C^k(\Omega)$.

Pour f dans $C^k(\Omega)$ et ν dans $\mathbb{N}^n$ et telle que $\sum_{i=1}^n \mu_i \leq k$, on note

$$\delta^\nu f = \frac{\delta^{|\nu|} f}{(\delta x_1)^{\nu_1} \dots (\delta x_n)^{\nu_n}}$$

L'ordre des dérivations importe peu, comme on l'a vu dans le chapitre de calcul différentiel.

Définition 727 (Opérations dans $\mathbb{N}^n$) *Etant donnés ν et η dans $\mathbb{N}^n$:*

- *on note $\nu! = \Pi_{i=1}^n(\nu_i)!$.*
- *on note $\nu \geq \eta$ si $\forall i \in [1,n]\nu_i - \eta_i \geq 0$*
- *si $\nu \geq \eta$ on note $\alpha = \nu - \eta$ avec $\forall i \in [1,n]\alpha_i = \nu_i - \eta_i$*
- *si $\nu \geq \eta$ on note $C_\nu^\eta = \frac{\nu!}{\eta!(\nu-\eta)!} = \Pi_{i=1}^n C_{\nu_i}^{\eta_i}$*
- *on note $|\nu| = \sum_{i=1}^n \nu_i$*
- *on note 0 l'élément $(0,...,0)$ de $\mathbb{N}^n$.*

Proposition 728 (Formule de Leibnitz)

$$\delta^\alpha(f.g) = \sum_{\beta \leq \alpha} C_\alpha^\beta \delta^\beta f \delta^\alpha g$$

⚠ Il s'agit d'un produit et pas d'une composition.

Démonstration : Récurrence facile, utilisant le corollaire 453.□

Définition 729 (Distance sur $C^k(\Omega)$) *On définit maintenant K_m comme étant l'intersection de la boule $\overline{B}(0,m)$ et de $\{x/d(x,\Omega^c) \geq \frac{1}{m}\}$. On définit ensuite $N_m(f)$, pour f dans $C^k(\Omega)$ par $N_m(f) = \sum_{\nu \in \mathbb{N}^d/|\nu| \leq k} sup_{K_m} \delta^\nu f(x)$.*

On définit ensuite sur $C^k(\Omega)$ la distance :

$$d(f,g) = \sum_{m>0} \frac{1}{2^m} \frac{N_m(f-g)}{1+N_m(f-g)}$$

Il est indispensable pour la suite de consulter les propriétés topologiques des K_m ainsi définis ; voir lemme 261.

Théorème 730 • *N_m est une semi-norme*

• *d est bien définie et est une distance*

• *La topologie définie pour cette distance a pour suites convergentes les suites de fonctions (f_n) de $C^k(\Omega)$ telles que pour tout ν tel que $|\nu| \leq k$ $\delta^\nu f_n$ converge uniformément sur tout compact K de Ω.*

• *$C^k(\Omega)$ est complet pour cette distance*

<u>Démonstration :</u>

• Le fait que N_m soit une semi-norme est évident (rappelons qu'un semi-norme a tout d'une norme à ceci près qu'une semi-norme n'est pas nécéssairement nulle seulement en 0)

• d est bien définie, car $\frac{1}{2^m}\frac{N_m(f-g)}{1+N_m(f-g)} \leq \frac{1}{2^m}$. Il est clair que $d(f,g) = 0 \iff f = g$, et que $d(f,g) = d(g,f)$. Il reste à voir l'inégalité triangulaire.

Pour cela soient f g et h dans $C^k(\Omega)$. Alors

$$N_m(f-g) \leq N_m(f-h) + N_m(h-g)$$

Par croissance de $x \mapsto \frac{x}{1+x}$,

$$\frac{N_m(f-g)}{1+N_m(f-g)} \leq \frac{N_m(f-h)+N_m(h-g)}{1+N_m(f-h)+N_m(h-g)}$$

$$\frac{N_m(f-g)}{1+N_m(f-g)} \leq \frac{N_m(f-h)}{1+N_m(f-h)} + \frac{N_m(h-g)}{1+N_m(h-g)}$$

Il ne reste qu'à sommer en pondérant par $1/2^m$ pour avoir le résultat désiré.

• Commençons par montrer qu'une suite convergente pour cette distance est bien convergente uniformément sur tout compact, ainsi que toutes ses dérivées. Ce résultat est en fait clair ; il suffit de voir que tout compact de K est inclus dans un K_i ; et que pour que $d(f,g)$ tende vers 0, il faut que $N_m(f,g)$ tende vers 0.

La réciproque est plus laborieuse.

Réciproquement, supposons que toutes les dérivées $\leq k$ de f_n convergent uniformément sur tout compact, notons f la fonction limite. Alors donnons nous $\epsilon > 0$. Soit m tel que $\sum_{i=m+1}^{\infty} 1/2^i < \epsilon$. Choisissons ensuite N tel que pour $n \geq N$ et tout $m' < m$ $N_{m'}(f_n - f) \leq \epsilon$. Alors on a bien $d(f_n, f) \leq \epsilon$ pour tout $n \geq N$.

• Il reste à montrer la propriété de complétude.

Donnons-nous f_n une suite de Cauchy pour la distance ainsi définie sur $C^k(\Omega)$.

Pour tout x dans Ω il existe un certain m tel que $x \in Int(K_m)$

Le fait que $(f_n)_n$ soit une suite de Cauchy nous permet de déduire que pour $\nu \in \mathbb{R}^n$ tel que $|\nu| \leq k$ $(f_n^\nu(x))_n$ est une suite de Cauchy. On note $f_{\infty,\nu}(x)$ la limite.

Pour tout m, on va montrer par récurrence sur $|\nu|$ que f est $C^{|\nu|}$ sur K_m, et que sur l'intérieur de K_m $f_{\infty,\nu} = \delta^\nu f_{\infty,0}$.

La propriété est claire pour $|\nu| = 0$; une limite uniforme de fonctions continues est continue.

On se donne alors ν en supposant la propriété vraie jusqu'à $|\nu| - 1$.

On définit ν' tel que $\delta^\nu = \frac{\delta}{\delta x_p}\delta^{\nu'}$.

Alors sur K_m tel que y appartienne à l'intérieur de K_m, intéressons-nous à la dérivée suivant δx_p de $\delta^{\nu'}$ $f_{\infty,0} = f_{\infty,\nu}$ (si on montre son existence et sa continuité, on aura conclu grâce au théorème 472).

Pour y' suffisamment proche de y pour être dans K_m et pour que le segment $[y, y']$ soit dans K_m, avec $\forall i \in [1, n] i \neq p \Rightarrow y_i = y'_i$ (c'est à dire que le point y' est juste déplacé suivant la coordonnée p.

$$\delta^{\nu'} f_i(y') - \delta^{\nu'} f_i(y) = \int_{y_p}^{y'_p} \frac{\delta}{\delta x_p}\delta^{\nu'} f_i(y_1, ..., y_{p-1}, u, y_{p+1}, ..., y_n)du$$

$$\delta^{\nu'} f_i(y') - \delta^{\nu'} f_i(y) = \int_{y_p}^{y'_p} \delta^{\nu} f_i(y_1, ..., y_{p-1}, u, y_{p+1}, ..., y_n)du$$

Et en faisant tendre i vers $+\infty$

$$f^{\nu'}(y') - f^{\nu'}(y) = \int_{y_p}^{y'_p} f_{\infty,\nu}(y_1, ..., y_{p-1}, u, y_{p+1}, ..., y_n)du$$

Donc la dérivée partielle $\frac{\delta}{\delta x_p}$ existe et est continue en x (elle vaut $f_{\infty,\nu}$).

Moralité de tout ça :

- Si on se donne un compact K et $\nu \leq k$

- Alors K est inclus dans l'intérieur d'un certain K_m

- Sur ce K_m il y a convergence uniforme de la dérivée $\delta^\nu f_n$ vers $\delta^\nu f_{\infty,0}$, puisqu'il y a convergence pour N_m.

- La limite est bien dans $C^k(\Omega)$.

Donc $C^k(\Omega)$ est bien complet pour la métrique que l'on a définie !□

Corollaire 731 (Autre façon de voir la topologie sur $C^k(\Omega)$) *La même topologie serait définie en définissant les fermés comme étant les sous-ensembles contenant les limites de toute suite convergente pour la topologie de la convergence uniforme de toutes les dérivées d'ordre total $\leq k$ sur tout compact.*

20.4 La topologie faible

Nous allons montrer ici quelques propriétés de la topologie faible.

Corollaire 732 *La topologie faible est séparée.*

Démonstration : C'est une application directe du théorème 698 ; les singletons sont compacts, on en prend deux, on les sépare au sens strict par un hyperplan fermé ; les deux demi-espaces ouverts restant sont des ouverts séparants les deux points...□

Définition 733 *On note $x_n \rightharpoonup x$ le fait que la suite x_n d'élément de E* **converge vers** *$x \in E$ pour la topologie faible.*

On note $x_n \to x$ la convergence de x_n vers x pour la topologie de la norme (ben oui, on est dans un espace vectoriel normé E), et $f_n \to f$ dans E' pour la convergence de f_n vers f pour la topologie forte. On pourra aussi qualifier de convergence forte la convergence dans E pour la norme.

Exemple :
Dans cette proposition, les $\forall f$ désignent $\forall f \in E'$.

$$(x_n \rightharpoonup x) \iff (\forall f, f(x_n) \to f(x))$$
$$(x_n \to x) \to (x_n \rightharpoonup x)$$
$$(x_n \rightharpoonup x) \to (\|x_n\| \text{ borné et } \|x\| \leq liminf\|x_n\|)$$
$$(x_n \rightharpoonup x \text{ et } f_n \to f) \to (f_n(x_n) \to f(x)(\text{dans } \mathbb{R}))$$
$$(x_n \to x \text{ et } f_n \rightharpoonup f) \to (f_n(x_n) \to f(x)(\text{dans } \mathbb{R}))$$

20.5 Liens entre topologie faible et topologie forte

20.5.1 En dimension finie

Théorème 734 *Si E est de dimension finie, alors la topologie faible sur E et la topologie forte sur E sont égales.*

Démonstration :

• La topologie forte est toujours plus fine que la topologie faible (évident au vu des définitions ; dans un ouvert (au sens de la topologie faible) , tout point possède un voisinage de la forme $\cap_{i \in I}\{x/f_i(x-x_0) < \epsilon_i\}$, et pour tout point x dans cette intersection on loge une boule ouverte centrée sur x de rayon $inf\{(\epsilon_i - \|x-x_0\|)/\|f_i\|_\infty | i \in I\})$: il s'agit là d'un ouvert pour la topologie forte.

• Réciproquement, soit x dans E, et U un ouvert (pour la topologie forte) contenant x. On cherche à construire un ouvert pour la topologie faible qui contienne x et qui soit inclus dans U. On peut naturellement se restreindre à $U = B(x, \epsilon)$, boule ouverte de centre x et de rayon r.

• On fixe $(e_1, ..., e_n)$ une base de E de vecteurs de norme 1.

• On note $(f_i)_{i \in [1,n]}$ la famille des applications telles que $\forall t \in E\, t = \sum_{i \in [1,n]} f_i(t)$.

• On peut alors écrire $\|t - x\| \leq \sum_{i \in [1,n]} |f_i(t-x)|$.

• Il suffit alors d'écrire $V = \{t/|f_i(t-x)| < \epsilon/n\}$ pour avoir un ouvert V pour la topologie faible inclus dans U et contenant x.$\square$

On verra en partie 20.7.1 que cette propriété est caractéristique de la dimension finie.

20.5.2 Dans le cas général

Théorème 735 *Soit E un espace vectoriel normé , et C convexe inclus dans E. Alors C est faiblement fermé (i.e. fermé pour la topologie faible) si et seulement si C est fortement fermé (i.e. fermé pour la topologie forte).*

On trouvera une application avec la proposition 742.

Démonstration :

• Il est clair que si C est faiblement fermé, alors il est fortement fermé. Il suffit donc de se préoccuper de la réciproque.

• Supposons maintenant C fortement fermé.

• On se donne un point x appartenant au complémentaire de C.

• D'après le théorème 698 il existe un hyperplan fermé qui sépare C et $\{x\}$ au sens strict.

• L'hyperplan délimites deux demi-espaces faiblement ouverts, dont l'un contient x et est inclus dans le complémentaire de C.

• C a donc un complémentaire faiblement ouvert, et C est donc faiblement fermé.$\square$

Théorème 736 *Soient E et F des espaces de Banach. On se donne T une application linéaire de E dans F. Alors T est continue pour E et F munis chacun de sa topologie faible si et seulement si T est continue pour E et F munis de leur topologie d'espaces vectoriels normés (ie la topologie forte).*

Démonstration : • Supposons tout d'abord que T est continue de E dans F pour la topologie forte, et montrons que T est continue pour la topologie faible.

- on va procéder en montrant que pour toute forme linéaire continue f sur F, l'application $f \circ T$ est continue (de E muni de la topologie faible dans $\mathbb{R}$).

- soit donc $f \in F'$.

- $f \circ T$ est continue pour la topologie forte, et linéaire.

- $f \circ T$ est donc continue pour la topologie faible aussi (puisque, par définition, la topologie faible rend continues toutes les formes linéaires continues).

• Supposons maintenant que T est continue de E dans F pour la topologie faible, et montrons que T est continue pour la topologie forte.

- le graphe de T est alors fermé dans le produit $E \times F$, muni de la topologie produit des topologies faibles de E et F.

- le graphe de T est donc aussi fermé pour le produit des topologies fortes car ce graphe est un convexe faiblement fermé de $E \times F$ de $E \times F$, et donc T est continue pour la topologie forte (re-utilisation du théorème du graphe fermé 258). $\square$

20.6 Espaces de Hölder

20.6.1 Espaces $Lip_\alpha(\Omega)$

Définition 737 *On dit qu'une application d'un métrique E dans $\mathbb{C}$ vérifie la* **condition de Hölder d'ordre** α *si il existe C dans $\mathbb{R}^+$ tel que pour tous x et y dans E $|f(x) - f(y)| \leq Cd(x,y)^\alpha$.*

Etant donné Ω un ouvert de $\mathbb{R}^n$ et α un réel appartenant à $]0,1]$, on note $Lip_\alpha(\Omega)$ l'ensemble des applications bornées de Ω dans $\mathbb{C}$ vérifiant la condition de Hölder d'ordre α sur Ω.

Etant donné f dans $Lip_\alpha(\Omega)$, on note $\|f\|_\alpha$ le réel $\|f\|_\infty + \sup_{x\neq y} \frac{|f(x)-f(y)|}{\|x-y\|^\alpha}$. Il s'agit d'une norme.

définir Lip_α pour $\alpha > 1$ serait peu intéressant, car on travaillerait sur des fonctions localement constantes sur un ouvert, c'est à dire, les composantes connexes d'un ouvert étant ouvertes et dénombrables, sur $\mathbb{R}^n$ ou $\mathbb{R}^{\mathbb{N}}$ ...

Proposition 738 • *Toute fonction dans $Lip_\alpha(\Omega)$ est uniformément continue.*

• *Toute fonction dans $Lip_\alpha(\Omega)$ se prolonge en une fonction continue sur $\overline{\Omega}$.*

• *Si $0 < \alpha \leq \beta \leq 1$, alors $Lip_\beta(\Omega) \subset Lip_\alpha(\Omega)$*

• *Toute fonction C^1 à dérivée bornée est dans Lip_α pour tout $\alpha \in]0,1]$.*

Démonstration : La plupart des points sont évidents ; le prolongement en une fonction continue utilise le fait que $\mathbb{C}$ est complet, le fait que Ω est dense dans $\overline{\Omega}$, l'uniforme continuité de toute fonction dans $Lip_\alpha(\Omega)$ et le théorème 246.□

Théorème 739 *$Lip_\alpha(\Omega)$ (muni de la norme $\|.\|_\alpha$)est un espace de Banach.*

Démonstration :

• Donnons-nous (f_m) une suite de Cauchy dans $Lip_\alpha(\Omega)$.

• (f_m) est aussi de Cauchy pour la norme $\|.\|_\infty$.

• Soit donc f la limite de la suite (f_m) pour la convergence uniforme.

• f est bien bornée, puisque limite uniforme de fonctions bornées.

• $|f_m(x) - f_m(y)| \leq \underbrace{\|f_p\|_\alpha}_{\text{bornée.}} \|x-y\|^\alpha$

• On fait alors tendre m vers ∞, et on constate que f vérifie la condition de Hölder d'ordre α.

• Vérifier que f_m tend vers f pour $\|.\|_\alpha$ est facile...□

Il existe une fonction qui soit dans $Lip_\alpha(\mathbb{R})$ pour tout α dans $]0,1[$ mais pas dans $Lip_1(\mathbb{R})$; par exemple la fonction définie au théorème 490. Montrer ce fait est toutefois fortement non trivial...

20.6.2 Espaces $C^{k,\alpha}(\Omega)$

Définition 740 (Espaces de Hölder) *Etant donné Ω un ouvert de $\mathbb{R}^n$, α dans $]0,1]$, $k \in \mathbb{N}$, on définit par récurrence sur k les espaces $C^{k,\alpha}(\Omega)$ par*

$$C^{0,\alpha}(\Omega) = Lip_\alpha(\Omega)$$

$$k \geq 1 \Rightarrow$$

$$C^{k,\alpha}(\Omega) = \{f \textit{ bornée de } \Omega \textit{ dans } C \ / \forall i \in [1,n] \frac{\delta f}{\delta x_i} \textit{ existe et appartient à } C^{k-1,\alpha}(\Omega)\}$$

Cette définition équivaut à (voir définition 727 pour les opérations sur $\mathbb{N}^n$) :

$$C^{k,\alpha}(\Omega) = \{f \in C^k(\Omega)/f \textit{ bornée } \wedge \forall \nu/|\nu| \leq k \Rightarrow D^\nu f \in Lip_\alpha(\Omega)$$

On munit $C^{k,\alpha}(\Omega)$ de la norme $f \mapsto \|f\|_{k,\alpha} = \sum_{|\nu|\leq k} \|D^\nu f\|_\alpha$.

De manière équivalente, $\|f\|_{k,\alpha} = \sum_{\nu \leq k}(\|D^\nu f\|_\infty + sup_{x\neq y} \frac{|f^{(\nu)}(x) - f^{(\nu)}(y)|}{|x-y|^\alpha})$ (c'est la même expression développée !) et la norme suivante est équivalente à celle-ci :

$$f \mapsto \|f\|'_{k,\alpha} = \sum_{|\nu|\leq k} \|D^\nu f\|_\infty + \sum_{|\nu|=k} \sup_{x\neq y} \frac{f^{(\nu)}(x) - f^{(\nu)}(y)}{|x-y|^\alpha}$$

Bien sûr il convient de vérifier l'équivalence des deux définitions.

Quelques résultats sans preuve :

Théorème 741 • *$C^{k,\alpha}(\Omega)$ est un espace de Banach.*

• *Les fonctions de $C^{k,\alpha}(\Omega)$ sont prolongeables par continuité sur $\overline{\Omega}$ en fonctions vérifiant la condition de Hölder pour toutes les dérivées $\leq k$.*

• *$k + \alpha \geq k' + \alpha'$ implique $C^{k,\alpha} \subset C^{k',\alpha'}$*

• *$\|uv\|_{k,\alpha} \leq \|u\|_{k,\alpha} \|v\|_{k,\alpha}$*

Pour plus d'informations sur les espaces de Hölder, on pourra consulter le livre [22].

20.7 Zoologie de l'analyse fonctionnelle

20.7.1 La topologie faible n'est pas la topologie forte en dimension infinie

Proposition 742 *Soit E un Banach de dimension infinie. Alors la topologie faible est différente de la topologie forte.*

Démonstration : On a vu au théorème 734 que si la dimension est finie, alors la topologie faible et la topologie forte sont égales. On a aussi vu que dans le cas général, la topologie forte est plus fine que la topologie faible. On va montrer ici que la topologie forte est strictement plus fine en dimension infinie, en exhibant un ouvert pour la topologie forte qui n'est pas ouvert pour la topologie faible, ou, ce qui revient au même par passage au complémentaire, un fermé pour la topologie forte qui n'est pas fermé pour la topologie faible.

• On considère la sphère unité S de E. Elle est fermée, comme image réciproque d'un singleton (donc un fermé) par une application continue (la norme).

• On va chercher à déterminer l'adhérence de S pour la topologie faible.

• Soit x de norme < 1.

• On se donne U un voisinage de x pour la topologie faible.

• Alors (propriété de base de la topologie faible), U contient une intersection d'un nombre fini de $\{t / |f_i(x - t)| < \epsilon_i\}$.

• Les f_i étant en nombre fini, l'intersection de leurs noyaux ne saurait être réduite à 0 (en effet sinon l'application qui à t associe $(f_1(t), ..., f_n(t))$ serait injective, et donc la dimension de E serait finie).

• On peut donc choisir y non nul tel que $f_i(y) = 0$ pour tout i.

• $x + \lambda.y$ est dans U pour tout λ.

• $\|x + \lambda.y\|$ est minoré par $|\ |\lambda|.\|y\| - \|x\|\ |$, et donc en faisant tendre λ vers $\pm\infty$ on conclut que U intersecte S.

• On en déduit d'un coup que la boule ouverte de rayon 1 n'est pas ouverte, que la sphère de rayon 1 n'est pas fermée, et que l'adhérence de la sphère de rayon 1 contient au moins la boule fermée de rayon 1. La boule unité fermée du dual d'un espace vectoriel normé étant compacte pour la topologie faible (voir théorème 193), cette boule est fermée pour la topologie faible[2] ; et donc l'adhérence de la sphère unité est bien la boule unité fermée.$\square$

20.8 Les topologies sur E'

Rappelons que E' est le dual de E, c'est à dire l'ensemble des formes linéaires continues sur E. En tant que dual d'un espace vectoriel normé , E' est un Banach, c'est à dire qu'il est normé complet.

E' est muni naturellement de deux topologies déjà vues ; d'une part la topologie forte (c'est à dire la topologie de la norme $\|.\|_\infty$, avec $\|f\|_\infty = sup_{x\in S}\|f(x)\|$ - S étant la sphère unité), d'autre part la topologie faible - définie par rapport à son dual $(E')'$, c'est à dire le bidual E'' de E.

On va introduire une troisième topologie, encore moins fine que la topologie faible ; la topologie faible-*.

[2]On peut aussi éviter l'utilisation de 193 en disant que $\overline{B}(0,1)$ est un convexe fermé pour la topologie forte, donc est fermée pour la topologie faible (théorème 735)

20.8.1 La topologie faible-*

Pour plus d'informations on consultera le livre [2].

Définition 743 *On définit une* **injection canonique de** E **dans son bidual** E'' *par* $x \mapsto (f \mapsto f(x))$. *A tout élement de* E *on associe donc une forme linéaire continue sur* E' *(il s'agit donc bien d'un élément de* E''. *On notera* ϕ_x, *pour* x *dans* E, *l'application qui à* f *dans* E' *associe* $f(x)$.

Proposition 744 • *Il s'agit bien d'une injection (voir résultat 732).*

• *Il s'agit d'une isométrie (voir corollaire 694).*

• *Il ne s'agit pas nécessairement d'une bijection ; c'est toutefois le cas lorsque* E *est de dimension finie ou est un espace de Hilbert. Par définition, l'espace* E *est dit* **réflexif** *lorsqu'il s'agit d'une bijection.*

Définition 745 *La* **topologie faible étoile**, *alias* **topologie faible-***, *est la topologie engendrée par la famille des* ϕ_x *pour* x *dans* E.

On notera $f_n \overset{*}{\rightharpoonup} f$ *la convergence de la suite* f_n *vers* f *dans* E' *pour la topologie faible *.*

Proposition 746 • *La topologie faible * est séparée.*

• $f_n \overset{*}{\rightharpoonup} f$ *si et seulement si pour tout* x $f_n(x) \to f(x)$.

• *Convergence forte* $\geq$ *convergence faible* $\geq$ *convergence faible-**

• *Si* $f_n \overset{*}{\rightharpoonup} f$ *alors* $\|f_n\|$ *est bornée et* $\|f\| \leq liminf \|f_n\|$

• *Si* $f_n \overset{*}{\rightharpoonup} f$ *et* $x_n \to x$ *alors* $f_n(x_n) \to f(x)$

20.8.2 Un résultat utilisant le théorème d'isomorphisme de Banach, le théorème d'Ascoli et le théorème de Riesz

Théorème 747 *Soit* E *l'espace vectoriel des applications continues de* $[0,1]$ *dans* $\mathbb{R}$ *muni de* $\|f\|_0 = max_{[0,1]}\|f(t)\|$. *Alors tout sous-espace vectoriel de* E *formé de fonctions* C^1 *et fermé (pour la topologie de* $(E, \|.\|_0)$*) est de dimension finie.*

Démonstration : Soit F un tel sous-espace.

• F est un Banach pour la norme $\|.\|_1$, où $\|f\|_1 = \|f\|_0 + \|f'\|_0$. Prouvons-le :

- Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $(F, \|.\|_1)$.

- f_n est aussi de Cauchy dans $(F, \|.\|_0)$.

- f'_n aussi.

- f_n et f'_n convergent donc vers deux fonctions, disons f et g respectivement, continues.

- pour tout t dans $[0,1]$, $\int_0^t f'_n(u)du \to \int_0^t g(u)du$.

- donc $f_n(t) - f_n(0) \to \int_0^t g$, donc $f(t) - f(0) = \int_0^t g$

- donc $g = f'$

- donc $f_n \to f$ dans $(F, \|.\|_1)$.

• La norme $\|.\|_1$, définie par $\|f\|_1 = \|f\|_0 + \|f'\|_0$ pour f C^1, est majorée sur F par $A.\|f_0\|$, pour un certain $A > 0$. Prouvons-le :

- Considérons l'application J identité de $(F, \|.\|_1)$ dans $(F, \|.\|_0)$, où $\|f\|_1 = \|f\|_0 + \|f'\|_0$.

- F est fermé dans $(E, \|.\|_0)$, donc $(F, \|.\|_0)$ est un Banach.

- F est aussi un Banach dans $(E, \|.\|_1)$ (voir le • précédent).

- J est linéaire, continue, bijective entre 2 Banach ; c'est donc un homéomorphisme.

- ainsi par le théorème d'isomorphisme de Banach (256), on a bien le résultat annoncé.

• Soit maintenant $B = \overline{\{f \in F / \|f\|_0 \leq 1\}}$. Alors B est équicontinue. Prouvons-le :

- $\|f\|_1 \leq A.\|f\|_0 \leq A$, pour tout $f \in B$, donc par l'inégalité des accroissements finis, B est équicontinue.

• Pour tout x dans $[0,1]$, l'ensemble des $f(x)$ pour f dans B est inclus dans $[-1,1]$, par définition de B ; donc cet ensemble est relativement compact.

• Par le théorème d'Arzéla-Ascoli, et grâce aux deux points précédents, B est relativement compacte. B est fermée par définition.

• Par le théorème de Riesz 192, F est donc de dimension finie.$\square$

Chapitre 21

Théorie des groupes

21.1 Les bases

21.1.1 Définition d'un groupe

Définition 748 *Un* **groupe** *est un ensemble G, muni d'une* **loi de composition interne** *(lci), c'est à dire une application de $G \times G \to G$, généralement notée par la concaténation ($(x, y) \mapsto xy$), vérifiant :*
• $(xy)z = x(yz)$
• $\exists 1 / \forall x \; x.1 = 1.x = x$ *; 1 est dit l'élément neutre*
• $\forall x \exists x^{-1} / xx^{-1} = x^{-1}x = 1$

Pour vérifier qu'un ensemble muni d'une loi est bien un groupe, il suffit de vérifier que les deux premiers • sont vérifiés, et que pour tout x il existe x^{-1} tel que $xx^{-1} = 1$.

Définition 749 *Un groupe G est dit* **commutatif ou abélien** *si $xy = yx$. Dans ce cas on note souvent additivement ; l'élément neutre est alors noté 0, et x^{-1} est noté $-x$.*

Définition 750 *G est un p-***groupe***, avec p premier, si G est de cardinal une puissance de p.*

21.1.2 Sous-groupe

Définition 751 (Sous-groupe) *$H \subset G$ est un* **sous groupe de** *G si et seulement si :*
• $1 \in H$
• $(x, y) \in H^2 \rightarrow xy \in H$
• $\forall x \ x^{-1} \in H$

NB : un sous-groupe est un groupe, et un groupe inclus dans un groupe (pour les mêmes lois bien sûr) est un sous-groupe de ce groupe.

On peut noter les conditions dessus plus simplement : $1 \in H \wedge HH \subset H \wedge H^{-1} \subset H$

Définition 752 *Deux sous-groupes A et B sont dits* **conjugués** *s'il existe g tel que $A = g.B.g^{-1}$.*
Etant donné H sous-groupe de G, le **normalisateur** *de H est $N_G(H) = \{g \in G / gHg^{-1} = H\}$.*
Un sous-groupe N est dit **distingué** *(ou* **normal***) si pour tout g $gNg^{-1} = N$; on note $N \lhd G$.*
Un sous-groupe N est dit **caractéristique** *si il est stable par tout automorphisme intérieur.*
Un groupe est dit **simple** *si ses seuls sous-groupes distingués sont $\{1\}$ et G.*
L'ensemble des x tels que x commute avec tout élément est appelé le **centre** *d'un groupe. Le centre est un sous-groupe. On note $Z(G)$ le centre de G.*

Il faut bien voir ce que dit la définition du normalisateur - le normalisateur de H "fait" de H un sous-groupe normal, au sens où H est normal dans son normalisateur. En fait le normalisateur est le plus grand sous-groupe contenant H dans lequel H est distingué.

Propriétés :
• Un sous-groupe est distingué si et seulement si son normalisateur est le groupe tout entier.
• Un sous-groupe est distingué si et seulement si il n'est conjugué à aucun autre sous-groupe.
• Un sous-groupe caractéristique est distingué (évident).
• Tout sous-groupe d'un groupe abélien est distingué ; par contre, en considérant H_8 le groupe des quaternions, on peut constater qu'il n'y a pas de réciproque (voir 21.10).
• $\{1\}$ et G sont toujours à la fois des sous-groupes distingués et caractéristiques.
• Le centre d'un groupe est caractéristique et distingué.

Exemples :
• $\mathbb{Z}/p\mathbb{Z}$ est simple (en effet ses seuls sous-groupes sont ses sous-groupes triviaux, donc ses seuls sous-groupes distingués sont ses sous-groupes triviaux...)
• $\mathcal{U}_n$ est simple (voir 21.10.11)

Définition 753 *On appelle* **commutateur** *de x et y l'élément $x.y.x^{-1}.y^{-1}$. On appelle* **groupe dérivé** *d'un groupe le sous-groupe engendré[a] par les commutateurs. On note $D(G)$ le groupe dérivé de G.*

[a]Voir paragraphe 21.1.5 pour la définition de sous-groupe engendré par une partie.

Il faut bien noter que l'ensemble des commutateurs n'est pas nécéssairement un groupe ; le groupe dérivé est le sous-groupe engendré par l'ensemble des commutateurs.

Propriétés :

• $D(G)$ est distingué et même caractéristique dans G.

21.1.3 Homomorphismes

Définition 754 *On appelle* **homomorphisme** *du groupe G dans le groupe G' une fonction ϕ telle que $\phi(xy) = \phi(x)\phi(y)$. On note $Hom(G, G')$ l'ensemble des homomorphismes de G dans G'.*

On montre que pour tout tel ϕ :

• $\phi(1) = 1$

• $\phi(x^{-1}) = \phi(x)^{-1}$

La fonction constante égale à 1 est un homomorphisme de G dans G' ; éventuellement ce peut être le seul.

L'inverse d'un homomorphisme bijectif est un homomorphisme bijectif.

Proposition 755 *L'ensemble des* **automorphismes**, *i.e. des* **endomorphismes** *bijectifs, i.e. homomorphismes de G dans G bijectifs, noté $Aut(G)$, est un groupe.*

Exemples :

a) G groupe, $x \in G$

$\phi_x \in Hom(\mathbb{Z}, G)$, avec $\phi_x(n) = x^n$; le plus petit n tel que $\phi_x(n) = 1$, s'il existe est appelé ordre de x.

b) G groupe, $g \in G$

La fonction α_g, $x \mapsto gxg^{-1}$ est un automorphisme de G, dit automorphisme **intérieur** associé à g, appelée aussi **conjugaison** par g. En outre la fonction $g \mapsto \alpha_g$ est un homomorphisme de G dans $Aut(G)$. Son noyau est le centre de G.

L'ensemble des automorphismes intérieurs d'un groupe est un sous-groupe de l'ensemble des automorphismes du dit groupe.

Quelques propriétés :

Proposition 756 *(G, G') groupes, $\phi \in Hom(G, G')$*
• $Ker\ \phi := \{g \in G / \phi(g) = 1\}$
est un sous-groupe distingué de G.
• $Im\ \phi$ *est un sous-groupe de G'.*
• ϕ *injectif* $\iff Ker\phi = \{1\}$

21.1.4 Extensions

Définition 757 *On appelle* **suite exacte** *un schéma comme suit :*

$$1 \to A \xrightarrow{i} B \xrightarrow{s} C \to 1$$

Cela signifie que A, B et C sont des groupes, et que
• *i est un homomorphisme injectif de A dans B*
• *s est un homomorphisme surjectif de B dans C*
• $Ker\ s = Im\ i$
(on note 0 au lieu de 1 lorsque les groupes sont notés additivement)
Lorsque i et s ne sont pas précisés, cela signifie simplement que l'on peut trouver de tels i et s.
On dit alors que B est une **extension** *de A par C. Si en outre il existe $\overline{C}$ sous-groupe de B tel que la restriction de s à $\overline{C}$ est un isomorphisme, alors on dit que $\overline{C}$ est un* **relèvement**. *Cela est équivalent à dire qu'il existe un homomorphisme t de C dans B tel que $s \circ t = Id_C$. S'il y a un relèvement, l'extension est dite* **scindée**. *t est appelée* **section** *de s.*

21.1.5 Sous-groupe engendré

Proposition 758 *Soit G un groupe, X inclus dans G.*

Il existe un plus petit sous-groupe H de G contenant X. On peut le définir de deux façons :
(i) H est l'intersection de tous les sous-groupes contenant X
(ii) H est l'ensemble des produits finis d'éléments de $X \cup X^{-1}$.

Démonstration :
(i) est évident car l'intersection de deux sous-groupes est un sous-groupe.
(ii) on procède en trois points :
• K ainsi défini est un sous-groupe
• $X \subset K$ donc par (i) $H \subset K$

• $K \subset H$ est clair□

Définition 759 *On note $H =< X >$, H est appelé groupe* **engendré** *par X, et X est appelée* **partie génératrice** *de H. Si X est réduit à un seul élément x on note souvent $H =< x >$ au lieu de $H =< \{x\} >$.*
Un groupe est dit **monogène** *s'il est engendré par un seul élément. On appelle groupe* **cyclique** *un groupe monogène fini.*
On appelle **ordre d'un élément** *le cardinal du groupe engendré par cet élément.*

Si deux homomorphismes coïncident sur une partie génératrice d'un groupe, alors ils coïncident sur l'ensemble du groupe.

Cela sera utile pour la proposition 789.

Définition 760 *On dit que G est* **de type fini** *si $\exists X$ fini qui engendre G.*

Ainsi $\mathbb{Z}$, $\mathbb{Z}^n$ sont de type fini, et tout groupe fini est de type fini.

tout groupe de type fini est dénombrable.
Il n'y a pas équivalence, car par exemple $(\mathbb{Q}^*, \times)$ n'est pas de type fini (preuve en considérant des générateurs et leurs décompositions en facteurs premiers)
$(\mathbb{Q}, +)$ non plus (considérer l'inf de l'intersection avec $\mathbb{R}^+$ d'un groupe de type fini, en réduisant au même dénominateur)

Proposition 761 *Le groupe engendré par un ensemble réduit à un élément x est commutatif, et est l'ensemble des x^n avec $n \in \mathbb{Z}$. Il est isomorphe à $\mathbb{Z}$ ou à $\mathbb{Z}/n\mathbb{Z}$.*

Démonstration : $\{x^n\}$ est un groupe et contient x, donc il est inclus dans $< x >$; s'il est fini alors il existe n tel que $x^p = x^{p+n}$, et donc $x^n = 1$, et donc $< x >= \{x^0, ..., x^{n-1}\}$.□

Proposition 762 *Tout sous-groupe d'un groupe cyclique est cyclique.*

Démonstration : Un tel sous-groupe H de G est évidemment fini. Notons ensuite a un générateur du groupe ; le groupe est donc de la forme $a^0, ..., a^{n-1}$. Soit $p > 0$ minimal tel que $a^p \in H$;

21.2 Groupe quotient

21.2.1 Rappel : ensemble quotient

Soit X un ensemble, et $\mathcal{R}$ une relation d'équivalence sur X ; l'ensemble des classes pour $\mathcal{R}$ est une partition de X. Cet ensemble de classes, noté $X/\mathcal{R}$, est appelé **ensemble quotient** de X par $\mathcal{R}$. La classe d'un élément est notée $\Pi(x)$, x est dit un représentant de $\Pi(x)$. Π est appelée **surjection canonique**.

Il y a en fait ainsi bijection entre l'ensemble des relations d'équivalence et l'ensemble des partitions en parties non vides. A toute relation d'équivalence $\equiv$ on peut associer une fonction f telle que $x \equiv y \iff f(x) = f(y)$ (il suffit pour le montrer de considérer la fonction Π).

Etant donnée une fonction définie sur X, on peut définir $\overline{f}$ **fonction quotient** si f est constante sur les classes d'équivalences, $\overline{f}$ étant alors définie par $\overline{f}(\Pi(x)) = f(x)$.

21.2.2 Le cas des groupes

Définition 763 *H sous-groupe de G*
On définit les **classes à gauche suivant** *H comme les xH, $x \in G$, et les* **classes à droite suivant** *H comme les Hx.*
On note G/H l'ensemble des classes à gauche, $H \setminus G$ l'ensemble des classes à droite.
On note $(G : H)$ le cardinal de G/H quand celui-ci est fini.

On travaille généralement sur G/H plutôt que sur $H \setminus G$.

Proposition 764 *Les classes à gauche déterminent une partition de G en parties non vides. Pareil pour les classes à droite.*

Proposition 765 *N sous-groupe de G, il y a équivalence entre les trois assertions suivantes :*
• *N distingué*
• *$gN = Ng$ pour tout g*
• *il existe une structure de groupe sur le quotient G/N telle que Π soit un homomorphisme.*

On voit donc que dans ce cas $G/H = H \setminus G$. Cette propriété d'un sous-groupe distingué est fondamentale : la partition en classes à droite est égale à la partition en classes à gauche. Ce fait est caractéristique des sous groupes distingués.

Proposition 766 *G et G' deux groupes, N sous-groupe distingué de G, $\phi \in Hom(G, G')$; alors les deux assertions suivantes sont équivalentes :*
- *Il existe $\overline{\phi}$ de G/N dans G' tel que $\phi = \overline{\phi} \circ \Pi$*
- *$N \subset Ker\ \phi$*

Dans ce cas $\overline{\phi}$ est unique, et est un homomorphisme de groupes de G/N dans G'.

En particulier, $\overline{\phi}$ est une injection si $N = Ker\ \phi$, et induit un isomorphisme de $G/Ker\ \phi$ dans $Im\phi$.

Les preuves de ces faits sont faciles, et sont logiques intuitivement ; si on quotiente par quelque chose de trop gros par rapport au noyau, alors on n'a plus la précision requise pour reconstruire la fonction...

$G/D(G)$ est un groupe abélien, c'est d'ailleurs son plus grand quotient abélien, et $D(G)$ est le seul sous-groupe à avoir cette propriété.

Théorème 767 (Factorisation d'homomorphismes) *Soit G un groupe, H un sous-groupe distingué de G, ϕ un homomorphisme de G vers un groupe G'. Alors si $H \subset Ker\ \phi$, alors il existe une application $\tilde{\phi}$[a] telle que*

$$\phi = \tilde{\phi} \circ p$$

avec p la projection canonique de G sur G/H.

[a] De G/H dans G'.

Cela servira par exemple pour le théorème 795.

Démonstration : Si facile que vous le prouver serait une injure... La fonction $\tilde{\phi}$ est bien définie, car si deux éléments on même image par p alors ils ont même image par ϕ, et l'application $\tilde{\phi}$ est bien un homomorphisme car ϕ en est un (la vérification de cette implication est facile).$\square$

21.2.3 Le théorème de Lagrange

Définition 768 *On appelle* **indice de H dans** *G, avec H un sous-groupe de G, le cardinal de G/H.*

Un théorème fondamental :

Théorème 769 (Théorème de Lagrange) *Soit G un groupe fini, et H un sous-groupe de G, alors*

$$|G| = |H|.|G/H|$$

Démonstration : Il suffit de montrer que chaque classe d'équivalence est de même cardinal, et que ce cardinal est $|H|$ (chose facile à prouver !).$\square$

On remarque qu'il n'est absolument pas nécéssaire que H soit distingué.

21.3 Opération d'un groupe sur un ensemble

Définition 770 *Avec G un groupe et X un ensemble, on appelle* **action à gauche** *de G sur X une application α de $G \times X$ dans X telle que :*
• $\alpha(1, x) = x$
• $\alpha(g, \alpha(h, x)) = (g.h, x)$
On dit aussi que G **opère à gauche** *sur X où que G est une* **opération à gauche** *sur X. Usuellement on note plus simplement $g.x$ au lieu de $\alpha(g, x)$. Les deux conditions deviennent alors :*
• *1.x=x*
• *(g.h).x=g.(h.x)*
On définit de manière symétrique une **action à droite***. Une* **action** *sans plus de précision désigne une action à gauche. On dit que X est un G-ensemble.*

Propriétés :
• $g.x = y \iff g^{-1}.y = x$
• Etant donné x et y dans X il n'est pas du tout nécéssaire qu'il existe un g tel que $g.x = y$.
• Si G opère sur X alors tout sous-groupe H de G opère sur X pour la loi restreinte.

L'équivalence suivante est fondamentale : se donner une action de G sur X revient à se donner un homomorphisme ϕ de G dans le groupe $\sigma(X)$ des bijections de X ($g.x = \phi(g).x$).

Un exemple fondamental est l'action d'un groupe sur lui-même ; l'action est en fait simplement la loi du groupe. Il est clair que les conditions sont vérifiées.

Pour le cas des actions à droite, il faut noter que si on a une action à droite $a_1(x, g)$, alors $a_2(g, x) \rightarrow a_1(x, g^{-1})$ est une action à gauche du groupe opposé (le groupe opposé à G étant G muni de $(x, y) \rightarrow yx$). On travaillera à peu près toujours avec des classes à gauche, les résultats étant les mêmes, et puisqu'on peut reformuler le problème en terme d'action à gauche.

Définition 771 *Etant donnés X et X' deux G-ensembles, on appelle* **G-homomorphisme** *de X vers X' une application ϕ de X dans X' telle que $\phi(g,x) = g.\phi(x)$ pour tous $x \in G$ et $g \in G$. On note $Hom(X,X')$ l'ensemble des homomorphismes de X sur X'. Comme d'habitude, un* **isomorphisme** *est un homomorphisme bijectif.*

Un exemple facile et classique :
Soit G un groupe et X un G-ensemble. L'application ϕ_x pour $x \in G$ qui à g dans G associe $g.x$ est un homomorphisme de G (en tant que G-ensemble) sur X (en tant que G-ensemble).

Démonstration : Soit g dans G et y dans G (y est pris dans G en tant que G-ensemble) alors $\phi_x(g.y) = g.y.x$ et $g.\phi_x(y) = g.y.x$.$\square$

Définition 772 *On note H_x ou G_x et on appelle* **stabilisateur** *ou* **fixateur** *de x l'ensemble des g tels que $g.x = x$. C'est un sous-groupe de G, qui n'est pas nécessairement distingué.*
On appelle **G-orbite** *de x appartenant à X (ou plus simplement* **orbite** *s'il n'y a pas de risque de confusion) et on note $\omega(x)$ ou $G.x$ la classe d'équivalence de x pour la relation $\mathcal{R}$ définie par $a\mathcal{R}b \iff \exists g \in G/g.a = b$ (il est facile de vérifier qu'il s'agit bien d'une relation d'équivalence).*
Un G-ensemble est dit **homogène** *s'il ne contient qu'une seule orbite.*
On dit que $x \in X$ est un **point fixe***, si l'orbite de x est réduite à $\{x\}$.*
On dit que G opère **transitivement** *si $\forall x \forall y \exists g / y = g.x$.*
On dit que G opère k **fois transitivement** *si $\forall (x_i)_{i\in\{1,\ldots,k\}} \forall (y_i)_{i\in\{1,\ldots,k\}} (i \neq j \rightarrow x_i \neq x_j \wedge y_i \neq y_j) \rightarrow \exists g \forall i \in \{1,...,k\}/y_i = g.x_i$.*
On dit que G opère **fidèlement** *si $\forall x\ g.x = x \rightarrow g = 1$.*

Proposition 773 *Lorsque G est fini, on a pour tout x dans X, $|\omega(x)|.|H_x| = |G|$.*

Démonstration : On constatera simplement que l'application qui à $\overline{g}$ associe $g.x$ de G/H_x dans $\omega(x)$ est une bijection.$\square$

La figure 21.1 tâche de montrer l'allure générale d'un G-ensemble.

Propriétés :
• Chaque orbite est un ensemble homogène.

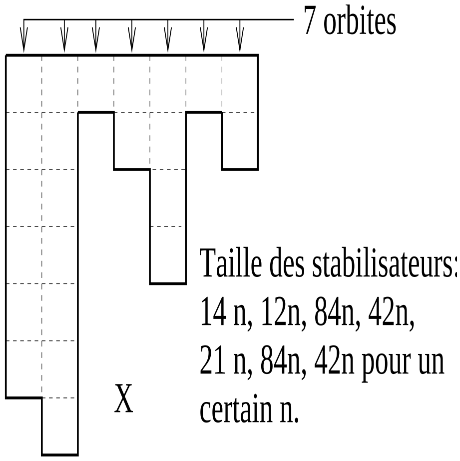


FIG. 21.1 – Exemple de G-ensemble X. Les séparations verticales sont les séparations entre les orbites, qui réalisent une partition de X. L'action n'étant pas nécessairement injective, les orbites ne sont pas nécessairement de même cardinal que G. A l'intérieur d'une même orbite, le stabilisateur est toujours le même à conjugaison près, et en particulier, les stabilisateurs dans une même orbite sont équipotents. Si le groupe et l'ensemble sont finis, le cardinal du groupe est le produit du cardinal de l'orbite par le cardinal d'un stabilisateur de cette orbite. On notera que le cardinal du groupe agissant sur cet ensemble est au moins 84 (ppcm des cardinaux des orbites).

Proposition 774 *G groupe, X et X' des G-ensembles homogènes, alors les assertions suivantes sont équivalentes :*
• $X \simeq X'$
• $\exists (x, x') \in X \times X' / H_x = H_{x'}$
• $\exists (x, x') \in X \times X' / H_x$ *est conjugué à* $H_{x'}$
• $\forall (x, x') \in X \times X' / H_x$ *est conjugué à* $H_{x'}$

Démonstration : laissée en exercice. □

Exemples classiques :
• Le groupe orthogonal $O(3, \mathbb{R})$ opère sur $\mathbb{R}^3$; les orbites sont les sphères de centre l'origine, le stabilisateur de 0 est $O(3, \mathbb{R})$ tout entier, et le stabilisateur d'un point quelconque autre que 0 est l'ensemble des rotations d'axe la droite vectorielle engendrée par ce point et des symétries par rapport à un sous-espace vectoriel passant par ce point. 0 est un point fixe.
• On peut faire opérer G sur ses sous-groupes par conjugaison, avec $g.H = gHg^{-1}$. Le stabilisateur d'un point (c'est à dire d'un sous-groupe) est alors le normalisateur de ce point (ie de ce sous-groupe).
• Si X est un espace topologique et est un G-ensemble tel que pour tout $g \in G$ l'application $y \mapsto g.y$ est un homéomorphisme, alors on dit que G **agit sur** X **par homéo-**

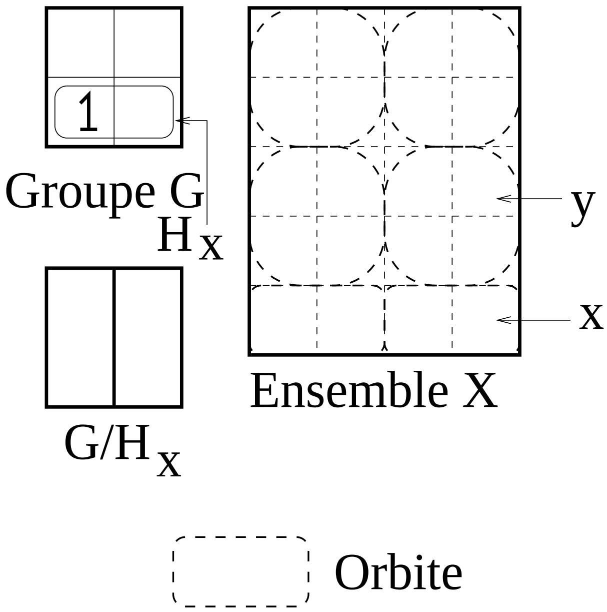


FIG. 21.2 – x est un élément de X. G est d'ordre 2, $\omega(x)$ est de cardinal 2, donc G/H_x est de cardinal 2 (il est en bijection avec $\omega(x)$), et H_x est de cardinal $\frac{4}{2} = 2$.

morphismes. La topologie quotient pour la relation d'équivalence "être dans la même orbite" vérifie des propriétés intéressantes (voir proposition 144 et théorème 1275).

Proposition 775 *Un G-ensemble homogène est isomorphe à un quotient G/H de G pour l'action de G sur G/H par translation à gauche.*

Pour bien voir l'intérêt de cette remarque, il faut se rappeler que tout G-ensemble est partitionné naturellement en orbites, qui sont des G-ensembles homogènes, et que donc on peut identifier à des actions par translation d'un groupe sur un groupe quotient.

Proposition 776 (Sur l'ensemble des points fixes) *Etant donnés G un p-groupe et X un ensemble sur lequel agit G, le cardinal de l'ensemble des points fixes de X pour G est congru au cardinal de X modulo p.*

Démonstration : Le cardinal des orbites divise le cardinal de G, donc le cardinal de l'union des orbites est congru au nombre d'orbites de cardinal 1 modulo p. Le cardinal de l'union des orbites est le cardinal de X.$\square$

21.4 Produits

Il faut bien noter que même si de nombreuses applications des résultats ci-dessous se font avec des groupes finis, ils sont valables pour des groupes quelconques.

21.4.1 Produit direct

Définition 777 (Produit direct de deux groupes) *On appelle* **produit direct** *de deux groupes N et H et on note $N \times H$ le produit cartésien des groupes N et H muni du produit terme à terme*

$$(n, h).(n', h') = (nn', hh')$$

La fonction p_2 qui à (n, h) associe h est appelée **projection** *de $N \times H$ sur H.*
La fonction p_1 qui à (n, h) associe n est appelée **projection** *de $N \times H$ sur N.*
On définit alors la généralisation à un produit d'un nombre quelconque de groupes par $\Pi_{i \in I} G_i$. La loi

$$(g_i)_{i \in I}(h_i)_{i \in I} = (g_i h_i)_{i \in I}$$

muni le produit d'une structure de groupe ; on appelle ce groupe le **groupe produit**.
On définit aussi le **produit restreint** *des G_i comme étant le sous-groupe du produit des G_i des éléments $(g_i)_{i \in I}$ ne comportant qu'un nombre fini de g_i différents de l'élément neutre. S'il s'agit d'un produit d'un nombre fini de groupes il est clair que le produit restreint est égal au produit.*

Propriétés :
• p est surjective, c'est un morphisme surjectif, son noyau est distingué et isomorphe à N. On a une suite exacte

$$1 \rightarrow N \xrightarrow{i} N \times H \xrightarrow{p} H \rightarrow 1$$

avec $i(n) = (n, 1)$.
• $N \times \{1\}$ est le noyau de p, il est distingué ; et $\{1\} \times H$ est distingué aussi.

21.4.2 Produit semi-direct

Définition 778 (Produit semi-direct) *Etant donnés deux groupes N et H, et un morphisme de groupe ϕ de H dans l'ensemble des automorphismes de N ; alors on appelle* **produit semi-direct de N et H relativement à ϕ** *et on note $N \rtimes H$ le produit cartésien $N \times H$ muni de la loi $(n, h).(n', h') = (n.\phi(h)(n'), hh')$.*

On notera que formellement il faudrait préciser $N \rtimes_\phi H$.

Proposition 779 • *Le produit semi-direct $N \rtimes H$ a une structure de groupe*
• *On a une suite exacte*

$$1 \to N \xrightarrow{i} N \rtimes H \xrightarrow{s} H \to 1$$

avec $i(n) = (n, 1)$ et $s(n, h) = h$.
• *$i(N)$, c'est à dire $N \times \{1\}$ est distingué, mais pas $\{1_N\} \times H$ (contrairement au cas du produit direct).*

Démonstration : Vérification facile.□

Remarque importante :
En identifiant $\overline{N}$ et $N \times \{1\}$ d'une part, H et $\{1\} \times H$ d'autre part, on constate qu'un produit semi-direct peut toujours s'écrire comme produit semi-direct de deux sous-groupes lié au morphisme ϕ de G dans $Aut(N)$ défini par $(\phi(h))(n) = hnh^{-1}$.

21.4.3 Identifier un produit direct ou semi-direct

Cette partie est fondamentale pour ramener l'étude d'un groupe à l'étude de groupes plus petits (tâche fondamentale en théorie des groupes !).

⊡ Identification d'un produit semi-direct

Proposition 780 (Décomposition en produit semi-direct) *Si on a une suite exacte*

$$1 \to N \xrightarrow{i} G \xrightarrow{s} H \to 1$$

(c'est à dire si i est injective, si s est surjective, et si $Ker\ s = Im\ i$) et si on a un sous-groupe $\overline{H}$ de G sur lequel la restriction de s est un isomorphisme vers H [a], alors G est isomorphe à $i(N) \rtimes \overline{H}$ relativement à la loi de l'automorphisme intérieur (voir la remarque de 21.4.2).
On peut donc aussi dire que G est isomorphe à $N \rtimes H$, i étant un isomorphisme de N sur $\overline{N}$, et s étant un isomorphisme de $\overline{H}$ sur H.

[a] C'est-à-dire un relèvement, une section, voir la partie 21.4.2

Démonstration : On considère $\overline{N}$ l'image de i, et $\overline{H}$ le sous-groupe de G sur lequel la restriction de s est un isomorphisme vers H.
Puisque $\overline{N} = Ker\ s$, $\overline{N} \lhd G$ (un noyau de morphisme de groupe est toujours distingué).Il est clair que :
• $\overline{N} \cap \overline{H} = \{1\}$
• $G = \overline{N}.\overline{H}$
Le premier point est évident, du fait que s est un isomorphisme depuis $\overline{H}$, et a donc un noyau nul.

Pour le deuxième point, soit $g \in G$, alors $s(g) = s(h)$ avec $h \in H$, et $s(g.h^{-1}) = s(g).s(h)^{-1} = 1$, donc $g.h^{-1} \in N$. L'écriture d'un élément de G comme produit d'un élément de N par un élément de H est unique (facile, au vu de $\overline{N} \cap \overline{H} = \{1\}$) ; G est donc ainsi en bijection avec $\overline{N} \times \overline{H}$, par $\phi(nh) = (n, h)$. On cherche maintenant à établir une loi sur $\overline{N} \times \overline{H}$ telle que cette bijection soit un isomorphisme.
Le produit de $n.h$ par $n'h'$ est $n.h.n'h'$, que l'on doit donc exprimer comme un produit d'un élément de $\overline{N}$ par un élément de $\overline{H}$; on peut réécrire $n.h.n'.h'$ sous la forme $n.(h.n'.h^{-1}).h.h'$; puisque N est distingué, il s'agit bien du produit de $n.h.n'.h^{-1}$ (élément de $\overline{N}$) par $h.h'$ (élément de $\overline{H}$).
On vérifie facilement que la loi $(n, h).(n', h') = (n.(h.n'.h^{-1}), h.h')$ fait de cette bijection un morphisme.□

Remarque importante : L'hypothèse revient exactement à avoir une extension scindée, c'est à dire une extension munie d'un relèvement (voir 21.1.4).

Proposition 781 *Si G est un groupe, si N et H sont des sous-groupes de G, si $N \lhd G$, si $N \cap H = \{1\}$ et si $G = N.H$, alors $G \simeq N \rtimes H$.*

Démonstration : Il suffit de reprendre la preuve ci-dessus.□

⊡ Identification d'un produit direct

En fait un produit direct est un cas particulier de produit semi-direct.
En reprenant les notations de la définition du produit semi-direct et des démonstrations ci-dessus, on a équivalences entre les assertions suivantes :
• $\phi(h) = Id_N$ pour tout h
• $\overline{H}$ est distingué
• la loi de groupe sur $N \rtimes H$ est celle du produit direct

On peut aussi raisonner sur les suites exactes. Lorsque l'on a une suite exacte avec relèvement, i.e. avec une section, i.e. si l'extension est scindée, ET si $\forall (n, h) \in N \times H \ nh = hn$.

21.4.4 Quelques remarques pour éviter les gaffes

La condition de la proposition 780 est suffisante mais non nécéssaire ; on peut avoir une extension sans relèvement, c'est à dire non scindée, c'est à dire sans qu'il y ait de section, sans pour autant que le groupe ne soit pas le produit semi-direct de N par H.

On peut très bien avoir $A \times B$ (produit direct) $\simeq A \rtimes_\phi B$, avec ϕ autre que $\phi(h) = Id_N$ pour tout h ; donc il ne suffit pas de décomposer un groupe comme produit semi-direct non trivial pour conclure qu'il n'est pas un produit direct. [14] cite ainsi $\sigma_3 \times \mathbb{Z}/2.\mathbb{Z}$.

21.5 Théorèmes de Sylow. Groupes de Sylow

Les deux théorèmes de Sylow sont extraits du classique [14].

Définition 782 *On appelle p-**sous-groupe de Sylow** ou plus simplement p-**Sylow** d'un groupe G de cardinal n, un sous-groupe de G d'ordre p^r avec p premier divisant n et $n = p^r.m$ et $p \not| m$.*

Proposition 783 *Un sous-groupe P de G est un p-sous-groupe de Sylow de G si :*
• P est un p-groupe
• $(G : P)$ est premier à p.

Démonstration : Pas difficile, y'a qu'à l'écrire.□

Théorème 784 (Théorème de Sylow) *G étant un groupe fini, et p un nombre premier divisant l'ordre de G, alors G admet au moins un p-sous-groupe de Sylow.*

Démonstration : On va procéder par étapes.
• Tout d'abord un cas particulier : $\mathbb{Z}/p\mathbb{Z}$ est un corps fini puisque p est premier, et $GL(n, \mathbb{Z}/p.\mathbb{Z})$ est d'ordre $\Pi_{i=0}^{n-1}(p^n - p^i)$, comme on peut s'en convaincre en comptant les bases de $(\mathbb{Z}/p.\mathbb{Z})^n$. Le cardinal de ce groupe est donc $m.p^{n.(n-1)/2}$, avec $p \not| m$. Un p-Sylow de ce groupe est alors l'ensemble des matrices de la forme

$$n \text{ lignes} \left\{ \underbrace{\begin{pmatrix} 1 & * & * & \dots & * & * \\ 0 & 1 & * & \dots & * & * \\ 0 & 0 & 1 & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & * & * \\ 0 & 0 & \dots & 0 & 1 & * \\ 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}}_{n \text{ colonnes}} \right.$$

• On a maintenant besoin d'un lemme :

Lemme 785 *Soit G un groupe d'ordre $p^\alpha.m$, avec $p \not| m$ et p premier, et soit H un sous-groupe de G et S un p-Sylow de G. Alors il existe $a \in G$ tel que $a.S.a^{-1} \cap H$ soit un p-sylow de H.*

Démonstration : G opère sur G/S par translation à gauche (voir 21.9.1) ; le stabilisateur d'un élément $g.S$ est $g.S.g^{-1}$.
D'autre part H opère sur G/S par translation à gauche aussi ; le stabilisateur d'un élément $g.S$ est $g.S.g^{-1} \cap H$.
Il est clair que tout $a.S.a^{-1}$ est bien un p-groupe, il reste à en trouver un qui soit bien

un p-Sylow. Il suffit pour cela que le quotient du cardinal de H par le cardinal de $H \cap a.S.a^{-1}$ soit premier avec p ; donc il suffit que le cardinal de $H/(H \cap a.S.a^{-1})$ soit premier avec p.
Or ce cardinal est en fait le cardinal de l'orbite de $a.S$ dans G/S sous l'action de H ; or toutes ces orbites ne peuvent être de cardinal un multiple de p, sinon le cardinal de G/S serait un multiple de p, ce qui contredirait le fait que S est un p-Sylow.
Ce lemme est donc prouvé.$\square$
• Maintenant on peut s'attaquer au cas général ; soit G un groupe vérifiant les hypothèses ; G est isomorphe à un sous-groupe de σ_n par le théorème de Cayley (voir 21.9.1). A son tour, σ_n est isomorphe à un sous-groupe de $GL(n, \mathbb{Z}/p.\mathbb{Z})$ (on considère la base canonique $(e_i)_{i\in[1,n]}$ de $(\mathbb{Z}/p.\mathbb{Z})^n$, et l'application qui à σ dans σ_n associe l'application linéaire qui à e_i associe $\sigma(e_i)$.
Par le premier point, ce groupe admet un p-Sylow ; et par le deuxième point, un sous-groupe d'un groupe admettant un p-Sylow admet un p-Sylow.$\square$

Corollaire 786 *Si G un est groupe de cardinal $p^r.m$ avec $p \not| m$ et p premier, alors G possède des sous-groupes d'ordre p^q pour tout $q \leq r$.*

<u>Démonstration :</u> G contient un p-Sylow, donc un sous-groupe de cardinal p^r. On peut donc se ramener au cas des p-groupes. Le centre d'un p-groupe est non trivial, comme on le montre en 21.10.1. On considère donc G un p-groupe, et $Z(G)$ son centre, de cardinal p^r. En appliquant l'hypothèse de récurrence à $Z(G)$, on a bien des groupes d'ordre p^q, pour $q \leq r$. On considère maintenant le groupe-quotient de G par $Z(G)$, il est de cardinal p^{r-q}, on peut donc lui appliquer l'hypothèse de récurrence et y trouver un groupe de cardinal p^t pour $t \leq r - q$. En considérant l'image inverse par la projection canonique sur le groupe quotient, on obtient alors un groupe de cardinal p^{t+q}, pour $t \leq r - q$, donc pour tout cardinal p^u avec $q \leq u \leq r$. $\square$

Théorème 787 (Deuxième théorème de Sylow) *Etant donné G un groupe, de cardinal $|G| = p^r.m$, avec $p \not| m$.*
• Tout p-groupe inclus dans G est inclus dans un p-Sylow de G.
• Les p-Sylow sont tous conjugués
• Les p-Sylow forment une orbite de G sous l'action de G par automorphisme intérieur
• Un p-Sylow est distingué si et seulement si il est l'unique p-Sylow
• Le nombre de p-Sylow est congru à 1 modulo p et divise $|G|$
• Le nombre de p-Sylow divise m

<u>Démonstration :</u>
• Démonstration de l'affirmation "Tout p-groupe inclus dans G est inclus dans un p-Sylow de G" :
Supposons H un p-groupe de G. Soit S un p-Sylow de G, dont l'existence est donnée par le premier théorème de Sylow. D'après le lemme 785, il existe a dans G tel que $a.S.a^{-1} \cap H$ soit un p-Sylow de H.
H étant un p-groupe, H est nécéssairement égal à $a.S.a^{-1} \cap H$. Donc H est bien inclus dans un Sylow.

• Pour montrer l'affirmation "Les p-Sylow sont tous conjugués", il suffit de faire le même raisonnement avec H un p-Sylow.
• Démonstration de l'affirmation "Les p-Sylow forment une orbite de G sous l'action de G par automorphisme intérieur" :
On a vu que les p-Sylow étaient tous conjugués ; si un autre élément leur est conjugué, c'est aussi un p-Sylow ; le résultat est donc en fait complètement évident.
• Démonstration de l'affirmation "Un p-Sylow est distingué si et seulement si il est l'unique p-Sylow" :
Si un p-Sylow est distingué et s'il n'est pas unique alors il est conjugué à l'autre... donc il n'est pas distingué.
Réciproquement si il est unique, alors s'il n'est pas distingué, alors il est conjugué à un autre p-Sylow - donc il n'est pas unique.
• Démonstration de l'affirmation "Le nombre de p-Sylow est congru à 1 modulo p et divise $|G|$" :
On rappelle que la proposition 776 affirme que le nombre de points fixes d'un ensemble X sous l'action d'un p-groupe G est congru au cardinal de X modulo p.
Il suffit alors de considérer l'ensemble des p-Sylow ; on peut faire agir dessus un p-Sylow S quelconque par conjugaison. Le nombre de p-Sylow est donc congru au nombre de points fixes de l'ensemble des p-Sylow sous l'action de S modulo p. Il reste donc à montrer qu'il y a un unique point fixe. L'existence d'un point fixe est évidente, il s'agit de S lui-même. Supposons que T soit un autre point fixe, T est donc un p-Sylow tel que pour tout s dans S, $sTs^{-1} = T$. On considère le groupe engendré par T et S, S et T sont des p-Sylow de ce groupe. Dans ce groupe toujours, T est distingué ; donc il est l'unique p-Sylow, donc il est égal à S. D'où le résultat.
• Démonstration de l'affirmation "Le nombre de p-Sylow divise m" :
Le nombre de p-Sylow est le cardinal d'une orbite, donc il divise le cardinal de G, or il est congru à 1 modulo p, donc il divise m.□

21.6 Applications des groupes de Sylow

21.6.1 Démontrer qu'un groupe n'est pas simple juste au vu de son cardinal

Cet exemple est tiré de l'excellent [14].

Proposition 788 *Un groupe d'ordre 63 ne peut être simple.*

Démonstration : on considère les 7-Sylow de G d'ordre 63 ; ce nombre de p-Sylow divise $\frac{63}{7}$ donc 9, et est congru à 1 modulo p ; donc il y a un unique 7-Sylow, donc il est distingué, donc G n'est pas simple.□

21.7 Groupes abéliens

On rappelle qu'un groupe abélien est un groupe commutatif.

Proposition 789 *Un groupe abélien G est de type fini si et seulement si il existe un homomorphisme surjectif de $\mathbb{Z}^n$ sur G pour un certain n, c'est-à-dire s'il est isomorphe à un quotient de $\mathbb{Z}^n$ par un de ses sous-groupes*[a]*.*
Plus précisément, G est alors engendré par n éléments, si n est minimal.

[a]Notez qu'il peut s'agir du quotient de $\mathbb{Z}^n$ par n'importe quel sous-groupe, puisque $\mathbb{Z}^n$ étant commutatif, tous ses sous-groupes sont distingués

Cela nous servira pour la proposition 795.

Démonstration : En effet, supposons que G est finiment engendré, par $g_1, ..., g_n$. Considérons alors l'application de $\mathbb{Z}^n$ dans G définie par $(p_1, ..., p_n) \mapsto p_1.g_1 + ... + p_n.g_n$. Puisque G est engendré par les g_i ET G est commutatif, cette application est surjective. Il est clair que c'est un morphisme puisque G est commutatif. Donc G est isomorphe au quotient de $\mathbb{Z}^n$ par le noyau de ce morphisme, d'où le résultat.

Réciproquement, supposons que l'on ait un morphisme surjectif de $\mathbb{Z}^n$ sur G ; alors il est égal à l'homomorphisme $(p_1, ..., p_n) \mapsto p_1.g_1 + ... + p_n.g_n$, avec g_i l'image de $(0, ..., 0_{i-1\text{fois}}, 1, 0, ..., 0)$ (voir la remarque de la partie 21.1.5). Il est donc clair que G est engendré par les g_i.$\square$

Définition 790 (Somme) *Soit $(A_i)_{i\in I}$ une famille de groupes abéliens. On note $\bigoplus_{i\in I} A_i$ l'ensemble des familles $(x_i)_{i\in I}$ avec $x_i \in A_i$ et les x_i presque tous nuls ; c'est un groupe abélien pour l'addition terme à terme ; on l'appelle* **somme** *des groupes A_i.*
Si $i_0 \in I$, on identifie A_{i_0} à l'ensemble des $(x_i)_{i\in I}$ tels que $i \neq i_0 \rightarrow x_i = 0$.
Si $\forall i\ A_i = A$ alors on note $A^{(I)} = \bigoplus_{i\in I} A_i$.

Proposition 791 (Propriété universelle des groupes abéliens) *Etant donnée une famille $(A_i)_{i\in I}$ de groupes abéliens, A' un groupe abélien, ϕ_i un homomorphisme de A_i sur A', alors il existe un unique homomorphisme de $\bigoplus A_i$ vers A' tel que la restriction de cet homomorphisme à A_i soit ϕ_i.*

Démonstration : Considérer $\phi(x) = \sum_i \phi_i(x_i)$.$\square$

Définition 792 (Somme directe) *A étant un groupe abélien, les A_i étant des sous-groupes de A, alors :*
• les A_i sont dits en **somme directe** *si l'application canonique de $\bigoplus A_i$ dans A qui à $(x_i)_{i\in I}$ associe $\sum_i x_i$ est injective. On identifie alors son image avec $\bigoplus_{i\in I} A_i$.*
• On dit que A est **somme directe** *des A_i si l'application est bijective. On note alors (abusivement) $A = \bigoplus_{i\in I} A_i$.*

Proposition 793 *A abélien, (A_i) famille de sous-groupes, alors les A_i sont en somme directe si $\sum_i x_i = 0$ avec $x_i \in A_i$ (support fini) implique $\forall i x_i = 0$.*

Définition 794 (Groupe de torsion) *Un élément d'un groupe est dit* **élément de torsion** *s'il est d'ordre fini.*
Un groupe abélien est dit **de torsion** *si tous ses éléments sont d'ordre fini.*
Etant donné p un nombre premier, un groupe abélien est dit **de p-torsion** *si tous ses éléments sont d'ordre une puissance de p.*
Un groupe abélien est dit **libre** *s'il est isomorphe à $\mathbb{Z}^n$ pour un certain $n \in \mathbb{N}$.*
On appelle **sous-groupe de torsion** *d'un groupe abélien G le sous-groupe constitué par les éléments de torsion[a].*

[a]On vérifie facilement qu'il s'agit bien d'un sous-groupe.

Proposition 795 *Un groupe de torsion[a] et de type fini est fini.*

[a]Sous-entendu : abélien (un groupe de torsion est abélien par définition).

Démonstration : En effet, si G est de type fini et abélien, alors c'est un quotient de $\mathbb{Z}^n$, par la proposition 789.

On considère alors $ppcm$ le ppcm des ordres des n générateurs donnés par la proposition 789. L'ordre de tout élément est alors un diviseur de $ppcm$. L'homomorphisme surjectif de $\mathbb{Z}^n$ dans son quotient a pour noyau un ensemble contenant $(k\mathbb{Z})^n$. Donc il se factorise à travers $(\mathbb{Z}/k\mathbb{Z})^n$ (voir le théorème 767). Donc le groupe G est de cardinal plus petit que k^n.□

Les deux théorèmes ci-dessous sont donnés sans démonstration (laissées aux lecteurs pour exercice !).

Théorème 796 *• Tout groupe abélien sans torsion de type fini est libre.*
• Tout sous-groupe d'un groupe libre est libre.
• Deux groupes libres $\mathbb{Z}^n$ et $\mathbb{Z}^p$ sont isomorphes si et seulement si $n = p$.
• Tout groupe abélien de type fini est produit d'un groupe libre et d'un groupe de torsion. Cette décomposition est unique à isomorphisme près.

Théorème 797 *Tout groupe abélien fini G s'exprime de manière unique sous la forme*

$$\mathbb{Z}/a_1\mathbb{Z} \times \mathbb{Z}/a_2\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z}$$

avec $\forall i \in [1, n-1] a_i | a_{i+1}$, et $a_1 > 1$. Les a_i sont appelés **facteurs invariants** *du groupe.*
La décomposition ainsi obtenue est appelée **décomposition cyclique** *du groupe G.*

Cette décomposition a de nombreuses conséquences :

Corollaire 798 *Soit G un groupe abélien fini ; il existe un élément d'ordre le ppcm des ordres des éléments du groupe.*

Démonstration :
• Soit G un groupe abélien fini.
• La décomposition cyclique nous permet d'écrire G sous la forme

$$\mathbb{Z}/a_1\mathbb{Z} \times \mathbb{Z}/a_2\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z}$$

• On considère un élément $x = (x_1, \ldots, x_n)$ de ce produit.
• L'ordre de x est le ppcm des ordres des x_i ; or l'ordre de x_i divise a_i, qui lui-même divise a_n.
• Le ppcm des ordres est donc en fait un diviseur de a_n, donc c'est a_n lui-même.
• L'élément $(0, \ldots, 0, 1)$ convient donc (on peut remplacer 1 par n'importe quel générateur de $\mathbb{Z}/a_n\mathbb{Z}$).□
Autre conséquence :

Corollaire 799 *Soit G un groupe abélien fini. Pour tout diviseur d de $Card(G)$, il existe un sous-groupe H de G d'ordre d.*

Démonstration :
• On écrit G sous forme :

$$\mathbb{Z}/a_1\mathbb{Z} \times \mathbb{Z}/a_2\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z}$$

• Décomposons d en facteurs premiers :

$$d = \Pi_{i=1}^m p_i$$

• Définissons :

$$d_1 = pgcd(d, a_1)$$

$$d_2 = pgcd(\frac{d}{d_1}, a_2)$$

$$d_3 = pgcd(\frac{d}{d_1 d_2}, a_3)$$

$$\ldots$$

$$d_n = pgcd(\frac{d}{d_1 \dots d_n}, a_n)$$

• Puisque $d|n$ et $n = \Pi_{i=1}^n a_i$, on arrive à se "débarasser" de chaque facteur premier de d dans l'un des d_i, et donc $\Pi_{i=1}^n d_i = d$.

• Pour tout i, d_i divise a_i.

• Dans un groupe cyclique, il existe un sous-groupe du cardinal de n'importe quel diviseur de l'ordre du groupe, donc dans $\mathbb{Z}/(a_i\mathbb{Z})$ il existe un sous-groupe H_i de cardinal d_i.

• Le produit des H_i est un sous-groupe de G de cardinal d.$\square$

Quelques autres corollaires, sans preuve :

Corollaire 800 *Soit G un groupe abélien fini. Soit $c = p_1^{n_1} p_2^{n_2} \dots p_l^{n_l}$ la décomposition de $c = card(G)$ en facteurs premiers. Alors pour tout $i \in [1, l]$ il existe un et un seul sous-groupe H_i de G de cardinal $p_i^{n_i}$. En outre, G est isomorphe au produit des H_i.*
Les H_i, uniques, sont appelés les composantes primaires du groupe commutatif G.

21.8 Exercices sur les groupes

21.8.1 Exemples de groupes

Les objets suivants sont-ils des groupes ?
$(\mathbb{C}^*, .)$, $(\mathbb{R}^*, .)$, $(\mathbb{C}, +)$, $(\mathbb{R}, +)$, $(\mathbb{R}, .)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}, .)$, $(\mathbb{N}, .)$, $(\mathbb{Z}, +)$, l'ensemble des translations du plan, l'ensemble des homothéties du plan, l'ensemble à la fois des translations et des homothéties (pour la composition) ? Ces groupes sont-ils commutatifs ?

Oui, sauf $(\mathbb{R}, .)$, $(\mathbb{Q}, .)$, $(\mathbb{N}, .)$ et l'ensemble des homothéties. Tous ces groupes sont commutatifs sauf le groupe à la fois des translations et des homothéties.$\square$

21.8.2 Conditions réduites pour un groupe

Si G a une lci[1] associative, un élément 1 tel que $\forall g \; ge = g$, et un élément g' pour tout g tel que $gg' = 1$, alors G est un groupe.

Démonstration : On multiplie gg' par g' à gauche, et on montre que $g'g = 1$ en utilisant le g'' tel que $g'g'' = 1$. Il est facile de voir que $eg = g$ en utilisant ceci.$\square$

21.8.3 Conditions suffisantes de commutativité

Si tout élément est son propre inverse, alors G est commutatif.

21.8.4 $\mathbb{Z}/n\mathbb{Z}$

Proposition 801 *pour $d|n$, $\mathbb{Z}/n\mathbb{Z}$ comporte un seul sous-groupe d'ordre d.*

Démonstration : Existence triviale, unicité en considérant l'ensemble des élé-

[1] lci=loi de composition interne.

ments x tels que $x^d = 1$.□

Proposition 802 *Etant donnés deux entiers n et k, on a équivalence entre les trois assertions suivantes :*
- $\overline{k}$ (dans $\mathbb{Z}/n\mathbb{Z}$) engendre $\mathbb{Z}/n\mathbb{Z}$.
- n et k sont premiers entre eux.
- $\overline{k}$ est inversible dans l'anneau $\mathbb{Z}/n\mathbb{Z}$.

21.8.5 Sous-groupes

• Tout sous-groupe d'un groupe cyclique est cyclique.

• G sous-groupe de $(\mathbb{R}, +)$ $\rightarrow G$ monogène ou dense.
Démonstration : Considérer l'inf de $\mathbb{R}^+ \cap G$.□

• Il existe des sous-groupes denses de $\mathbb{R}$ de type fini.
Démonstration : $\mathbb{Z} + \sqrt{2}\mathbb{Z}$□

• L'union de deux sous-groupes est un groupe si et seulement si l'un est inclus dans l'autre.

• Le produit élément par élémentde deux sous-groupes A et B est un sous-groupe si et seulement si $AB = BA$.
Démonstration : : Supposons que AB soit un sous-groupe. Alors soit $a \in A$ et $b \in B$. $a^{-1}b^{-1} \in AB \rightarrow ba \in AB$ donc $BA \subset AB$
ab admet un inverse dans AB donc $a'b'ab = 1$, donc $b'ab = a'^{-1}$, donc $ab = b'^{-1}a'^{-1}$, donc $ab \in BA$, donc $AB \subset BA$
Réciproquement, supposons $AB = BA$, alors l'inverse de ab, $b^{-1}a^{-1}$, appartient bien à AB ; en outre il est immédiat que AB est stable par produit.□

• L'image d'un sous-groupe distingué par un homomorphisme est un sous-groupe distingué de l'image de l'homomorphisme. L'image réciproque d'un sous-groupe distingué par un homomorphisme est un sous-groupe distingué.

• L'intersection de deux sous-groupes distingués est un sous-groupe distingué.

• Tout sous-groupe d'un groupe abélien est distingué ; mais on peut avoir cette propriété sans que le groupe soit abélien ; considérer par exemple $\{1, i, j, k, -1, -i, -j, -k\}$, muni des opérations $ij = k, ji = -k, jk = i, kj = -i, ki = j, ik = -j$ (notons que ce groupe possède aussi la propriété de n'avoir que des sous-groupes propres abéliens).

21.8.6 Divers

• Dans un groupe, $(ab)^n = 1$; montrer que $(ba)^n = 1$.

• Montrer que $\mathbb{Q}$ n'est pas de type fini.

Démonstration : Considérer un nombre fini d'éléments de $\mathbb{Q}$, et un dénominateur commun de ces éléments.$\square$

• L'ensemble des automorphismes de $\mathbb{Z}/n\mathbb{Z}$ muni de la composition est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, \times)$.

• Un sous-groupe additif de $\mathbb{R}$ est soit dense, soit de la forme $a.\mathbb{Z}$. De même, $\mathbb{R}^{+*}$ muni de la multiplication n'admet que des sous-groupes denses ou de la forme $a^{\mathbb{Z}}$.

Démonstration : Considérer la borne inf de l'intersection du sous-groupe et de $\mathbb{R}^{+*}$. Le cas multiplicatif s'obtient en considérant le log, qui est un isomorphisme de groupe, et qui préserve la densité.$\square$

• $\mathbb{Z} + b\mathbb{Z}$ dans $(\mathbb{R}, +)$ est de la forme $a.\mathbb{Z}$ si b est rationnel, et dense si b est irrationnel.

21.9 Zoologie des opérations d'un groupe sur un ensemble

21.9.1 Opération d'un groupe G sur lui-même par translation à gauche

On associe à tout élément g de G la fonction qui à $x \in G$ associe $g.x$.

Cette opération est transitive (il y a une seule orbite) et fidèle.

Théorème 803 (Théorème de Cayley) *Si G est fini, alors G est isomorphe à un sous-groupe du groupe des permutations de G.*

Démonstration : L'application qui à g associe l'application $x \mapsto g.x$ est un homomorphisme injectif ; donc G est isomorphe à son image par cette application, qui est donc un sous-groupe du groupe des permutations de G.$\square$

Pour "fixer les idées", on peut se représenter $\mathbb{Z}/n\mathbb{Z}$ isomorphe à l'ensemble des applications qui à $\overline{x}$ associe $\overline{x} + \overline{p}$.

21.9.2 Opération d'un groupe G sur le groupe G/H par translation à gauche

Définition 804 (Action par translation à gauche) *A un élément g et une classe $g'.H$ on associe la classe $g.g'.H$. On vérifie facilement que cette opération est bien définie et définie bien une opération d'un groupe sur un ensemble. L'opération est clairement transitive, par contre elle n'est pas fidèle en général. Le noyau de l'application ϕ associée (voir définition d'un opération d'un groupe sur un ensemble) est égal à*

$$\cap_{g\in G} g.H.g^{-1}$$

(par définition du noyau, il suffit de l'écrire)

21.9.3 Opération d'un groupe sur lui-même par automorphismes intérieurs

Définition 805 (Action par automorphismes intérieurs) *A $g \in G$ on associe l'automorphisme intérieur $x \mapsto g.x.g^{-1}$.*

Proposition 806 *• Les orbites sont exactement les classes d'équivalence pour la relation de conjugaison. • Le stabilisateur d'un élément x est l'ensemble des g tels que $x = g.x.g^{-1}$, c'est à dire $x.g = g.x$; c'est donc l'ensemble des éléments qui commutent avec x, on l'appelle* **centralisateur** *de x. On généralise cette définition en l'élargissant aux parties de G ; le* **centralisateur** *d'une partie est l'ensemble des éléments qui commutent avec tous les éléments de cette partie.*
• Le centralisateur de G tout entier est donc le centre de G, c'est à dire l'ensemble des éléments qui commutent avec tous les autres.
• Les éléments d'une classe de conjugaison ont même ordre et même nombre de points fixes.
On peut par exemple considérer le groupe $GL(n, \mathbb{K})$; les classes de conjugaison, c'est à dire les orbites, sont alors les classes d'équivalence pour la relation "être semblable à".

21.10 Zoologie des groupes

On trouvera une étude des groupes $\mathbb{Z}/n\mathbb{Z}$ et $(\mathbb{Z}/n\mathbb{Z})^*$ dans la partie22.5.

21.10.1 Les p-groupes

Rappelons qu'un p-groupe est un groupe de cardinal (donc d'ordre) p^r avec p un nombre premier.

Proposition 807 (Le centre d'un p-groupe non trivial est non trivial) *Si G est un p-groupe de cardinal > 1 alors son centre est de cardinal > 1.*

Démonstration : On fait agir G sur lui-même par automorphismes intérieurs, comme indiqué en 21.9.3, et on applique la proposition 21.10.1. On en déduit que le centre est de cardinal congru à p, or il est non réduit à 0.□

21.10.2 Groupe linéaire et groupe spécial linéaire

Définition 808 *Etant donné $\mathbb{K}$ un corps commutatif quelconque et E un $\mathbb{K}$-espace vectoriel de dimension finie, on appelle* **groupe linéaire** *$GL(E)$ le groupe des automorphismes de E.*

$GL(E)$ est isomorphe à $GL(n, \mathbb{K})$, groupe des matrices inversibles de taille $n \times n$, à coefficients dans $\mathbb{K}$, avec n la dimension de E.

On notera que deux matrices semblables A et B vérifient qu'il existe P tel que $A = P^{-1}.B.P$; cela revient donc à dire que A et B sont conjuguées.

Définition 809 *Le noyau de l'homomorphisme qui à f associe son déterminant est par définition l'ensemble des automorphismes de déterminant 1 ; on l'appelle* **groupe spécial linéaire** *et on le note $SL(E)$.*

$SL(E)$ est isomorphe à $SL(n, \mathbb{K})$, groupe des matrices de $GL(n, \mathbb{K})$ de déterminant 1.

Proposition 810 *On a une suite exacte :*

$$1 \to SL(E) \to GL(E) \xrightarrow{\text{dét}} \mathbb{K}^* \to 1$$

En outre, le groupe linéaire $GL(E)$ est isomorphe au produit semi-direct du groupe spécial linéaire $SL(E)$ par $\mathbb{K}^*$.

$\mathbb{K}^*$ désignant $\mathbb{K} - \{0\}$.

Démonstration : Il suffit de prendre pour injection de $SL(E)$ dans $GL(E)$ la simple identité, et de considérer le sous-groupe H de $GL(E)$ des matrices de la forme

$$\begin{pmatrix} \lambda & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

avec λ non nul.

Le déterminant induit bien une bijection de H sur $\mathbb{K}^*$, on a bien $H \cap SL(E)$ réduit à l'élément neutre, on a bien $SL(E).H = GL(E)$, et $SL(E)$ est clairement distingué. □

Proposition 811 (Générateurs de $GL(E)$ et $SL(E)$) • *$GL(E)$ est engendré par l'ensemble des dilatations de E.*
• *$SL(E)$ est engendré par l'ensemble des transvections de E.*

Démonstration : Je ne détaillerai pas intégralement la preuve, laborieuse, mais peu difficile. Il suffit de montrer les points suivants, dans cet ordre :

• Toute matrice de la forme $I + \lambda E_{i,j}$ pour $i \neq j$, avec $\lambda \in \mathbb{K}$ et $E_{i,j}$ la matrice définie par $(E_{i,j})_{k,l} = 1$ si $i = k$ et $j = l$ et 0 sinon, est la matrice d'une transvection. L'inverse d'une matrice de transvection, est une matrice de transvection.

• Une matrice de déterminant 1 est égale à un produit de matrices de transvections. Pour le prouver, on considère M une telle matrice, on la multiplie par des matrices de transvection pour se ramener à une matrice n'ayant qu'un seul élément non nul sur la première ligne, pour que cet élément soit l'élément en haut à gauche, et pour qu'il soit égal à 1. Il suffit alors de procéder par récurrence en considérant un produit de matrices par bloc.

(ce point est exactement le deuxième point annoncé)

• Une matrice appartenant à $GL(E)$ est le produit d'une matrice appartenant à $SL(E)$ et d'une matrice de dilatation (voir proposition 810).□

21.10.3 Groupe orthogonal et groupe spécial orthogonal

⊡ Cas général

Définition 812 *On appelle* **groupe orthogonal d'un espace euclidien** *E l'ensemble des automorphismes orthogonaux de E muni de la composition $\circ$; on le note $O(E)$.*
On appelle **groupe spécial orthogonal d'un espace euclidien** *E l'ensemble des automorphismes orthogonaux de E de déterminant 1 muni de la composition $\circ$; on le note $SO(E)$ ou $O^+(E)$.*
On note $O^-(E)$ le complémentaire de $SO(E)$ dans $O(E)$.
On note en outre $O_n(\mathbb{R})$ l'ensemble $O(\mathbb{R}^n)$.
On note en outre $SO_n(\mathbb{R})$ l'ensemble $SO(\mathbb{R}^n)$.

Ces espaces sont isomorphes aux espaces décrits en 21.10.4, donc je n'approfondis pas plus ici pour le moment, à part les cas spéciaux des dimensions 1, 2 et 3.

⊡ **Dimension** 1

Ce cas est de peu d'intérêt ; les seules transformations orthogonales sont $x \mapsto x$ et $-x \mapsto -x$...

⊡ **Dimension** 2

Un rapide calcul montre que les matrices des transformations orthogonales en dimension 2 sont de l'une des deux formes suivantes :

$$\begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \begin{pmatrix} cos(\theta) & sin(\theta) \\ sin(\theta) & -cos(\theta) \end{pmatrix}$$

La matrice de gauche représente une transformation du groupe spécial orthogonal (c'est à dire de déterminant 1, et donc dans $SO(E) = O^+(E)$), celle de droite une transformation qui n'est pas de ce groupe (c'est à dire que celle-ci est de déterminant -1, et donc dans $O^-(E)$).

(le calcul est facile, il suffit de se souvenir que $x^2+y^2 = 1 \rightarrow \exists\theta/x = cos(\theta)$ et $y = sin(\theta)$)

> **Définition 813** *On appelle* **rotation d'angle** θ *un endomorphisme associé à la matrice :*
>
> $$\begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$$

Toujours par des calculs sans grande difficulté on montrerait que $SO_2(\mathbb{R})$ commute, et est en fait isomorphe à $\mathbb{R}/(2.\Pi.\mathbb{Z})$; les seules transformations orthogonales de déterminant 1 sont en fait les rotations. On note r_θ la rotation d'angle θ.

En étudiant la matrice de droite, on constate qu'elle est symétrique, donc diagonalisable (voir la partie 29.5.2) ; son polynôme caractéristiques est $X^2 - 1$; elle est semblable à $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Il s'agit donc en fait d'une symétrie par rapport à un hyperplan (ici hyperplan = droite car on est en dimension 2). Ainsi les transformations orthogonales de déterminant -1 sont en fait des symétries par rapport à des droites. On note que les symétries ne commutent pas, elles. On note s_θ la symétrie correspondant θ (θ est en fait le double de l'angle de l'axe invariant avec le premier axe).

On notera que $s_\theta \circ s_{\theta'} = r_{\theta-\theta'}$.

⚠ l'angle n'est défini qu'à $2.\Pi$ près pour les rotations et les symétries.

⊡ **Dimension** 3

> **Proposition 814** *En dimension* 3*,* $O^+(E)$ *comporte :*
> • *les rotations axiales*
> • *l'identité, qui est un cas particulier de rotation axiale*
> • *la symétrie par rapport à une droite, qui est un cas particulier de rotation axiale*
> *En dimension* 3*,* $O^-(E)$ *comporte :*
> • *les symétries orthogonales par rapport à un plan*
> • *les composées d'une rotation autour d'un axe et d'une symétrie par rapport au plan orthogonal à cet axe*

On se donne f un endomorphisme orthogonal de E euclidien de dimension 3, et on considère I l'ensemble $\{x/f(x) = x\}$ (ensemble des invariants par f). On va classer les f possibles suivant la dimension de I.

$\diamond$ $dim\ I = 3$
Pas drôle : f est l'identité, et donc $f \in SO(E) = O^+(E)$.

$\diamond$ $dim\ I = 2$
Alors l'orthogonal de I est de dimension 1 ; la restriction de f à cet espace est un endomorphisme orthogonal (rappelons que si un espace est stable pour un endomorphisme orthogonal, alors son orthogonal aussi). Ce n'est pas l'identité puisque f n'est pas l'identité, donc il s'agit de $x \mapsto -x$ (si un endormorphisme est orthogonal, ses seules valeurs propres possibles sont 1 et -1). f est donc une symétrie par rapport à un plan. $f \in O^-(E)$.

$\diamond$ $dim\ I = 1$
La restriction de f à l'orthogonal de I (rappelons que si un espace est stable pour un endomorphisme orthogonal, alors son orthogonal aussi) est un endomorphisme orthogonal et n'a pas de vecteur invariant ; donc c'est une rotation. Donc f est une rotation autour d'un axe. Sa matrice est semblable à la matrice

$$\begin{pmatrix} cos(\theta) & -sin(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

f est de déterminant 1, et donc appartient à $SO(E) = O^+(E)$.

$\diamond$ $dim\ I = 0$
En dimension 3, tout endomorphisme admet au moins une valeur propre (tout polynôme de degré impair admettant au moins une racine sur $\mathbb{R}$).
f admet donc nécéssairement une valeur propre. Or un endomorphisme orthogonal ne peut avoir pour valeur propre que 1 ou -1 ; donc -1 est valeur propre.
On va maintenant considérer O, l'ensemble des x tels que $f(x) = -x$, et on va raisonner sur la dimension de O.

$dim\, O = 3$ On a alors f la symétrie par rapport à 0 ; f est dans $O(E)$ et pas dans $SO(E)$; f est dans $O^-(E)$.

$dim\, O = 2$ **: cas impossible** Supposons $dim\, O = 2$.

Alors l'orthogonal de O est stable par f ; donc soit la restriction de f à cet orthogonal est l'identité, soit c'est moins l'identité ; puisque $dim\ I = 0$ il s'agit de moins l'identité. Donc en fait $dim\, O = 3$, d'où contradiction. Donc ce cas ne peut se produire.

$dim\, O = 1$

On considère alors la restriction de f à l'orthogonal de O. Il s'agit d'un endomorphisme orthogonal en dimension 2, sans valeur propre ; donc une rotation qui n'est pas une symétrie par rapport à un point, ni l'identité. f est de déterminant -1, et donc est dans $O(E)$ mais pas dans $SO(E)$; $f \in O^-(E)$.

21.10.4 Groupe orthogonal réel et groupe spécial orthogonal réel

Définition 815 *On appelle* **groupe orthogonal réel d'ordre** n *l'ensemble des matrices* M *réelles de type* (n, n) *telles que* ${}^tM.M = I$*, on le note* $O_n(\mathbb{R})$ *; il s'agit d'un sous-groupe du groupe linéaire réel d'ordre* n*.*
On appelle **groupe spécial orthogonal réel d'ordre** n *l'ensemble des matrices* M *réelles de type* (n, n) *telles que* ${}^tM.M = I$ *et* $det\ M = 1$*, on le note* $SO_n(\mathbb{R})$ *; il s'agit d'un sous-groupe du groupe orthogonal réel d'ordre* n *et d'un sous-groupe du groupe spécial linéaire d'ordre* n*.*
On appelle **matrice orthogonale** *une matrice appartenant à* $O_n(\mathbb{R})$ *pour un certain* n*.*

Proposition 816 (Propriété des matrices orthogonales réelles) *Une matrice est orthogonale si et seulement si sa transposée l'est.*
Une matrice est orthogonale si et seulement si ses vecteurs colonnes forment une famille orthonormale de $\mathbb{R}^n$*.*
Une matrice est orthogonale si et seulement si ses vecteurs lignes forment une famille orthonormale de $\mathbb{R}^n$*.*
Une matrice est orthogonale si et seulement si il s'agit d'une matrice de changement de bases orthonormales.
Une matrice orthogonale est de déterminant 1 *ou* -1*.*
Une valeur propre de matrice orthogonale est soit 1 *soit* -1*.*
Une matrice orthogonale M *vérifie* $com(M) = det(M).M$*.*

21.10.5 Groupe affine d'un espace affine

Définition 817 *On appelle* **groupe affine d'un espace affine** X *l'ensemble des applications affines bijectives de* X *dans lui-même muni de la composition ; c'est un groupe. On le note* $GA(X)$.
On appelle **groupe spécial affine d'un espace affine** X *l'ensemble des applications affines bijectives* f *de* X *dans lui-même telles que* $det\ \overrightarrow{f} = 1$, *muni de la composition ; c'est un groupe. On le note* $SA(X)$.

Le fait qu'il s'agisse d'un groupe est facile à voir. La proposition suivante est évidente :

Proposition 818 *L'application* $f \mapsto \overrightarrow{f}$ *est un morphisme de* $GA(X)$ *dans* $GL\overrightarrow{X}$.

Etudions maintenant la structure du groupe $GA(X)$.

⊡ **Générateurs de** $GA(X)$ **et** $SA(X)$

Proposition 819 • $GA(X)$ *est engendré par l'ensemble des dilatations affines de* X
• $SA(X)$ *est engendré par l'ensemble des transvections affines de* X

Démonstration : Simple conséquence de la proposition 811.□

⊡ **Sous-groupes remarquables de** $GA(X)$

◇ **Le sous-groupe des symétrie**
L'ensemble des symétries est un sous-groupe distingué de $GA(X)$. En effet, avec $s_{A,\overrightarrow{B}}$ la symétrie par rapport à A parallèlement à $\overrightarrow{B}$, on a

$$g \circ s_{Y,\overrightarrow{Z}} \circ g^{-1} = s_{g(Y),\overrightarrow{g}(\overrightarrow{Z})}$$

◇ **Le sous-groupe des translations**
L'ensemble $T(X)$ des translations de l'espace affine X est un groupe pour la composition ; ce groupe est distingué. On le voit en constatant que c'est le noyau du morphisme qui à f dans $GA(X)$ associe $\overrightarrow{f}$ dans $GL(\overrightarrow{X})$.

Le groupe quotient de X par $T(X)$ est isomorphe à $\overrightarrow{X}$.

◇ **Le sous-groupe des homothéties-translations**
L'ensemble des homothéties et des translations d'un espace affine X est stable par composition et contient l'identité ; or il est inclus dans $GA(X)$. Donc c'est un sous-groupe de $GA(X)$. Il est généré par les homothéties (toute translation s'exprime comme composée de deux homothéties de rapport inverse).

Ce sous-groupe est exactement l'ensemble des bijections de X transformant toute droite en une droite parallèle.

$\diamond$ **Le sous-groupe des applications affines bijectives de X laissant une partie donnée invariante**

On fixe une partie P de X, et on considère G l'ensemble des f appartenant à $GA(X)$ telles que $f(P) \subset P$; G est stable par composition et contient l'identité, c'est donc un sous-groupe de $GA(X)$.

$\diamond$ **En dimension finie, le sous-groupe des applications affines laissant fixe un repère**

On suppose que X est de dimension finie n. On se donne alors un repère affine $A_0, A_1, ..., A_n$.

Nécéssairement, une bijection affine f laissant invariant un repère a pour restriction à la partie $\{A_0, ..., A_n\}$ une permutation σ. Une application affine étant entièrement déterminée par l'image d'un repère affine, on en déduit que l'ensemble des applications affines bijectives laissant invariant le repère $A_0, A_1, ..., A_n$ est un groupe isomorphe à σ_{n+1}

$\boxdot$ **Le groupe affine comme produit semi-direct**

On a vu que $T(X)$, ensemble des translations de X, est distingué dans $GA(X)$, puisque noyau du morphisme $f \mapsto \overrightarrow{f}$. On a une suite exacte

$$1 \to T(X) \to GA(X) \xrightarrow{f \mapsto \overrightarrow{f}} GL(\overrightarrow{X}) \to 1$$

On se donne O appartenant à X donné ; l'application $f \mapsto \overrightarrow{f}$ induit une bijection de l'ensemble des bijections affines de X laissant O invariant sur $\overrightarrow{X}$; on a donc un relèvement de $GL(\overrightarrow{X})$.

Donc $GA(X) = T(X) \rtimes GL(\overrightarrow{X})$, avec pour action de $GL(X)$ dans $T(X)$ $\overrightarrow{f}.t = f_0^{-1} \circ t \circ f_0$ avec f_0 l'application affine laissant O invariant et associée à $\overrightarrow{f}$ (ce qui revient à $\overrightarrow{f}.t_{\overrightarrow{a}} = t_{\overrightarrow{f}(\overrightarrow{a})}$, en notant $t_{\overrightarrow{u}}$ la translation de vecteur $\overrightarrow{u}$).

En considérant l'isomorphisme évident entre $T(X)$ et $\overrightarrow{X}$ (c'est-à-dire en remplaçant une translation par le vecteur de cette translation) on peut aussi écrire

$$GA(X) = \overrightarrow{X} \rtimes GL(\overrightarrow{X})$$

Et quel que soit O dans X on peut écrire toute application bijective affine f de X dans X sous la forme $f = t \circ u_0$ avec u_0 application affine bijective laissant 0 invariant.

21.10.6 Groupe projectif d'un espace vectoriel de dimension finie

Définition 820 - Proposition *On se donne E un $\mathbb{K}$-espace vectoriel de dimension finie. L'ensemble des homographies de $P(E)$ dans $P(E)$ forme un groupe pour $\circ$, appelé* **groupe projectif de** *E, noté $PGL(E)$. Ce groupe est isomorphe à $GL(E)/(\mathbb{K} \setminus \{0\}.I)$, avec I l'identité de E dans E.*
On note usuellement $PGL_n(\mathbb{K})$ pour $PGL(\mathbb{K}^n)$.

Démonstration : Seul l'isomorphisme mérite d'être détaillé.

Considérons l'application H, qui à un endomorphisme de E associe l'homographie associée à cet endomorphisme (on se donne bien entendu pour cela un repère projectif de E).

Son noyau est l'ensemble des applications linéaires de E dans E qui laissent toute droite invariante. Il faut donc montrer qu'un endomorphisme laissant toute droite invariante est une homothétie.

Lemme 821 *Un endomorphisme d'un espace vectoriel de dimension finie laissant invariante toute droite est une homothétie.*

Démonstration : On procède par récurrence sur la dimension n de l'espace vectoriel .

Pour $n = 1$ le résultat est clair.

Pour $n > 1$, on considère f un endomorphisme de E, tel que pour tout x il existe un scalaire λ_x tel que $f(x) = \lambda_x.x$.

Il est clair que si x et y de E sont liés, alors $\lambda_x = \lambda_y$.

Considérons maintenant x et y linéairement indépendants.

Alors $f(x + y) = \lambda_{x+y}.x + \lambda_{x+y}.y = f(x) + f(y) = \lambda.x + \lambda.y$, donc $\lambda_x = \lambda_{x+y} = \lambda_y$.

On a donc montré le résultat souhaité.□

Du coup, grâce à ce lemme, la preuve de la proposition est achevée.□

21.10.7 Groupe unitaire et groupe spécial unitaire d'un espace hermitien

Définition 822 *On appelle* **groupe unitaire de** E *et on note* $U(E)$ *avec* E *un espace hermitien (voir partie29.6) l'sensemble des automorphismes unitaires de* E*, c'est-à-dire des automorphismes* f *de* E *tels que* $f^{-1} = f^*$*, muni de la composition.*
On appelle **groupe spécial unitaire de** E*, et on note* $SU(E)$*, avec* E *un espace hermitien, le sous-groupe de* $U(E)$ *constitué des automorphismes unitaires de* E *de déterminant* 1*.*

Ces groupes sont isomorphes aux groupes dont il est question ci-dessous.

On note bien que le déterminant d'un élément de $U(E)$ peut être n'importe quelle valeur du cercle unité, et pas seulement 1 et -1 comme dans le cas des endomorphismes orthogonaux d'un espace euclidien.

21.10.8 Groupe unitaire complexe d'ordre n et groupe spécial unitaire complexe d'ordre n

Définition 823 *L'ensemble des matrices M de type (n, n) à coefficients dans $\mathbb{C}$ telles que ${}^t\overline{M}.M = I$ est un groupe pour $\circ$; on l'appelle* **groupe unitaire complexe d'ordre** n*, et on le note $U_n(\mathbb{C})$.*
L'ensemble des matrices M de type (n, n) à coefficients dans $\mathbb{C}$ telles que ${}^t\overline{M}.M = I$ et $det\ M = 1$ est un groupe pour $\circ$; on l'appelle **groupe spécial unitaire complexe d'ordre** n *; on le note $SU_n(\mathbb{C})$, c'est un sous-groupe de $U_n(\mathbb{C})$.*

21.10.9 Groupe des similitudes d'un espace euclidien

Définition 824 *On appelle* **groupe des similitudes d'un espace euclidien** E *et on note $GO(E)$ l'ensemble des similitudes d'un espace euclidien E, muni de la composition $\circ$. On appelle* **groupe des similitudes d'un espace euclidien** E *et on note $GO(E)$ l'ensemble des similitudes d'un espace euclidien E, muni de la composition $\circ$.*

Il s'agit d'un groupe, sous-groupe de $GL(E)$ (groupe linéaire de E, ensemble des automorphismes de E.

Il est isomorphe à $\mathbb{R}_+^* \times O(E)$, avec $O(E)$ l'ensemble des automorphismes orthogonaux de E.

21.10.10 Groupe des quaternions

Définition 825 *On le note H_8. Ses éléments sont $1, -1, i, j, k, -i, -j, -k$, et la multiplication est définie par la table suivante :*

	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

On peut aussi résumer la loi de multiplication par

$$ij = k, jk = i, ki = j$$

$$ji = -k, kj = -i, ik = -j$$

$$i^2 = j^2 = k^2 = -1$$

$$D(H_8) = \{1, -1\}$$

$$Z(H_8) = \{1, -1\}$$

Ce groupe n'est pas commutatif. Ses sous-groupes sont $\{1\}$, $\{1, -1\}$, $\{1, -1, i, -i\}$ et H_8 lui-même ; ils sont tous distingués. Cela montre d'ailleurs que la propriété des groupes abéliens d'avoir tous leurs sous-groupes distingués n'est pas une condition suffisante pour que le groupe soit abélien.

21.10.11 Groupe symétrique

Définition 826 *On appelle* **permutation** *d'un ensemble une bijection de cet ensemble sur lui-même.*
On appelle **support** *d'une permutation sur un ensemble tout élément de cet ensemble qui n'est pas invariant par cette permutation.*
On appelle **cycle** *d'un ensemble une bijection f telle qu'il existe $a_1, ..., a_n$ (en nombre fini et distincts) tels que $f(a_i) = a_{i+1}$ pour $i < n$, $f(a_n) = a_1$ et $f(b) = b$ si b n'est aucun des a_i. n est l'ordre du cycle ; il ne s'agit pas d'une définition, car cet ordre colle à la notion d'ordre sur les éléments d'un groupe. n est aussi appelé* **longueur** *du cycle (cette fois-ci c'est bien une définition !).*
*On appelle n-***cycle** *un cycle d'ordre n.*
On appelle **transposition** *une permutation qui «échange» deux éléments. On note (a, b) la transposition qui échange a et b. Une transposition est un cycle est longueur 2.*
On appelle **groupe symétrique** *d'un ensemble E l'ensemble des permutations de cet ensemble.*
*On note σ_n et on appelle n-***ième groupe symétrique standard** *le groupe symétrique de $\{1, 2, ..., n\}$. Tous les groupes symétriques sur des ensembles de cardinal n sont isomorphes à σ_n.*
Pour un n donné on appelle **signature** *l'unique homomorphisme ϵ de σ_n dans $\{1, -1\}$ tel que $\epsilon(\tau) = -1$ lorsque τ est une permutation.*
*On appelle n-***ième groupe alterné** *le noyau de ϵ (dans σ_n). On le note U_n.*
On appelle **matrice associée la permutation** *σ* **de** *σ_n la matrice M telle que $M_{i,j} = \delta_{i,\sigma(j)}$.*

Remarques et propriétés :
• On parle aussi, au lieu de n-ième groupe symétrique standard, de **groupe symétrique d'ordre** n ; il faut bien voir que ce groupe n'est *PAS* d'ordre n mais d'ordre $n!$.
• Pour bien faire il faudrait démontrer que l'on caractérise bien ici la signature. Cela serait fait dans la partie 21.10.11.

Proposition 827 • *$|\sigma_n| = n!$*
• *$Z(\sigma_n) = 1$ si $n \geq 3$.*
• *σ_n est engendré par les transpositions*
• *σ_n est engendré par les transpositions de la forme $(a, a + 1)$, avec $a \in [1, n - 1]$.*
• *σ_n est engendré par les transpositions de la forme $(1, a)$, avec $a \in [1, n]$.*
• *σ_n est engendré par la transposition $(1, 2)$ et le cycle $(1, 2, ..., n)$.*
• *U_n est engendré par les cycles d'ordre 3.*
• *Des cycles de supports disjoints commutent.*

Proposition 828 *Soit p une permutation de E fini. L'orbite d'un point x pour p est l'ensemble des $p^n(x)$ avec $n \in \mathbb{N}$. p est un cycle s'il existe une orbite et une seule qui soit de cardinal > 1.*

Démonstration : Si on a un cycle, il est clair qu'il existe une seule orbite de cardinal > 1 ; si on a une seule orbite dans ce cas, alors on considère les éléments de l'orbite, la suite est évidente.$\square$

Théorème 829 *Toute permutation peut s'écrire comme produit de cycles de supports deux à deux disjoints. La décomposition est unique à l'ordre près des facteurs.*

Démonstration :
Considérons une permutation σ de $[1, n]$. L'unicité de sa décomposition sous la forme annoncée découle immédiatement de l'étude des orbites de l'action de $\{Id, \sigma\}$ sur $[1, n]$ (on considère ce qu'il se passe sur chaque orbite).

Pour l'existence, on se restreint aussi à une telle orbite. Il est clair que σ se comporte sur cette orbite comme un cycle. D'où le résultat.$\square$

Proposition 830 *Le centre de σ_n est trivial dès que $n \geq 3$.*

Démonstration : Soit σ élément non neutre de σ_n. Il existe alors i tel que $\sigma(i) = j \neq i$. On prend alors k différent à la fois de i et de j, et on constate que $\sigma \circ (j\ k)(i) \neq (j\ k) \circ \sigma(i)$$\square$

⊡ La conjugaison dans σ_n

On considère l'opération de σ_n sur σ_n par automorphisme intérieur, comme étudié en 21.9.3.

Proposition 831 • *Cette opération est transitive, et même k transitive pour tout k. Elle est fidèle pour $n \geq 3$, puisqu'alors le centre est trivial.*
• *Si on se limite à U_n, cette opération est $n - 2$ fois transitive.*

Démonstration : Le premier point est évident. Pour le second, on considère $k \leq n - 2$ éléments de $\{1, ..., n\}$, et on ajoute deux autres points ; on considère la permutation qui affecte nos k points correctements, et qui, si la permutation obtenue est impaire, permute les deux points supplémentaires.$\square$

Proposition 832 • *Pour tout m, l'ensemble des cycles d'ordre m est une orbite (c'est à dire une classe de conjugaison).*
• *Si $n \geq 5$ les cycles d'ordre 3 sont conjugués dans U_n.*

Démonstration : Remarquons tout d'abord que si $f = (x_1, ..., x_k) \in \sigma_n$ et $g \in \sigma_n$, alors $g.f.g^{-1} = (g(x_1), ..., g(x_k))$. Pour montrer le premier point, il suffit alors, étant donnés deux cycles de même longueur $(x_0, ..., x_k)$ et $(y_0, ..., y_k)$ de considérer la permutation p qui à x_i associe y_i ; on a bien $p.x.p^{-1} = y$.
Le deuxième point est plus délicat, et utilise la proposition 831. Etant donnés deux 3-cycles (x_0, x_1, x_2) et (y_0, y_1, y_2), on considère la permutation p de U_n qui à x_i associe y_i ; on a bien $x = p.y.p^{-1}$.□

◎ Les matrices de permutations

Définition 833 *L'application ϕ qui à une permutation associe la matrice associée à cette permutation est un morphisme injectif dans $GL_n(\mathbb{K})$ (ensemble des matrices inversibles de type (n, n)).*
On a la propriété $\phi(s)^{-1} = \phi(\sigma^{-1}) =^t \phi(\sigma)$.
Le déterminant de $\phi(s)$ est égal à la signature de s.

◎ La signature

Proposition 834 (Différentes caractérisations de la signature) *On peut définir la signature ϵ sur σ_n de l'une des façons suivantes :*
1) On appelle **inversion** *d'une permutation p, une paire (i, j) d'éléments tels que $(j - i).(p(j) - p(i)) < 0$. On définit $\epsilon(p) = (-1)^{Inv(p)}$, avec $Inv(p)$ le nombre d'inversions.*
2) Il existe un unique morphisme ϵ de σ_n sur $\{-1, 1\}$ tel que $\epsilon(t) = -1$ si t est une permutation.
3) $\epsilon(p)$ est égal à $(-1)^s$ avec s le nombre de transpositions dans une décomposition de p en produit de transpositions.

Démonstration : Pas très très dur... Pour voir que 1 entraine 2 il faut voir que $\epsilon(p)$ est le produit des $\Pi_{i<j} \frac{p(j)-p(i)}{j-i}$, le reste est facile.□
Il y a en outre une caractérisation de la signature, donnée en 21.10.11.

◎ Simplicité de U_n pour $n > 4$; conséquences

Cette preuve est tirée de [14, p. 28], fort bon livre en algèbre, pour ceux qui connaissent déjà les bases du moins.

Théorème 835 *U_n est simple (i.e. sans sous-groupe distingué non trivial) si $n \geq 5$.*

Démonstration : On procède en deux étapes :

• Le cas $n = 5$

- Le groupe U_5 se décompose en 60 éléments ; l'identité, 15 éléménts d'ordre 2, qui sont des produits de deux transpositions disjointes, 20 éléments d'ordre 3, qui sont des 3-cycles, et 24 d'ordre 5, qui sont des 5-cycles. On va se préoccuper des classes de conjugaison de U_5.
- les éléments d'ordre 2 sont conjugués (facile).
- les 3-cycles sont conjugués.

Supposons H sous-groupe de U_5, et $H \lhd U_5$, et $H \neq \{1\}$.
- S'il contient un élément d'ordre 3 il les contient tous, puisqu'il est distingué et que les éléments d'ordre 3 sont conjugués.
- S'il contient un élément d'ordre 2 il les contient tous, puisqu'il est distingué et que les éléments d'ordre 2 sont conjugués.
- S'il contient un éléments x d'ordre 5, alors il contient aussi le 5-Sylow engendré par x (voir les théorèmes de Sylow, 21.5). Les 5-Sylow étant tous conjugués, il les contient donc tous ; tout élément d'ordre 5 étant inclus dans un 5-sylow, tout élément d'ordre 5 est alors inclus dans H.
- S'il contient donc un seul type d'éléments parmi les éléments ci-dessus en plus de l'unité, alors alors son cardinal serait soit $1 + 20$, soit $1 + 24$, soit $1 + 15$; or ces nombres ne divisent pas 60. Donc il contient au moins deux types de ces éléments. Donc son cardinal est au moins $1 + 15 + 20$, et comme il divise 60, H est en fait égal à U_5. Le résultat est donc prouvé dans le cas de U_5.

• Le cas $n > 5$

- On considère $H \lhd U_n$, $H \neq \{1\}$; on considère σ dans H, $\sigma \neq 1$.
- Par hypothèse on a un certain a tel que $b = \sigma(a) \neq a$.
- on peut choisir c différent à la fois de a, de b et de $\sigma(b)$.
- on considère τ le 3-cycle (acb). $\tau^{-1} = abc$.
- On note ρ la permutation $(\tau.\sigma.\tau^{-1}).\sigma^{-1} = (acb)(\sigma.a, \sigma.b, \sigma.c)$.
- L'ensemble $\{a, b, c, \sigma.a, \sigma.b, \sigma.c\}$ ayant au plus 5 élément (car $\sigma.a = b$), on le complète par des éléments quelconques pour avoir un ensemble F de 5 éléments contenant $\{a, b, c, \sigma.a, \sigma.b, \sigma.c\}$.
- ρ est l'identité en dehors de F, et $\rho(F) = F$.
- On constate que ρ est différent de l'identité car $\rho(b) \neq b$.
- U_F, ensemble des permutations paires de F est isomorphe à U_5 ; on a un morphisme injectif ϕ de U_F dans U_n en considérant pour une permutation t de U_F la permutation dont la restriction à F est t et la restriction à F^c est l'identité.
- On considère H' l'intersection de H et de U_F.
- H' est distingué dans U_F, clairement.
- il est clair que ρ_F appartient à U_F, et que ρ_F n'est pas l'élément neutre.
- Par simplicité de U_F, on sait alors que H' est égal à U_F.
- On considère alors un 3-cycle c de F, il est dans H', donc $\phi(c)$ est dans H.
- H contient donc un 3-cycle, or puisqu'il est distingué il contient aussi sa classe de conjugaison, donc il contient tous les 3-cycles (les 3-cycles étant tous conjugués). Donc il contient le groupe engendré par les 3-cycles, c'est-à-dire U_n.

Ceci termine la preuve.□

> **Corollaire 836** *$D(U_n) = U_n$ pour $n \geq 5$ et $D(\sigma_n) = U_n$ pour $n \geq 2$.*

Démonstration : • Première preuve (en utilisant le théorème) :
$D(U_n)$ est distingué, donc il ne saurait être plus petit que U_n, puisque U_n est simple, donc $D(U_n) = U_n$.
$D(\sigma_n)$ est le sous-groupe engendré par les commutateurs de σ_n, or il est clair que ces commutateurs appartiennent à U_n (considérer leurs signatures). Donc $D(\sigma_n)$ est inclus dans U_n, or puisqu'il est distingué, il ne saurait être inclus strictement.
• Deuxième preuve (élémentaire) :
- on montre facilement que tout commutateur de σ_n est dans U_n.
- on en déduit que $D(U_n) \subset D(\sigma_n) \subset U_n$
- on montre alors que tout 3-cycle de U_n s'écrit comme commutateur d'éléments de U_n ; en effet avec f un tel 3-cycle, f et f^2 sont conjugués dans U_n (vrai pour toute paire de 3-cycles), donc $f^2 = t.f.t^{-1}$, et donc $f = t.f.t^{-1}.f^{-1}$.□

> **Corollaire 837** *Les sous-groupes distingués de σ_n sont $\{1\}, U_n, \sigma_n$.*

Démonstration : Supposons H sous-groupe distingué de σ_n.
• $H \cap U_n$ est égal à 1 ou U_n.
• Si $H \cap U_n = U_n$, alors si $H \neq U_n$, alors H contient un produit impair de transpositions ; en multipliant par l'inverse du produit des transpositions sauf une on constate que H contient une transposition. Etant données deux transpositions, on constate qu'elles sont conjuguées par les éléments de U_n ; donc H contient en fait tout σ_n.
• Si $H \cap U_n = \{1\}$, alors ϵ est un isomorphisme de H sur $\epsilon(H)$. Donc H contient en fait un seul autre élément au plus. S'il en contient deux alors soit τ l'autre élément ; il doit commuter avec n'importe quel élément puisque H est distingué et puisque τ n'est pas conjugué à l'unité ; or le centre de σ_n est trivial.□

> **Corollaire 838** *Sois G un sous-groupe de σ_n d'indice n. Alors G est isomorphe à σ_{n-1}.*

Démonstration : On rappelle que l'indice d'un sous-groupe, on appelle indice le cardinal du groupe quotient. Un sous-groupe d'indice n de σ_n est donc en fait un sous-groupe de cardinal $(n-1)!$.
Le cas $n \leq 4$ s'obtient facilement. Pour $n \geq 5$, on constate que σ_n ou G opère à gauche sur l'ensemble quotient (par translation à gauche, voir 21.9.2). On a donc un homomorphisme ϕ de σ_n dans l'ensemble des permutations de σ_n/H, qui est isomorphe à σ_n.Il reste maintenant à voir que cet homomorphisme est injectif (le caractère surjectif se déduisant alors des cardinaux). Son noyau est l'intersection des $a.G.a^{-1}$ pour a dans σ_n, et donc il est de cardinal au plus le cardinal de G, donc $(n-1)!$; or il est distingué, et on a montré que les seuls sous-groupes distingués de σ_n étaient $\{1\}$, U_n et σ_n ; donc il s'agit de 1, d'où le résultat. □

Admettons enfin sans preuve la proposition ci-dessous :

Proposition 839 *Si $n \neq 4$ et $n \neq 6$ tous les G vérifiant ces hypothèses sont conjugués. En fait, avec les mêmes hypothèses que ci-dessus, il existe i tel que G soit l'ensemble des permutations laissant i invariant.*

⊡ Décomposition de σ_n

On a une suite exacte

$$1 \to U_n \to \sigma_n \xrightarrow{\epsilon} \{-1, 1\} \to 1$$

avec ϵ la signature. Avec τ une transposition (c'est à dire une permutation de deux éléments) alors $\{Id, \tau\}$ est un groupe qui est une section pour ϵ, donc on a

$$\sigma_n \simeq U_n \rtimes \{\tau, Id\} \simeq U_n \rtimes \{-1, 1\} \simeq U_n \rtimes \mathbb{Z}/2\mathbb{Z}$$

En outre, σ_n est isomorphe à l'ensemble des automorphismes intérieurs lorsque $n \geq 3$; en effet le centre est alors trivial. On verra plus bas que l'ensemble des automorphismes intérieurs est lui-même égal à l'ensemble des automorphismes lorsque $n \neq 6$.

⊡ Automorphismes de σ_n

◇ Les automorphismes intérieurs sont des automorphismes

Les automorphismes intérieurs sont les automorphismes de la forme $t \to u.\tau.u^{-1}$, avec u une permutation quelconque. Les automorphismes intérieurs forment un sous-groupe du groupe des automorphismes, de manière évidente.

◇ Les automorphismes sont des automorphismes intérieurs lorsque $n \neq 6$

Proposition 840 *Un automorphisme de σ_n transformant toute transposition en transposition est un automorphisme intérieur.*

Démonstration : On considère les transpositions $t_i = (1, i)$ pour $i > 1$. Ces transpositions engendrent toutes les transpositions. Il suffit donc de montrer que ϕ, qui transforme toutes ces transpositions en transpositions, coïncident avec un automorphisme intérieur.
Pour cela on constate que :
• les $\phi(t_i)$ ne sont pas disjointes deux à deux.
• $\phi(t_i)$ et $\phi(t_j)$ ont même élément commun que $\phi(t_i)$ et $\phi(t_k)$.
• on peut donc noter $\phi(t_i)$ sous la forme (z_1, z_i).
• z est la permutation recherchée, tel que l'automorphisme intérieur correspondant corresponde à ϕ.□

Proposition 841 *On suppose $n = 1.k_1 + 2.k_2 + ... + n.k_n$ et que σ est une permutation produit de $\sum_i k_i$ cycles disjoints, k_1 d'ordre 1, k_2 d'ordre 2, k_3 d'ordre 3, ... ,k_n d'ordre k_n. Alors le cardinal du centralisateur de σ est égal à*

$$|c(s)| = \Pi_{i=1}^{n} k_i!.i^{k_i}$$

Démonstration : • Tout d'abord on montre le résultat pour un seul cycle, d'ordre n.
Le centralisateur est alors tout simplement de cardinal n ; il s'agit du sous-groupe engendré par ce cycle. Pour le voir on se ramène à un cycle $(1, 2, ..., n)$; pour que τ commute avec ce cycle, il faut que $\tau(n+1) = \tau(n)+1$, c'est-à-dire que $\tau(n+1)-\tau(n) = 1$, donc que $\tau(n) = \tau(0) + n$ (on compte modulo n) ; on a donc un élément dans le centralisateur pour tout élément de $[1, n]$.
• On le généralise ensuite à k cycles de même ordre i.
Alors en se restreignant aux permutations laissant invariants chacun des supports, on a i possibilités, on obtient donc i^k. Mais il reste la possibilité d'intervertir les supports, il faut donc multiplier par $k!$. Il est clair que toutes les permutations ainsi construites sont bien dans le centralisateur ; pour la réciproque, il suffit de supposer que a et $a + 1$ appartenant au support du même cycle (supposés de la forme $(j, j + 1, .., j + i - 1)$, et qu'ils ne sont pas envoyés dans un même support ; on constate alors que notre permutation ne saurait commuter avec notre produit de cycles.
• On le généralise enfin au cas le plus général.
Facile ! Il suffit de faire comme ci-dessus et de constater que quand deux supports ont pas la même taille il est impossible de mettre tous les éléments de l'un dans l'autre...□

Théorème 842 *Si $n \neq 6$ alors tout automorphisme de σ_n est un automorphisme intérieur.*

Démonstration : L'image d'une transposition par un automorphisme ϕ est d'ordre 2, et donc est un produit de k cycles disjoints. Par la proposition 841 le cardinal de son centralisateur est alors $2^k.k!.(n - 2.k)!$; ce cardinal est aussi le cardinal du centralisateur de notre transposition initiale, donc $2.(n - 2)!$. Si $n = 6$, on a une solution avec $n = 6$ et $k = 3$, si $n \neq 6$, on a une seule solution pour $k = 1$. Donc l'image d'une transposition par ϕ est une transposition ; donc par la proposition 840 ϕ est un automorphisme intérieur.□

21.10.12 Groupes en géométrie

⊡ **Groupe diédral D_n**

Définition 843 (Groupe diédral) *On appelle* **groupe diédral d'ordre** n *et on note* D_n *le groupe des isométries du plan conservant un polygone régulier à* n *côtés. Il contient* $2.n$ *éléments, comme on pourra s'en convaincre en distinguant le cas* n *pair et le cas* n *impair ;* n *rotations et* n *symétries.*
On note R_n *l'ensemble des* n *rotations de* D_n.

Proposition 844 *On a* $R_n \lhd D_n$*, et donc*

$$1 \to R_n \to \mathbb{D}_n \xrightarrow{p} \mathbb{Z}/2\mathbb{Z} \to 1$$

en effet D_n *étant d'ordre* $2.n$*, le quotient de* D_n *par* R_n *est d'ordre* 2*, et ne peut donc être qu'isomorphe à* $\mathbb{Z}/2.\mathbb{Z}$*.*
Etant donnée $r \in D_n \setminus R_n$*,* $\{r, Id\}$ *fournit une section ; donc on a* $D_n = R_n \rtimes \{r, Id\}$*, donc* $D_n \simeq \mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$*.*

On pourra consulter la partie 34.3.1 pour plus d'informations sur le groupe diédral.

21.11 Application des groupes à la géométrie

Ci-dessous une liste non exaustive d'applications des groupes en géométrie :
- **cercle unité complexe** $(\mathbb{U}, \times)$, groupe des nombres complexes de module 1, permettant de définir les angles. Isomorphe à $O_2^+(\mathbb{R})$.
- **groupe linéaire** $GL(E)$ des applications linéaire d'un espace vectoriel dans lui-même. Voir 21.10.2.
- **groupe affine** $GA(\xi)$ des bijections affines d'un espace euclidien ξ dans lui-même. Le centre de $GA(\xi)$ est réduit à l'identité. On remarquera notamment que si le groupe additif d'un espace vectoriel agit librement et transitivement sur un ensemble ξ, celui-ci est muni par cette opération d'une structure d'espace affine. Voir 21.10.5.
- **groupe des isométries** d'un ensemble (voir partie34.3)
- **groupe des similitudes** d'un espace euclidien (voir partie 21.10.9)
- **groupe orthogonal** d'un espace euclidien, $O(E)$, voir 21.10.3.
- **groupe projectif** d'un espace vectoriel de dimension finie. Voir 21.10.6.

Chapitre 22

Anneaux

22.1 Définitions

Définition 845 (Anneau) *Un* **anneau** *est un triplet* $(A, +, \times)$ *tel que*
• A *est un ensemble non vide*
• $+$ *est une loi de composition interne (c'est à dire une application de* $A \times A$ *dans* A*), telle que* $(A, +)$ *est un groupe commutatif.*
• $\times$ *est une loi de composition interne associative, ayant un élément neutre, distributive par rapport à* $+$

On appelle **unité** *de* $(A, +, \times)$ *tout élément inversible pour* $\times$*.*
Si en outre $\times$ *est commutative, l'anneau est dit* **commutatif**.
On note 0 *l'élément neutre pour l'addition,* 1 *l'élément neutre pour la multiplication, le symétrique de* $a \in A$ *pour* $+$ *est noté* $-a$*, et le symétrique, lorsque* a *est une unité, de* a *pour* $\times$ *est noté* a^{-1}*.*

$a \times b$ *sera souvent abrégé* $a.b$ *ou même* ab*.*
a *et* b *appartenant à* A *sont dits* **associés** *si* $a = b.x$ *pour un certain* x *unité. La relation d'association est une relation d'équivalence.*
On dit que a **divise** b*, ou que* a *est un* **diviseur** *de* b*, ou que* b *est un* **multiple** *de* a*, pour* a *et* b *dans* A*, s'il existe* x *tel que* $b = a.x$*.*
On dit que a *est un* **plus grand commun diviseur** *ou* **pgcd** *des éléments* $a_1, ..., a_n$*, si pour tout* i*,* $d|a_i$ *et si pour tout* d' $\forall i$ $d'|a_i$ *implique* $d'|d$*. On dit que* a *est un* **plus petit commun multiple** *ou* **ppcm** *des éléments* $a_1, ..., a_n$*, si pour tout* i*,* $a_i|d$ *et si pour tout* d'*,* $\forall i$ $a_i|d'$ *implique* $d|d'$*.* $a \in A$ *est dit* **irréductible** *si* a *n'est pas une unité et si* $b|a$ *implique que* b *est une unité ou que* b *est associé à* a*.*

Les notions de ppcm et pgcd seront surtout utilisées dans le cadre d'anneaux principaux (voir partie 22.2), bien que leur définition puisse être utilisée dans un cadre plus général.

Proposition 846 • *$a \in A$ est irréductible si et seulement si a n'est pas une unité et si $b.c = a$ implique b ou c est une unité.*
• *Dans $\mathbb{Z}$ les éléments irréductibles sont les nombres premiers.*

J'ai ici imposé l'existence d'un élément neutre pour la multiplication ; selon les terminologies ce n'est pas toujours le cas. Si l'on ne suppose pas l'existence d'un élément neutre pour la définition d'un anneau, alors un anneau vérifiant en outre cette propriété sera appelé anneau **unitaire**. Dans la vie de tous les jours, les anneaux sont toujours unitaires. L'hypothèse de commutativité est très classique, mais ici cette hypothèse sera précisée quand elle est nécessaire.

Exemples : :
• $(\mathbb{Z}, +, \times)$ est un anneau.
• $(\mathcal{P}(E), \Delta, \cap)$ est un anneau commutatif, avec Δ la différence symétrique, c'est à dire $A \Delta B = A \cup B - A \cap B$.

Propriétés :(notez que na, pour $n \in \mathbb{N}$ et $a \in A$, désigne $a + a + a + \dots + a$ (a n fois), et a^n désigne $a \times a \times a \times \dots \times a$ (a n fois).
• $1 \neq 0$, à moins que le cardinal de A soit 1.
• $a.0 = 0.a = 0$ pour tout $a \in A$.
• $-(a.b) = (-a).b = a.(-b)$ pour tous $(a, b) \in A^2$
• $(na).b = n.(ab) = a.(nb)$ pour tous $(a, b) \in A^2$ et $n \in \mathbb{N}$.
• L'ensemble des unités forme un groupe pour $\times$.

Proposition 847 (Formule du binôme de Newton) *Soit a et b dans un anneau A. Si a et b commutent, alors*

$$(a + b)^n = \sum_{k \in [0,n]} C_n^k a^k b^{n-k}$$

Démonstration : Par une récurrence sans difficulté, en se rappelant que $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$ □

Exemple Maple
$> \ (x + y)\hat{}\ 3$ $$(x+y)^3$$ $> \ expand(\%);$ $$x^3 + 3x^2y + 3xy^2 + y^3$$

Définition 848 (Diviseurs de 0, anneaux intègres, éléments nilpotents) *1. Un élément a est dit* **diviseur à gauche de** 0 *s'il existe $b \neq 0$ tel que $b.a = 0$.*
Un élément a est dit **diviseur à droite de** 0 *s'il existe $b \neq 0$ tel que $a.b = 0$.*
Un élément est dit **diviseur de** 0 *s'il est à la fois diviseur à gauche de 0 et diviseur à droite de 0.*
Un anneau est dit **sans diviseur de** 0 *s'il n'admet pas de diviseur à gauche de 0 ou de diviseur à droite de 0 autre que 0 lui-même.*

2. Un anneau est dit **intègre** *si :*
• il est de cardinal > 1
• il est commutatif
• il est sans diviseur de 0

3. Un élément a est dit **nilpotent** *s'il existe $n \in \mathbb{N}$ tel que $a^n = 0$. On appelle alors* **indice de nilpotence de** *a le plus petit n convenable non nul.*

Remarques :
• $(\mathbb{Z}, +, \times)$ est un anneau intègre.
• Tout anneau comporte un diviseur de 0 à gauche, un diviseur de 0 à droite, et un diviseur de 0 tout court ; il s'agit de 0 lui-même. Un anneau sans diviseur de 0 ne signifie donc pas que l'anneau ne comporte pas de diviseur de 0.
• Un anneau est sans diviseur de 0 s'il n'admet pas de diviseur à gauche de 0 autre que 0. En effet, si A n'admettant pas de diviseur à gauche de 0 admet un diviseur à droite de 0 autre que 0, alors $0 = ab$ pour a et b non nul, ce qui contredit le fait que 0 n'ait pas de diviseur à gauche.
• De même, un anneau est sans diviseur de 0 s'il n'admet pas de diviseur à droite de 0 autre que 0.
• Un anneau est sans diviseur de 0 si $ab = 0 \Rightarrow a = 0$ ou $b = 0$.

Définition 849 (Morphisme d'anneaux) *Une application f d'un anneau $(A, +, \times)$ vers un anneau $(B, +, \times)$ est un* **morphisme d'anneaux** *(ou* **homomorphisme***) si :*
• f est un morphisme du groupe $(A, +)$ vers le groupe $(B, +)$
• $f(x.y) = f(x).f(y)$ pour tout $(x, y) \in A^2$
• $f(1_A) = 1_B$
On appelle alors **noyau** *de f l'ensemble $ker\ f$ des $x \in A$ tels que $f(x) = 0$.*

Remarques :
• Le noyau d'un morphisme d'anneaux est le noyau du morphisme de groupes sous-jacent.

• 0 appartient au noyau de tout morphisme d'anneaux.
• L'image de l'inverse est l'inverse de l'image, pour chacune des deux lois.

Définition 850 (Produit d'anneaux) *On appelle* **produit de deux anneaux** *leur produit cartésien muni de l'addition terme à terme et de la multiplication terme à terme.*

On vérifie facilement qu'un produit d'anneaux est un anneau.

Définition 851 (Sous-anneau) *Etant donné $(A, +, \times)$ un anneau, une partie B de A est un* **sous-anneau** *de A si*
• $1 \in B$
• $(B, +)$ *est un sous-groupe de* $(A, +)$
• B *est stable par multiplication*

Propriétés :
• Un sous-anneau est un anneau, mais un anneau inclus dans un anneau n'en est pas nécessairement un sous-anneau ; en effet il faut considérer la condition $1 \in B$.
Par exemple l'ensemble des matrices de la forme

$$\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix}$$

est un anneau inclus dans l'anneau des matrices 2×2, mais n'en est pas un sous-anneau.
• L'image réciproque d'un sous-anneau par un morphisme d'anneaux est un sous-anneau.
• L'image d'un sous-anneau par un morphisme d'anneaux est un sous-anneau.

Théorème 852 *Pour tout anneau $(A, +, \times)$, il existe un unique morphisme d'anneaux de $(\mathbb{Z}, +, \times)$ dans $(A, +, \times)$. Il est défini par $\phi(n) = 1_A + ... + 1_A$ n fois et $\phi(-n) = -1_A - 1_A \cdots - 1_A$ n fois pour $n > 0$.*

Démonstration : On vérifie aisément que ϕ ainsi définit est bien un morphisme d'anneau. $\phi(1)$ est nécessairement égal à 1 et $\phi(0)$ à 0. Par récurrence, les propriétés des anneaux permettent de vérifier que les autres éléments sont aussi définis de manière unique.□

Remarque : ceci montre que tout anneau contient un sous-anneau minimal qui est $\phi(\mathbb{Z})$.

22.2 Idéaux, anneaux quotients

Définition 853 (Idéal à gauche, idéal à droite) *On se donne $(A, +, \times)$ un anneau et I une partie non vide de A.*
I est un **idéal à gauche (resp. à droite)** *de $(A, +, \times)$ si*
• I est stable pour l'addition
• $A.I$ est inclus dans I (resp. $I.A$ est inclus dans I)

I est un **idéal** *(parfois on dit* **idéal bilatère***) si I est à la fois un idéal à gauche et un idéal à droite. A et $\{0\}$ sont toujours des idéaux de A ; on les appelle* **idéaux triviaux** *de A. Les autres idéaux sont appelés idéaux non triviaux (on dit parfois aussi* **idéaux propres***) de A.*

Exemples : Dans $\mathcal{M}_n(\mathbb{R})$, l'ensemble des matrices à première colonne nulle est un idéal à gauche, l'ensemble des matrices à première ligne nulle est un idéal à droite.

Propriétés :
• Un idéal contenant 1 ou toute autre unité de l'anneau est l'anneau tout entier.
• La réunion d'une suite croissante d'idéaux est un idéal.
• I idéal de A et J idéal de B ; alors $I \times J$ est un idéal de $A \times B$.
• L'intersection d'une famille d'idéaux est un idéal.

Proposition 854 *• Le noyau d'un morphisme est un idéal.*
• L'image réciproque d'un idéal par un morphisme est un idéal.
• L'image d'un idéal par un morphisme est un idéal de l'image de l'anneau (et pas nécessairement de l'anneau dans lequel l'image est incluse...).

Définition 855 *Une intersection d'idéaux étant un idéal, on peut définir l'***idéal engendré par une partie** *de A comme l'intersection de tous les idéaux contenant cette partie. C'est donc aussi le plus petit idéal contenant cette partie. On note (E) l'idéal engendré par E.*

Définition 856 *On appelle* **idéal principal** *un idéal I d'un anneau commutatif engendré un singleton $\{x\}$. On note abusivement (x) pour $(\{x\})$.*
On appelle **anneau principal** *un anneau intègre tel que tout idéal est principal.*
Un idéal I d'un anneau commutatif est dit **idéal maximal** *s'il est différent de l'anneau tout entier et si tout idéal incluant I est égal à I ou à l'anneau lui-même.*
On appelle **somme** *d'une famille d'idéaux $(I_k)_{k \in K}$ l'ensemble des $\sum_{i \in J} x_i$ avec J fini inclus dans K et $x_i \in I_i$.*
Un idéal est dit **de type fini** *s'il est somme d'un nombre fini d'idéaux principaux.*

Remarques :

• Un anneau principal est donc commutatif, non réduit à $\{0\}$, sans diviseur de 0 ; et tout idéal de cet anneau est principal.
• On notera bien qu'un idéal maximal n'est pas un idéal qui est maximal... Il est en fait maximal parmi les idéaux propres.
• Dans un anneau commutatif A, $(x) = \{x.a/a \in A\}$.
• Une somme d'idéaux est un idéal.
• La somme des idéaux I_k avec $I_k = (x_k)$ est l'idéal engendré par la famille des x_k.

Nb : Un idéal de type fini est donc un idéal engendré par un nombre fini d'éléments.

Proposition 857 *Si a et b sont associés alors $(a) = (b)$.*
Dans un anneau intègre il y a réciproque.

Démonstration : Facile au vu de la dernière remarque.□

⚠ Il n'y a pas de réciproque dans le cas général !

Théorème 858 (Théorème de Bezout) *A est supposé principal.*
• un générateur de $I = (a_1) + (a_2) + ... + (a_n)$ est un pgcd des a_i.
• d, diviseur commun des a_i, est pgcd des a_i si et seulement s'il existe une famille $(\lambda_i)_{i\in[1,n]}$ tels que $d = \sum \lambda_i a_i$ **(relation de Bezout)***.*
• un générateur de $I = (a_1) \cap (a_2) \cap ... \cap (a_n)$ est un ppcm des a_i.

Démonstration :

Le premier • est simple : un tel générateur d doit nécessairement diviser tous les a_i, et il doit nécessairement être dans l'idéal I, et donc tout élément qui divise tous les a_i, étant lui même un générateur de I, doit diviser d.

Le second • est une simple traduction du fait que d soit bien dans I et soit un générateur de I.

Pour le troisième • , donnons-nous p un tel générateur ; il appartient à I, et donc est un multiple de chaque a_i ; et si p' est un autre multiple des a_i, alors il est dans tous les (a_i), et donc appartient à I, et donc est un multiple de p.□

Dans $A = \mathbb{Z}$ ou $A = \mathbb{K}[X]$ avec $\mathbb{K}$ un corps, il est utile de disposer d'un algorithme pratique permettant de découvrir une relation de Bezout entre a et b si une telle relation existe. Pour cela, il suffit de constater que a et b ont même pgcd que a et $a - qb$, pour tout q dans A, par exemple avec q le quotient dans la division euclidienne de a par b. Si a est divisible par b, le pgcd de a et b est simplement b ; sinon, on effectue une division euclidienne. Considérons un exemple pratique, cherchons le pgcd de 42 et 30.

$$30 \not| 42$$

$$42 = 1 \times 30 + 12$$

$$12 \not| 30$$

$$30 = 2 \times 12 + 6$$

$$12|6 \text{ et } 12 = 2 \times 6$$

Donc

$$6 = 30 - 2 \times 12 = 30 - 2 \times (42 - 30) = 3 \times 30 - 2 \times 42$$

ce qui est bien la relation de Bezout attendue. Cet algorithme est appelé **algorithme d'Euclide**.

Définition 859 *Un idéal I est dit* **premier** *si et seulement si A/I est intègre. Un élément d'un anneau est dit* **premier** *si et seulement l'idéal engendré par cet élément est premier.*

Proposition 860 • *Un idéal I de A est premier si et seulement s'il est différent de A et si $a.b \in I$ implique $a \in I$ ou $b \in I$.*
• *L'image réciproque d'un idéal premier par un homomorphisme d'anneaux est un idéal premier.*

La première de ces deux propriétés est fondamentale car c'est généralement celle que l'on utilise pour montrer qu'un idéal est premier.

Proposition 861 *Un anneau commutatif est intègre si et seulement si (0) est un idéal premier.*

Démonstration :

(0) idéal premier
si et seulement si
$a.b \in (0) \to a \in (0)$ ou $b \in (0)$
si et seulement si
$a.b = 0 \to a = 0$ ou $b = 0$
si et seulement si
A intègre$\square$

Lemme 862 *Soit A un anneau. A est un corps si et seulement si A est non réduit à $\{0\}$ et ses seuls idéaux sont $\{0\}$ et A.*

Démonstration : • Supposons que les seuls idéaux de A soient $\{0\}$ et A.
Soit x dans A, $x \neq 0$, $x.A$ est un idéal, autre que $\{0\}$, donc il contient tout A, donc en particulier il contient 1, donc il est inversible.
• Réciproquement si A est un corps, alors soit x non nul appartenant à un idéal I, alors I contient $x.A$, donc $x.x^{-1}.A$, donc A.$\square$

Proposition 863 *Dans un anneau principal, les idéaux premiers sont (0) et (p), avec p irréductible.*

Démonstration : • Soit I un idéal premier et $p \neq 0$ un élément de A tel que $I = (p)$. Supposons que $p = a.b$. Alors $a.b \in (p)$, et donc puisque I est premier,

$a \in (p)$ ou $b \in (p)$; on suppose $a \in (p)$. Alors $a = p.a'$. On a alors $p.a'.b = p$, donc $p(1 - a'b) = 0$, or A est intègre, donc $a'.b = 1$, donc b est une unité.□

Proposition 864 *Dans un anneau principal, pour tout p irréductible, (p) est un idéal maximal.*

Démonstration : Soit $I = (p)$, avec p irréductible. Supposons $I \subset J$, avec J inclus dans A. Alors $J = (q)$, et $p = q.a$. Mais p étant irréductible, soit $q = p.x$ avec x unité, soit q est une unité. Dans le premier cas, $J = I$, et dans le deuxième cas, $J = A$. □

On va maintenant étudier la notion d'anneau quotient.
Cette notion n'est étudiée que dans le cas d'anneaux commutatifs.

Définition 865 *Etant donné I un idéal de A, on définit une relation d'équivalence $\mathcal{R}_I$ par*

$$a\mathcal{R}b \iff a - b \in I$$

Alors l'ensemble quotient pour cette relation, muni des opérations induites par les opérations sur I, est un anneau ; on l'appelle **anneau quotient** *de A par l'idéal I, et on le note A/I.*

Il convient de vérifier que la relation est bien compatible avec les opérations définies sur l'anneau (vérification aisée).

22.3 Décomposition d'un homomorphisme d'anneaux et utilisation des idéaux

Définition 866 (Factorisation d'un homomorphisme) *On dit que f homomorphisme d'un anneau A vers un anneau B se factorise par A/I avec I idéal de A si et seulement s'il existe g homomorphisme de A/I dans B tel que $f(x) = g(\overline{x})$.*

Théorème 867 *Soit f un homomorphisme d'anneaux de A vers B. Alors pour tout I idéal inclus dans $Ker\ f$, on définit $x \to \overline{x}$ la projection canonique de A sur A/I, et on a les propriétés suivantes :*
- *Il existe un unique homomorphisme g de $A/IdansB$ tel que $\forall x\ f(x) = g(\overline{x})$*
- *$Im\ f \simeq A/Ker\ f$*
- *g est injectif si et seulement si $I = Ker\ f$*
- *g est surjectif si et seulement si f est surjectif*

Proposition 868 (Image et image réciproque d'un idéal par un homomorphisme)
• *L'image réciproque d'un idéal par un homomorphisme est un idéal*
• *Si f est un homomorphisme surjectif, alors l'image d'un idéal par f est un idéal.*

Proposition 869 *Soit I idéal de A. Alors l'application ϕ qui à un idéal J avec $I \subset J \subset A$ associe la projection $\overline{J}$ de J sur A/I est une bijection de l'ensemble des idéaux de A contenant I vers l'ensemble des idéaux de A/I.*

Démonstration : • $x \mapsto \overline{x}$ étant surjectif, il est clair que ϕ associe bien un idéal à un idéal.
• Pour montrer que ϕ est bijective, on considère l'application ψ qui à un idéal K de A/I associe $\{a/\overline{a} \in K\}$. □

Proposition 870 *Un idéal I d'un anneau A est maximal si et seulement si A/I est un corps.*

Démonstration : On utilise le lemme 862 et la propriété ci-dessus.□

Corollaire 871 *Tout idéal maximal est premier.*

Démonstration : Supposons que I, idéal maximal, contient $a.b$, avec $a \notin I$ et $b \notin I$.
Alors la classe de a et la classe de b dans A/I sont non nulles, et leur produit est nul, d'où contradiction.□

Théorème 872 (Krull) *Pour tout idéal I de A, I différent de A, il existe un idéal maximal de A contenant I.*

Démonstration : Cette preuve nécéssite l'axiome du choix, via le théorème de Zorn (voir le lemme 36).
• On considère l'ensemble des idéaux différents de A contenant I idéal de A, ordonné par l'inclusion.
• Cet ensemble est inductif. En effet étant donnée une chaîne, on considère la réunion, c'est un idéal différent de A (en effet il ne contient pas 1 par exemple).
• On peut donc considérer un élément maximal pour l'inclusion, et conclure que cet idéal est maximal.□

22.4 Anneaux commutatifs

Dans la vie de tous les jours, comme je l'écrivais un peu plus haut, les anneaux sont généralement supposés commutatifs. On va maintenant étudier des cas particuliers d'anneaux commutatifs, avec des cas de plus en plus riches. Tout d'abord les anneaux euclidiens, puis les anneaux noethériens, puis les anneaux intègres, puis les anneaux factoriels, puis les anneaux principaux. On a les implications (principal $\Rightarrow$ factoriel) et (factoriel $\Rightarrow$ intègre).

22.4.1 Anneaux euclidiens

Cette notion n'est étudiée que dans le cas d'anneaux commutatifs.

Définition 873 *Un anneau A commutatif est dit* **euclidien** *pour une application f de $A \setminus \{0\}$ dans $\mathbb{N}$, si pour tout a dans A et tout b dans $A \setminus \{0\}$ il existe $(q, r) \in A^2$ tels que $a = b.q + r$ et $r = 0$ ou $f(r) < f(b)$.*
Un anneau A commutatif est dit **euclidien** *s'il existe une application pour laquelle il est euclidien.*

Proposition 874 • *$\mathbb{Z}$ est euclidien.*
• *$\mathbb{K}[X]$ est euclidien.*

Démonstration : Considérer respectivement :
• $f(z) = |z|$
• $f(P) = deg(P)$ (voir la démonstration de la division euclidienneen 25.2)□

Proposition 875 *Etant donnée f une application multiplicative (i.e. $f(a.b) = f(a).f(b)$) de $A \setminus \{0\}$ dans $\mathbb{N} \setminus \{0\}$, avec A anneau intègre, on prolonge f multiplicativement sur le corps des fractions de A en posant $f(a/b) = f(a)/f(b)$ (f est maintenant à valeurs dans $\mathbb{Q}$). Alors A est euclidien pour f si et seulement si pour tout x dans le corps des fractions il existe a dans A tel que $f(x-a) < 1$.*

Démonstration : Facile...□

Proposition 876 *Tout anneau euclidien est principal.*

Démonstration : Soit A un tel anneau (commutatif, euclidien). On se donne P_0 dans $I \setminus \{0\}$ tel que $f(P_0)$ soit minimal. On note I' l'idéal engendré par P_0, c'est à dire l'ensemble des $P_0.P$ pour $P \in \mathbb{K}[X]$.
Pour tout P dans I, on utilise la définition $P = P_0.Q + R$; alors $P \in I$, $P_0 \in I$, donc $P_0.Q \in I$ (par définition d'un idéal), et donc $R \in I$; or $f(R) < f(P_0)$ si R est non nul, ce qui contredit la définition de P_0, donc R est nul, donc $P \in I'$, donc $I = I'$ et

donc l'anneau est principal.□

Proposition 877 *$\mathbb{Z}[i]$ et $\mathbb{Z}[\sqrt{2}]$ sont euclidiens.*

Démonstration : Dans les deux cas on utilise la caractérisation de la proposition 875.
Dans le premier cas on choisit $f(a+i.b) = |a+i.b|$ pour a et b dans $\mathbb{Q}$.
Dans le second cas on utilise $f(a+b.\sqrt{2}) = |a^2 - 2.b^2|$ si a et b dans $\mathbb{Z}$.
Ce second choix est particulièrement instructif ; $f(a+b.\sqrt{d}) = |a^2 - d.b^2|$ sera souvent utile.□

22.4.2 Anneaux noethériens

Définition 878 (Anneau noethérien) *Un anneau commutatif dont tout idéal est de type fini est dit* **noethérien***.*

Proposition 879 *Un anneau commutatif est noethérien si et seulement si toute suite croissante d'idéaux est stationnaire à partir d'un certain rang.*

Démonstration : Trop facile pour que nous le prouvions ! Exercice pour le lecteur.□

Proposition 880 *Un anneau commutatif A est noethérien si et seulement si tout ensemble non vide d'idéaux de A admet un élément maximal pour l'inclusion.*

Démonstration : Rappelons juste qu'un élément maximal n'est pas nécessairement le plus grand élément, l'existence d'un élément maximal n'entraîne pas même celle d'un plus grand élément (voir les définitions en partie2).□

Proposition 881 • *Tout anneau quotient d'un anneau noethérien est noethérien.*
• *Un anneau principal est noethérien.*

Démonstration : • La proposition 869 montre qu'un idéal du quotient est la projection d'un idéal, ce dernier étant de type fini, le projeté est de type fini.
• Facile, tout idéal d'un anneau principal est engendré par un seul élément, donc par un nombre fini d'éléments.□

⚠ La propriété annoncée pour les anneaux quotients n'est pas vraie pour les sous-anneaux.

Théorème 882 (Théorème de Hilbert) *Si A est un anneau noethérien, alors pour tout n $A[X_1, ..., X_n]$ est aussi un anneau noethérien.*

<u>Démonstration :</u> Admis (preuve difficile).□

:

• Tout corps est un anneau noethérien, donc tout $K[x_1, \ldots, x_n]$ aussi.
• $\mathbb{Z}[x_1, \ldots, x_n]$ est noethérien.

22.4.3 Anneaux intègres

On a déjà vu les définitions, mais voici un rappel : Un anneau est dit intègre si :
• il est de cardinal > 1
• il est commutatif
• il est sans diviseur de 0

Définition 883 (Définitions dans les anneaux intègres) *a et b dans A anneau intègre sont dits* **premiers entre eux** *si*

$$\forall x \in a\ x|a \text{ et } x|b \to x \text{ est une unité}$$

De même les éléments d'une famille $(a_i)_{i\in[1,n]}$ sont dits premiers entre eux si un élément divisant tous les a_i est nécessairement une unité.

Corollaire 884 (Théorème de Bezout) *Dans un anneau principal des éléments a_i sont premiers entre eux si et seulement s'il existe une famille λ_i d'éléments de A telle que $\sum \lambda_i a_i$ soit une unité.*

Proposition 885 *Soit A un anneau intègre, et ϕ l'application qui à x dans A quotienté par la relation d'association $\mathcal{R}$ associe l'idéal engendré par x. ϕ est un isomorphisme d'ordre entre $A/\mathcal{R}$ muni de la divisibilité et l'ensemble des idéaux principaux de A muni de l'inverse de l'inclusion.*

<u>Démonstration :</u> • Tout d'abord il est clair que ϕ est bien définie, car deux éléments associés engendrent évidemment le même idéal.
• L'application est surjective, par définition, puisqu'on considère l'ensemble des idéaux principaux.
• Montrons que l'application est injective : si deux éléments a et b engendrent le même idéal alors $b = b'.a$ et $a = a'.b$ et donc $b = b'.a'.b$ et donc b' et a' sont des unités (car A est intègre), et donc $\overline{b} = \overline{a}$.
• Montrons qu'il s'agit d'un morphisme d'ordres :
- Si $a|b$ alors $b = a.c$ donc $b.A = a.c.A \subset a.A$ et $(b) \subset (a)$ clairement.

- Si $(b) \subset (a)$ alors $b = a.c$ pour un certain c et donc $a|b$.□

22.4.4 Anneaux factoriels

Cette notion n'est étudiée que dans le cadre d'anneaux commutatifs.

Définition 886 (Anneau factoriel) *Un anneau A est dit* **factoriel** *si :*
• *il est intègre*
• *tout a dans A s'écrit de manière unique à association près et à permutation près $a = a'.p_1.p_2.....p_n$ avec a' unité et p_i irréductible pour tout i.*

Etant donné un élément p irréductible de A un anneau factoriel, on appelle **valuation p-adique de** *A pour a dans A le nombre d'occurences d'un élément associé à p dans la décomposition de a sous forme $a = a'.p_1.....p_n$. On note généralement $v_p(A)$ la valuation p-adique de a.*

Proposition 887 *Etant donné A un anneau factoriel, on peut choisir un élément dans chaque classe d'équivalence de A pour la relation d'association $\mathcal{R}$. L'ensemble de ces éléments permet de simplifier la décomposition de $a \in A$ en $a = a'.\Pi_{i \in A/\mathcal{R}} p_i^{v_{p_i}(a)}$, le support de $i \mapsto v_{p_i}(a)$ étant fini.*

Démonstration : Evident.□

Voyons maintenant quelques propriétés intéressantes des anneaux factoriels.

Proposition 888 *Dans un anneau factoriel un élément est irréductible si et seulement s'il est premier.*

Le lemme et le théorème qui suivent se démontrent très facilement, simplement en considérant les décompositions de x, y et éventuellement z pour conclure.

Lemme 889 (lemme d'Euclide) *Si A est un anneau factoriel, alors si p est irréductible et divise $x.y$, alors p divise x ou p divise y.*

Théorème 890 (Théorème de Gauss) *Si z divise $x.y$ et si z est premier avec x alors z divise y.*

Un exemple d'application (parmi beaucoup d'autres) est le théorème 942.

Proposition 891 *Un anneau intègre noethérien vérifiant le lemme d'Euclide ou le théorème de Gauss est factoriel.*

Démonstration : Admis.□

Dans l'exemple ci-dessous on utilise le fait que $\mathbb{Z}$ est factoriel.

Exemple Maple
$> ifactor(200!);$
$(2)^{197}(3)^{97}(5)^{49}(7)^{32}(11)^{19}(13)^{16}(17)^{11}(19)^{10}(23)^{8}(29)^{6}(31)^{6}(37)^{5}(41)^{4}$ $(43)^{4}(47)^{4}(53)^{3}(59)^{3}(61)^{3}(67)^{2}(71)^{2}(73)^{2}(79)^{2}(83)^{2}(89)^{2}(97)^{2}(101)(103)$ $(107)(109)(113)(127)(131)(137)(139)(149)(151)(157)(163)(167)(173)(179)$ $(181)(191)(193)(197)(199)$

22.4.5 Anneaux principaux

On rappelle tout d'abord la définition d'un anneau principal : il s'agit d'un anneau intègre dont tout idéal est principal.

Je donne sans démonstration (voir [19, p156]) le résultat important suivant :

Proposition 892 *Tout anneau principal est factoriel.*

On peut préciser aussi que sur le corps des fractions rationnelles à coefficients dans un anneau principal, on dispose de la décomposition en éléments simples.

22.5 Zoologie des anneaux

22.5.1 Nilpotence (d'une somme de deux éléments nilpotents qui commutent)

Proposition 893 *La somme de deux éléments nilpotents qui commutent est nilpotente.*

Démonstration : Considérer deux tels éléments a et b, et développer par le binome de Newton 847 la puissance $(a+b)^n$ avec n la somme de leurs indices de nilpotence respectifs.$\square$

22.5.2 $\mathbb{Z}/n\mathbb{Z}$

⊡ Généralités

Etant donné $m \in \mathbb{N}$ on note $\overline{m}$ la classe de m dans $\mathbb{Z}/n\mathbb{Z}$ (la relation d'équivalence considérée étant la congruence modulo n - x et y sont équivalents si n divise $x-y$).

Proposition 894 *On a équivalence entre les propriétés suivantes :*
• m *est premier avec* n
• $\overline{m}$ *est générateur du groupe* $(\mathbb{Z}/n\mathbb{Z}, +)$
• $\overline{m}$ *est inversible dans l'anneau* $(\mathbb{Z}/n\mathbb{Z}, +, \times)$

Démonstration : Facile, en application du théorème de Bezout.□

Définition 895 (Fonction d'Euler) *On appelle* **fonction d'Euler** *la fonction* ϕ *telle que* $\phi(n)$ *soit le nombre d'entiers* x *tels que* $1 \leq x \leq n$ *et* $x \wedge n = 1$.

Proposition 896 • *Si* n *est premier* $\phi(n) = n - 1$ *et* $\phi(n^r) = n^{r-1}.(n-1)$ *si* $r > 0$
• $\phi(n)$ *est le nombre d'éléments inversibles de* $(\mathbb{Z}/n\mathbb{Z}, +, \times)$

Démonstration : le premier point est clair ; il suffit de voir qu'un élément est premier avec p^n si et seulement s'il n'est pas divisible par p.
Le second point est un corollaire de la proposition précédente.□

⊡ Lemme chinois

Lemme 897 (Lemme Chinois) *Si* p *et* q *sont premiers entre eux alors*

$$(\mathbb{Z}/pq\mathbb{Z}, +) \simeq (\mathbb{Z}/p\mathbb{Z}, +) \times (\mathbb{Z}/q\mathbb{Z}, +)$$

Démonstration : Il s'agit des groupes additifs usuels. L'égalité des cardinaux montre qu'il suffit de trouver un morphisme de groupes injectif. Pour cela on associe à la classe de n dans $\mathbb{Z}/pq\mathbb{Z}$ la classe de n dans $\mathbb{Z}/p\mathbb{Z}$ et la classe de n dans $\mathbb{Z}/q\mathbb{Z}$. Il est clair que si deux entiers ont la même classe modulo pq alors ils ont la même classe modulo p et modulo q, donc l'application est bien définie.
Le fait que cette application soit un morphisme est clair.
L'application est injective, car si deux entiers ont la même classe modulo p et q, alors ils ont la même classe modulo pq.□

Corollaire 898 *Si $n = \Pi_i p_i^{\alpha_i}$, avec les p_i premiers distincts et les $\alpha_i > 0$, alors*

$$(\mathbb{Z}/n\mathbb{Z}, +, \times) \simeq \Pi_i(\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}, +, \times)$$

$$(\mathbb{Z}/n\mathbb{Z})^* \simeq \Pi_i(\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})^*$$

$$\phi(n) = \Pi_i\phi(p_i^{\alpha_i}) = n.\Pi_i(1 - 1/P_i)$$

Démonstration : Le premier point découle de l'utilisation récurrente du lemme chinois, le deuxième et le troisième sont des conséquences immédiates du premier.□

⊡ Automorphismes de $\mathbb{Z}/n\mathbb{Z}$

Proposition 899 *L'ensemble des automorphismes de $\mathbb{Z}/n\mathbb{Z}$ est isomorphe à $(\mathbb{Z}/n\mathbb{Z})^*$.*

Démonstration : Il suffit de considérer l'application ψ qui à un élément inversible m associe l'automorphisme $x \mapsto m.x$;
• il est clair que c'est un morphisme injectif de $(\mathbb{Z}/n\mathbb{Z})^*$ dans $Aut(\mathbb{Z}/n\mathbb{Z})$
• étant donné un automorphisme f de $\mathbb{Z}/n\mathbb{Z}$ on montre facilement qu'il est égal à $\psi(f(1))$.□

Corollaire 900 *$Aut(\mathbb{Z}/n\mathbb{Z})$ est un groupe abélien d'ordre $\phi(n)$.*

⊡ Forme des groupes $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$

p désigne un nombre premier.

Lemme 901 $(\mathbb{Z}/p\mathbb{Z})^* \simeq (\mathbb{Z}/(p-1)\mathbb{Z}, +)$

Démonstration : $(\mathbb{Z}/p\mathbb{Z})$ est un corps fini (voir le chapitre sur la théorie des groupes).

On sait (voir proposition 921) que le groupe multiplicatif d'un corps fini est cyclique, donc isomorphe à un certain $(\mathbb{Z}/n\mathbb{Z})$.

Il suffit donc de se rappeler que le cardinal de $(\mathbb{Z}/p\mathbb{Z})^*$ est $p-1$ pour conclure.□

Lemme 902 *Si $k > 0$, alors $(1+p)^{p^k} = 1 + \lambda.p^{k+1}$, avec $\lambda > 0$ et $\lambda \wedge p = 1$.*

Démonstration : Par récurrence sur k :
• $k = 1$
$(1+p)^p = \sum_{i=0}^{p} C_p^i p^i$

donc $(1+p)^p = 1 + p^2 + m.p^3 = 1 + p^2.(1+m.p)$
• k quelconque
On écrit $(1+p)^{p^{k+1}} = ((1+p)^{p^k})^p = (1+\lambda.p^{k+1})^p$; il suffit alors de développer en utilisant le binôme de Newton la puissance k-ième en isolant le premier et le dernier terme.$\square$

Corollaire 903 *$1+p$ est d'ordre $p^{\alpha-1}$ dans $\mathbb{Z}/p^\alpha\mathbb{Z}$.*

Démonstration : il est d'ordre au plus $p^{\alpha-1}$ au vu du lemme précédent.
En outre $(1+p)^{p^{\alpha-2}} = 1 + \lambda.p^{\alpha-1}$, et donc ne saurait être congru à 1 modulo p^α.$\square$

Lemme 904 *Si m et t sont premiers entre eux, et si a et b commutent, et si a est d'ordre m et b est d'ordre t, alors $a.b$ est d'ordre $m.t$.*

Démonstration : • Il est facile de voir que ab est d'ordre au plus $m.t$, puisque a et b commutent.
• Réciproquement, si $(ab)^n = 1$, alors $a^{n.t}.b^{n.t} = 1$, donc $a^{n.t} = 1$, puisque $b^{n.t} = 1$. Donc t divise l'ordre de ab. De même m divise l'ordre de ab ; donc $m.t$ divise l'ordre de ab, puisque m et t sont premiers entre eux.$\square$

Proposition 905 *Si p premier > 2, $m \geq 2$, alors*

$$(\mathbb{Z}/p^\alpha\mathbb{Z})^* \simeq \mathbb{Z}/\phi(p^\alpha)\mathbb{Z} \simeq \mathbb{Z}/p^{\alpha-1}.(p-1)\mathbb{Z}$$

Démonstration : On va utiliser les lemmes précédents.
• On considère tout d'abord l'application ψ définie par

$$\psi : (\mathbb{Z}/p^\alpha\mathbb{Z})^* \to (\mathbb{Z}/p\mathbb{Z})^*$$

ψ étant la fonction induite par l'identité (il convient de bien vérifier que ψ est bien définie et est un morphisme de groupes surjectif)
• par le lemme 901 tout élément dont l'image par ψ est non égal à $\overline{1}$ engendre $(\mathbb{Z}/p\mathbb{Z})^*$; donc son ordre est un multiple de $p-1$ (voir le lemme 901).
• étant donné x un tel élément, il existe y appartenant au groupe engendré par x tel que y est d'ordre $p-1$.
• On applique alors le lemme 904, $y.(p+1)$ est d'ordre le produit des ordres de y et de $p+1$; or y est d'ordre $p-1$ comme on vient de le voir, et $p+1$ est d'ordre $p^{\alpha-1}$ par le corollaire 903 ; $y.(p+1)$ est donc d'ordre $(p+1).p^{\alpha-1}$; le groupe engendré par y est donc nécessairement $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ tout entier, d'où le résultat.$\square$

On vient donc par cette proposition de détailler la forme des $(\mathbb{Z}/p^\alpha\mathbb{Z})^*$ dans le cas où $p > 2$. Il convient de considérer le cas $p = 2$.

Lemme 906
$$k > 0 \rightarrow 5^{2^k} = 1 + \lambda.2^{k+2}$$
avec $\lambda \wedge 2 = 1$ (i.e. λ impair)

Démonstration : Facile, par récurrence.□

Proposition 907 $(\mathbb{Z}/2.\mathbb{Z})^* = \{1\}$, $(\mathbb{Z}/4.\mathbb{Z})^* = \{1, -1\} \simeq \mathbb{Z}/2.\mathbb{Z}$, *et ensuite (pour $\alpha \leq 3$)* $(\mathbb{Z}/2^\alpha\mathbb{Z})^* \simeq \mathbb{Z}/2.\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z}$.

Démonstration :
• On considère le morphisme surjectif ψ induit par l'identité de $(\mathbb{Z}/2^\alpha\mathbb{Z})^*$ sur $(\mathbb{Z}/4.\mathbb{Z})^* = \{1, -1\} \simeq \mathbb{Z}/2.\mathbb{Z}$.
• Le noyau de ψ est d'ordre $2^{\alpha-2}$
• 5 appartient au noyau de ψ.
• par le lemme 906, 5 est d'ordre $2^{\alpha-2}$ (l'ordre est une puissance de 2, et $5^{2^{\alpha-3}}$ ne peut être congru à 1)
• le noyau de ψ est donc cyclique (au vu des trois affirmations précédentes)
• On a alors la suite exacte (voir chapitre sur la théorie des groupes)

$$1 \rightarrow \mathbb{Z}/2^{\alpha-2} \rightarrow (\mathbb{Z}/2^\alpha\mathbb{Z})^* \xrightarrow{\psi} \mathbb{Z}/2.\mathbb{Z} \rightarrow 1$$

• Le sous-groupe $\{-1, 1\}$ de $(\mathbb{Z}/2^\alpha\mathbb{Z})^*$ est une section de ψ, et il est distingué puisque notre groupe $(\mathbb{Z}/2^\alpha\mathbb{Z})^*$ est abélien. Donc on a un produit direct

$$(\mathbb{Z}/2^\alpha\mathbb{Z})^* \simeq \mathbb{Z}/2.\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z}$$

ce qui conclut la preuve (précisons que le produit direct $A \times B$ est isomorphe au produit direct $B \times A$...).□

22.5.3 Idéaux étrangers

Définition 908 *Soit A un anneau et c et d deux idéaux bilatèresde A. Les anneaux c et d sont dits étrangers (ou comaximaux) si $c + d = A$.*

Exemples : deux idéaux maximaux sont toujours étrangers.

Proposition 909 *Si $a_1, \ldots, a_n$, $b_1, \ldots, b_m$ sont des idéaux bilatères de A, et si pour tout $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ a_i et b_j sont étrangers, alors les idéaux $a_1 \times a_2 \times \cdots \times a_n$ et $b_1 \times b_2 \times \cdots \times b_m$ sont étrangers.*

Démonstration : On fait d'abord la preuve pour $m = 1$ en utilisant des égalités : $x_i + y_i = 1$ avec $x_i \in a_i$ et $y_i \in b_1$. On multiplie terme à terme. Puis on fait pareil avec $a_1, \ldots, a_n$ et chaque b_i. □

Remarque : dans $\mathbb{Z}$, a et b sont premiers entre eux si et seulement si (a) et (b) sont étrangers.

Chapitre 23

Corps

23.1 Définitions de base

Définition 910 *Un anneau* $(K, +, .)$ *est un* **corps** *si et seulement si le groupe des unités est* $K - \{0\}$.
Un corps est dit **commutatif** *si l'anneau sous-jacent est commutatif, c'est à dire si la multiplication est commutative.*

Propriétés :
• Un anneau commutatif non nul est un corps si et seulement si ses seuls idéaux sont les idéaux triviaux.
• Un anneau intègre fini est un corps.

23.2 Extensions de corps

Définition 911 *Un sous-anneau* L *de l'anneau sous-jacent à un corps* K *est un* **sous-corps** *de* K *si c'est un corps pour les lois induites.*
Si L *est un sous-corps de* K*, on dit que* K *est un* **sur-corps** *ou une* **extension** *de* L.
Avec L *sous-corps de* K*, et* $A \subset K$*, on dit que* A **engendre** K *sur* L *si* K *est le plus petit sous-corps de* K *contenant* A *et* L*. On note alors* $K = L(A)$*. Si* A *est fini on note* $K = L(a_1, ..., a_n)$*. L'extension est dite monogène si* A *contient un seul élément.*

Théorème 912 *Etant donné un anneau intègre* A*, il existe un unique corps* K *(à isomorphisme près) contenant un anneau intègre* B *isomorphe à* A *et tel que tout sous-corps de* K *contenant* B *soit* K *lui-même.*

Démonstration : On procède selon les étapes suivantes pour montrer l'existence :
• On considère les classes d'équivalences sur $A \times A$ pour la relation $\mathcal{R}$ définie par $(x,y)\mathcal{R}(x',y') \iff xy' = x'y$ (intuitivement les classes d'équivalence sont les fractions). Appelons K l'ensemble quotient ainsi obtenu.
• On considère ensuite l'addition sur ces classes, facile à retrouver au vu de la considération sur les fractions ; il s'agit de $(x,y)+(x',y') = (xy'+x'y, yy')$. De même la multiplication est définie par $(a,b).(a',b') = (aa',bb')$. Il est facile de voir que ces lois vérifient toutes les propriétés souhaitées, et qu'elles sont bien définies dans la structure quotient. On trouve un élément $(0,1)$ neutre pour l'addition, et un élément $(1,1)$ neutre pour la multiplication.
• L'application qui à x associe $(x,1)$ est un morphisme injectif de A dans K. C'est donc un isomorphisme de A sur son image A'.
• Etant donné un sous-corps de K contenant A', il contient nécéssairement les quotients d'éléments de A', et donc K tout entier.
• Il ne reste plus qu'à vérifier l'unicité de K, à isomorphisme près. Cette tâche est laissée au lecteur.□

Applications :
• Construction de $\mathbb{Q}$ à partir de $\mathbb{Z}$.
• Construction du corps des fractions rationnelles, à partir de l'anneau des polynômes.

Proposition 913 • *Si L est un sous-corps de K, alors K est un L-espace vectoriel.*
• *Si la dimension de K en tant que L-espace vectoriel est finie alors on l'appelle* **degré** *de K pour L et on le note $[K : L]$.*
• *Si K et L sont finis, alors $|K| = |L|^{[K:L]}$.*

Démonstration : Le premier point est clair.
Le second point est une définition.
Le troisième point est clair.□

Théorème 914 (Théorème des bases téléscopiques) *Si $M \subset L \subset K$ (tous trois des corps) alors si e_i est une base de K en tant que L-espace vectoriel et si f_j est une base de L en tant que M-espace vectoriel , alors $e_i.f_j$ est une base de K en tant que M-espace vectoriel . Donc $[K : M] = [K : L].[L : M]$.*

Démonstration : Facile.□

Définition 915 (Différentes extensions de corps) *Si L est une extension du corps K, alors un élément a de L est dit* **algébrique sur** *K s'il existe un polynôme P à coefficients dans K tel que $P(a) = 0$. Un nombre réel est souvent dit simplement* **algébrique** *s'il est algébrique sur $\mathbb{Q}$. L'ensemble des éléments de L algébriques sur K est appelée extension algébrique de K dans L.*
Etant donné K un corps et $P \in K[X]$, on appelle **corps de rupture de** *P un sur-corps L de K dans lequel P admet une racine a et tel que $L = K(a)$.*
Etant donné K un corps et $P \in K[X]$, on appelle **corps de décomposition de** *P un sur-corps L de K dans lequel P est scindé et $L = K(Z)$, avec Z l'ensemble des zéros de P dans L.*
Etant donné K un corps, on appelle **cloture algébrique** *de K une extension de K algébriquement close et dont tous les éléments soient algébriques sur K.*

On a existence du corps de décomposition, et existence du corps de rupture lorsque le polynôme est irréductible. Dans les deux cas, on a unicité à isomorphisme près. Le théorème de Steinitz (difficile) montre que tout corps admet une cloture algébrique, unique à isomorphisme près.

Démonstration : (de l'existence du corps de rupture) Le corps $K(X)/(P)$ convient (ie le quotient de K par l'idéal engendré par P). □

23.3 Corps finis

Proposition 916 *Un anneau intègre fini est un corps.*

Démonstration : Si un anneau est intègre, l'application $x \mapsto yx$ est bijective pour tout y. En particulier, il existe x tel que $yx = 1$.□

Théorème 917 *Un corps fini n'est jamais algébriquement clos.*

Démonstration : Facile ; il suffit de considérer le polynôme $\Pi_{k\in K}(X - k) + 1$.□

Théorème 918 *Quel que soit p premier, quel que soit n dans $\mathbb{N}$ non nul, il existe un unique corps, à isomorphisme près, de cardinal p^n. Tout corps fini est de cette forme.*

Théorème 919 (Wedderburn) *Tout corps fini est commutatif.*

Démonstration : Ces résultat, non triviaux, ne seront pas prouvés ici. On pourra consulter [14] pour une preuve compréhensible.□

Enfin deux résultats (non triviaux) donnés sans preuve :

Proposition 920 *Le groupe des automorphismes d'un corps fini de cardinal* p^n *est cyclique, d'ordre* n*, engendré par* $x \mapsto x^n$.

Proposition 921 *Le groupe multiplicatif d'un corps fini est cyclique.*

Chapitre 24

Quelques résultats supplémentaires d'arithmétique et théorie des nombres

24.1 Sous-groupes additifs de $\mathbb{R}$

Proposition 922 *Tout sous-groupe G du groupe $(\mathbb{R}, +)$ vérifie l'une et une seule des deux conditions suivantes :*

$$\exists x / G = x\mathbb{Z}$$

G est dense dans $\mathbb{R}$

Démonstration :
• On considère $\alpha = inf G \cap R^{+*}$.
• On distingue les deux cas $\alpha > 0$ et $\alpha = 0$.$\square$

24.2 Représentation p-adique des réels

Définition 923 *On se donne un entier $p > 1$. On appelle représentation p-adique du réel x la suite d'entiers $(c_n)_{n\in\mathbb{N}}$ définie par*

$$c_n = \left\{ \begin{array}{c} E(x) \text{ si } n = 0 \\ \frac{1}{p^n} E(p^n x) - pE(p^{n-1}x) \text{ sinon} \end{array} \right.$$

($E(y)$ désignant la partie entière de y).

□

Théorème 924 (Caractérisation du développement p-adique) Théorème 925 *Le développement p-adique de $x \in \mathbb{R}$ est périodique à partir d'un certain rang si et seulement si x est rationnel.*

<u>**Démonstration :**</u>

• Supposons tout d'abord le développement périodique.

Alors x est somme des $c_n p^{-n}$ pour $n \in \mathbb{N}$. Vue la périodicité, cette somme se réécrit comme somme d'un rationnel et de $\sum_{n\geq N} \frac{a}{(p^{-k})^n}$, avec a dans $\mathbb{N}$, et donc x est somme d'un rationnel et de $\frac{ap^{-kN}}{1-p^{-k}}$, et donc x est rationnel.

• Réciproquement supposons que x soit rationnel.

- On peut écrire $x = a/b$ avec a et b dans $\mathbb{N}$ (on se limite au cas $x > 0$, les autres cas étant similaires)

- On définit $x_0 = a$, et par récurrence $x_{n+1} = (x_n - bc_n)p$, avec c_n le quotient dans la division euclidienne de x_n par b.

- On montre facilement par récurrence que $0 \leq x_i < bp$ pour tout i et que les c_i sont le développement p-adique de x.

- les x_i étant bornés, on passe nécéssairement deux fois par la même valeur ; à partir de ce moment, le développement est clairement périodique.□

24.3 Fractions continues

Définition 926 (Fractions continues) *Une* **fraction continue** *est un objet de la forme suivante :*

$$[a_0, a_1, \ldots, a_n, \ldots] = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots}}}$$

Elle est caractérisée par une suite d'entiers qui est finie ou infinie.
On appelle **convergents** *d'une fraction continue la suite de numérateurs p_n et de dénominateurs q_n définis par :*
– $p_0 = a_0$, $q_0 = 1$
– $p_1 = a_0 a_1 + 1$, $q_1 = 1$
$\vdots$
– $\vdots$,$\vdots$
– $p_n = a_n p_{n-1} + p_{n-2}$,$q_n = a_n q_{n-1} + q_{n-2}$

Propriétés :

– A tout nombre réel on peut associer un et un seul développement en fraction continue.
– Tout nombre rationnel peut être représenté par une fraction continue finie (ex $\frac{1}{3} = 0 + \frac{1}{2+1}$).
– Seuls les nombres rationnels peuvent être représentés par une fraction continue finie.
– Un nombre est quadratique (ie solution d'une équation du second degré à coefficients dans $\mathbb{Z}$) si et seulement si son développement en fraction continue est périodique.
– Une fraction continue est liée à ses convergents par les relations $[a_0, \ldots, a_n] = p_n/q_n$ et $[a_0, \ldots, a_n, \ldots] = \lim_n \frac{p_n}{q_n}$. En outre avec

$$[a_0, \ldots, a_n] = \frac{p_n}{q_n},$$

$$|[a_0, \ldots, a_n] - [a_0, \ldots, a_n, \ldots]| < 1/q_n^2.$$

Théorème 927 (Formule d'Euler) *Supposons les a_i tous non nuls.*
Alors $1/a_1 - 1/a_2 + 1/a_3 - 1/a_4 + ... + (-1)^n/a_n =$

$$\cfrac{1}{a_1 + \cfrac{a_1^2}{a_2 - a_1 + \cfrac{a_1^2}{a_3 - a_2 + \cfrac{a_3^2}{\ddots + \cfrac{\ddots}{a_{n-1} - a_{n-2} + \frac{a_{n-1}^2}{a_n - a_{n-1}}}}}}}$$

<u>Démonstration :</u>

• Pour $n = 1$, le résultat est clair.
• Au rang 2, un calcul rapide montre que le résultat est encore valable.
• On procède ensuite par récurrence, en supposant l'égalité vraie pour $n - 1$ et les rangs inférieurs.
• Dans l'égalité pour $n - 1$, on remplace a_n par $\frac{a_n.a_{n+1}}{a_{n+1}-a_n}$
• Le résultat en découle tout seul...$\square$

24.4 Cryptographie à clé révélée : RSA

Précisons que l'on parle aussi de clé publique .

L'objectif de la **cryptographie** est de permettre de communiquer par des messages codés, qui ne pourront être lus que par leur destinataire.

Pour cela, un "superviseur" donne à chaque receveur potentiel un "décodeur" et une "clé". La clé, comme son nom ne l'indique peut-être pas, est quelque chose qui peut être diffusé à tout le monde.

Pour envoyer un message M crypté à un individu I, il suffit de passer le message M par la moulinette de la clé correspondante à I. Cela n'est pas difficile, puisque I diffuse abondamment sa clé, à tous ses correspondants éventuels. Lorsque I reçoit un message, il peut alors utiliser son décodeur, qu'il est seul à posséder, pour transformer le message crypté en le message original.

La difficulté est que, formellement, il est toujours possible de reconstruire le message initial à partir du message crypté, pourvu que vous ayez la clé. Pour cela, il suffit de tester tous les messages possibles, l'un après l'autre (ils sont bien en bijection avec $\mathbb{N}$, comme on peut s'en convaincre facilement en considérant l'ordre lexicographique sur les messages possibles), et de les passer par la moulinette de la clé jusqu'à ce que l'on retrouve le message crypté. Mais il reste un espoir de fabriquer une cryptographie efficace, car bien sûr, cette méthode prendrait un temps énorme. La cryptographie est ainsi basée sur l'hypothèse de base que certaines tâches, faciles à faire dans un sens (le sens du cryptage par une clé), est difficile à faire dans l'autre (décryptage à l'aide d'une clé).

On note bien que la difficulté réside dans le fait que la fonction "clé" est publique. Si on cache la clé, il est très facile de réaliser des cryptographies parfaites. Par exemple, on peut utiliser le protocole suivant pour que A envoie un message à B :

- A signale à B qu'il veut lui envoyer un message, que l'on supposera constitué uniquement de 0 et de 1 (par un codage quelconque on peut facilement se ramener à cela), et de longueur 1000.

- B fournit à A une liste L de 1000 chiffres 0 ou 1, 0 et 1 étant équiprobables.

- le protocole recommence à l'étape précédente jusqu'à ce que la liste de chiffres soit passée sans être interceptée ; on tire au sort une nouvelle liste de 1000 chiffres à chaque nouvel essai.

- A transforme le message M en un message M', par $M' = M + L$ dans $\mathbb{Z}/2\mathbb{Z}$.

- A envoit M' à B ; si M' est intercepté, il ne sera pas décodable, puisque L n'est pas connu.

- B décode M' par $M = M' + L$ dans $\mathbb{Z}/2\mathbb{Z}$.

Aucune interception ne permet de décoder le message ; mais les étapes 2 et 3 peuvent prendre du temps ou n'être pas réalisables. Il est indispensable de changer de liste L à chaque nouveau message, ou du moins régulièrement - sinon, en considérant les fréquences des 0 et des 1, un observateur des différents M' pourrait finir par reconstituer L.

L'algorithme **RSA**, du nom de ses inventeurs, Rivest, Shamir et Adleman, est basé sur la difficulté de la factorisation d'un nombre entier en nombres premiers.

Supposons que A souhaite envoyer des messages cryptés RSA à B. Alors B se donne deux grands nombres premiers p et q. Maple permet aisément de construire de tels nombres, par exemple ; il suffit par exemple de tirer des nombres au sort et de recommencer jusqu'à ce qu'ils soient premiers, grâce à un algorithme permettant de dire si oui ou non un nombre est premier. En fait les algorithmes utilisés pour cela sont généralement probabilistes, c'est-à-dire qu'ils ont une probabilité non nulle de se tromper ; mais les erreurs sont extrêmement rares. La fonction Maple "isprime" permet de tester la primalité d'un nombre ; par exemple, "isprime(1234567890123456789012345 67)" renvoit "false", donc "1234567890123456789012345 67" n'est pas premier.
"nextprime(1234567890123456789012345 67)" renvoit
"1234567890123456789012346 51", qui est donc le nombre premier le plus petit plus grand que celui-ci. On constate donc que Maple permet très rapidement de trouver de grands nombres premiers ; les exemples ici fournis ne sont pas du tout à la limite du faisable, on peut aller largement au delà.

Ces deux nombres premiers seront notés p et q. La première partie de la clé publique, notée c, sera le produit de p et q. A choisit alors un nombre d, premier avec $\phi(c) = (p-1)(q-1)$, avec ϕ la fonction d'Euler, c'est-à-dire le nombre de nombres premiers plus petits que c, donc $(p-1)(q-1)$. Il est facile de choisir un nombre qui soit premier avec un autre : il suffit d'en piocher un au hasard, et de recommencer jusqu'à ce que l'algorithme de Bezout confirme que ces nombres sont premiers entre eux. On peut aussi déterminer facilement d^{-1}, inverse de d dans $(\mathbb{Z}/c\mathbb{Z})^* \simeq \mathbb{Z}/\phi(c)\mathbb{Z}$ (il y a isomorphisme car $(\mathbb{Z}/c\mathbb{Z})^* \simeq (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/q\mathbb{Z})^*$ par le corollaire 898 et isomorphisme entre $(\mathbb{Z}/p\mathbb{Z})^*$ et $\mathbb{Z}/(p-1)\mathbb{Z}$ (resp. $(\mathbb{Z}/q\mathbb{Z})^*$ et $\mathbb{Z}/(q-1)\mathbb{Z}$) par le lemme 901) : il suffit d'utiliser l'algorithme de Bezout.

Cryptage :

- On suppose le message suffisamment court pour être codable par un élément inversible de $\mathbb{Z}/c\mathbb{Z}$, ce qui est possible en remplaçant le message par des tranches successives suffisamment petites (si on a un alphabet de taille α, il suffit de prendre des tranches de longueur l avec $\alpha^l < \phi(c)$). Cette méthode de codage n'a pas à être compliquée ni à être cachée. Il suffit donc d'avoir une injection de $[1, \alpha^l]$ dans $(\mathbb{Z}/c\mathbb{Z})^*$.

- chaque message de A est donc remplacé (par A) par un élément n inversible dans $\mathbb{Z}/c\mathbb{Z}$.

- A envoie alors à B le nombre $e = n^d$ dans $\mathbb{Z}/c\mathbb{Z}$.

Décryptage :

- B, qui dispose de d et de d^{-1}, effectue simplement le calcul de $e^{d^{-1}}$, qui lui donne n, et donc le message initial.

D'autres systèmes de cryptographie à clé publique font intervenir des structures plus complexes que $\mathbb{Z}/n\mathbb{Z}$, comme par exemple les courbes elliptiques.

Chapitre 25

Polynômes à une indéterminée

25.1 Généralités

Définition 928 *Soit A un anneau commutatif unitaire (resp. $\mathbb{K}$ un corps).*
L'ensemble des suites nulles à partir d'un certain rang d'éléments de A, noté $A^{(\mathbb{N})}$, est un A-module (resp. un $\mathbb{K}$-espace vectoriel) pour l'addition et la multiplication par un scalaire usuelles. En le munissant en outre du produit suivant :

$$\times : (u, v) \mapsto w \text{ avec } w_n = \sum_{i+j=n} u_i.v_j$$

On obtient une A-algèbre (resp. $\mathbb{K}$-algèbre), notée $A[X]$ (resp. $\mathbb{K}[X]$).
Les éléments de $\mathbb{K}[X]$ sont appelés **polynômes***.*
Deux polynômes P et Q non nuls sont dits **associés** *s'il existe λ inversible tel que $P = \lambda.Q$.*
On identifie A $\mathbb{K}$ et l'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ avec $u_n = 0$ pour tout $n > 0$, par l'isomorphisme canonique $x \mapsto (u_n)_{n\in\mathbb{N}}$ avec $u_0 = x$ et $n > 0 \rightarrow u_n = 0$.
On note X l'élément $(u_n)_{n\in\mathbb{N}}$ avec $u_0 = 0$, $u_1 = 1$, et $u_n = 0$ pour $n > 1$.
La famille des X^i pour $i \in \mathbb{N}$ constitue la base canonique du module libre $A^(\mathbb{N})$ (resp. du $\mathbb{K}$-espace vectoriel $K^{(N)}$.
Etant donné P un polynôme, on appelle **degré de** *P et on note $\deg P$ le plus grand n tel que P_n est non nul. On appelle* **coefficient dominant de** *P le coefficient de $X^{\deg P}$ (que l'on peut voir comme $X^{deg(P)^*}(P)$ si l'on travaille avec un corps, voir la partie 27.7) ; on le note $coef(P)$.*
Un polynôme non nul est dit **unitaire** *si son coefficient dominant est 1.*
On appelle **support d'un polynôme** *P l'ensemble des $n \in \mathbb{N}$ tels que $X^{n*}(P) \neq 0$.*
Par définition d'un polynôme, son support est fini.
Le degré d'un polynôme P est donc aussi le sup de son support.
On appelle **valuation de** *P et on note $val(P)$ l'inf du support de P.*
On appelle **composé de deux polynômes** *P* **et** *Q et on note $P \circ Q$ le polynôme $\sum P_n Q^n$ (que l'on peut aussi voir comme $\sum_{n\in\mathbb{N}} X^{i^*}(P).Q^i$ si l'on travaille avec un corps).*

Proposition 929 • 1 *est élément neutre pour la multiplication,* X *élément neutre pour* $\circ$, 0 *élément neutre pour l'addition.*
• *Les éléments inversibles de* $\mathbb{K}[X]$ *sont les éléments identifiés aux éléments de* $\mathbb{K} \setminus \{0\}$.
• $deg(0) = -\infty$ *et* $val(0) = +\infty$
• $deg(1) = 0$
• $deg(X^i) = i$ *et* $val(X^i) = i$
• $deg(P.Q) = deg(P) + deg(Q)$ *et* $val(P.Q) = val(P) + val(Q)$
• $deg(P + Q) \leq sup(deg(P), deg(Q))$ *avec égalité si* $deg(P) \neq deg(Q)$ *ou si les coefficients dominants de* P *et* Q *ne sont pas opposés.*
• $val(P + Q) \leq sup(val(P), val(Q))$ *avec égalité si* $val(P) \neq val(Q)$ *ou si* $P_{val(P)}$ *et* $Q_{val(Q)}$ *ne sont pas opposés.*
• *si* A *est intègre alors* $A[X]$ *est un anneau intègre ; c'est à dire que le produit de deux polynômes est nul si et seulement si l'un des deux polynômes est nul.*
• $(P+Q)\circ R = P\circ R+Q\circ R$ *mais en général* $P\circ(Q+R) \neq P\circ Q+P\circ R$

25.2 Division euclidienne

Théorème 930 (Division euclidienne) *Soient* A *et* B *deux polynômes, avec* $coef(B)$ *inversible. Alors*

$$\exists (Q, R) \text{ polynômes } / A = B.Q + R$$

avec

$$deg\ R < deg\ B$$

<u>Démonstration :</u> • Unicité :
Supposons $B.Q + R = B.Q' + R'$ avec les conditions données sur le degré. Alors

$$B.(Q - Q') = R' - R$$

$$deg(B) + deg(Q - Q') = deg(R' - R)$$

donc

$$deg(Q - Q') = -\infty$$

et $Q = Q'$ et $R = R'$.
• Existence :
On distingue deux cas.
- Si le degré de A est inférieur strictement au degré de B, le résultat est clair avec $Q = 0$ et $R = A$.
- Sinon on procède par récurrence sur le degré de A, en considérant

$$A - \frac{coef(A)}{coef(B)}.X^{deg(A)-deg(B)}$$

...□

Il est à remarquer que c'est par la méthode de la récurrence que l'on pratique la division euclidienne.

Définition 931 *Q est appelé* **quotient** *de A par B, et R est appelé* **reste** *de A par B.*

Un exemple en Maple :

Exemple Maple

$> \ rem(x^4 + x^3 + x^2 + x + 1, x^3, x);$

$$1 + x^2 + x$$

$> \ quo(x^4 + x^3 + x^2 + x + 1, x^3, x);$

$$x + 1$$

Corollaire 932 *Soit $\mathbb{K}$ un corps. $\mathbb{K}[X]$ est un anneau euclidien, donc un anneau principal.*

Démonstration : La division euclidienne en est la preuve ; dans un corps, tout polynôme non nul a son coefficient dominant inversible.□

25.3 Fonction associée, racines d'un polynôme

Définition 933 *Etant donnée B une A-algèbre associative commutative et unitaire, on peut identifier P à une application de A dans A, dite* **application polynômiale associée à** *A, noté $\tilde{P}$, et définie par*

$$\tilde{P}(x) = \sum_{n \in \mathbb{N}} P_n \, x^n$$

Cela est notamment valable pour $B = A$; implicitement $\tilde{P}$ désignera généralement une fonction de A dans A.

Proposition 934 *Soit a dans A et P un polynôme appartenant à $A[X]$.*
Le reste de la division euclidienne de P par $(X - a)$ est $P(a)$.

En considérant la division euclidienne de P par $(x - a)^n$, on peut énoncer la définition suivante :

Proposition 935 - Définition
Soit P un polynôme, a un élément de A et n un entier naturel.
Les conditions suivantes sont équivalentes :
- $(X-a)^n|P$ et $(X-a)^{n+1} \not| P$
- Il existe un polynôme Q tel que $P=(X-a)^nQ$ et $\tilde{Q}(a)$ non nul.
On dit alors que a est **racine** *de P d'***ordre de multiplicité** *n.*

25.4 Cas où $A = \mathbb{K}$ est un corps

Théorème 936 *Si $\mathbb{K}$ est un corps commutatif alors $\mathbb{K}[X]$ est un anneau euclidien, donc un anneau principal, donc un anneau factoriel.*

Démonstration : Le fait que $\mathbb{K}[X]$ est un anneau euclidien a été prouvé en proposition 874.$\square$

Proposition 937 - Définition *On dit qu'un corps $\mathbb{K}$ est* **algébriquement clos** *si et seulement si l'une des trois propositions suivantes est vérifiée :*
- tout polynôme de $\mathbb{K}[X]$ est **scindé**, *c'est-à-dire produit de polynômes de degré* 1
- tout polynôme non constant a une racine dans $\mathbb{K}$
- tout polynôme irréductible est de degré 1

Dans $\mathbb{K}[X]$ avec $\mathbb{K}$ corps algébriquement clos, tout polynôme de degré n s'écrit de manière unique à l'ordre près des facteurs sous la forme :

$$c(X-k_1)(X-k_2)\dots(X-k_n)$$

avec c inversible, et les k_i les racines (pas forcément distinctes !) du polynôme.

$\mathbb{C}$ est algébriquement clos, comme énoncé dans la proposition 503.

Un exemple en Maple :

Exemple Maple
$> P := X- > x\hat{}4 - 1;$
$$P := X^4 - 1$$
$> factor(P);$
$$(x-1)(x+1)(x^2+1)$$
$> factor(P, complex);$
$$(x+1.)(x+1.I)(x-1.I)(x-1.000000000)$$

25.5 Zoologie des polynômes

En dehors des paragraphes ci-dessous, on pourra consulter la partie30.8.1 sur l'interpolation par les polynômes de Lagrange, d'ailleurs généralisable à un cadre plus vaste que les polynômes réels, et le théorème 417, d'approximation par les polynômes de Bernstein.

25.5.1 Relations entre les racines et les coefficients d'un polynôme - localisation des racines d'un polynôme

On se donne pour l'ensemble de cette partie un polynôme $P \in \mathbb{C}[X]$, de degré n, P non nul. On définit $P = \sum_{i=0}^{n} p_i X^i$.

⊡ Relations entre les racines et les coefficients d'un polynôme

On utilisera ici les polynômes symétriques élémentaires $\Sigma_i = \Sigma_{i,n}$ définis en partie 26.3.1.

> **Théorème 938 (Relations entre racines et coefficients d'un polynôme)**
> *Notons $\sigma_i = \Sigma_i(r_1, ..., r_n)$, avec r_1,...,r_n des complexes.*
> *On a $P = \lambda \Pi_{i=1}^{n}(X - r_i)$ pour un certain λ si et seulement si $\sigma_i = (-1)^i \frac{p_{n-i}}{p_n}$.*

Démonstration : On écrit simplement l'égalité

$$\sum_{i=0}^{n} p_i X^i = \lambda \Pi_{i=1}^{n}(x - r_i)$$

On en déduit que $\lambda = p_n$, et les relations souhaitées en développant d'un côté et de l'autre du signe $=$.□

⊡ Localisation des racines d'un polynôme

◇ **Premières informations** Le théorème de Rolle 575 permet de montrer que le polynôme dérivé d'un polynôme scindé est scindé.

⋄ **Méthode itérative** On peut par exemple utiliser la méthode de Newton, trouvable dans tout bon ouvrage d'analyse numérique. On pourra par exemple consulter [9]. FLEMMARD verifier

⋄ **Méthode algébrique** La théorie du résultant donne quelques résultats intéressants sur la localisation de racines ; voir théorème 942.

25.5.2 Polynômes irréductibles

Définition 939 *Un polynôme P appartenant à $\mathbb{K}[X]$ est dit* **polynôme irréductible** *si il est irréductible en tant qu'élément de l'anneau $\mathbb{K}[X]$, c'est-à-dire s'il n'est pas inversible et si tout diviseur de P est une unité ou est produit de P par une unité.*

Pour plus d'informations sur la recherche de facteurs irréductibles communs à deux polynômes, on consultera le théorème 942.

Théorème 940 • *Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré* 1.
• *Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré* 1 *et les polynômes $aX^2 + bX + c$, avec $a \neq 0$ et $b^2 - 4ac < 0$.*

Démonstration : Il est évident que les polynômes considérés sont bel et bien irréductibles, dans les deux cas réel et complexe. Réciproquement, le théorème de D'Alembert-Gauss 503 donne le résultat dans le cas complexe. Dans le cas réel, on procède comme suit :

• Supposons $P \in \mathbb{R}[X]$ irréductible dans $\mathbb{R}[X]$.

• Par simplicité et sans perte de généralité, on va supposer P unitaire.

• Si x est racine de P dans $\mathbb{C}$, alors $\overline{x}$ l'est aussi, avec le même ordre de multiplicité.

• P n'a pas de racine réelle r, sinon $X - r$ diviserait P et P ne serait pas irréductible.

• P peut donc s'écrire $P = \Pi_{i=1}^{r}(X - r_i)^{n_i}(X - \overline{r}_i)^{n_i}$.

• $(X - r_i)(X - \overline{r}_i)$ est alors un polynôme réel (on le voit simplement en le développant), donc P est un produit de polynômes de discriminants négatifs strictement (là aussi il suffit de développer pour le voir). Il n'est irréductible que s'il contient un et un seul tel terme.

D'où le résultat.□

25.5.3 Résultant. Discriminant

Définition 941 *Etant donnés P et Q deux polynômes de $\mathbb{K}[X]$, on appelle* **résultant de** P **et** Q *le déterminant de la matrice suivante :*

$$\begin{pmatrix}
P_0 & 0 & 0 & 0 & \dots & 0 & Q_0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\
P_1 & P_0 & 0 & 0 & \dots & 0 & Q_1 & Q_0 & 0 & 0 & \dots & 0 & 0 & 0 \\
P_2 & \dots & P_0 & 0 & \dots & 0 & Q_2 & \dots & Q_0 & 0\dots & 0 & 0 & 0 & \\
P_3 & \dots & \ddots & \ddots & \ddots & 0 & Q_3 & \dots & \ddots & \ddots & \ddots & \ddots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & Q_q & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & Q_q & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & 0 & Q_q & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & 0 & Q_q & \vdots & \vdots & \vdots & \vdots \\
P_p & P_{p-1} & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & 0 & Q_q & \vdots & \vdots & \vdots & \vdots \\
0 & P_p & \ddots & \ddots & \ddots & \ddots & 0 & \ddots & 0 & Q_q & \ddots & \ddots & \ddots & \ddots \\
0 & 0 & P_p & \ddots & \ddots & \ddots & 0 & \ddots & 0 & Q_q & \ddots & \ddots & \ddots & \ddots \\
0 & \dots & 0 & P_p & \ddots & \ddots & 0 & \ddots & 0 & Q_q & \ddots & \ddots & \ddots & \ddots \\
0 & \dots & \dots & \dots & \dots & P_p & 0 & \dots & \dots & \dots & \dots & \dots & 0 & Q_q
\end{pmatrix}$$

avec $P = \sum_{k=0}^{p} P_p X^k$ et $Q = \sum_{k=0}^{q} Q_q X^k$.
On appelle **discriminant** *d'un polynôme P le résultant de P et de P' son polynôme dérivé.*

Théorème 942 *Le résultant de P et Q est nul si et seulement si P et Q ont au moins un facteur irréductible en commun.*
Le discriminant d'un polynôme P est nul si et seulement si il a au moins un facteur irréductible en commun avec son polynôme dérivé.

Démonstration :

La deuxième affirmation n'est naturellement qu'une spécialisation de la première. On se contentera donc de prouver la première.

• Supposons tout d'abord que P et Q aient un facteur irréductible commun R.

- Alors $P = RS$ et $Q = RT$, avec $deg\ T = deg\ Q - deg\ R$ et $deg\ S = deg\ P - deg\ R$, T et S non nuls.

- $PT = QS$, ou $PT - QS = 0$. Ceci exprime très exactement l'existence d'un vecteur X tel que la matrice M donnée dans l'énoncé vérifie $MX = 0$, avec X non nul ; donc la matrice n'est pas la matrice d'une bijection, donc son déterminant est nul.

• Supposons maintenant qu'il existe X tel que MX soit nul et X soit $\neq 0$. Alors, il existe P et Q vérifiant $PT = QS$, avec $deg\ T = deg\ Q - deg\ R$ et $deg\ S = deg\ P - deg\ R$.

- Supposons alors que P et Q n'aient pas de facteur irréductible en commun. Alors, par le lemme de Gauss 890, P divise S, ce qui est impossible car $deg\ S < deg\ P$.□

En particulier, si les polynômes sont scindés, ils ont une racine commune si et seulement si leur résultant est nul. Si P est scindé, son discriminant est nul si et seulement si il admet une racine double.

25.5.4 Division suivant les puissances croissantes

Théorème 943 (Division suivant les puissances croissantes) *Soit $n \in \mathbb{N}$, C et D des polynômes à une indéterminée sur un même anneau A commutatif et unitaire.*
On suppose que $D(0)$ (en tant qu'élément de A), est inversible.
Alors il existe deux polynômes Q et R vérifiant

$$C = DQ + X^{n+1}R$$

$$\deg Q \leq n$$

Q et R sont appelés respectivement **quotient et reste de la division suivant les puissances croissantes de C par D à l'ordre n.**

<u>**Démonstration :**</u> La preuve se fait par récurrence inverse sur la valuation de C. Si cette valuation est supérieure ou égale à $n+1$, le résultat est clair. La suite est facile...□

Cela servira notamment pour les développements limités de quotients (voir proposition 628, ou pour la décomposition de fractions rationnelles en éléments simples (voir proposition ....).

Voyons un exemple concret, la division suivant les puissances croissantes de $X^3 + 2X^2 + 2$ par $X + 2$:

$$\begin{array}{cr|l} & X^3+2X^2+2 & X+2 \\ \cline{3-3} - & X+2 & 1 \\ \cline{2-2} & X^3+2X^2-X & \end{array}$$

Donc $X^3 + 2X^2 + 2 = (X + 2)(1) + (X^2 + 2X - 1)X$, c'est la division suivant les puissances croissantes à l'ordre 0. Continuons :

$$\begin{array}{cr|l} & X^3+2X^2+2 & X+2 \\ \cline{3-3} - & X+2 & 1-\frac{X}{2} \\ \cline{2-2} & X^3+2X^2-X & \\ - & -\frac{X^2}{2}-X & \\ \cline{2-2} & X^3+\frac{5}{2}X^2 & \end{array}$$

Donc $X^3 + 2X^2 + 2 = (X + 2)(1 - \frac{X}{2}) + (X + \frac{5}{2})X^2$, c'est la division suivant les puissances croissantes à l'ordre 1. Continuons encore :

$$\begin{array}{cr|l} & X^3+2X^2+2 & X+2 \\ \cline{3-3} - & X+2 & 1-\frac{X}{2}+\frac{5}{4}X^2 \\ \cline{2-2} & X^3+2X^2-X & \\ - & -\frac{X^2}{2}-X & \\ \cline{2-2} & X^3+\frac{5}{2}X^2 & \\ - & \frac{5}{4}X^3+\frac{5}{2}X^2 & \\ \cline{2-2} & -\frac{1}{4}X^3 & \end{array}$$

Donc, division suivant les puissances croissantes à l'ordre 2, $X^3 + 2X^2 + 2 = (X + 2)(1 - \frac{X}{2} + \frac{5}{4}X^2) - \frac{1}{4}X^3$.

25.5.5 Polynômes orthogonaux

Définition 944 *On suppose donné $(a, b) \in \overline{\mathbb{R}}^2$, $a < b$. On suppose donnée une fonction w de $]a, b[$ dans $\mathbb{R}_+^*$ continue. Enfin on suppose que pour tout n, $\int_a^b x^n w(x)dx$ est convergente [a]. On note alors E l'ensemble des fonctions de $]a, b[$ dans $\mathbb{R}$ telles que*

$$\|f\|_2 := \sqrt{\int_a^b |f(x)|^2 w(x)dx} < \infty$$

L'ensemble des polynômes est inclus dans E, E muni du produit scalaire suivant :

$$< f, g >= \int_a^b f(x)g(x)w(x)dx$$

est un espace de Hilbert.
Il existe alors une suite de polynômes $(P_n)_{n\in\mathbb{N}}$, telle que $degP_n = n$, et telle que les P_n forment une famille orthogonale.

[a] Un cas classique est $(a, b) \in \mathbb{R}^2$ et $w = 1$, ou bien a et b quelconques et $w(x) = e^{-x^2}$.

Démonstration : Le résultat découle simplement de l'orthogonalisation de Schmidt (proposition 1115) appliquée à $1, X, X^2, X^3, ...$ Le fait que le degré de P_n est $\leq n$ provient simplement des propriétés de l'orthogonalisation de Schmidt, ie le fait que P_n appartient à l'espace engendré par $1, X, ..., X^n$. Le fait que ce degré est $\geq n$ provient simplement du fait que s'il existait un P_n de degré $< n$, alors la famille $P_0, ..., P_n$ serait une famille libre (puisqu'orthogonale) et située dans un espace de dimension n ; ce qui contredit le lemme de Steinitz 1143.

Les polynômes orthogonaux ont de multiples applications, que l'on pourra trouver par exemple dans le livre [9].

On pourra consulter l'exemple Maple qui se trouve suite à l'orthogonalisation de Schmidt (voir proposition 1115).

25.5.6 Polynômes de Tchebycheff de première espèce

Théorème 945

$$\forall n \in \mathbb{N} \exists T_n \in \mathbb{R}[X] / \deg\, T_n \leq n \wedge \forall t \cos(nt) = T_n(cos(t))$$

T_n est appelé **polynôme de Tchebycheff de permière espèce**.

Démonstration :
• Soit $n \in N$. On a

$$e^{inx} = \cos(nx) + i\sin(nx) = (\cos(x) + i\sin(x))^n = \sum_{k=0}^{n} C_n^k \cos(x)^{n_k} \sin(x)^k (i)^k$$

• En prenant les parties réelles :

$$cos(nx) = \sum_{k=0}^{k \leq n/2} C_n^{2k} cos(x)^{n-2k} (-1)^k (1 - cos(x)^2)^k$$

D'où le résultat.□

Proposition 946

$$T_n = 2^{n-1} \pi_{k=0}^{n-1} (X - \cos(\frac{2k+1}{2n}\pi))$$

Démonstration : Il suffit de vérifier que le coefficient dominant est le bon, que le degré est le bon, et que les $\cos(\frac{2k+1}{2n}\pi)$ sont bien des racines.□

25.5.7 Tout polynôme positif est somme de deux carrés

Théorème 947 *Soit P un polynôme appartenant à $\mathbb{R}[X]$.*
On suppose en outre que P est positif sur $\mathbb{R}$.
Alors P est somme de deux carrés.

Démonstration : • Soit $Cool$ l'ensemble des polynômes qui s'expriment comme somme de deux carrés.

• Alors $Cool$ contient tous les polynômes de la forme $(x - a)^n$, pour n pair.

• Tout polynôme irréductible unitaire de degré 2 est dans $Cool$. En effet, $X^2 + b.X + c$ est égal à $(X - \frac{b}{2})^2 + (c - \frac{b^2}{4})$, qui est bien une somme de deux carrés si $c - \frac{b^2}{4}$

• Du coup, tout polynôme irréductible de degré 2 à coefficient dominant > 0 est dans $Cool$.

• Si P et Q sont dans $Cool$, alors PQ est dans $Cool$ ($Cool$ est stable par multiplication). En effet, avec $P = A^2 + B^2$ et $Q = C^2 + D^2$:

$$(A^2 + B^2).(C^2 + D^2) = (AD + BC)^2 + (AC - BD)^2$$

• $Cool$ contient les polynômes positifs dépourvus de racine. En effet, soit P sans racine ; il est produit de polynômes irréductibles. Le coefficient dominant est positif, au vu de l'équivalent en $\pm\infty$, donc on peut l'exprimer comme produit de polynômes irréductibles de degré 2 à coefficients dominants positifs.

• Soit P un polynôme positif. On a déjà vu que s'il n'admet pas de racine il est dans $Cool$. On suppose maintenant qu'il admet des racines, par exemple une racine a. Soit n maximal tel que $(X - a)^n$ divise P. $P = (X - a)^n Q$ est équivalent en a à $Q(a)(X - a)^n$; donc n doit être pair pour que le signe de P puisse être positif. En supposant par récurrence que pour les degrés inférieurs à celui de P le résultat est acquis, on conclut que Q et $X - a$ sont dans $Cool$, et donc que P est dans $Cool$.□

Chapitre 26

Polynômes à plusieurs indéterminées

Cette partie sera délibérément très peu détaillée ; beaucoup de démonstrations sont calquées sur le cas des polynômes à une indéterminée. Ceux qui préfèrent un cadre simple et utile peuvent se limiter à la partie25, ou l'on travaillera avec des polynômes à une seule indéterminée, et fournissant les méthodes permettant de s'attaquer à cette partie plus abstraite.

Cette partie doit être complétée par la partie25, qui par contre peut être lue indépendamment de celle-ci.

26.1 Généralités

Définition 948 *Soit A un anneau commutatif unitaire.*
On appelle **polynôme à n indéterminées à coefficients dans** *A l'ensemble des applications presques nulles de $\mathbb{N}^n$ dans A. On note $A[X_1, \ldots, X_n]$ l'ensemble des polynômes à n indéterminées à coefficients dans A.*
On dit que $P \in \mathbb{K}[X_1, \ldots, X_n]$ est **de degré** *d si d est le max des $|\nu|$ tels que P_ν est non nul (voir définition 727 pour les rappels sur les opérations dans $\mathbb{N}^n$).*
On dit que P est de degré d en X_i si le $\sup$ des i tels qu'il existe ν tel que ν_i est non nul est d.
On note X_i l'élément de $A[X_1, \ldots, X_n]$ nul partout sauf en $\nu = (\delta_{i,j})_{j\in[1,n]}$, avec $X_\nu = 1$.
Etant donnés P et Q deux polynômes, on note $R = P \times Q$ le **produit de** *P* **et** *Q avec*

$$R_\nu = \sum_{\alpha+\beta=\nu} P_\alpha Q_\beta$$

(pour les opérations dans N^n, voir définition 727).
On appelle **monôme** *un polynôme dont un seul élément est non nul.*
On appelle **dérivé formel** *d'un polynôme P par D^ν pour $\nu \in \mathbb{N}^n$ le polynôme*

$$\sum_{\alpha\in\mathbb{N}^n} \frac{(\nu+\alpha)!}{\alpha!} P_{\alpha+\nu}$$

On note parfois $\frac{\delta}{\delta X_i}$ D^ν avec $\nu_j = (\delta_{i,j})_{j\in[1,n]}$.

Proposition 949 *On identifie $A[X_1, \ldots, X_n]$ à $A[X_1, \ldots, X_{n-1}][X_n]$, ainsi que $A[X_1, \ldots, X_p][X_{p+1}, \ldots, X_n]$ à $A[X_1, \ldots, X_n]$.*
$A[X_1, \ldots, X_n]$ est intègre si et seulement si A est intègre.
$A[X_1, \ldots, X_n]$ est muni naturellement d'une structure de A-module. Muni de la multiplication définie plus haut, il s'agit d'une A-algèbre.
L'ensemble des monômes unitaires est une base de $A[X_1, \ldots, X_n]$.
Etant donnée B une A-algèbre associative commutative unitaire, $P \in A[X_1, \ldots, X_n]$ et $x_1, ..., x_n$ n éléments de B, on appelle **valeur de** *P* **en** *$(x_1, \ldots, x_n)$ l'élément de B $\sum_{\nu\in\mathbb{N}^n} P_\nu x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$. On note cet élément $\tilde{P}(x_1, \ldots, x_n)$. On constate ainsi qu'un polynôme P s'identifie naturellement à une application $\tilde{P}$ de B^n dans B. On note $A[x_1, \ldots, x_n]$ l'ensemble des $P(x_1, \ldots, x_n)$ pour $P \in A[x_1, \ldots, x_n]$.*
Si $(x_1, \ldots, x_n)$ vérifie $\tilde{P}(x_1, \ldots, x_n) = 0$, on dit que $(x_1, \ldots, x_n)$ est un zéro de P.
Etant donné $(x_1, \ldots, x_n)$ n éléments de B, l'ensemble des polynômes P vérifiant $P(x_1, \ldots, x_n) = 0$ est un idéal de $A[X_1, \ldots, X_n]$, engendré par les $(X_i - a_i)$ pour $i \in [1, n]$.

26.2 Si A est un corps $\mathbb{K}$

Proposition 950 *Si $\mathbb{K}$ est un corps, $\mathbb{K}[X_1, \ldots, X_n]$ est naturellement muni d'une structure de $\mathbb{K}$-espace vectoriel .*
Formule de Taylor, si $\mathbb{K}$ est un corps de caractéristique nulle : soit $P \in \mathbb{K}[X]$, alors

$$P = \sum_{\nu \in \mathbb{N}^n} \frac{1}{\nu!}(D^\nu P)(0)X^\nu$$

26.3 Zoologie des polynômes à plusieurs indéterminées

26.3.1 Polynômes symétriques

A est supposé ici anneau commutatif et unitaire.

Définition 951 *Soit $P \in A[X_1, \ldots, X_n]$. P est dit* **polynôme symétrique** *si et seulement si pour tout σ permutation de $[1, ..., n]$, $P(X_1, ..., X_n) = P(X_{\sigma(1)}, X_{\sigma(2)}, ..., X_{\sigma(n)})$.*
On appelle **polynômes symétriques élémentaires** *les polynômes de la forme*

$$\Sigma_{k,n} = \sum_{1 \leq a_1 \leq a_2 \leq \ldots \leq a_k \leq n} X_{a_1} X_{a_2} ... X_{a_k}$$

pour $1 \leq k \leq n$.
On appelle k-**ième polynôme de Newton** *le polynôme $N_k = \sum_{i=1}^n X_i^k$.*

Les polynômes symétriques élémentaires sont de la forme suivante, dans le cas $n = 3$:

$$\Sigma_{1,3} = X_1 + X_2 + X_3$$

$$\Sigma_{2,3} = X_1X_2 + X_2X_3 + X_1X_3$$

$$\Sigma_{3,3} = X_1X_2X_3$$

Je ne donnerai pas ici de preuve des résultats énoncés ; on pourra se référer à [19]. On a les propriétés suivantes :

• Les polynômes symétriques élémentaires sont symétriques (évident)

• Les polynômes de Newton sont symétriques (évident)

• Si Q est un polynôme, alors $P = Q(\Sigma_{1,n}, \Sigma_{2,n}, ..., \Sigma_{n,n})$ est un polynôme symétrique (facile)

• Si $P \in A[X_1, ..., X_n]$ est symétrique, alors il existe un polynôme Q tel que $P = Q(\Sigma_{1,n}, \Sigma_{2,n}, ..., \Sigma_{n,n})$ (pas évident du tout, récurrence sur le nombre d'indéterminées et sur le degré du polynôme.

• **Relations de Newton** : Si $1 \leq k \leq n$ on a

$$N_k = \sum_{i=1}^{k-1} (-1)^i N_{k-i} \Sigma_{i,n} + (-1)^k k \Sigma(k, n)$$

Si $n \leq k$ on a

$$N_k = \sum_{i=1}^{n} (-1)^i N_{k-i} \Sigma_{i,n}$$

Chapitre 27

Algèbre linéaire

27.1 Généralités

Définition 952 (Espace vectoriel) $(E, +, .)$ *est un* $\mathbb{K}$**-espace vectoriel** *si*
• $(E, +)$ *est un groupe abélien*
• *. est une application de* $K \times E$ *dans* E
• $\forall(\lambda, \mu, x, y)\ (\lambda + \mu)x = \lambda.x + \mu.x \wedge \lambda.(x + y) = \lambda.x + \lambda.y \wedge (\lambda.\mu).x = \lambda.(\mu.x) \wedge 1.x = x$
Les éléments de E *sont appelés* **vecteurs**, *les éléments de* $\mathbb{K}$ *sont appelés* **opérateurs** *ou* **scalaires**. *Le neutre pour* $+$ *est noté* 0. *« . » est appelé* **produit externe**.

Des exemples classiques sont : les vecteurs dans le plan ou l'espace usuel, le corps K lui-même avec pour produit externe le produit usuel, les polynômes sur K, éventuellement à plusieurs indéterminées ou de degré borné ; comme on le verra un peu plus loin avec les espaces produits, $\mathbb{K}^n$ muni des lois que l'on verra est un $\mathbb{K}$-espace vectoriel ; l'ensemble $\mathbb{R}^{\mathbb{N}}$ des suites réelles (resp. complexes) aussi.

Les propriétés suivantes sont importantes :
• $0.x = 0$
• $\lambda.0 = 0$
• $\lambda.x = 0 \rightarrow \lambda = 0 \vee x = 0$
• $(-\lambda).x = \lambda.(-x) = -(\lambda.x)$
• $\lambda.(x - y) = \lambda.x - \lambda.y$
• $(\lambda - \mu)x = \lambda.x - \mu.x$

Définition 953 (Différentes notions dans un espace vectoriel) *On appelle* **segment** *d'extrémités x et y dans un $\mathbb{R}$-espace vectoriel l'ensemble des $t.x + (1-t).y$ pour $t \in [0,1]$. (la définition s'étend au cas d'un $\mathbb{C}$-espace vectoriel en utilisant t réel dans $[0,1]$).*
Une partie A d'un $\mathbb{K}$-espace vectoriel (avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$) est dite **convexe** *si tout segment d'extrémités dans A est inclus dans A.*
Etant donnée A une partie convexe d'un $\mathbb{K}$-espace vectoriel une application f de A dans $\mathbb{R}$ est dite **convexe** *si étant donnés x et y dans A et t dans $[0,1]$ on a $f(t.x + (1-t).y) \leq t.f(x).(1-t).f(y)$. $-f$ est alors dite* **concave***.*

Propriétés :
• L'intersection de deux convexes est un convexe.
• Une boule ouverte est convexe.
• Une boule fermée est convexe.
• Un segment est convexe.
• Une application f est convexe si l'ensemble des $(x, f(x))$ est convexe dans le produit $A \times \mathbb{R}$.
• Une application convexe sur un intervalle $]a, b[$ est continue
• $x \mapsto e^x$, $x \mapsto x^n$ sont convexes.
• $x \mapsto ln(x)$ est concave.

La notion de partie convexe est nécessaire pour définir les fonctions convexes, dont on verra une foultitude d'applications en partie11.5.1. Cela servira aussi par exemple pour les résultats 661 (théorème de Cauchy), et surtout pour les théorème de Hahn-Banach(20.1.1) (aux multiples applications !). Une utilisation amusante de la convexité stricte sera donnée avec le résultat 292.

Proposition 954 (Connexité dans un $\mathbb{R}$-espace vectoriel normé) *Une partie ouverte d'un $\mathbb{R}$-espace vectoriel normé est connexe si et seulement si elle est connexe par arcs si et seulement si entre tout x et tout y de cette partie il existe une ligne brisée.*

Démonstration : Il est clair que l'existence d'une ligne brisée implique l'existence d'un chemin (donc la connexité par arcs) qui implique à son tour la connexité. Il suffit donc de montrer que la connexité implique l'existence d'une ligne brisée.
• On se donne U une telle partie, connexe, supposée non vide (sinon le problème est trivial)
• On se donne x_0 dans U
• On note V l'ensemble des x que l'on peut joindre à x_0 par une ligne brisée.
• V est non vide (il contient x_0)
• V est ouvert (si x est dans V, alors $B(x, \epsilon) \subset U$ pour ϵ assez petit, et donc $B(x, \epsilon) \subset V$ - car dans toute boule est convexe).
• V est fermé ; en effet soit x dans U adhérent à V, alors il existe une boule ouverte centrée sur x intersectant V, donc x est dans V - car toute boule est convexe.
• V, fermé, ouvert, non vide d'un connexe U, est égal à U.$\square$

La figure 27.1 illustre cette démonstration.

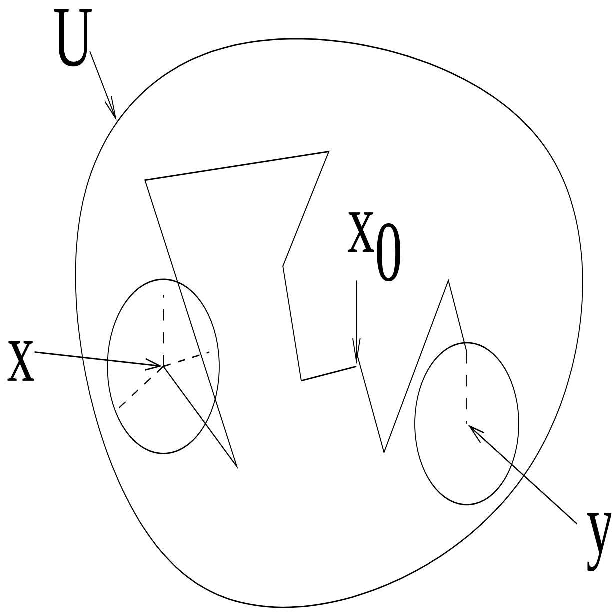


FIG. 27.1 – Illustration de la proposition 954. x est supposé dans V, y dans l'adhérence de V.

Proposition 955 (Distance dans un connexe ouvert d'un $\mathbb{R}$-espace vectoriel normé) *Soit U un ouvert connexe d'un $\mathbb{R}$-espace vectoriel normé . Si on note $d(x,y)$ l'inf des longueurs des lignes brisées joignant x à y, alors d est une distance et définit la même topologie que la norme.*

Démonstration : pas difficile du tout !□

Définition 956 (Espace vectoriel produit) *Le produit de n espaces vectoriels, muni de l'addition terme à terme et avec pour multiplication $\lambda.(x_1,...,x_n) = (\lambda.x_1,...,\lambda.x_n)$ est un espace vectoriel. On l'appelle* **espace vectoriel produit**.

Théorème 957 *L'ensemble des fonctions de A dans E, noté E^A, avec E $\mathbb{K}$-espace vectoriel, est un $\mathbb{K}$-espace vectoriel. La somme de deux fonctions est la fonction qui à un élément associe la somme des deux images, et le produit d'un scalaire par une fonction est la fonction qui à un vecteur associe le produit du scalaire par l'image de ce vecteur.*

Démonstration : La preuve est pas bien difficile...□

Définition 958 (Sous-espaces vectoriels) *Une partie d'un espace vectoriel est un* **sous-espace vectoriel** *de cet espace lorsqu'elle est non vide, stable par addition et stable par multiplication par un scalaire.*

On note qu'on peut se contenter de vérifier que la partie est non vide et que si λ etμ sont des scalaires et si x et y lui appartiennent, alors $\lambda.x + \mu.y$ appartient aussi à l'ensemble.

Proposition 959 *Un sous-espace vectoriel est un espace vectoriel, et un espace vectoriel inclus dans un espace vectoriel (et muni des lois induites bien sûr) est un sous-espace vectoriel.*

Les polynômes de degré plus petit que n sont un sous-espace de l'ensemble des polynômes. L'espace des fonctions bornées sont un sous-espace vectoriel de l'espace des fonctions de A dans $\mathbb{R}$. L'ensemble des suites bornées de $\mathbb{K}$ est un sous-espace de l'espace des suites de $\mathbb{K}$. L'ensemble des suites presque nulles de $\mathbb{K}$ est aussi un sous-espace vectoriel de l'espace des suites de $\mathbb{K}$.

Proposition 960 *Une intersection quelconque de sous-espaces vectoriels est un espace vectoriel.*

Cela permet d'introduire la définition suivante :

Définition 961 (Sous-espace vectoriel engendré) *On appelle* **sous-espace vectoriel** *engendré par une partie A l'intersection de tous les sous-espaces vectoriels contenant A. On la note $Vect\ (A)$.*

27.2 Applications linéaires

Définition 962 (Application linéaire) *Une application f d'un $\mathbb{K}$-espace vectoriel vers un autre est une* **application linéaire** *si :*
• $f(x+y) = f(x) + f(y)$
• $f(\lambda.x) = \lambda.f(x)$
Une application linéaire est aussi appelée **morphisme d'espaces vectoriels**, *ou* **morphisme algébrique**. *C'est en particulier un morphisme de groupes. Une application linéaire bijective est appelée* **isomorphisme**, *une application linéaire de E dans E est appelée* **endomorphisme**. *Un endomorphisme qui est aussi un isomorphisme est appelé* **automorphisme**. *L'inverse d'un isomorphisme est un isomorphisme, appelé* **isomorphisme réciproque**.
On note $\mathcal{L}(E,F)$ l'ensemble des morphismes de E dans F ; c'est un sous-espace vectoriel de F^E. On note $\mathcal{L}(E)$ l'ensemble des endomorphismes de E. On note $Isom(E,F)$ l'ensemble des isomorphismes de E dans F. On note $Aut(E)$ l'ensemble des automorphismes de E ; il est noté $GL(E)$ une fois qu'on l'a muni de la composition. $GL(E)$ est un groupe, appelé **groupe linéaire**.
La notation $E \simeq F$ signifie qu'il existe un isomorphisme de E dans F.
L'image réciproque d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel. On note $Ker\ f$ et on appelle **noyau** *de f l'image réciproque de $\{0\}$, c'est un sous-espace vectoriel. Une application linéaire est injective si et seulement si son noyau est $\{0\}$. On notera que le noyau d'une application linéaire est le noyau du morphisme de groupes correspondant.*
On note Id la fonction identité de E ; c'est une application linéaire et un isomorphisme. On note $Invf$ **l' ensembles des invariants de** *f ; $Invf = Kerf - Id$. On note $Oppf$* **l'ensemble des vecteurs changés en leur opposé** *; $Oppf = Kerf + Id$.*
On appelle **homothétie de rapport λ d'un espace vectoriel** *E l'application $x \mapsto \lambda.x$.*

Quelques propriétés :
• On peut se contenter pour vérifier la linéarité d'une application d'assurer que $f(\lambda.x+\mu.y) = \lambda.f(x) + \mu.f(y)$
• L'image de 0 par une application linéaire est 0.
• L'identité est un automorphisme.
• L'image d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel. On note $Im\ f$ l'image de f, c'est un sous-espace vectoriel.
• La composée de deux applications linéaires est une application linéaire.
• L'application qui à f associe $g \circ f$ (resp. $f \circ g$) est un morphisme d'espaces vectoriels.
• $Ker\ f \subset Ker\ g \circ f$
• $Im g \circ f = Im g$
• $g \circ f = 0$ si et seulement si $Im\ f \subset Ker\ g$
• L'application qui à un polynôme associe son polynôme dérivé est un endomorphisme.

Définition 963 *On appelle* n-**ième itérée de** f *la fonction* f^n. f *est dit* **nilpotent** *si il existe* n *tel que* $f^n = 0$. *Le plus petit* n *convenable est appelé* **indice de nilpotence** *de* f. *On convient que l'indice de nilpotence d'une fonction non nilpotente est* $+\infty$.

Propriétés :
Si f est un endomorphisme nilpotent, alors $f - Id$ et $f + Id$ sont des automorphismes.

Définition 964 (Définitions dans le cas d'un espace vectoriel produit)
On appelle k-**ième projection canonique** *d'un espace vectoriel produit* $E_1 \times ... \times E_n$ *l'application qui à* x *dans le produit associe sa composante dans* E_k.
On appelle k-**ième injection canonique** *d'un espace vectoriel produit* $E_1 \times ... \times E_n$ *l'application qui à* x *dans* E_k *associe* $(0, ..., 0, x, 0, ...0)$.

On notera que si f_i est linéaire de E_i dans F_i, alors $(x_1, ..., x_n) \to (f_1(x_1), ..., f_n(x_n))$ est une application linéaire ; son noyau est le produit des noyaux.

27.3 Somme de sous-espaces vectoriels

27.3.1 Généralités

Définition 965 (Somme de sous-espaces vectoriels) *On se donne* $F_1, ..., F_n$ *des sous-espaces vectoriels de l'espace vectoriel* E. *L'application* ϕ *qui à* $(x_1, ..., x_n)$ *associe* $(x_1 + x_2 + ... + x_n)$ *est une application linéaire sur l'espace produit* $F_1 \times ... \times F_n$. *L'image de* ϕ *est appelée* **somme** *des sous-espaces vectoriels* $F_1, ..., F_n$, *et est notée* $F_1 + ... + F_n$ *ou* $\sum_{i \in [1,n]} F_i$.

Propriétés :
• La somme des F_i contient tous les F_i.
• L'espace vectoriel engendré par l'union des F_i est leur somme.

Définition 966 (Somme directe de sous-espaces vectoriels) *On dit que la somme de* $F_1, ..., F_n$ *est une* **somme directe** *lorsque la fonction* ϕ *est injective. Au lieu de noter* $\sum F_i$ *on peut alors noter* $\bigoplus F_i$

Propriétés :
• Cela revient à dire que tout vecteur de l'espace vectoriel somme s'écrit comme somme d'un élément de F_1, d'un élément de F_2, ... , d'un élément de F_n, cette décomposition étant unique.
• La somme est directe si et seulement si pour tout j $F_j \cap \sum_{i \neq j} F_i = \{0\}$.
• On peut encore simplifier ce critère ; la somme est directe si et seulement si pour tout

j $F_j \cap \sum_{i=1..j-1} F_i = \{0\}$.

27.3.2 Espaces supplémentaires

⊡ Généralités

Définition 967 *Deux sous-espaces vectoriels d'un espace E sont dits* **supplémentaires** *lorsque leur somme est directe et égale à E.*

Propriétés :
• Deux espaces sont supplémentaires si et seulement si leur intersection est $\{0\}$ et si leur somme est E.
• Deux espaces sont supplémentaires si l'application ϕ (voir la définition d'une somme de sous-espaces vectoriels) est bijective.
• Deux sous-espaces vectoriels qui sont supplémentaires d'un même sous-espace vectoriel sont isomorphes - preuve en considérant la projection sur le supplémentaire en question par rapport à l'un des deux sous-espaces ; un espace supplémentaire du noyau est isomorphe à l'image, d'où le résultat - les définitions nécessaires arrivent plus bas).
Exemples :
Dans $\mathbb{K}^n$, l'espace vectoriel engendré par un vecteur $(a_1, ..., a_n)$ est supplémentaire de l'espace vectoriel des $(x_1, ..., x_n)$ tels que $a_1.x_1 + ... + a_n.x_n = 0$.

⊡ Applications aux applications linéaires. Projections, symétries.

Théorème 968 *Deux applications linéaires ayant les mêmes restrictions à deux espaces vectoriels supplémentaires sont égales.*
Etant données deux applications linéaires définies respectivement de F_1 dans G et de F_2 dans G, avec F_1 et F_2 deux supplémentaires de E, il existe une et une seule application linéaire dont les restrictions à F_1 et F_2 soient ces applications.

Proposition 969 *Etant donné un sous-espace vectoriel H, f une application linéaire, alors $Ker\ f_{|H} = Ker\ f \cap H$*

Théorème 970 (Théorème noyau-image) *$Im\ f$ est isomorphe à tout supplémentaire de $Ker\ f$.*

Démonstration : Il suffit de considérer la restriction de f à un supplémentaire de $Ker\ f$, et de montrer l'injectivité et la surjectivité.□

Définition 971 (Projection) *On a vu que dans le cas où F et G étaient des espaces supplémentaires de E, pour tout x il existait un unique $(f, g) \in F \times G$ tel que $x = f + g$. On appelle* **projection sur** F **parallèlement à** G *l'application qui à x associe f.*

Définition 972 *On dit qu'un endomorphisme f est* **idempotent** *lorsque $f \circ f = f$. Un endomorphisme idempotent est aussi appelé* **projecteur**.

Propriétés :
Etant donné p projecteur, on a :
- $E = Ker\ p \oplus Im\ p$
- $Invp = Im\ p$

On note qu'une projection est un projecteur.

Proposition 973 *Un projecteur p est en fait la projection sur $Im\ p$ parallèlement à $Ker\ p$.*

__**Démonstration :**__ Il suffit de considérer les propriétés ci-dessus, elles-mêmes facilement démontrables.□

Proposition 974 *$Id - f$ est un projecteur si et seulement si f est un projecteur.*

__**Démonstration :**__ Il suffit de développer $(Id - f)^2$.□

Définition 975 (Projecteurs associés) *Deux projecteurs sont dits* **projecteurs associés** *lorsque leur somme est l'identité.*

Proposition 976 *Si f et g sont deux projecteurs associés, alors $Im\ f = Ker\ g$ et $Im\ g = Ker\ f$.*

Définition 977 (Symétrie) *A et B étant supplémentaires, on appelle* **symétrie** *par rapport à A parallèlement à B l'endomorphisme s tel que $s_{|A} = Id$ et $s_{|B} = -Id$.*

On a vu plus haut qu'un endomorphisme pouvait être défini par ses restrictions sur deux espaces supplémentaires.

Définition 978 (Involution) *Un endomorphisme f est une* **involution** *lorsque $f \circ f = Id$.*

Définition 979 *Une symétrie s et un projecteur p sont dits* **associés** *lorsque $s = 2.p - Id$.*

C'est le cas lorsque s et p se font par rapport au même espace et parallèlement au même espace.

Propriétés :

• Une symétrie est une involution

• Etant donnée s symétrie, $E = Inv\ s \oplus Opp\ s$

• s est la symétrie par rapport à $Inv\ s$ et parallèlement à $Opp\ s$.

• Toute symétrie est associée à un projecteur, et tout projecteur est associée à une symétrie.

Dilatations, transvections

Définition 980 *Soit H un hyperplan d'un espace vectoriel E, et D une droite supplémentaire de H. On appelle* **dilatation d'hyperplan** *H***, de direction** *D* **et de rapport** *λ l'application linéaire dont la restriction à H est l'identité et dont la restriction à D est l'homothétie de rapport λ.*
Soit H un hyperplan d'un espace vectoriel E, et h une forme linéaire de noyau H. On appelle **transvection d'hyperplan** *H une application de E dans E de la forme*

$$x \mapsto x + h(x).u$$

pour un certain u dans H.

⚠ Ne pas se méprendre sur la notion de transvection d'hyperplan H ; il existe plusieurs transvections différentes d'hyperplan H.

Rappelons que la proposition 811 signale que $GL(E)$ est engendré par les transvections et les dilatations, et que $SL(E)$ est engendré par les transvections.

27.4 Espace vectoriel quotient

27.4.1 Généralités

Définition 981 *Soit F un sous-espace vectoriel de E. Alors la relation définie par*

$$x\mathcal{R}y \iff x - y \in F$$

est une relation d'équivalence compatible avec l'addition et le produit externe. L'ensemble quotient est un espace vectoriel pour les lois induites ; il est appelé **espace vectoriel quotient** *et est noté $E_{/F}$.*

Propriétés :

• La surjection canonique qui à x associe $\overline{x}$ est linéaire. Son noyau est F.

• Si F et G sont supplémentaires, alors G est isomorphe à $E_{/F}$.

27.4.2 Application aux applications linéaires

Théorème 982 *Etant donné une application linéaire f de E dans F, alors f s'écrit de manière unique $f = i \circ b \circ s$, avec s la surjection canonique de E dans $E_{/Ker\ f}$, i l'injection canonique de $Im\ f$ dans F, et g un isomorphisme de $E_{/Ker\ f}$ dans Imf.*

Démonstration : Facile ! □

27.5 Translations - espaces affines

Pour plus d'informations on pourra consulter la partie32.

Définition 983 (Translations) *On appelle* **translation** *de vecteur a l'application d'un espace affine X dans lui-même qui à x associe $x + a$. On note $\mathcal{T}(E)$ l'ensemble des translations de E.*
On appelle **sous-espace affine** *de E l'image d'un sous-espace vectoriel de E par une translation de E.*
On appelle **direction d'un sous-espace affine** *l'ensemble des $x - y$ pour x et y dans l'espace affine.*
On dit qu'un sous-espace affine A est **parallèle** *à un sous-espace affine B si et seulement si A est vide ou la direction de A est incluse dans la direction de B.*
On dit que deux sous-espaces affines sont **parallèles** *s'ils sont non vides et ont même direction.*
On dit que deux sous-espaces affines sont **strictement parallèles** *s'ils sont non vides et ont même direction et sont distincts.*

Propriétés :
• $(\mathcal{T}(E), \circ)$ est un groupe commutatif. $(\mathcal{T}(E), \circ) \simeq (E, +)$.
• Un sous-espace affine de E contient 0 si et seulement si c'est un sous-espace vectoriel (en le munissant des lois induites).
• Si un espace affine A est de la forme $a + F$, alors il est aussi de la forme $x + F$ pour tout x de A.
• Le sous-espace affine A est égal à $a + F$, avec a quelconque dans A, et F la direction de A.
• Etant donnés A et B deux espaces affines, $a \in A$ et $b \in B$, de direction respectives F et G, alors $A \cap B \neq \emptyset \iff b - a \in F + G$
• Etant donnés deux espaces affines d'intersection non vide, l'intersection est un espace affine de direction l'intersection de leurs directions.
• Si A est parallèle à B alors $A \cap B = \emptyset$ ou $A \subset B$.

27.6 Familles libres. Familles génératrices. Bases

27.6.1 Généralités

Définition 984 *Un élément x de l'espace vectoriel E est dit* **combinaison linéaire** *d'une famille $(x_i)_{i\in I}$ d'éléments de E si il existe un entier n, un n-uplet $(\lambda_1, ..., \lambda_n)$ d'éléments de $\mathbb{K}$, et un n-uplet $(i_1, ..., i_n)$ d'éléments de I tels que $x = \sum_{1\leq j\leq n} \lambda_j x_{i_j}$.*
Un élément x est dit **combinaison linéaire** *d'un sous-ensemble A de E si x est combinaison linéaire d'une famille d'éléments de A.*
Une famille d'éléments d'un sous-espace vectoriel F est dite **famille génératrice** *de F si l'espace vectoriel engendré par cette famille est F.*
Une famille d'éléments de E est dite **famille libre** *si une combinaison linéaire de cette famille est nulle si et seulement si chaque λ_i est nul.*
Une famille est dite **famille liée** *quand elle n'est pas libre.*

Propriétés :
• L'image d'une famille génératrice de F par une application linéaire f est une famille génératrice de $f(F)$.
• L'ensemble des combinaisons linéaires de A est l'espace vectoriel engendré par A.
• Toute famille contenant le vecteur nul est liée.
• Les éléments d'une famille libre sont deux à deux distincts.
• Une famille est liée si et seulement si un de ses éléments est combinaison linéaire des autres.
• Une famille est liée si et seulement si un de ses éléments appartient à l'espace vectoriel engendré par les autres.
• Toute sous-famille d'une famille libre est libre.
• Toute sur-famille d'une famille liée est liée.

Théorème 985 • *L'image d'une famille liée par une application linéaire est une famille liée.*
• *L'image d'une famille libre par une application injective est une famille libre.*

Le théorème suivant, facile à démontrer, est fondamental pour la suite.

Théorème 986 • *Etant donnée une famille libre et un élément appartenant à l'espace vectoriel engendré par cette famille, alors cet élément s'écrit de manière unique comme combinaison linéaire d'éléments de la famille.*

Définition 987 (Base) *Une base d'un espace vectoriel E est une famille libre et génératrice.*

Proposition 988 *Une famille est une base si et seulement si c'est une famille libre maximale (au sens de l'inclusion).*

Démonstration : Si elle est pas maximale, on peut ajouter un élément tout en la laissant libre ; on en déduit que ce n'est pas une famille génératrice. Si elle n'est pas génératrice, alors on peut ajouter un élément n'appartenant pas à l'espace engendré ; cette nouvelle famille est aussi libre, donc la précédente n'était pas maximale.$\square$

Proposition 989 *Une famille est une base si et seulement si c'est une famille génératrice minimale.*

Démonstration : Même principe que la preuve ci-dessus.$\square$

Définition 990 *Etant donnée une base (e_i) d'un espace vectoriel F et un élément x de F, il existe une unique famille (λ_i) de support fini telle que $x = \sum \lambda_i e_i$. Les λ_i sont appelés* **coordonnées** *du vecteur x.*

27.6.2 Applications aux applications linéaires

Théorème 991 *Etant donnée $(f_i)_{i\in I}$ une base de F et $(g_i)_{i\in I}$ une famille de G, il existe une unique application linéaire f de F dans G telle que $f(f_i) = g_i$ pour tout $i \in I$.*
- *La famille (g_i) est libre si et seulement si f est injective.*
- *La famille (g_i) est génératrice si et seulement si f est surjective.*
- *La famille (g_i) est une base de G si et seulement si f est bijective.*

27.6.3 Applications aux sous-espaces vectoriels

Proposition 992 *L'ensemble $\mathcal{L}(E)$ des endomorphismes de E muni de l'addition, de la composition, et de la multiplication par un scalaire, est une algèbre.*

Attention, cette algèbre n'est pas commutative (en général).

27.7 Dualité

27.7.1 Définitions et premières propriétés. Le dual et le bidual

Définition 993 *On appelle* **forme linéaire** *sur un K-espace vectoriel E une application linéaire de E dans K.*
On appelle **dual** *d'un K-espace vectoriel E l'ensemble E^* des formes linéaires sur cet espace vectoriel.*
On appelle **hyperplan** *tout sous-espace vectoriel admettant une droite pour supplémentaire.*
On appelle **bidual** *d'un $\mathbb{K}$-espace vectoriel le dual de son dual. On le note E^{**}.*
On appelle **application linéaire canonique** *de E dans E^{**} l'application qui à x associe $\phi_x : u \mapsto u(x)$ (on vérifiera facilement que c'est une application linéaire).*

Propriétés :
• Une forme linéaire non nulle est surjective.
• Un sous-espace vectoriel est un hyperplan si et seulement si tout droite vectorielle passant par un vecteur appartenant à son complémentaire est un supplémentaire de cette droite.
• Un sous-espace vectoriel est un hyperplan si et seulement si c'est le noyau d'une forme linéaire non nulle.
• Deux formes linéaires non nulles sont liées (dans le dual) si et seulement si elles ont même noyau.
• Etant données u et v des formes linéaires sur E, il existe $\lambda \in \mathbb{K}$ tel que $u = \lambda v$ si et seulement si $Ker\ u \subset Ker\ v$.

27.7.2 Orthogonalité

Définition 994 *Etant donné x dans E et u dans E^*, on dit que x et u sont* **orthogonaux** *si et seulement si $u(x) = 0$*
Etant donnée une partie non vide A de E, on appelle **orthogonal** *de A dans E^* et on note $A^\perp$ l'ensemble des u orthogonaux à tous les éléments de A.*
Etant donnée une partie non vide A de E^* on appelle **orthogonal** *de A dans E et on note A^o l'ensemble des x orthogonaux à tous les éléments de A.*

Propriétés :
J'utilise le terme orthogonal sans plus de précision lorsque la propriété vaut à la fois dans le cas de l'orthogonal d'une partie de E dans le dual et dans le cas de l'orthogonal d'une partie du dual E^* dans E.
Il n'y a pas ici de bidualité ; l'orthogonal d'une partie du dual E' est à considérer dans E.
• L'orthogonal d'une partie est l'orthogonal de l'espace engendré par cette partie.

• L'orthogonal d'une partie est un sous-espace vectoriel.
• Toute partie est incluse dans l'orthogonal de son orthogonale.
• $A \subset B$ alors l'orthogonal de B est inclus dans l'orthogonal de A.
• L'intersection des orthogonaux est l'orthogonal de l'union.

⚠ Attention, ne pas confondre l'orthogonal de l'orthogonal de A $A^{\perp^o}$ et l'orthogonal de l'orthogonal de A $A^{\perp^\perp}$.

27.8 Transposition

Définition 995 *Etant donnée f une application linéaire de E dans F on appelle* **transposée de** *f l'application de F^* dans E^* qui à v associe $v \circ f$. C'est une application linéaire, et on la note ${}^t f$.*

Propriétés :
• L'application qui à f associe ${}^t f$ est une application linéaire.
• ${}^t(g \circ f) = {}^t f \circ {}^t g$
• ${}^t Id_E = Id_{E^*}$
• Si f est un isomorphisme, ${}^t f$ est un isomorphisme, et ${}^t(f^{-1}) = ({}^t f)^{-1}$
• $Ker\ {}^t f = (Im\ f)^\perp$
• f est surjective si et seulement si ${}^t f$ est injective.

27.9 $\mathbb{K}$-algèbres

27.9.1 Définitions et généralités

Définition 996 ($\mathbb{K}$-algèbre) $(A, +, \times, .)$ *est une* $\mathbb{K}$**-algèbre** *si*
• $(A, +, .)$ *est un $\mathbb{K}$-espace vectoriel*
• $(A, +, \times)$ *est un anneau*
• $(\lambda.a) \times b = a \times (\lambda.b) = \lambda.(a \times b)$
La $\mathbb{K}$-algèbre est en outre **commutative** *lorsque $\times$ est commutative.*

L'ensembles des suites à valeurs dans $\mathbb{K}$ est un $\mathbb{K}$-algèbre commutative. L'espace vectoriel des applications de A dans $\mathbb{K}$ est une $\mathbb{K}$ algèbre commutative.

Définition 997 (Morphisme de $\mathbb{K}$-algèbres) *Un* **morphisme d'algèbre** *est une application qui est à la fois un morphisme d'anneaux sur les anneaux sous-jacents et un morphisme d'espaces vectoriels sur les espaces vectoriels sous-jacents.*

27.10 Résultats d'algèbre linéaire

• L'union de deux sous-espaces vectoriels est un sous-espace vectoriel si et seulement si l'un des deux est inclus dans l'autre.

• Etant donnés deux projecteurs, la somme est un projecteur si et seulement si les composés de ces projecteurs (dans les deux sens, la composition n'étant pas commutative) sont nulles. Dans ce cas le noyau de la somme est l'intersection des noyaux, et l'image de la somme est la somme directe des images.

• Etant donnés F et G des espaces vectoriels et f et g des vecteurs alors $f + F \subset g + G$ si et seulement si $F \subset G \wedge f - g \in G$.

Définition 998 *On appelle* **homothétie vectorielle de rapport** k *l'application qui à* x *associe* $k.x$ *; un endomorphisme* f *est une homothétie vectorielle si et seulement si pour tout* x *la famille* $(x, f(x))$ *est liée.*

Démonstration : On procède par étapes :
• tout d'abord montrer que pour tout x il existe $k_x \in K$ tel que $f(x) = k_x.x$.
• ensuite montrer que k_x est constant sur toute droite vectorielle.
• développer $f(x + y)$ avec (x, y) une famille libre. On en déduit $k_x = k_y$.□

• La famille des fonctions f_a de $\mathbb{R}$ dans $\mathbb{R}$ qui à x associe 1 si $x > a$ et 0 si $x \leq a$ est libre.

Démonstration : On suppose qu'une telle fonction est combinaison linéaire d'un nombre fini d'autres fonctions, et on constate que la fonction doit être continue là où justement elle ne l'est pas...□

27.11 Exercices d'algèbre linéaire

Quelques résultats pour s'habituer aux manipulations sur les polynômes d'endomorphismes et aux sommes d'espaces :

$$f^2 - 5.f + 6.Id = 0 \to E = Ker(f - 2.Id) \oplus Ker(f - 3.Id)$$

(f est un endomorphisme)

Pour s'habituer aux manipulations sur les noyaux et les images :
(f est un endomorphisme)

$$Ker\ f = Ker\ f^2 \iff Im\ f \cap Ker\ f = \{0\}$$
$$Im\ f = Im\ f^2 \iff Im\ f + Ker\ f = \{0\}$$
$$n \mapsto Ker\ f^n \text{ est croissante}$$
$$n \mapsto Im\ f^n \text{ est décroissante}$$
$$\exists n / Ker\ f^n = Ker\ f^{n+1} \to \forall k > n\ Ker\ f_k = Ker\ f_n$$
$$\exists n / Im\ f^n = Im\ f^{n+1} \to \forall k > n\ Im\ f_k = Im\ f_n$$

Chapitre 28

Algèbre multilinéaire

Pour la suite il est utile de connaître quelques bases sur σ_n ; que l'on trouvera à la partie21.10.11. Les premières sections, de type fondements théoriques pour la suite, sont un juste rapidement ébauchées.

28.1 Généralités

Définition 999 (Application multilinéaire) *Soient E_1, ... , E_n et F des $\mathbb{K}$-espace vectoriel , alors $f : \Pi_{i\in[1,n]}E_i \to F$ est n-**linéaire** si pour tout (x_i) dans ΠE_i et tout j l'application $x \to f(x_1,...,x_{j-1},x,x_{j+1},...,x_n)$ est linéaire.*
*Si $E_1 = E_2 = ...E_n$ ont dit que f est une **application n-linéaire** sur E.*
*Si F est le corps $\mathbb{K}$, alors f est dite **forme n-linéaire**.*
On note $L_n(E,F)$ l'ensemble des applications n-linéaires de E dans F.
On note $L_n(E)$ l'ensemble des formes n-linéaires sur E, c'est à dire $L_n(E,\mathbb{K})$.
Etant donné $f \in L_n(E,F)$ et $\sigma \in \sigma_n$ on note f_σ l'application n-linéaire $(x_1,...,x_n) \mapsto f(x_{\sigma(1)},x_{\sigma(2)},...,x_{\sigma(n)})$.
*Une application n-linéaire est dite **symétrique** si pour tout σ $f_\sigma = f$.*
*Une application n-linéaire est dite **antisymétrique** si pour tout σ $f_\sigma = \epsilon(s).f$.*
*Une application n-linéaire est dite **alternée** si $i \neq j$ et $x_i = x_j$ implique $f(x_1,...,x_n) = 0$.*

On note $S_n(E,F)$ l'ensemble des applications n-linéaires symétriques de E dans F et $A_n(E,F)$ l'ensemble des applications n-linéaires alternées de E dans F.
On note $S_n(E)$ l'ensemble des formes n-linéaires symétriques sur E et $A_n(E)$ l'ensemble des formes n-linéaires alternées sur E.

⚠ Une application p-linéaire n'est pas une application linéaire.

Proposition 1000 *] • $L_p(E,F)$ est un $\mathbb{K}$-espace vectoriel , sous-espace vectoriel de F^{E^p}.*
• f est symétrique si et seulement si pour toute transposition τ $f_\tau = f$.
• f est antisymétrique si et seulement si pour toute transposition τ $f_\tau = -f$.

Démonstration : Premier point évident, les deux points suivants demandent juste de se rappeler que les transpositions engendrent σ_n. On pourrait en fait utiliser d'autres familles génératrices.□

Proposition 1001 *Une application n-linéaire alternée est antisymétrique ; si $\mathbb{K}$ n'est pas de caractéristique 2, la réciproque est vraie aussi.*

Démonstration : Supposons f n-linéaire, et $x_i = x_j$, avec $i \neq j$.

$$0 = f(x_1,...,x_i + x_j,...,x_j + x_i,...,x_n)$$

(car f est alternée)

$$0 = f(x_1,...,x_i,...,x_i,...,x_n)$$
$$+f(x_1,...,x_i,...,x_j,...,x_n)$$

$$+f(x_1, ..., x_j, ..., x_j, ..., x_n)$$

$$+f(x_1, ..., x_j, ..., x_i, ..., x_n)$$

(car f est n-linéaire) or $f(x_1, ..., x_i, ..., x_i, ..., x_n) = f(x_1, ..., x_j, ..., x_j, ..., x_n) = 0$ puisque f est alternée
donc $f(x_1, ..., x_i, ..., x_j, ..., x_n) = -f(x_1, ..., x_j, ..., x_i, ..., x_n)$.

Par la proposition 1000, le résultat alterné $\rightarrow$ antisymétrique est donc prouvé. La réciproque est évidente, en utilisant la même caractérisation de l'antisymétrie par la proposition 1000.$\square$

On suppose désormais que $\mathbb{K}$ est un corps de caractéristique $\neq 2$.

Théorème 1002 *Soit ϕ une forme n-linéaire antisymétrique sur E $\mathbb{K}$-espace vectoriel de dimension n, $e_1, ..., e_n$ une base de E, $e_1^*, ..., e_n^*$ sa base duale et $x_1, ..., x_n$ une famille de n éléments de E.*
Alors :

$$\phi(x_1, ..., x_n) = (\sum_{\sigma \in \sigma_n} \epsilon(s).\Pi_{i=1}^n e_{\sigma(i)}^*(v_i)).\phi(e_1, ..., e_n)$$

et

$$\phi(x_1, ..., x_n) = (\sum_{\sigma \in \sigma_n} \epsilon(s).\Pi_{i=1}^n e_i^*(v_{\sigma(i)})).\phi(e_1, ..., e_n)$$

Démonstration : On procède en quatre étapes :
• On remplace x_j par $\sum_i e_i^*(x_j).e_i$ dans $\phi(x_1, ..., x_n)$.
• On développe en utilisant la linéarité suivant chaque composante, on obtient donc

$$\sum_{(i_1, ..., i_n) \in [1,n]^n} e_{i_1}^*(x_1).e_{i_2}^*(x_2).....e_{i_n}^*(x_n).\phi(e_{i_1}, ..., e_{i_n})$$

• On supprime tous les termes tels que le cardinal de $\{i_1, ..., i_n\}$ soit différent de n, ce qui permet d'introduire les permutations.
• Il ne reste plus qu'à remplacer $\phi(e_{i_1}, ..., e_{i_n})$ par $\epsilon(\sigma).\phi(e_1, ..., e_n)$.
La deuxième formule s'obtient simplement en remplaçant σ par σ^{-1}.$\square$

Définition 1003 (Symétrisé et antisymétrisé d'une forme n-linéaire) *Soit f une forme n-linéaire sur un $\mathbb{K}$-espace vectoriel E.*
Alors l'application $S(f)$ égale à $(x_1, ..., x_n) \mapsto \sum_{\sigma \in \sigma_n} f_\sigma(x_1, ..., x_n)$ est appelée **symétrisée** *de f ; elle est symétrique.*
Alors l'application $A(f)$ égale à $(x_1, ..., x_n) \mapsto \sum_{\sigma \in \sigma_n} \epsilon(\sigma).f_\sigma(x_1, ..., x_n)$ est appelée **antisymétrisée** *de f ; elle est alternée.*
L'application $f \mapsto S(f)$ est appelée **opérateur de symétrisation**.
L'application $f \mapsto A(f)$ est appelée **opérateur d'antisymétrisation**.

Définition 1004 (Produit tensoriel, produit symétrique, produit extérieur) *Soient $f_1, ..., f_n$ des formes linéaires sur le $\mathbb{K}$-espace vectoriel E.*
On appelle **produit tensoriel de** $(f_1, ..., f_n)$ *l'application qui à $(x_1, ..., x_n)$ associe $f_1(x_1) \times f_2(x_2) ... \times f_n(x_n)$. On le note $f_1 \otimes f_2 \otimes ... \otimes f_n$.*
L'aplication symétrisée du produit tensoriel est appelée **produit symétrique** *de $(f_1, ..., f_n)$; on la note $f_1.f_2. \ldots .f_n$.*
L'application antisymétrisée du produit tensoriel est appelée **produit extérieur** *de $(f_1, ..., f_n)$; on le note $f_1 \wedge f_2 \wedge ... \wedge f_n$.*
Une application n-linéaire exprimable comme produit tensoriel est dite décomposable dans $L_n(E)$.
Une application n-linéaire exprimable comme produit symétrique est dite décomposable dans $S_n(E)$.
Une application n-linéaire exprimable comme produit extérieur est dite décomposable dans $A_n(E)$.

Proposition 1005 • *Le produit symétrique de n formes linéaires est symétrique.*
• *Le produit extérieur de n formes linéaires est antisymétriques.*
• *L'application qui à n formes linéaires associe leur produit tensoriel est n-linéaire de ${E^*}^n$ dans $L_n(E)$.*
• *L'application qui à n formes linéaires associe leur produit symétrique est n-linéaire symétrique.*
• *L'application qui à n formes linéaires associe leur produit extérieur est n-linéaire alternée.*

Quelques théorèmes donnés sans démonstration :

Théorème 1006 *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, $(e_1, ..., e_n)$ une base de E, $(e_1^*, ..., e_n^*)$ sa base duale. On note $\mathcal{F}$ l'ensemble des applications de $[1, p]$ dans $[1, n]$. Pour tout f dans $\mathcal{F}$ on note $e_f = e^*_{f(1)} \otimes e^*_{f(2)} \otimes \cdots \otimes e^*_{f(p)}$. Alors la famille des e_f pour $f \in \mathcal{F}$ est une base de $L_p(E)$.*

Définition 1007 *La base donnée par le théorème précédent est appelée* **base associée à** $(e_1, ..., e_n)$.

Corollaire 1008 *La dimension de* $L_p(E)$ *est* $(dim\ E)^n$.

Théorème 1009 *Soit* E *un* $\mathbb{K}$*-espace vectoriel de dimension finie* n, $(e_1, ..., e_n)$ *une base de* E, $(e_1^*, ..., e_n^*)$ *sa base duale.*
$\mathbb{K}$ *est supposé de caractéristique nulle.*
On note A *l'ensemble des applications* f *de* $[1, n]$ *dans* $[0, p]$ *telles que* $\sum_{i=1}^n f(i) = p$. *Alors on note* $e_f = e_1^{f(1)}.e_2^{f(2)}.\ldots.e_n^{f(n)}$.
Alors la famille des e_f *pour* $f \in A$ *forme une base de* $S_p(E)$.

Corollaire 1010 *La dimension de* $S_p(E)$ *est égale à* $C^p_{dim\ E+p-1}$.

Théorème 1011 *Soit* E *un* $\mathbb{K}$*-espace vectoriel de dimension finie* n, $(e_1, ..., e_n)$ *une base de* E, $(e_1^*, ..., e_n^*)$ *sa base duale.*
$\mathbb{K}$ *est supposé de caractéristique nulle.*
On note A *l'ensemble des applications* f *strictement croissantes de* $[1, p]$ *dans* $[1, n]$. *Alors on note* $e_f = e_{f(1)}.e_{f(2)}.\ldots.e_{f(n)}$.
Alors la famille des e_f *pour* $f \in A$ *forme une base de* $S_p(E)$.

Corollaire 1012 *Si* E *est un* $\mathbb{K}$*-espace vectoriel de dimension finie* n *et si* $\mathbb{K}$ *est de caractéristique nulle, alors* $dim\ A_p(E) = C_n^p$.

28.2 Algèbre multilinéaire et topologie

Définition 1013 *Etant donnés* E_1, ... , E_n *et* F *des espaces vectoriels normés , on note* $\mathcal{L}(E_1, ..., E_n; F)$ *l'espace des applications* n*-linéaires continues de* $E_1 \times ... \times E_n$ *dans* F. *On le norme par* $f \mapsto sup_{\forall i \|x_i\| \leq 1} \|f(x_i)$ *; on obtient ainsi un espace vectoriel normé .*

Théorème 1014 (Quelques théorèmes (pas durs !) sans preuve) • *Soient $E_1, ..., E_n$ et F des espaces vectoriels normés , et soit f application n-linéaire de $E_1, ..., E_n$ dans F. Alors :*
- f est continue si et seulement si f est continue en 0
- f est bornée sur le produit des boules unités des E_i

• *Si F est un espace de Banach, alors $\mathcal{L}(E_1, ..., E_n; F)$ est un espace de Banach.*

• *Etant donnée f application n-linéaire continue de $E_1 \times ... \times E_n$ dans F la différentielle de f en $(x_1, x_2, ..., x_n)$ est $(h_1, ..., h_n) \mapsto f(h_1, x_2, ..., x_n) + f(x_1, h_2, x_3, ..., x_n) + ... + f(x_1, x_2, ..., x_{n-1}, h_n)$.*

• $\mathcal{L}(E_1, E_2; F) \simeq \mathcal{L}(E_1; \mathcal{L}(E_2; F))$.

28.3 Déterminants

28.3.1 Déterminant d'une famille de vecteurs

On suppose E $\mathbb{K}$-espace vectoriel de dimension finie n.

Définition 1015 *On appelle* **déterminant** *d'une famille $(x_1, ..., x_n)$ d'éléments de E dans une base $(e_1, ..., e_n)$ de E la somme :*

$$\sum_{\sigma \in \sigma_n} \epsilon(\sigma) \Pi_{i=1}^{n} e^*_{\sigma(i)}(x_i)$$

On le note $det_{(e_1,...,e_n)}(x_1, ..., x_n)$.

Proposition 1016 *Si f est n-linéaire sur E, alors $\sum_{\sigma \in \sigma_n} \epsilon(\sigma).f_\sigma$ est n-linéaire et antisymétrique.*

Démonstration : La somme est trivialement n-linéaire, et l'antisymétrie se montre facilement (on rappelle juste que la signature du produit de deux permutations est le produit des signatures de ces deux permutations).□

On suppose pour la suite que $\mathbb{K}$ n'est pas de caractéristique 2.

Théorème 1017 • *Etant donnée une base B de E, il existe une et une seule forme n-linéaire alternée égale à 1 sur B ; c'est $det_B(.)$.*
• *Les formes n-linéaires alternées sur E sont égales à multiplication par un scalaire près.*

Démonstration : • Il est clair que $det_B(B) = 1$
• La proposition précédente nous assure que $det_B(.)$ est n-linéaire alternée.
• Le théorème 1002 nous assure le deuxième point.□

En conclusion, on a les propriétés élémentaires suivantes du déterminant dans une base B :

Proposition 1018 • *Le déterminant est n-linéaire alterné.*
• *On ne change pas le déterminant des x_i en ajoutant à un des x_i une combinaison linéaire des autres x_j.*
• *Le déterminant de $(x_1, ..., x_n)$ est égal à $\epsilon(\sigma)$ fois le déterminant de $(x_{\sigma(1)}, ..., x_{\sigma(n)})$.*
• $det_B = det_B(B').det_{B'}$
• $det_B(B').det_{B'}(B) = 1$
• *La famille des x_i est une base si et seulement si $det_B(x_1, ..., x_n) \neq 0$*

Démonstration : (simplement du dernier point, les autres se montrant facilement l'un après l'autre)
Si c'est une base, alors on applique la formule juste au dessus pour conclure que le déterminant est non nul.
Si le déterminant est non nul, alors supposons la famille liée, on peut ajouter à un vecteur une combinaison linéaire des autres et on a ainsi une famille dont un vecteur est nul, et donc le déterminant devrait être nul.□

28.3.2 Déterminant d'un endomorphisme

E est toujours un $\mathbb{K}$-espace vectoriel de dimension finie n sur un corps $\mathbb{K}$ de caractéristique différente de deux.

Définition 1019 *On appelle* **déterminant de l'endomorphisme** *f le déterminant de $f(B)$ dans la base B ; on le note $det\ f$.*
On appelle **groupe spécial linéaire de** *E l'ensemble des endomorphismes de E de déterminant 1 ; on le note $SL(E)$.*

Démonstration : Il convient pour que la définition ait un sens de démontrer que ce déterminant ne dépend pas de la base B. On suppose donc données deux bases B et B', et on montre que $det_B(f(B)) = det_{B'}(f(B'))$.

• La fonction qui à un n-uple $x \in E^n$ associe $det_B(f(x))$ est n-linéaire alternée, donc c'est $\lambda.det_B$.
• En spécialisant en B cette égalité, on obtient

$$det_B(f(x)) = det_B(f(B)).det_B(x)$$

• Or par les propriétés du déterminant on a

$$det_B(x) = det_B(B').det_{B'}(x)$$

$$det_B(f(x)) = det_B(B').det_{B'}(f(x))$$

• Des deux points précédents on déduit

$$det_B(B').det_{B'}(f(x)) = det_B(f(B)).det_B(B').det_{B'}(x)$$

c'est-à-dire $det_{B'}(f(x)) = det_B(f(B)).det_{B'}(x)$, et donc $det'_B(f(B')) = det_B(f(B))$ en spécialisant en B'.$\square$

Proposition 1020 • *$det\ f \circ g = det\ f.det\ g$*
• *f est un automorphisme si et seulement si $det\ f \neq 0$*
• *Si $det\ f \neq 0$, alors f est un automorphisme et $det\ f^{-1} = \frac{1}{det\ f}$*
• *det est un morphisme de groupe entre $GL(E)$ et $\mathbb{K} \setminus \{0\}$.*
• *$SL(E)$, puisqu'il est le noyau du déterminant, est un sous-groupe de $GL(E)$*

28.3.3 Déterminant d'une matrice

On travaille encore dans un corps $\mathbb{K}$ de caractéristique différente de 2.

Définition 1021 *On appelle* **déterminant** *d'une matrice carrée M le déterminant de l'endomorphisme canonique associé à M dans $\mathbb{K}^n$. On le note $det\ M$ ou $|M|$.*
On appelle **groupe spécial linéaire d'ordre** *n et on note $SL_n(K)$ l' ensemble des matrices de déterminant égal à 1.*

Proposition 1022 *Le déterminant d'une matrice est aussi le déterminant de ses vecteurs-colonnes ou de ses vecteurs-lignes dans la base canonique de $\mathbb{K}^n$.*

Les propriétés suivantes se déduisent facilement des propriétés équivalentes chez les endomorphismes ou les familles de vecteurs :
• Une matrice est inversible si et seulement si son déterminant est non nul.
• Le déterminant du produit est le produit des déterminants.
• $SL_n(K)$ est un sous-groupe de $GL_n(K)$, noyau du déterminant en tant que morphisme de groupe de $GL_n(K)$ vers $\mathbb{K} \setminus \{0\}$.

• Deux matrices semblables ont même déterminant.

Proposition 1023 *Le déterminant d'une matrice triangulaire est égal au produit des éléments diagonaux.*

Démonstration : Il suffit de voir que seule la permutation identité est telle que pour tout i $M_{\sigma(i),i}$ soit non nul.$\square$

28.3.4 Pratique du calcul d'un déterminant ; développement suivant une ligne ou une colonne

Le point de vue adopté ici est celui du calcul du déterminant d'une matrice de type (n, n) sur un corps $\mathbb{K}$; bien sûr il faut bien voir qu'il en va de même du calcul du déterminant d'un endomorphisme ou d'une famille de vecteurs dans une base.

Pour la suite il est nécessaire d'avoir lu le début de la partie 30.3.10.

Proposition 1024 (Développement suivant une colonne)

$$\forall j \in [1, n] det\ M = \sum_{i\in[1,n]} \gamma_{i,j}.M_{i,j}$$

Démonstration : Il suffit de se rappeler que le déterminant est n-linéaire.$\square$

Proposition 1025 (Développement suivant une ligne)

$$\forall i \in [1, n] det\ M = \sum_{j\in[1,n]} \gamma_{i,j}.M_{i,j}$$

Démonstration : Il suffit de se rappeler que $det\ M = det\ {}^tM$.$\square$

Proposition 1026

$$det\ \tilde{M} = det\ com(M) = (det M)^{n-1}$$

Démonstration : • Si $\tilde{M}$ est inversible et pas M, alors $\tilde{M}.M = (det\ M).I = 0$ et donc $M = 0$, et donc $\tilde{M} = 0$, d'où contradiction.

• Si $\tilde{M}$ et M ne sont pas inversibles ni l'une ni l'autre, alors les déterminants sont égaux à 0, et l'égalité annoncée est vérifiée.

• Si M est inversible, alors

$$\tilde{M}.M = (det\ M).I$$

$$det\ (\tilde{M}.M) = det((det\ M).I)$$

$$det\ (\tilde{M}).det\ M = (det\ M)^n$$

$$det\ \tilde{M} = (det\ M)^{n-1}$$

D'où le résultat annoncé.□

28.4 Algèbre bilinéaire

On travaillera avec un corps $\mathbb{K}$ sympathique, comme $\mathbb{R}$ ou $\mathbb{C}$.

28.4.1 Formes bilinéaires

⊡ Le cas général

Définition 1027 *On appelle* **forme bilinéaire sur** E *une forme multilinéaire de* $\mathcal{L}_2(E)$.
Etant donnée ϕ *une forme bilinéaire on note* ${}^t\phi$ *l'application* $(x, y) \to \phi(y, x)$.

Il est immédiat que ϕ est symétrique si $\phi =^t \phi$, et que ϕ est antisymétrique si ${}^t\phi = -\phi$.

Proposition 1028 • *L'application qui à* ϕ *associe* ${}^t\phi$ *est un automorphisme involutif de* $\mathcal{L}_2(E)$.
• $\mathcal{L}_2(E)$ *est somme directe des deux sous-espace vectoriel de* $\mathcal{L}_2(E)$ *respectivement constitués des formes bilinéaires symétriques et des formes bilinéaires antisymétriques. On a en fait* $\phi = a + s$ *avec* $s = \frac{{}^t\phi+\phi}{2}$, $a = \frac{\phi-{}^t\phi}{2}$, a *antisymétrique, et* s *symétrique.*

⊡ Le cas de la dimension finie - expression matricielle

On travaille maintenant avec E $\mathbb{K}$-espace vectoriel de dimension finie n. On se donne une base $(e_1, ...e_n)$ une base de E.

Définition 1029 *Etant donné ϕ une forme bilinéaire sur E, on appelle* **matrice de ϕ dans la base $\mathcal{B}$** *la matrice M définie par*

$$M_{i,j} = \phi(e_i, e_j)$$

On la note $Mat_B(\phi)$.
Réciproquement, on appelle **forme bilinéaire sur E associée à la matrice M et à la base B** *l'application ϕ définie par*

$$\phi(x, y) =^t X.M.Y$$

avec X le vecteur défini par $X_i = e_i^*(x)$ et Y le vecteur défini par $Y_i = e_i^*(y)$.
La forme bilinéaire canoniquement associée à une matrice M de type (n, n) est la forme bilinéaire associée à cette matrice dans $\mathbb{K}^n$ pour la base canonique.

Proposition 1030 *Avec M la matrice de ϕ dans la base B, avec X le vecteur défini par $X_i = e_i^*(x)$ et Y le vecteur défini par $Y_i = e_i^*(y)$, on a*

$$\phi(x, y) =^t X.M.Y$$

Démonstration : Il n'y a qu'à l'écrire.□

Corollaire 1031 *La matrice de ${}^t\phi$ est la transposée de la matrice de ϕ.*

Proposition 1032 *Etant donnée B une base de E, l'application qui à une matrice associe la forme bilinéaire associée sur E pour B est un isomorphisme.*

Corollaire 1033 $dim\ \mathcal{L}_2(E) = n^2$

Proposition 1034 *Etant donnée B et B' deux bases de E, alors*

$$Mat_{B'}(\phi) =^t P_{B,B'}.Mat_B(\phi).P_{B,B'}$$

Démonstration : Evident.□

Au vu de ce résultat, on comprend l'intérêt d'introduire la définition suivante :

Définition 1035 *Deux matrices P et Q sont dites* **congruentes** *si il existe M inversible telle que $P =^t MQM$.*

Proposition 1036 • *La congruence est une relation d'équivalence*
• *Deux matrices sont congruentes si et seulement si elles représentent la même forme bilinéaire dans deux bases différentes*
• *Deux matrices congruentes ont même rang*

⚠ Deux matrices congruentes n'ont pas nécéssairement même déterminant.

28.4.2 Formes quadratiques

Le cas général

Définition 1037 *On appelle* **forme quadratique associée à la forme bilinéaire** *ϕ l'application $x \mapsto \phi(x, x)$.*
Une application de E^2 dans $\mathbb{K}$ est une **forme quadratique** *sur E si et seulement si c'est la forme quadratique associée à une certaine forme bilinéaire.*

Proposition 1038 *Soit q une forme quadratique, alors*

$$q(x + y) + q(x - y) = 2(q(x) + q(y))$$

(voir figure 28.1)

Démonstration : Il suffit de développer la formule en considérant ϕ à laquelle est associé q.□

Proposition 1039 *L'application qui à une forme bilinéaire associe la forme quadratique qui lui est associée est une application linéaire de $\mathcal{L}_2(E)$ dans l'ensemble des fonctions de E dans $\mathbb{K}$. Son noyau est l'ensemble des applications bilinéaires antisymétriques, et elle induit un isomorphisme de l'ensemble des applications bilinéaires symétriques sur E sur l'ensemble des formes quadratiques.*

Démonstration : • Le fait que cette application soit linéaire est évident.
• Montrons que si sa forme quadratique associée est nulle, alors ϕ est antisymétrique. Pour tout x et tout y $\phi(x + y, x + y) = \phi(x, x) + \phi(y, y) + \phi(x, y) + \phi(y, x)$; donc si pour tout z $\phi(z, z) = 0$, alors $\phi(x, y) + \phi(y, x) = 0$. D'où le résultat.□

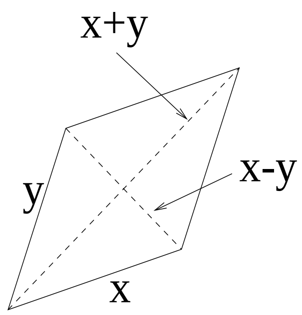


FIG. 28.1 – Formule du parallélogramme. La somme des carrés des diagonales est égale à la somme des carrés des côtés.

Définition 1040 *Etant donnée q une forme quadratique q sur E, l'unique forme bilinéaire symétrique ϕ telle que pour tout x $q(x) = \phi(x,x)$i est appelée* **forme polaire de** q.

Proposition 1041 *Les formules suivantes permettent de déterminer la forme polaire ϕ associée à une forme quadratique q :*
• $\phi(x,y) = \frac{1}{2}(q(x+y) - q(x) - q(y))$
• $\phi(x,y) = \frac{1}{4}(q(x+y) - q(x-y))$

Définition 1042 (Orthogonalité) *Etant données q une forme quadratique et ϕ sa forme polaire :*
• *x et y appartenant à E sont* **orthogonaux** *si et seulement si $\phi(x,y) = 0$*
• *deux parties X et Y de E sont dites* **orthogonales** *si et seulement si tout x dans X et tout y dans Y sont orthogonaux.*
• *On appelle* **orthogonal** *d'une partie X de E et on note $X^\perp$ l'ensemble des éléments orthogonaux à tous les éléments de X.*
• *On appelle* **noyau** *de q l'orthogonal de E (à ne pas confondre avec le cône isotrope de q) ; on le note $N(q)$.*
• *On appelle* **cône isotrope de** *q et on note $C(q)$ l'ensemble des x tels que $q(x) = 0$ (à ne pas confondre avec le noyau de q). Un élément du cône isotrope est appelé* **vecteur isotrope**.
• *Un sous-espace vectoriel de E est dit* **isotrope** *si la restriction de q à ce sous-espace vectoriel est dégénérée.*
• *Un sous-espace vectoriel de E est dit* **totalement isotrope** *si la restriction de q à ce sous-espace vectoriel est nulle.*
• *Une forme quadratique est dite* **dégénérée** *si son noyau n'est pas réduit à $\{0\}$.*
• *Une forme quadratique est dite* **définie** *si son cône isotrope est réduit à $\{0\}$.*
• *Une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E est dite* **positive** *(resp.* **négative***) lorsque pour tout x on a $q(x,x) \geq 0$ (resp. $q(x,x) \leq 0$).*

Proposition 1043 • *Un sous-espace vectoriel est isotrope si et seulement si il a une intersection non réduite à $\{0\}$ avec son orthogonal.*
• *Un sous-espace vectoriel est totalement isotrope si et seulement si il est inclus dans son orthogonal.*
• *L'orthogonal d'une partie de E est l'orthogonal du sous-espace vectoriel engendré par cette partie.*
• *Avec X et Y des parties de E, $X \subset Vect(Y) \rightarrow Y^\perp \subset X^\perp$*

espace vectoriel (il contient le sous-espace vectoriel noyau de q).

⚠ La restriction d'une forme quadratique non-dégénérée à un sous-espace vectoriel n'est pas nécéssairement non-dégénérée.

Proposition 1045 *Le cône isotrope est un cône, c'est à dire que x isotrope $\rightarrow$ $\lambda.x$ isotrope pour tout λ dans $\mathbb{K}$.*

Définition 1046 (Familles orthogonales et orthonormales) *Une famille (x_i) de vecteurs de E est dite* **orthogonale** *si $i \neq j$ implique $\phi(x_i, x_j) = 0$.*
Une famille (x_i) de vecteurs de E un $\mathbb{C}$-espace vectoriel est dite **réduite** *si elle est orthogonale et si $\phi(x_i, x_i) = \chi_{1,rg(q)}(i)$.*
Une famille (x_i) de vecteurs de E un $\mathbb{R}$-espace vectoriel est dite **réduite** *si elle est orthogonale et si $\phi(x_i, x_i) = \chi_{1,p}(i) - \chi_{p+1,rg(q)}(i)$ pour un certain p dans $[0, rg(q)]$.*
Une famille (x_i) de vecteurs de E est dite **orthonormale** *si $\phi(x_i, x_j) = \delta_{i,j}$.*
Une matrice réelle ou complexe de type (n, n) est dite **orthogonale** *si la famille de ses vecteurs colonnes forme une famille orthonormale de $\mathbb{K}^n$.*

Proposition 1047 • *Une famille orthogonale sans vecteur isotrope est libre. En particulier si q est définie une famille orthogonale de vecteurs non nuls est libre.*
• *Une famille orthonormale est libre.*

Démonstration : $\phi(\sum_i \lambda_i.x_i, x_i) = \lambda_i.\phi(x_i, x_i)$, etc... □

◻ Le cas de la dimension finie - expression matricielle

On suppose que $B = (e_1, ..., e_n)$ est une base de E.

Définition 1048 *On appelle* **matrice d'une forme quadratique dans une base** *B la matrice de sa forme polaire dans la base B.*
On note $Mat_B(q)$ la matrice de la forme quadratique q dans la base B.
On appelle **rang** *de q la rang de sa matrice dans une base quelconque (le rang est indépendant de la base).*
On appelle **discriminant** *d'une forme quadratique q dans une base B le déterminant de la matrice de q dans la base B.*

⚠ Le discriminant dépend de la base.

Le troisième point appelle une preuve, que voici ci-dessous :

Proposition 1049 $Mat_{B'}(q) = ^t P_{B,B'}.Mat_B(q).P_{B,B'}$

Démonstration : Il suffit d'aller voir la démonstration équivalente pour les formes bilinéaires, en 28.4.1.□

Deux matrices congruentes ayant même rang, la définition ci-dessus est donc cohérente.

Proposition 1050 *Etant donnés x dans E et X le vecteur défini par $X_i = e_i^*(x)$, on a $q(x) = ^t X.Mat_B(q).X$.*

Démonstration : Il suffit de considérer la forme polaire de q et de consulter la partie 28.4.1.□

Proposition 1051 *Un polynôme de degré ≤ 2 en $x_1, ..., x_n$ est une forme quadratique sur $\mathbb{K}^n$.*
Pour obtenir sa forme polaire, on remplace chaque $x_i.x_i$ par $x_i.y_i$, et chaque $x_i.x_j$ par $\frac{1}{2}(x_i.y_j + x_j.y_i)$; le polynôme en (x_i) et (y_i) est une forme bilinéaire symétrique sur $\mathbb{K}^n$, et est la forme polaire du polynôme.

Proposition 1052 *Le noyau d'une forme quadratique en dimension finie est le noyau de l'endomorphisme ayant même matrice (dans la même base).*
La dimension de E est la somme du rang de q et de la dimension du noyau de q.

Démonstration : Evident, il n'y a qu'à l'écrire.□

Proposition 1053 • *Une base est orthogonale si et seulement si la matrice de q dans cette base est diagonale.*
• *Une base est orthonormale si et seulement si la matrice de q dans cette base est l'identité.*

Démonstration : Evident !□

Le théorème suivant est très important :

Théorème 1054 *Pour toute forme quadratique sur E de dimension finie, il existe une base de E orthogonale pour q.*

Démonstration : On montre ce résultat par récurrence :
• Le cas $n = 1$ est trivial.
• Si q est nulle, le résultat est clair ; sinon on choisit e_1 non isotrope, et on note H l'orthogonal de e_1. Tout vecteur x s'écrit

$$x = \underbrace{\in \mathbb{K}.e_1}\frac{\phi(e_1, x)}{q(e_1)}.e_1 + \underbrace{\in e_1^\perp} x - \frac{\phi(e_1, x)}{q(e_1)}.e_1$$

donc $E = \mathbb{K}.e_1 \oplus H$. Il suffit alors d'appliquer l'hypothèse de récurrence sur H.□

Corollaire 1055 *Pour tout forme quadratique q sur E $\mathbb{K}$-espace vectoriel de dimension n, il existe $p \in [1, n]$ et $\lambda_1, ..., \lambda_p$ dans $\mathbb{K} \setminus \{0\}$ et f_1,...,f_p dans E^* tel que*
• les f_i forment une famille libre
• $q(x) = \sum_{i=1}^p \lambda_i.f_i(x)^2$
En outre, p est unique et est égal au rang de q.

Démonstration : Il s'agit simplement de la traduction du théorème précédent.□

On en déduit les deux corollaires suivants, l'un dans le cas $\mathbb{K} = \mathbb{C}$, l'autre dans le cas $\mathbb{K} = \mathbb{R}$:

Corollaire 1056 (Cas $\mathbb{K} = \mathbb{C}$) *Pour tout forme quadratique q sur E $\mathbb{K}$-espace vectoriel de dimension n, il existe $p \in [1, n]$ et f_1,...,f_p dans E^* tel que*
• les f_i forment une famille libre
• $q(x) = \sum_{i=1}^p f_i(x)^2$
En outre, p est unique et est égal au rang de q.

Démonstration : Il suffit de voir dans le corollaire précédent que l'on peut multiplier f_i par une racine carrée de $\frac{1}{\lambda_i}$.□

Corollaire 1057 (Cas $\mathbb{K} = \mathbb{R}$) *Pour tout forme quadratique q sur E $\mathbb{K}$-espace vectoriel de dimension n, il existe $r \in [1, n]$ et p dans $[1, r]$ et f_1,...,f_p dans E^* tel que*
• les f_i forment une famille libre
• $q(x) = \sum_{i=1}^p f_i(x)^2 - \sum_{i=p+1}^r f_i(x)^2$
En outre, r est unique et est égal au rang de q, et p est unique.

Démonstration : Il suffit de voir que l'on peut multiplier f_i par une racine carrée de $\frac{1}{|\lambda_i|}$; il ne reste alors plus qu'à montrer l'unicité de p. Pour cela, on considère s le p maximal possible, et u le p minimal possible ; on note $t = q - u$.
Alors • q est définie positive sur un espace de dimension s
• q est définie négative sur un espace de dimension t

On déduit facilement que :
• ces deux espaces sont en somme directe, donc $s + t \leq r$
• $p \leq s$
• $r - p \leq t$
Donc si $p < s$, $p+r-p < r$, et si $r-p < t$, $p+r-p < r$, donc nécéssairement, $p = s$. □

Encore un corollaire dans le cadre d'une forme quadratique définie positive :

Corollaire 1058 (Cas $\mathbb{K} = \mathbb{R}$ et q définie positive) *Pour tout forme quadratique q sur E $\mathbb{K}$-espace vectoriel de dimension n, il existe f_1,...,f_n dans E^* tel que*
• *les f_i forment une famille libre*
• $q(x) = \sum_{i=1}^{n} f_i(x)^2$
On en déduit aussi l'existence d'une base orthonormale.

28.4.3 Formes quadratiques réelles

Pour toute la durée de cette section, on se place dans le cadre de E un $\mathbb{R}$-espace vectoriel .

Le cas général

On rappelle la définition suivante :

Définition 1059 *Une forme quadratique q sur un $\mathbb{R}$-espace vectoriel E est dite* **positive** *(resp.* **négative***) lorsque pour tout x on a $q(x,x) \geq 0$ (resp. $q(x,x) \leq 0$).*

Théorème 1060 (Inégalité de Schwartz) • *Soit q une forme quadratique positive sur un $\mathbb{R}$-espace vectoriel E, et soit ϕ sa forme polaire. Alors pour tout x et tout y dans E*

$$\phi(x,y)^2 \leq q(x).q(y)$$

• *Soit q une forme quadratique définie positive sur un $\mathbb{R}$-espace vectoriel E, et soit ϕ sa forme polaire. Alors pour tout x et tout y dans E*

$$\phi(x,y)^2 \leq q(x).q(y)$$

et

$$\phi(x,y)^2 = q(x).q(y) \Longrightarrow (x,y) \text{ est une famille liée.}$$

Démonstration : • Il suffit de considérer le discriminant du polynôme en t $q(x.t + y)$ (ce polynôme est toujours positif puisque la forme quadratique est positive).
• L'égalité implique que $q(y - \frac{\phi(x,y)}{q(x)}.x) = 0$.□

On remarque que l'inégalité de Schwartz implique qu'une forme bilinéaire symétrique positive est continue.

Corollaire 1061 (Noyau et cône isotrope d'une forme quadratique positive)
• *Si q est positive alors $N(q) = C(q)$.*
• *Si q est positive alors q est définie si et seulement si elle est non-dégénérée.*

Proposition 1062 *Une forme quadratique q sur un $\mathbb{R}$-espace vectoriel qui est définie est nécéssairement soit positive soit négative.*

Démonstration : On suppose qu'il existe x et y avec $q(x) > 0$ et $q(y) < 0$. Alors l'application qui à t dans $[0,1]$ associe $q(t.x + (1-t).y)$ est continue (il suffit de développer pour le voir), donc par le théorème des valeurs intermédiaires 209 elle s'annule en un certain t_0. Nécéssairement $t.x + (1-t)y = 0$ (puisque q est définie), donc x et y sont linéairement dépendants ($t \neq 0$ et $t \neq 1$). Donc $q(x)$ et $q(y)$ sont de même signe, d'où contradiction.$\square$

Une autre formulation des inégalités de Schwartz est donnée dans le corollaire ci-dessous :

Corollaire 1063 (Inégalités de Minkowski) • *Soit q une forme quadratique positive sur un $\mathbb{R}$-espace vectoriel E. Alors pour tout x et tout y dans E*

$$\sqrt{q(x+y)} \leq \sqrt{q(x)} + \sqrt{q(y)}$$

• *Soit q une forme quadratique définie positive sur un $\mathbb{R}$-espace vectoriel E, alors pour tout x et tout y dans E*

$$\sqrt{q(x+y)} = \sqrt{q(x)} + \sqrt{q(y)} \Longrightarrow (x,y) \text{ est une famille positivement liée.}$$

Démonstration : • Par l'inégalité de Schwartz

$$\phi(x,y)^2 \leq q(x).q(y)$$

et donc

$$\rightarrow \frac{q(x+y) - q(x) - q(y)}{2} \leq \sqrt{q(x).q(y)}$$

$$\rightarrow q(x+y) \leq q(x) + q(y) + 2.\sqrt{q(x).q(y)}$$

$$\rightarrow \sqrt{q(x+y)} \leq \sqrt{q(x)} + \sqrt{q(y)}$$

• Même principe.$\square$

Proposition 1064 • *Une forme quadratique q est positive si et seulement si $-q$ est négative*
• *Une forme quadratique sur un $\mathbb{R}$-espace vectoriel est convexe si et seulement si elle est positive*
• *Une forme quadratique sur un $\mathbb{R}$-espace vectoriel est concave si et seulement si elle est négative*

⊡ Le cas de la dimension finie - expression matricielle

On suppose maintenant que l'on travaille sur E un $\mathbb{R}$-espace vectoriel de dimension finie n. q est une forme quadratique.

Proposition 1065 (Propriété fondamentale des formes quadratiques définies positives) *Si E est une forme quadratique définie positive sur un $\mathbb{R}$-espace vectoriel E de dimension finie, alors il existe une base de E orthonormale pour q.*

Démonstration : Il suffit de considérer le corollaire 1058.□

Proposition 1066 (Quelques propriétés sur l'orthogonalité sur un $\mathbb{R}$-espace vectoriel de dimension finie)
• *Pour F sous-espace vectoriel de E, on a $dim\ F + dim\ F^{\perp} \geq n$*
• *Soit F sous-espace vectoriel de E, avec $q_{|F}$ définie, alors $E = F \oplus F^{\perp}$.*

Démonstration : • Soit $(f_i)_{i\in[1,f]}$ une base de F, complétée en $(f_i)_{i\in E}$ une base de E.
Soit p l'application de E dans E définie par $p(x) = \sum_{i\in[1,f]} \phi(x, f_i).f_i$. Cette application est linéaire ; on peut donc écrire

$$dim\ E = rg(p) + dim\ Ker\ p$$

Or

$$rg(p) \leq dim\ F$$

et

$$dim\ Ker\ p = dim\ F^{\perp}$$

donc

$$dim\ E \leq dim\ F + dim\ Ker\ p$$

• $F \cap F^{\perp} = \emptyset$ car q est définie sur F. L'inégalité précédente donne $dim\ F + dim\ F^{\perp} \geq n$, d'où le résultat.□

Définition 1067 (Signature d'une forme quadratique sur un $\mathbb{R}$-espace vectoriel de dimension finie) *On appelle* **signature d'une forme quadratique** *q le couple (s, t) avec s la dimension maximale d'un sous-espace vectoriel de E sur lequel q est définie positive et t la dimension maximale d'un sous-espace vectoriel de E sur lequel q est définie négative.*

Le théorème suivant, très important, permet de cerner l'intérêt de la notion.

Théorème 1068 (Théorème d'inertie de Sylvester) *Pour toute base q-orthogonale e_i, l'ensemble des i tels que $q(e_i) > 0$ a même cardinal, l'ensemble des i tels que $q(e_i) = 0$ a même cardinal, l'ensemble des i tels que $q(e_i) > 0$ a même cardinal.*
Le sous-espace vectoriel engendré par l'ensemble des i tels que $q(e_i) > 0$ est un sous-espace vectoriel F de dimension maximale tel que $q_{|F}$ soit définie positive.
Le sous-espace vectoriel engendré par l'ensemble des i tels que $q(e_i) < 0$ est un sous-espace vectoriel F de dimension maximale tel que $q_{|F}$ soit définie négative.
L'ensemble des i tels que $q(e_i) = 0$ est égal à la dimension de E moins le rang de q.

Démonstration : Il suffit de consulter la preuve du corollaire 1057.□

Corollaire 1069 • *Toute matrice symétrique est congruente à une matrice diagonale dont les termes diagonaux sont $(1, ..., 1, -1, ... -1, 0, ...0)$.*
• *Deux formes quadratiques ont la même signature si et seulement si on passe de l'un à l'autre en composant par un automorphisme.*

Proposition 1070 *Une matrice M de type (n, n) est antisymétrique si et seulement si pour tout vecteur X de $\mathbb{K}^n$ on a ${}^tX.M.X = 0$.*

Démonstration : La forme polaire de la forme quadratique associée à M est $(X, Y) \mapsto {}^t X.(\frac{1}{2}(M +^t M)).Y$. Elle est nulle si et seulement si M est antisymétrique (voir proposition 1039).□

⊡ Le cas d'un espace euclidien

Pour plus d'informations sur les espaces euclidiens on consultera la partie 29.5. Un espace euclidien étant réel et de dimension finie, ce qui vient d'être dit est encore valable.

On va noter $Q(E)$ l'espace des formes quadratiques. Les notations usuelles seront utilisées :

• $(e_1, ..., e_n)$ est une base de E
• $q \in Q(E)$
• M la matrice (symétrique) associée à q pour la base des e_i
• ϕ la forme polaire de q (symétrique, de matrice M dans la base des e_i)
• X désigne le vecteur colonne des coordonnées de x dans la base des e_i
• Y désigne le vecteur colonne des coordonnées de y dans la base des e_i

◇ **Formulaire**

$$q(x_1.e_1, x_2.e_2, ..., x_n.e_n) = \sum_{(i,j)\in[1,n]^2} M_{i,j} x_i.x_j$$

$$\phi((x_1.e_1, x_2.e_2, ..., x_n.e_n), (y_1.e_1, y_2.e_2, ..., y_n.e_n)) = \sum_{(i,j)\in[1,n]^2} M_{i,j} x_i.y_j$$

$$M_{i,j} = \phi(e_i, e_j)$$

$$q(e_i) = M_{i,i}$$

$$q(x) =^t X.M.X$$

$$\phi(x, y) =^t X.M.X$$

et avec $(f_1, ..., f_n)$ une autre base,

$$Mat_{(f_i)}(\phi) = Mat_{(f_i)}(q) =^t P_{(e_i),(f_i)}.M.P_{(e_i),(f_i)}$$

◇ **Résultats divers**

Théorème 1071 *L'application F de $L(E)$ dans l'ensemble des applications de E dans $\mathbb{R}$ définie par $F(f) = (y \mapsto < f(y)|y >$ induit un isomorphisme de $S(E)$ (ensemble des endomorphismes symétriques de E) sur $Q(E)$ (ensemble des formes quadratiques sur E).*

Démonstration :

• $F(f)$ appartient à $Q(E)$ pour tout f dans $L(E)$ se voit en considérant la forme bilinéaire $\phi = (x, y) \mapsto \frac{1}{2}(< f(x)|y > + < x|f(y) >)$.

• $< f(x)|x >= 0$ pour tout $x \iff f$ antisymétrique (si vous n'en êtes pas convaincu, revoyez la partie 29.5.2).

• Avec n la dimension de E, on sait alors que l'image de F est de dimension $n^2 - \frac{1}{2}n.(n-1) = \frac{1}{2}n.(n+1) = dim\ Q(E)$, donc F a bien pour image $Q(E)$; $S(E)$ étant un supplémentaire du noyau de F, F induit bien un isomorphisme de $S(E)$ sur $Q(E)$.$\square$

Corollaire 1072 • *Toute forme quadratique q s'écrit $x \mapsto < f(x)|x >$ pour un certain endomorphisme symétrique f (on peut d'ailleurs aussi écrire $x \mapsto < x|f(x) >$, puisque f est symétrique).*
• *Dans la même base, q, ϕ et f ont même matrice.*

Définition 1073 *f (défini comme précédemment) est appelé* **endomorphisme symétrique associé à la forme quadratique** q.

Ceci nous permet de donner quelques résultats, conséquences immédiates de résultats connus sur les endomorphismes symétriques :

Théorème 1074 • *Soit q une forme quadratique sur E euclidien ; alors il existe une base orthonormale de E dans laquelle la matrice de l'endomorphisme associée à q est diagonale ; c'est à dire que cette base est orthogonale pour q aussi.*
• *Si on a deux formes quadratiques sur un $\mathbb{R}$-espace vectoriel E de dimension finie dont l'une (au moins) est définie, alors il existe une base orthogonale pour les deux formes quadratiques (il suffit de considérer l'espace euclidien engendré par la forme définie (ou sa négation si elle est négative) pour conclure).*
• *Une forme quadratique sur E euclidien est positive si et seulement si toutes les valeurs propres de l'endomorphisme symétrique associé sont positives*
• *Une forme quadratique sur E euclidien est définie si et seulement si toutes les valeurs propres de l'endomorphisme symétrique associé sont non nulles et de même signe*
• *Une forme quadratique sur E euclidien est négative si et seulement si toutes les valeurs propres de l'endomorphisme symétrique associé sont négatives*

28.4.4 Formes quadratiques complexes

⊡ **Le cas général**

Le cadre le plus général est simplement celui d'un $\mathbb{C}$-espace vectoriel .

Définition 1075 *On appelle* **forme quadratique hermitienne** *sur un $\mathbb{C}$-espace vectoriel E une application q de E dans $\mathbb{C}$ telle qu'il existe une forme sesquilinéaire hermitienne ϕ telle que pour tout x on ait $q(x) = \phi(x, x)$. Cette forme sesquilinéaire hermitienne est unique (à vérifier plus bas) ; on l'appelle* **forme polaire** *de q.*

Proposition 1076 *Soit q une forme quadratique hermitienne. Alors :*

$$\forall x q(x) \in \mathbb{R}$$

en effet $\phi(x, x) = \overline{\phi(x, x)}$ car ϕ est hermitienne

$$\forall (x, y) \in E^2 \ \phi(x, y) = \frac{1}{4}(\sum_{j=1}^{4} \overline{i^j}.q(x + i^j.y))$$

cela se montre simplement en développant chacun des 4 termes de droite
L'ensemble des formes quadratiques hermitiennes sur E noté $QH(E)$ est un $\mathbb{R}$-espace vectoriel .

Ce n'est pas un $\mathbb{C}$-espace vectoriel , comme on s'en convainc facilement en considérant une forme quadratique hermitienne non nulle multipliée par i...

On note au passage que le deuxième résultat de cette proposition donne l'unicité recquise dans la définition de la forme polaire ci-dessus.

⊡ **Le cas de la dimension finie - expression matricielle**

Définition 1077 *Etant donnée $(e_1, ..., e_n)$ une base de E espace hermitien et q une forme quadratique hermitienne sur E de forme polaire ϕ, on définit la* **matrice** M **associée à** q *ou* **matrice associée à** ϕ *par $M_{i,j} = \phi(e_i, e_j)$. On note $M = Mat_{(e_i)}(\phi)$ ou $M = Mat_{(e_i)}(q)$.*

Proposition 1078 *La matrice M associée à une forme quadratique hermitienne est hermitienne c'est-à-dire que $M =^t \overline{M}$.*
Avec X le vecteur colonne des coordonnées de x dans une base donnée, Y le vecteur colonne des coordonnées de y dans la même base, M la matrice associée à q ou ϕ dans la même base, on a

$$\phi(x,y) =^t \overline{X}.M.Y$$

$$q(x) =^t \overline{X}.M.X = \sum_{i=1}^{n} M_{i,i}.|X_i|^2 + 2.Re(\sum_{i<j} M_{i,j}\overline{X}_i.X_j)$$

Si $(e_i)_{i\in[1,n]}$ et $(f_i)_{i\in[1,n]}$ sont deux bases de E, alors $Mat_{(e_i)}(q) =^t \overline{P_{(e_i),(f_i)}}.Mat_{(f_i)}(q).P_{(e_i),(f_i)}$.

⊡ Formes quadratiques sur un espace hermitien

Définition 1079 *On note $QH(E)$ le $\mathbb{R}$-espace vectoriel des formes quadratiques hermitienne sur E espace hermitien.*
On note $H(E)$ le $\mathbb{R}$-espace vectoriel des endomorphismes hermitiens de E, espace hermitien.
Etant donné $f \in H(E)$, la forme quadratique $x \mapsto < f(x)|x >$ est appelée **forme quadratique hermitienne associée à l'endomorphisme hermitien** *f ; réciproquement f est appelée* **endomorphisme hermitien associée à cette forme quadratique** *(voir unicité ci-dessous).*

Théorème 1080 *L'application F de $H(E)$ dans $QH(E)$ défini par $F(f) = (x \mapsto < f(x)|x >)$ est un isomorphisme.*

Démonstration : • Le fait que pour tout f dans $H(E)$ $F(f)$ soit une forme quadratique est clair (considérer la forme sesquilinéaire $(x,y) \mapsto < f(x)|y >$).
• Le fait que f est un morphisme est facile à prouver.
• La surjectivité :
Il suffit, étant donné une forme quadratique, de considérer sa matrice dans une base orthonormale quelconque, et l'endomorphisme associé à la même matrice dans la même base ; l'image de cet endomorphisme par f.
• L'injectivité :
Supposons $F(f) = 0$. L'application $(x,y) \mapsto < f(x)|y >$ est la forme polaire de $F(f)$ (elle est bien sesquilinéaire et hermitienne) ; donc cette forme sesquilinéaire est nulle par unicité de la forme polaire. Donc $< f(x)|y >$ est nul pour tout x et tout y, d'où le

résultat en spécialisant par $y = f(x)$.$\square$

Théorème 1081 *Pour toute forme quadratique hermitienne il existe une base orthonormale (pour le produit scalaire hermitien) qui est orthogonale pour cette forme quadratique.*

Démonstration : Il suffit de considérer la proposition 1140 appliqué à l'endomorphisme associé à une forme quadratique.$\square$

Quelques liens entre une forme quadratique q et l'endomorphisme hermitien associé f :

• q est définie si et seulement si toutes les valeurs propres de f sont de même signe et non nulles

• q est positive si et seulement si toutes les valeurs propres de f sont positives

• q est négative si et seulement si toutes les valeurs propres de f sont négatives

28.5 Zoologie des déterminants

28.5.1 Déterminant d'ordre 2

$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a.d - b.c$

28.5.2 Déterminant d'ordre 3

$$\begin{vmatrix} M_{1,1} & M_{1,2} & M_{1,3} \\ M_{2,1} & M_{2,2} & M_{2,3} \\ M_{3,1} & M_{3,2} & M_{3,3} \end{vmatrix}$$

$$= M_{1,1}.M_{2,2}.M_{3,3} + M_{1,2}.M_{2,3}.M_{3,1} + M_{2,1}.M_{3,2}.M_{2,3}$$

$$-M_{3,1}.M_{2,2}.M_{1,3} - M_{2,1}.M_{1,2}.M_{3,3} - M_{3,2}.M_{2,3}.M_{1,1}$$

Une façon usuelle de retenir ce résultat pas très élégant vu comme ça est le schéma 28.2.

On suit les flèches marquées d'un $+$ pour retrouver les produits à compter positivement, et les flèches marquées d'un $-$ pour retruver les produits à compter négativement.

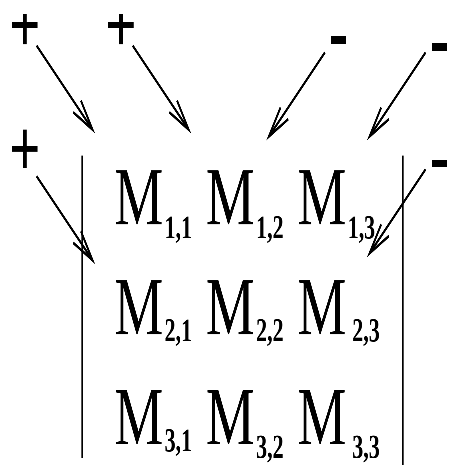


FIG. 28.2 – La règle de Sarrus

28.5.3 Déterminant de Vandermonde

Définition 1082 *On appelle* **déterminant de Vandermonde associé à un n-uple** $(x_1, ..., x_n)$ *le déterminant*

$$\begin{vmatrix} x_1^0 & x_1^1 & x_1^2 & x_1^3 & \dots & x_1^n \\ x_2^0 & x_2^1 & x_2^2 & x_2^3 & \dots & x_2^n \\ x_3^0 & x_3^1 & x_3^2 & x_3^3 & \dots & x_3^n \\ x_4^0 & x_4^1 & x_4^2 & x_4^3 & \dots & x_4^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^0 & x_n^1 & x_n^2 & x_n^3 & \dots & x_n^n \end{vmatrix}$$

Proposition 1083 (Déterminant de Vandermonde) *Le déterminant de Vandermonde associé à* $(x_1, ..., x_n)$ *est*

$$\Pi_{i<j} x_j - x_i$$

<u>Démonstration :</u>

On note W_n le déterminant

$$\begin{vmatrix} x_1^0 & x_1^1 & x_1^2 & x_1^3 & \dots & x_1^n \\ x_2^0 & x_2^1 & x_2^2 & x_2^3 & \dots & x_2^n \\ x_3^0 & x_3^1 & x_3^2 & x_3^3 & \dots & x_3^n \\ x_4^0 & x_4^1 & x_4^2 & x_4^3 & \dots & x_4^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^0 & x_n^1 & x_n^2 & x_n^3 & \dots & x_n^n \end{vmatrix}$$

On note P_n le polynôme $\Pi_{i\in[1,n-1]}(X-x_i)$. On a

$$P = X^{n-1} + \sum_{k=1}^{n-1} p_k.X^{k-1}$$

On ajoute alors à la dernière colonne la somme des $p_k.c_k$ pour $k \in [1, n-1]$ avec c_k la colonne k. Sur la dernière colonne, on a maintenant seulement des 0, sauf pour la dernière ligne où l'on a $P(x_n)$.

Par récurrence, on en déduit le résultat annoncé.□

28.5.4 Déterminant d'une matrice de permutation

Définition 1084 (Matrice d'une permutation) *On appelle* **matrice de la permutation** $\sigma \in \sigma_n$ *la matrice M de type (n, n) définie par $M_{i,j} = \delta_{i,\sigma(j)}$.*

Proposition 1085 *Le déterminant de la matrice de la permutation σ est égal à la signature de la permutation $\epsilon(\sigma)$.*

Démonstration : Il suffit de revenir à la définition du déterminant d'une famille de vecteurs dans une base, et de voir qu'il n'y a qu'une permutation qui n'annule pas le produit correspondant dans la formule.□

28.5.5 Déterminant circulant droit

Définition 1086 *On appelle* **matrice circulante associé au n-uple** $(x_1, ..., x_n)$ *la matrice M définie par $M_{i,j} = x_{j-i\,(\text{modulo } n)}$, c'est à dire*

$$\begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ x_n & x_1 & x_2 & \ddots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \ddots & x_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \dots & x_1 \end{pmatrix}$$

On trouvera la définition d'une matrice circulante droite en 30.7.

Proposition 1087 $C_n = \Pi_{i=1..n}P(y_i)$ *avec* $P(X) = \sum_{i=0}^{n-1} x_i.X^{i-1}$ *et* $y_i = e^{\frac{2.i.\Pi}{n}}$.

Démonstration : On note Y_i le vecteur $(y_i^0, ..., y_i^{n-1})$.
On constate que $M.Y_i = P(y_i).Y_i$.
On note Y la matrice dont les vecteurs colonnes sont $Y_0, ..., Y_{n-1}$.
On a alors $M.Y = M.diag(P(y_0), ..., P(y_{n-1}))$.
Donc comme le déterminant de M est non nul (voir 28.5.3), le déterminant de Y est $\Pi_{i=1..n} P(y_i)$.□

Quand même, il fallait y penser, à multiplier par la transposée de la matrice de Vandermonde associée aux racines n-ièmes de l'unité...

On définit de même les matrices circulantes gauche, dont on calcule le déterminant en utilisant une permutation bien choisie sur les lignes...

28.5.6 Déterminant de $M_{i,j} = inf\{i, j\}$

Le déterminant est le suivant :

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix}$$

Puisqu'on peut à volonté sans changer le déterminant ajouter à une colonne une combinaison linéaire des autres colonnes, on peut en particulier soustraire à une colonne la colonne précédente ; on obtient alors le déterminant plus facile :

$$\begin{vmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & \dots & \dots & \dots & \dots & 1 \end{vmatrix}$$

Ce déterminant est donc égal à 1.

28.6 Zoologie de l'algèbre bilinéaire

28.6.1 Procédé d'orthogonalisation de Gauss

Théorème 1088 (Orthogonalisation de Gauss) *Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n, et q une forme quadratique sur E. Alors, avec r le rang de q, il existe r formes linéaires indépendantes sur E, f_1,...,f_r tels que $q = \sum_{i=1}^r f_i^2$.*

Démonstration : Il s'agit simplement d'opérations sur les lignes et les colonnes ; voir la méthode de Gauss, 1196.□

Chapitre 29

Espaces préhilbertiens et espaces de Hilbert

29.1 Espaces préhilbertiens réels

Il est recommandé de lire au préalable la partie 28.4 et la partie 5.1.2.

Définition 1089 *Etant donné E un $\mathbb{R}$-espace vectoriel , on appelle* **produit scalaire euclidien sur** E *une application $< .|. >$ de E^2 dans $\mathbb{R}$ telle que :*
• $< .|. >$ *est bilinéaire*
• $< x|y >=< y|x >$ *pour tout x et tout y*
• $\forall x \neq 0 < x|x >> 0$ *(i.e. la forme quadratique associée à ϕ est une forme bilinéaire définie positive)*

On appelle **espace préhilbertien réel** *un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire euclidien.*
Un sous-espace vectoriel F d'un espace préhilbertien réel E muni d'un produit scalaire euclidien, muni de la restriction du produit scalaire euclidien à F, est appelée **sous-espace préhilbertien** *de E (c'est un espace préhilbertien).*

Etant donné un produit scalaire euclidien $< .|. >$, on définit une **norme euclidienne** *; il s'agit de l'application $x \mapsto \|x\| = \sqrt{< x|x >}$. On verra plus loin qu'il s'agit d'une norme (facile au vu des résultats de la partie 28.4).*

⚠ On n'a à aucun moment imposé que la dimension soit finie.

Un produit scalaire euclidien sur un $\mathbb{R}$-espace vectoriel est donc une forme bilinéaire symétrique associée à une forme quadratique définie positive.

Exemples :

• Le **produit scalaire euclidien canonique** sur $\mathbb{R}^n$ est défini par

$$(x|y) = \sum_{i\in[1,n]} x_i.y_i.$$

• Le **produit scalaire euclidien canonique** sur le sous-ensemble de $\mathbb{R}^{\mathbb{N}}$ des suites sommables (i.e. des $(u_n)_{n\in\mathbb{N}}$ telles que $\sum_{n\in\mathbb{N}} |u_n|$ converge) est défini par

$$< (u_n)_{n\in\mathbb{N}}|(v_n)_{n\in\mathbb{N}} >= \sum_{n\in\mathbb{N}} u_n.v_n.$$

Il est important de rappeler le théorème 1060 et le corollaire 1063. Ils stipulent que :
• $< x|y >^2 \leq < x|x > . < y|y >$ (inégalité de Schwartz)
• $| < x|y > | \leq \|x\|.\|y\|$ (inégalité de Schwartz, en passant à la racine)
• $\|x+y\| \leq \|x\| + \|y\|$ (inégalité de Minkowski = inégalité triangulaire)
• Le produit scalaire est continu (conséquence de Schwartz)

La notation $< x|y >$ peut être remplacée par $(x|y)$, $< x,y >$, (x,y), ou même $x.y$.

29.2 Espaces préhilbertiens complexes

Définition 1090 *Une application d'un $\mathbb{C}$-espace vectoriel E dans un $\mathbb{C}$-espace vectoriel F est dite* **semi-linéaire** *si*
• $\forall (x,y) \in E^2$ $f(x+y) = f(x) + f(y)$
• $\forall (x,\lambda) \in E \times \mathbb{C}$ $f(\lambda.x) = \overline{\lambda} f(x)$
Une application semi-linéaire est un **semi-isomorphisme** *si et seulement si elle est semi-linéaire et bijective.*
Une **forme semi-linéaire** *est une application semi-linéaire d'un $\mathbb{C}$-espace vectoriel dans $\mathbb{C}$.*
Etant donnés E et F des $\mathbb{C}$-espace vectoriel une application ϕ de $E \times F$ est dite **forme sesquilinéaire** *sur $E \times F$ si*
• *$\forall x$ l'application $y \mapsto \phi(x,y)$ est une forme linéaire sur F*
• *$\forall y$ l'application $x \mapsto \phi(x,y)$ est une forme semi-linéaire sur E*
Une forme sesquilinéaire sur $E \times E$ est dite **hermitienne** *lorsque en outre $\forall (x,y) \in E^2$ $\phi(x,y) = \overline{\phi(y,x)}$.*
Une forme sesquilinéaire hermitienne ϕ sur E^2 est dite **produit scalaire hermitien sur** *E si $\forall x \in E \setminus \{0\}$ $\phi(x,x) \in \mathbb{R}^+_*$. On note généralement alors $< x|y >= \phi(x,y)$*
Etant donné un produit scalaire hermitien $< .|. >$, on définit une **norme hermitienne** *; il s'agit de l'application $x \mapsto \|x\| = \sqrt{< x|x >}$. On verra plus loin qu'il s'agit d'une norme.*
On appelle **espace préhilbertien complexe** *un $\mathbb{C}$-espace vectoriel muni d'un produit scalaire hermitien. Un sous-espace vectoriel F d'un espace préhilbertien complexe E muni d'un produit scalaire hermitien, muni de la restriction du produit scalaire hermitien à F, est appelée* **sous-espace préhilbertien** *de E (c'est un espace préhilbertien).*

⚠ On n'a à aucun moment imposé que la dimension soit finie.

⚠ La notation $< x|y >$ peut être remplacée par $(x|y)$, $< x,y >$, (x,y), ou même $x.y$.

Remarquons que le fait que pour une forme sesquilinéaire hermitienne ϕ on ait $\forall x$ $\phi(x,x) \in \mathbb{R}$ découle du fait que ϕ est hermitienne ; il suffit de vérifier que $\phi(x,x) > 0$.

⚠ Une forme linéaire n'est pas nécéssairement une forme semi-linéaire

⚠ Une forme semi-linéaire n'est pas nécéssairement une forme linéaire

Une forme sesquilinéaire est donc semi-linéaire par rapport à la première variable et linéaire par rapport à la seconde.

Exemples : Sur $\mathbb{C}^n$ le produit scalaire hermitien canonique est défini par $< x|y >= \sum_{i=1..n} \overline{x_i}.y_i$.

Les inégalités de Schwartz et de Minkowski montrées dans la partie28.4 sont valables ici aussi ; mais la démonstration, basée sur la bilinéarité et utilisant les formes quadratiques, n'est plus valable.

Lemme 1091 (Egalité utile)

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2.Re(<x|y>)$$

avec $Re(u)$ la partie réelle de u.

Démonstration : Evidente, en utilisant $<x|y> = \overline{<y|x>}$.$\square$

Théorème 1092 (Inégalité de Cauchy-Schwartz) *Dans un espace préhilbertien complexe*

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

Il y a égalité si et seulement si la famille est liée.

Démonstration : Soit θ l'argument de $<x|y>$, alors pour tout t réel, au vu du lemme 1091 :

$$\|t.e^{i\theta}.x + y\|^2 = t^2.\|x\|^2 + \|y\|^2 + 2.t.| <y|x> |$$

On en déduit donc que le discriminant de $t \mapsto t^2.\|x\|^2 + \|y\|^2 + 2.t.| <y|x> |$ est négatif ou nul, ce qui donne l'inégalité annoncée. Le cas d'égalité est le cas où le discriminant est nul.$\square$

Corollaire 1093 (Inégalité de Minkowski) *Dans un espace préhilbertien complexe*

$$\forall (x,y) \in E^2 \; \|x+y\| \leq \|x\| + \|y\|$$

Il y a égalité si $y = \lambda.x$ ou $x = \lambda.y$ avec $\lambda > 0$.

Démonstration : Par le lemme 1091, on a

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2.Re(<x|y>)$$

$$\leq \|x\|^2 + \|y\|^2 + 2.| <x|y> |$$

(par Cauchy-Schwartz ci-dessus)

$$= (\|x\| + \|y\|)^2$$

D'où le résultat. Le cas d'égalité se montre facilement... $\square$

Proposition 1094 *Dans le cas euclidien, retrouver le produit scalaire à partir de la norme était facile ; dans le cas hermitien c'est un peu plus compliqué :*

$$< x|y >= \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 - i\|x+i.y\|^2 + i.\|x-i.y\|^2)$$

$$= \frac{1}{4}(\sum_{n=0}^{3} e^{-2.i.\pi.n/4}\|x + e^{2.i.\pi.n/4}.y\|)$$

⚠ La dernière ligne est un bon moyen mnémotechnique, mais il faut bien penser que l'on a un signe moins dans le coefficient de l'exponentiel en dehors de $\|.\|$ et un signe plus à l'intérieur.

29.3 Espaces préhilbertiens

On se place ici dans le cadre de E espace préhilbertien, réel ou complexe. On ne suppose absolument pas E de dimension finie.

Définition 1095 *x et y appartenant à E sont dits* **orthogonaux** *si $< x|y >=< y|x >= 0$.*
Deux parties X et Y de E sont dites **orthogonales** *si x et y sont orthogonaux pour tout (x, y) dans $X \times Y$.*

On appelle **orthogonal** *d'une partie X et on note $X^\perp$ l'ensemble des y tels que $< x|y >= 0$ pour tout x dans X.*

Une famille $(x_i)_{i\in I}$ est dite **orthogonale** *si $i \neq j \rightarrow < x_i|x_j >= 0$.*

Une famille $(x_i)_{i\in I}$ est dite **orthonormale** *si $< x_i|x_j >= \delta_{i,j}$.*

Remarques :
• $< x|y >= 0 \iff < y|x >= 0$
• toute partie est orthogonale à son orthogonal
• Etant donné F un sous-espace vectoriel de E, la somme de F et de $F^\perp$ est toujours directe, *mais* non nécéssairement égale à E
• Pour toute partie X, $X^\perp$ est un sous-espace vectoriel fermé de E (car c'est une intersection de fermés, par définition)

⚠ Ne pas confondre "être orthogonal à" et "être l'orthogonal de".

Proposition 1096 *Toute famille orthogonale de vecteurs non nuls est libre.*

Démonstration : Supposons donnée une combinaison linéaire nulle $\sum_i \lambda_i x_i$, et considérons le produit scalaire avec x_{i_0}. On en déduit immédiatement que λ_{i_0} est nul. D'où le résultat.$\square$

Proposition 1097 (Orthogonalité et espaces supplémentaires) • *Si F et G sont supplémentaires, alors F et G sont orthogonaux si et seulement si G est l'orthogonal de F.*
• *Si F et $F^\perp$ sont supplémentaires, alors $F^{\perp^\perp} = F$.*

Démonstration : En exercice !$\square$

On va maintenant considérer quelques résultats de géométrie :

Théorème 1098 (Théorème de Pythagore) • *Si les x_i sont une famille finie orthogonale alors $\| \sum_i x_i\|^2 = \sum_i \|x\|^2$.*

Démonstration : Evident par récurrence.$\square$

On note que dans le cas d'un espace préhilbertien réel, $\|x\|^2 + \|y\|^2 = \|x + y\|$ équivaut à x et y orthogonaux.

Pas valable dans le cas complexe !

Proposition 1099 (Quelques résultats de géométrie) • *Formule du triangle :*

$$\|y-z\|^2 = \|y-x\|^2 + \|z-x\| - 2.Re(< z-x|y-x >)$$

• *Formule de la médiane :*

$$\|y-z\|^2 + 4\|x - \frac{1}{2}(y+z)\|^2 = 2.\|x-y\|^2 + 2.\|x-z\|^2$$

• *Formule du triangle pour un espace préhilbertien réel :*

$$\|y-z\|^2 = \|y-x\|^2 + \|z-x\| - 2 < z-x|y-x >$$

• *Formule du parallélogramme, pour un espace préhilbertien réel :*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Démonstration : Facile, facile ; il suffit de développer. Illustration en figure 29.1.□

Théorème 1100 • *Dans un espace préhilbertien E de dimension finie, tout sous-espace vectoriel F est supplémentaire à son orthogonal, ie $E = F \oplus F^\perp$. On a alors $dim\ E = dim\ F + dim\ F^\perp$.*
• *Tout espace préhilbertien de dimension finie admet une base orthonormale.*
• *Dans un espace préhilbertien de dimension finie, toute famille orthonormale peut être complétée en une base orthonormale.*
• *Si F est un sous-espace vectoriel de E, avec E espace préhilbertien, et F est de dimension finie, alors on a $E = F \oplus F^\perp$.*

Démonstration : Pour le premier • , on considère l'application f qui à x dans E associe $(< u_1|x >, < u_2|x >, < u_3|x >, ..., < u_p|x >)$. Le noyau est $F^\perp$, le rang est $\leq$ à la dimension de F. Donc $dim\ F^\perp \geq dim\ E - dim\ F$, donc $E = F \oplus F^\perp$ (rappelons qu'il est toujours vrai que $F \cap F^\perp = \{0\}$). On a donc bien $dim\ F^\perp + dim\ F = dim\ E$.

Pour le second • , on raisonne par récurrence. Pour E préhilbertien de dimension ≤ 1, le résultat est évident. Il suffit ensuite de considérer un vecteur quelconque e_n de E de norme 1, et une base $(e_1, ..., e_{n-1})$ de $(\mathbb{K}.u)^\perp$. $(e_1, ..., e_n)$ convient.

Le troisième • est immédiat ; il suffit de considérer F engendré par la famille orthonormale $(e_1, ..., e_p)$ considérée, et une base $(e_{p+1}, ..., e_n)$ de $F^\perp$, réunie en $(e_1, ..., e_n)$, base orthonormale de E.

Le quatrième • est un peu plus long :
- on considère une base $(e_1, ..., e_p)$ orthonormale de F.
- on considère l'application f qui à x associe $\sum_{i=1^p} < e_i|x > e_i$.
- $\forall x \in E\ x = \underbrace{f(x)}_{\in F} + \underbrace{(x - f(x))}_{\in F^\perp}$

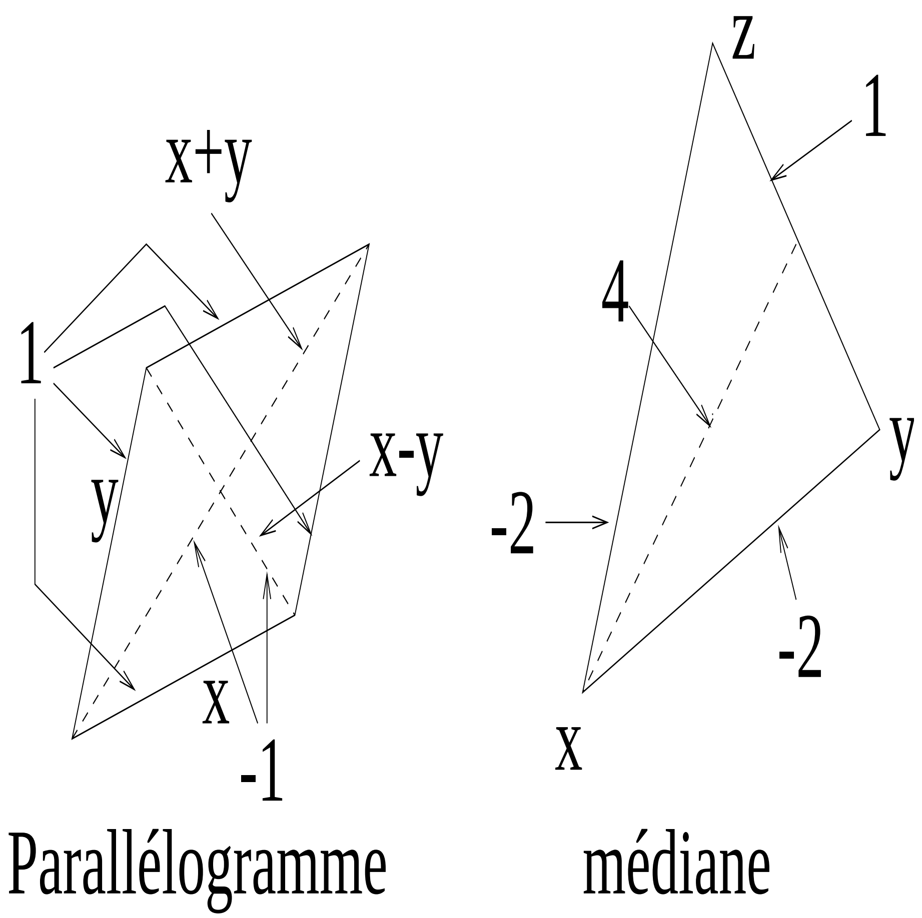


FIG. 29.1 – Illustrations de la formule de la médiane, du parallélogramme. Les coefficients attribués servent de moyen mnémotechnique ; en sommant les carrés des longueurs indiquées, pondérées par leur coefficient, on obtient zéro.

D'où le résultat.□

29.4 Espaces de Hilbert

Définition 1101 *On appelle* **espace de Hilbert** *un espace préhilbertien complet.*

29.4.1 Projection dans un espace de Hilbert

Définition 1102 *Soit H un espace de Hilbert, et E une partie convexe fermée non vide de H. Alors étant donné x appartenant à H on appelle* **projeté de x sur E** *un élément y de E tel que $\|x - y\|$ soit minimal, c'est-à-dire $y = argmin_{z \in E} \|z - x\|$.*
Un **isomorphisme d'espaces de Hilbert** *est un isomorphisme entre les espaces vectoriels sous-jacents qui préserve la norme et le produit scalaire.*

Théorème 1103 *Le projeté de x sur E existe et est unique, et y dans E est le projeté de x sur E si et seulement si pour tout e dans E $Re(< e - y|x - y >) \leq 0$.*

Il est clair que dans le cas d'un espace de Hilbert réel, on pourrait simplement formuler $< e - y|x - y >< 0$.

Démonstration : • Existence d'un projeté de x sur E.
On se donne une suite y_n telle que $\|x - y_n\|$ tende vers la distance d entre x et E.
Par la formule de la médiane appliquée aux points y_n, y_m et x, on a alors

$$\|y_m - y_n\|^2 = 2.(\|x - y_n\|^2 + \|x - y_m\|^2) - 4.\|x - \frac{1}{2}(y_n + y_m)\|^2$$

par convexité de E on a alors $\frac{1}{2}(y_n + y_m) \in E$ et donc

$$\|x - \frac{1}{2}(y_n + y_m)\|^2 \geq d^2$$

et donc

$$\|y_m - y_n\|^2 \to 0$$

Donc la suite y_n est de Cauchy, donc par complétude de H elle converge. Sa limite y est dans E parce que E est fermé.
• Supposons que y est dans E et que pour tout e dans E $Re(< y - e|y - x >) \leq 0$. Alors pour tout e dans E on a

$$\|x - e\|^2 = \|x - y\|^2 + \|e - y\|^2 - 2.Re(< e - y|x - y >)$$

On obtient d'un coup d'un seul que y est la borne inf, et que y est unique (car s'il y a égalité, $\|e - y\| = 0$).
• Supposons maintenant que y soit un projeté de x sur E, et montrons que $Re(< e - y|x - y >) \leq 0$ pour tout e dans E.
On a :

$$\|x - e\|^2 = \|x - y\|^2 + \|e - y\|^2 - 2.Re(< e - y|x - y >)$$

et donc

$$Re(< e - y|x - y >) = \frac{1}{2}(\|x - e\|^2 - \|x - y\|^2 + \|e - y\|^2)$$

par définition de y, on a $\|x-e\|^2 - \|x-y\|^2 \leq 0$, et donc

$$Re(<e-y|x-y>) \leq \frac{1}{2}(\|e-y\|^2)$$

valable pour tout e.
On se donne maintenant z dans E, et on considère $e = t.z + (1-t).y$. L'inégalité précédente s'écrit alors

$$Re(<t.z-t.y|x-y>) \leq \frac{1}{2}(\|t.z-t.y\|^2)$$

et donc

$$Re(<z-y|x-y>) \leq \frac{t}{2}\|z-y\|$$

En faisant tendre t vers 0, on en déduit

$$Re(<z-y|x-y>) \leq 0$$

La preuve est ainsi complète.□

Corollaire 1104 *Dans tout ensemble non vide fermé convexe il existe un unique élément de norme minimale.*

Proposition 1105 • *Le projeté de x sur un sous-espace vectoriel fermé E (qui est évidemment convexe) est le point y tel que $y-x$ est orthogonal à $e-x$ pour tout e dans E.*
• *Soit $(x_i)_{i\in[1,n]}$ une famille orthonormale ; alors l'espace vectoriel engendré par les x_i est fermé (car il y en a un nombre fini, voir le corollaire 189), convexe. La projection sur ce sous-espace vectoriel de H est l'application qui à x associe $\sum_{i\in[1,n]} <x_i|x>.x_i$.*
On a alors $\|x\|^2 = \sum_i |<x_i|x>|^2 + \|x-y\|^2$ avec $y = \sum_{i\in[1,n]} <x_i|x> .x_i$.
• *$\|x\|^2 \geq \sum_i <x_i|x>^2$ **(inégalité de Bessel)**.*

Démonstration : Exercice facile, utilisant les résultats que l'on vient de voir dans le théorème précédent...□

Définition 1106 *Soit E un sous-espace vectoriel fermé de H. On appelle* **projection orthogonale sur** *E l'application qui à x dans H associe son projeté sur E ; il s'agit d'un projecteur, et la symétrie associée à ce projecteur est appelée* **symétrie orthogonale par rapport à** *E (voir27.3.2).*

La projection orthogonale sur E est en fait la projection sur E suivant $E^\perp$.

29.4.2 Bases hilbertiennes

Définition 1107 (Base hilbertienne) *On appelle base hilbertienne d'un espace de Hilbert H une famille $(x_i)_{i \in I}$ telle que :*
• *la famille des x_i est orthonormale*
• *pour tout x dans H $x = \sum_{i \in I} < x_i | x > x_i$*

La seconde condition est une somme éventuellement infinie, en fait il s'agit d'une famille sommable, i.e. pour tout ϵ il existe J fini inclus dans I, tel que pour tout K fini compris entre J et I la somme sur K des $< x_i | x > .x_i$ est à une distance $< \epsilon$ de x.

Théorème 1108 • *Une famille orthonormale $(x_i)_{i \in I}$ est une base hilbertienne si et seulement si le sous-espace vectoriel engendré par les x_i est dense dans H.*
• *(***Relation de Parseval***) Une famille orthonormale $(x_i)_{i \in I}$ est une base hilbertienne si et seulement si $\sum_{i \in I} | < x_i | x > |^2 = \|x\|^2$.*

La relation du deuxième • est appelée **relation de Parseval**.

Démonstration : • Supposons le sous-espace vectoriel E engendré par les x_i dense dans H. Alors il existe y dans E tel que $\|x - y\| \leq \epsilon$, et $y = \sum \lambda_i x_i$. On en déduit que la famille est une base hilbertienne.
• Si la relation de Parseval est vérifiée, alors le sous-espace vectoriel engendré est dense dans H, comme on le voit en considérant une sous-famille de I suffisamment vaste.
• Si on suppose que la famille est une base hilbertienne, alors la relation de Parseval est vérifiée, au vu des résultats de la proposition 1105.□

Attention ! Une base hilbertienne n'est pas nécéssairement une base au sens des espaces vectoriels !

Théorème 1109 • *Une famille orthonormale $(x_i)_{i \in I}$ est une base hilbertienne si et seulement si pour tout x et tout y on a $< x | y >= \sum_i < x_i | x > . < x_i | y >$.*
• *Une famille orthonormale $(x_i)_{i \in I}$ est une base hilbertienne si et seulement si elle est maximale pour l'inclusion.*

Démonstration : Pas très dur, dans le même style que le théorème précédent, j'ai pas la patience de détailler...□

La proposition suivante permet de se ramener à une forme plus "visuelle" des espaces de Hilbert.

Théorème 1110 (Riesz-Fischer - Isomorphisme sur un l^2) *Soit H un espace de Hilbert, et $(x_i)_{i\in I}$ une base hilbertienne de H. Alors l'application $x \mapsto (< x_i|x >)_{i\in I}$ est un isomorphisme de H sur $l^2(I)$.*

Démonstration : • L'application $\phi = x \mapsto (< x_i|x >)_{i\in I}$ est bien à valeurs dans $l^2(I)$, vu la relation de Parseval (théorème 1108).
• ϕ est linéaire, conserve la norme.
• ϕ conservant la norme, ϕ est injective.
• ϕ conservant la norme, ϕ conserve les distances, et donc son image est un sous-espace vectoriel complet, donc fermé de $l^2(I)$.
• $\phi(x_i)$ est la fonction caractéristique du singleton i ; or les combinaisons linéaires de ces fonctions sont denses dans $l^2(I)$, donc $\phi(H)$ est dense.
• $\phi(H)$ est dense et fermé, donc il est égal à $l^2(I)$.
• Le produit scalaire est continu, donc le fait qu'il soit conservé sur les x_i suffit à montrer qu'il est conservé partout.□

Un résultat utérieur donnera toute sa puissance à ce résultat, en montrant que tout espace de Hilbert possède une base hilbertienne.

⚠ Il ne s'agit pas d'un isomorphisme ramenant tous les espaces de Hilbert possédant une base à un seul, car I peut avoir des cardinaux différents.

Lemme 1111 *Soit H un espace de Hilbert . Soit E un sous-espace vectoriel fermé de H. Si E n'est pas égal à H, alors il existe x dans H de norme 1 orthogonal à E.*

Démonstration : Il suffit de considérer $x - y$, pour $x \in E^c$ et y projeté de x sur E.□

Théorème 1112 *Tout espace de Hilbert possède une base hilbertienne .*

⚠ Nécéssite l'usage de l'axiome du choix !

Démonstration : On considère l'ensemble des familles orthonormales, munies de l'inclusion. Il s'agit bien d'un ensemble inductif (si l'on considère une chaîne de familles orthonormales, leur réunion est bien un majorant), et donc par le lemme de Zorn (voir le lemme 36) il admet un élément maximal. On considère l'adhérence du sous-espace vectoriel engendré par cette famille, et si l'on n'obtient par H lui-même, on applique le lemme 1111, et on contredit la maximalité de notre famille orthonormale.□

Corollaire 1113 (Corollaire des deux théorèmes précédents) *Tout espace de Hilbert est isomorphe à $l^2(I)$ pour un certain I.*

Rappelons que $\mathcal{L}(E,F)$ désigne l'ensemble des applications linéaires continues de E dans F, avec E et F des espaces vectoriels , et que pour E un $\mathbb{K}$-espace vectoriel on note E' le dual topologique de E, c'est à dire $\mathcal{L}(E,\mathbb{K})$.

Théorème 1114 *Soit H un espace de Hilbert réel (resp. complexe). Alors l'application $\phi : H \to H'$, $x \mapsto \phi(x) = (y \mapsto < x|y >)$ est une bijection linéaire (resp. semi-linéaire).*

Démonstration : • La linéarité (resp. semi-linéarité) est claire.
• L'injectivité n'est pas bien difficile ; il suffit de voir que si $\phi(x) = \phi(y)$, alors pour tout $z < x|z > = < y|z >$, et on conclut en considérant une base hilbertienne, ou bien en considérant $< x-y|x-y > = < x|x-y > - < y|x-y > = 0$...
• La surjectivité est plus intéressante. Soit f une forme linéaire continue sur H, autre que la forme linéaire nulle. Son noyau est un sous-espace vectoriel fermé E de H. On considère F l'orthogonal de E ; E et F sont en somme directe, car E est de codimension 1 en tant que noyau d'une forme linéaire, et F est de dimension au moins 1 (voir lemme 1111), et bien sûr $E \cap F$. On considère alors y dans F, non nul, et $x = \frac{f(y)}{\|y\|^2}.y$. On vérifie que $\phi(x) = f$ sur F, sur E, et donc sur H tout entier.

Proposition 1115 (Procédé d'orthonormalisation de Schmidt) *Etant donnée une famille $(x_i)_{i\in[0,N[}$ avec $N \in \mathbb{N} \cup \{+\infty\}$ libre de H (H espace de Hilbert), il existe une famille $(y_i)_{i\in[0,N[}$ orthonormale telle que le sous-espace vectoriel engendré par $(y_i)_{i\in[1,k]}$ soit égal au sous-espace vectoriel engendré par $(x_i)_{i\in[1,k]}$ pour tout $k < N$.*

Démonstration : Il suffit de procéder comme suit :

$$y_0 = x_0/\|x_0\|$$

et pour $n > 0$

$$z_n = x_n - \sum_{i=0..n-1} < y_i|x_n > .y_i$$

et

$$y_n = z_n/\|z_n\|$$

Et ça marche tout seul.□

Exemple Maple

Ci-dessous une implémentation dans le cadre des polynômes orthogonaux (voir partie 25.5.5).

$> scalaire := (f, g)- > int(f * g, t = 0..1);$

$> norme := f- > sqrt(scalaire(f, f));$

$>$

$N := 3; x := array(0..N);\ for\ i\ from\ 0\ by\ 1\ to\ N\ do\ x[i] := t^i;\ od;\ y := array(0..N);\ z := array(0..N);$

$> y[0] := x[0]/norme(x[0]); for\ i\ from\ 1\ by\ 1\ to\ N\ do\ z[i] := x[i];\ for\ j\ from\ 0\ by\ 1\ to\ (i-1)\ do\ z[i] := z[i] - scalaire(x[i], y[j]) * y[j];\ od;\ y[i] := simplify(z[i]/norme(z[i])); od;$

$> A := linalg[matrix](N + 1, N + 1); for\ i\ from\ 0\ by\ 1\ to\ N\ do\ for\ j\ from\ 0\ by\ 1\ to\ N\ do\ A[i+1, j+1] := simplify(scalaire(y[i], y[j])); od; od; evalm(A);$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(une partie des réponses de Maple est supprimée par économie de place)

Proposition 1116 *Si $N < +\infty$ alors la matrice de passage de la base des y_i à la base des x_i est triangulaire de déterminant $\Pi_{i=1..n} < x_i | y_i >$.*

Démonstration : Il suffit de voir que $P_{(y_i),(x_i)} =< x_i | y_j >$. $\square$

29.4.3 Quelques utilisations des espaces de Hilbert

On pourra aller voir par exemple le théorème 290 (une version de théorème de point fixe). La partie 10, au niveau des séries de Fourier, donne une application très classique des espaces de Hilbert et des bases hilbertiennes.

29.5 Espaces euclidiens

29.5.1 Les bases

Définition 1117 (Espace euclidien) *On appelle* **espace euclidien** *un espace préhilbertien réel de dimension finie non nulle.*
Un endomorphisme d'un espace euclidien est dit **orthogonal** *si l'image d'une certaine base orthonormale est une base orthonormale.*
On appelle **similitude** *d'un espace euclidien un endomorphisme égal à la composée d'une homothétie (i.e. une application du type $E \to E$, $x \mapsto \lambda.x$) et d'un automorphisme orthogonal.*
On appelle **rapport** *d'une similitude le rapport d'une homothétie de la décomposition de cette similitude en une homothétie et un automorphisme (le rapport est unique).*

Un espace euclidien est donc en particulier est espace préhilbertien et un espace de Hilbert ; on pourra donc consulter la partie 29 et plus spécialement la partie 29.4 pour avoir les bases.

Proposition 1118 • *L'image de toute base orthonormale par un endomorphisme orthogonal est une base orthonormale.*
• *Un endomorphisme est orthogonal si et seulement si sa matrice <u>dans une base orthonormale</u> est une matrice orthogonale.*
• *Un espace euclidien est un espace de Hilbert.*
• *$\mathbb{R}^n$ muni du produit scalaire euclidien canonique (i.e. $< x|y >= \sum_{i\in[1,n]} x_i.y_i$) est un espace euclidien.*
• *Un espace euclidien est isomorphe à $\mathbb{R}^n$ muni du produit scalaire euclidien canonique pour un certain n.*

<u>**Démonstration :**</u> Un espace euclidien est de dimension finie, en dimension finie toutes les normes sont équivalentes, donc l'espace est isomorphe au sens des espaces vectoriels normés à $\mathbb{R}^n$ muni d'une norme usuelle et est donc complet. Donc c'est un espace de Hilbert. Donc il admet une base orthonormale. Le tour est joué.$\square$

Soit E un espace euclidien, n sa dimension.
Les résultats suivants sont trop faciles pour valoir une démonstration (notez toutefois qu'ils n'apparaissent faciles qu'au vu des parties précédentes) :
• L'application $x \mapsto (y \mapsto < x|y >)$ est un isomorphisme de E sur E^*.
• Toute forme linéaire sur E s'écrit $< x|. >$ pour un certain x de E.
• Il existe une base orthonormale de E, qui est d'ailleurs une base hilbertienne aussi.

Définition 1119 *Il y a plusieurs notions d'***angles** *à définir :*
On définit l'angle entre deux vecteurs non nuls x et y comme étant le réel θ de $[0, \Pi]$ tel que $cos(\theta) = \frac{<x|y>}{\|x\|.\|y\|}$.
On définit l'angle entre deux droites par l'angle entre un vecteur d'une base de l'une et un vecteur d'une base de l'autre.
On définit l'angle entre deux hyperplans comme l'angle entre les droites qui leurs sont orthogonales.
On définit l'angle entre une droite et un hyperplan comme l'angle entre la droite et la droite orthogonale à l'hyperplan.
On dit que deux sous-espaces vectoriels de E sont **perpendiculaires** *s'ils sont orthogonaux.*

Proposition 1120 *Un endomorphisme est une similitude si et seulement si les deux conditions suivantes sont vérifiées :*
• *il est bijectif*
• *il conserve l'écart angulaire*

Démonstration : Il est très facile de vérifier qu'une similitude vérifie bien les deux propriétés annoncées.

Réciproquement, supposons les deux propriétés vérifiées par un endomorphisme f. Alors :

• Considérons l'application $g : x \mapsto \frac{\|f(x)\|}{\|x\|}$, pour x non nul.
• Conservant l'écart angulaire, l'application f conserve aussi l'orthogonalité.
• Etant donnés x et y distincts, $z = y - \frac{<x|y>}{\|x\|^2} x$ est orthogonal à x.
• $f(z)$ est donc orthogonal à $f(x)$.
• En développant $< f(x)|f(z) >$ puis en factorisant par $< x|y > \frac{\|f(x)\|}{\|x\|}$, on obtient que $g(x) = g(y)$.
• g est donc constante, égale à μ.
• $\frac{1}{\mu} f$ conserve donc la norme, et donc est orthogonal d'après la proposition 1123. Par définition, f est donc une similitude. □

La preuve ci-dessus montre qu'en fait suffisant qu'un endomorphisme bijectif conserve l'orthogonalité pour qu'on puisse conclure qu'il s'agit d'une similitude.

29.5.2 Endomorphisme adjoint

Théorème 1121 *Pour tout endomorphisme f d'un espace euclidien E, il existe un et un seul endomorphisme f^* tel que pour tout x et tout y de E, $< f(x)|y > = < x|f^*(y) >$.*

Démonstration : • Existence et unicité de f^* : pour tout y, l'application $x \mapsto < f(x)|y >$ est une forme linéaire ; donc par l'isomorphisme décrit plus haut entre E et son dual il existe un unique $f^*(y)$ tel que pour tout $x < f(x)|y >=< x|f^*(y) >$.
• Linéarité de f^* : soit

$$< f(x)|\lambda_1.y_1 + \lambda_2.y_2 >= \lambda_1. < f(x)|y_1 > +\lambda_2. < f(x)|y_2 >$$

$$= \lambda_1. < x|f^*(y_1) > +\lambda_2. < x|f^*(y_2) >$$

$$=< x|\lambda_1.f^*(y_1) > + < x|\lambda_2.f^*(y_2) >$$

$$=< x|\lambda_1.f^*(y_1) + \lambda_2.f^*(y_2) >$$

pour tout x et donc

$$f^*(\lambda_1.y_1 + \lambda_2.y_2) = \lambda_1.f^*(y_1) + \lambda_2.f^*(y_2)$$

ce qui conclut la preuve.□

Définition 1122 • f^* *est appelé* **endomorphisme adjoint** *de* f.
• f *est dit* **orthogonal** *si* $f^*.f = f.f^* = Id$.
• f *est dit* **symétrique** *si* $f^* = f$.
• *On note* $S(E)$ *l'ensemble des endomorphismes symétriques de* E.
• f *est dit* **antisymétrique** *si* $f^* = -f$.
• *On note* $A(E)$ *l'ensemble des endomorphismes antisymétriques de* E.
• *On appelle* **groupe orthogonal de** E *et on note* $O(E)$ *l'ensemble des endomorphismes orthogonaux de* E *; c'est un sous-groupe de* $GL(E)$*, ensemble des automorphismes de* E.
• *On appelle* **groupe spécial orthogonal de** E *et on note* $SO(E)$ *l'ensemble des endomorphismes orthogonaux de* E *de déterminant* 1 *; c'est un sous-groupe de* $O(E)$.

Quelques propriétés faciles de E euclidien :
• Un endomorphisme est orthogonal si et seulement si son adjoint l'est.
• L'application $f \mapsto f^*$ est un endomorphisme de $L(E)$; c'est un endomorphisme involutif, c'est à dire que f^{**} est égal à f.
• $Ker\ f^* = (Im\ f)^\perp$
• $Im\ f^* = (Ker\ f)^\perp$
• $Ker\ f = (Im\ f^*)^\perp$
• $Im\ f = (Ker\ f^*)^\perp$
• $(g \circ f)^* = f^* \circ g^*$
• $Mat_B(f^*) =^t Mat_B(f)$
• f est diagonalisable $\iff$ f^* est diagonalisable.
• si f est inversible, alors f^* aussi et $(f^*)^{-1} = (f^{-1})^*$.
• F sous-espace vectoriel de E est stable par f si et seulement si $F^\perp$ est stable par f^*.
• Un endomorphisme et son adjoint ont même polynôme caractéristique.

A peine plus dur :

Proposition 1123 • *Un endomorphisme est orthogonal si et seulement si il conserve le produit scalaire.*
• *Un endomorphisme est orthogonal si et seulement si il conserve la norme.*

Démonstration : Il est vite vu que les conditions en question sont équivalentes entre elles, car la norme s'exprime en fonction du produit scalaire, et le produit scalaire en fonction de la norme ; le calcul est facile. On va donc se contenter de montrer les deux implications suivantes :
• Si f est orthogonal, alors f conserve la norme : évident car

$$\|x\|^2 =< f(x)|f(x) >=< x|f^{-1}(f(x)) >=< x|x >= \|x\|^2$$

• Si f conserve le produit scalaire, alors f est orthogonal :
- f est injectif puisqu'il conserve la norme
- on est en dimension finie, donc f est inversible
- $< f(x)|y >=< f(x)|f(f^{-1}(y)) >=< x|f^{-1}(y) >$ et donc $f^{-1} = f^*$.□

On peut aussi noter le résultat amusant suivant :

Proposition 1124 • *Une application de E dans E avec E euclidien conservant le produit scalaire est linéaire.*
• *Une application de E dans E avec E euclidien conservant la norme est linéaire.*

Démonstration : • Il suffit de développer $\|f(\lambda.x + \mu.y) - \lambda.f(x) - \mu.f(y)\|$ et de calculer un peu.
• Un peu de calcul sur les formules liant produit scalaire et norme suffit...□

Corollaire 1125 *f est orthogonal si et seulement si l'image d'une base orthonormale est une base orthonormale.*
L'image d'une base orthonormale par un endomorphisme orthogonal est une base orthonormale.

Corollaire 1126 *Une application de E dans E avec E euclidien est une isométrie si et seulement si c'est la composée d'une translation et d'un endomorphisme orthogonal.*

Démonstration : Il est évident que la composée d'une translation et d'un endomorphisme orthogonal est une isométrie. Réciproquement, on procède comme suit :
• On se donne f une isométrie de E dans E

• On considère $g = (x \mapsto x - f(0) \circ f$
• On montre que g conserve la norme
• On conclut par la proposition 1124.□

Quelques propriétés faciles des endomorphismes orthogonaux, utilisant les résultats ci-dessus ;

Proposition 1127 *Soit $f \in O(E)$:*
• Le spectre de f est inclus dans $\{-1, 1\}$
• Le déterminant de f est -1 ou 1
• F est stable par $f \iff F^{\perp}$ est stable par F • Un endomorphisme orthogonal a toutes ses valeurs propres égales à 1 ou -1
• Un endomorphisme orthogonal diagonalisable est une symétrie orthogonale (voir27.3.2
• $E = Im\ (f - I) \oplus Ker\ (f - I)$

Quelques résultats sur les endomorphismes symétriques et antisymétriques (E un espace euclidien) :

Proposition 1128 *• Un endomorphisme est symétrique (resp. antisymétrique) si et seulement si sa matrice dans une base orthonormale est symétrique (resp. antisymétrique)*
• L'ensemble $L(E)$ des endomorphismes de E est somme directe de l'ensemble $S(E)$ des endomorphismes symétriques et de l'ensemble $A(E)$ des endomorphismes antisymétriques ; $L(E) = S(E) \oplus A(E)$.
• $dim\ L(E) = n^2$ avec $n = dim\ E$
• $dim\ S(E) = \frac{1}{2}n.(n+1)$ avec $n = dim\ E$
• $dim\ A(E) = \frac{1}{2}n.(n-1)$ avec $n = dim\ E$
• si f est un endomorphisme antisymétrique de E, alors $f \circ f$ est un endomorphisme symétrique
• Un endomorphisme f de E est antisymétrique si et seulement si $\forall x \ < f(x)|x >= 0$
• La seule valeur propre possible d'un endomorphisme antisymétrique est 0
• L'image et le noyau de f symétrique ou antisymétrique sont supplémentaires orthogonaux
• Le rang d'un endormophisme antisymétrique est pair (en effet sa restriction à son image est une bijection et est antisymétrique, donc il n'a pas de valeur propre puisque 0 n'est pas de valeur propre, donc le polynôme caractéristique n'a pas de racine, donc l'image est de dimension paire)
• Les sous-espaces propres d'un endomorphisme symétrique sont supplémentaires et orthogonaux.
• Le polynôme caractéristique d'un endomorphisme symétrique est scindé sur $\mathbb{R}$ (c'est à dire que la somme des multiplicités de ses racines sur $\mathbb{R}$ est égal à son degré).
• Un endomorphisme symétrique est diagonalisable dans une base orthonormale

Démonstration : La seule chose qui justifie que l'on fasse une preuve est l'avant-dernier point, qui entraîne le dernier.
• On se donne f un endomorphisme symétrique, et M sa matrice dans une base orthonormale.
• Soit λ une racine dans $\mathbb{C}$ du polynôme caractéristique de M.
• On peut trouver un vecteur colonne C (à coefficients complexes non nuls) tel que $M.C = \lambda.C$
• ${}^t\overline{X}.M.X = \lambda.{}^t\overline{X}.X$
• en conjuguant et en transposant on obtient ${}^t\overline{X}.(\tau\overline{M}).X = \overline{\lambda}{}^t\overline{X}.X$ c'est à dire ${}^t\overline{X}.M.X = \overline{\lambda}\,{}^t\overline{X}.X$
• on a donc $\lambda.{}^t\overline{X}.X = \overline{\lambda}.{}^t\overline{X}.X$
• ${}^t\overline{X}.X$ est non nul donc $\lambda \in \mathbb{R}$ et c'est fini.$\square$

29.5.3 Orientation d'un espace euclidien

On pourra consulter au préalable la partie28.

Définition 1129 *On appelle* **espace euclidien orienté** *un couple (E, C) où E est un espace euclidien et C une classe d'équivalence sur l'ensemble des bases de E pour la relation d'équivalence $\mathcal{R}$ définie par*

$$B\mathcal{R}B' \iff det\ P_{B,B'} > 0$$

*L'***orientation de** *F* **sous-espace vectoriel de dimension** *$n-p$* **de** *(E, C)* **espace euclidien orienté de dimension** *n* **suivant** *$(e_1, ..., e_p)$* **une base d'un supplémentaire de** *F est l'espace euclidien orienté*

$$(F, \{(e_{p+1}, ..., e_n)/(e_1, ..., e_n) \in C\})$$

(il s'agit d'un espace euclidien orienté)
On appelle **base directe** *de l'espace euclidien orienté (E, C) une base appartenant à C.*
On appelle **base indirecte** *de l'espace euclidien orienté (E, C) une base n'appartenant pas à C.*

Il n'y a que deux orientations possibles d'un même espace euclidien.

L'espace euclidien orienté usuel correspondant à $\mathbb{R}^n$ est donné par la base (directe par définition)

$$((1, 0, ..., 0), (0, 1, 0, ..., 0), (0, 0, 1, 0, ..., 0), ..., (0, 0, ..., 0, 1, 0), (0, ..., 0, 1)).$$

Propriétés faciles : (E est un espace euclidien donné, de dimension n)
• Si σ est une permutation paire (resp. impaire) et si $e_1, ..., e_n$ est une base de E espace euclidien, alors $(e_i)_{i\in[1,n]}$ et $(e_{\sigma(i)})_{i\in[1,n]}$ sont dans la même classe d'équivalence (resp. ne sont pas dans la même classe d'équivalence) (voir 28.5.4 pour en être complètement convaincu).

• Si $(e_i)_{i\in[1,n]}$ et $(f_i)_{i\in[1,n]}$ sont deux bases de E avec $f_i = f(e_i)$ (f est donc un automorphisme), alors les bases $(e_i)_{i\in[1,n]}$ et $(f_i)_{i\in[1,n]}$ sont dans la même classe d'équivalence si et seulement si $det\ f > 0$ (en effet f est alors l'endomorphisme de la matrice de passage de la base des e_i à la base des f_i.
• si B et B' sont dans la même classe alors $det_B(.) = det_{B'}(.)$.

Ceci permet d'introduire de nouvelles définitions :

Définition 1130 *On appelle* **produit mixte** *d'un espace euclidien* (E, C) *de dimension* n *l'application* det_B *pour une base* $B \in C$ *quelconque ; on le note* $(x_1, ..., x_n) \mapsto [x_1, ..., x_n] = det_B(x_1, ..., x_n)$*.*
Si E *est un espace euclidien de dimension* 3*, alors étant donnés* a *et* b *dans* E*, l'application qui à* x *dans* E *associe* $[a, b, x]$ *est linéaire, donc elle est égale à* $x \mapsto < c|x >$ *pour un certain* $c \in E$ *(voir le théorème 1114) ; on note* $a \wedge b$ *l'élément* c *de* E*, et on l'appelle* **produit vectoriel de** a **et** b*.*

Le produit vectoriel n'est pas commutatif !

Intuitivement, le produit mixte de $(x_1, ..., x_n)$ est le volume algébrique (lié à l'orientation) de $\{O + t_1.x_1 + t_2.x_2 + ... + t_n.x_n / (t_1, ..., t_n) \in [0, 1]^n\}$ (indépendamment de O).

Proposition 1131 *Propriétés du produit mixte et du produit vectoriel (celles du produit mixte sont valables en toute dimension ; le produit vectoriel n'est défini qu'en dimension* 3*) (dans les deux cas rien n'est possible en dehors d'un espace euclidien orienté) :*
• $a \wedge b \in Vect(a,b)^{\perp}$
• $a \wedge b = -b \wedge a$
• $a \wedge b = 0 \iff a$ *et* b *sont liés*
• $(a,b) \mapsto a \wedge b$ *est bilinéaire*
• (a,b) *libre* $\rightarrow (a,b,a \wedge b)$ *est une base directe*
• *Le produit mixte d'une base est non nul*
• *Le produit mixte d'une base directe est strictement positif*
• *Le produit mixte d'une base indirecte est strictement négatif*
• *Une famille est une base orthonormale directe si et seulement si son produit mixte est* 1
• *Une famille est une base orthonormale indirecte si et seulement si son produit mixte est* -1
• $[f(x_1),...,f(x_n)] = (det\ f)[x_1,...,x_n]$
• *Dans* $\mathbb{R}^3$*, le produit vectoriel de* (x,y,z) *par* (x',y',z') *est égal à*

$$(y.z' - z.y', z.x' - x.z', x.y' - -y.x')$$

moyen mnémotechnique : ce sont les cofacteurs de la troisième colonne dans la matrice suivante :

$$\begin{pmatrix} x & x' & ? \\ y & y' & ? \\ z & z' & ? \end{pmatrix}$$

• $(a \wedge b) \wedge c = < a|c > .b - < b|c > .a$
• $a \wedge (b \wedge c) = < a|c > .b - < a|b > .c$
• *Si* E *est euclidien orienté de dimension* 3*, alors l'application* ϕ *de* E *dans* $L(E)$ *définie par* $\phi(x) = (y \mapsto x \wedge y)$ *est à valeurs dans l'ensemble des endomorphismes antisymétriques de* E *; en restreignant l'espace d'arrivée à l'ensemble des endomorphismes antisymétriques de* E *c'est un isomorphisme de* E*.*
La matrice de $\phi(u)$ *avec* $u = (x,y,z)$ *est* $\begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$ *dans la même base.*

Toutes ces propriétés s'obtiennent facilement en considérant simplement la définition du produit mixte ou du produit vectoriel.

Proposition 1132 (Identité de Lagrange)

$$\|a \wedge b\|^2 = \|a\|^2.\|b\|^2 - < a|b >^2$$

Démonstration : On suppose a et b libres, sinon le résultat est clair par Cauchy-Schwartz (voir la partie 29.1).

$$[a, b, a \wedge b]^2 = \begin{vmatrix} < a|a > & < a|b > & 0 \\ < a|b > & < b|b > & 0 \\ 0 & 0 & < a \wedge b|a \wedge b > \end{vmatrix} = \|a \wedge b\|^2.(\|a\|^2.\|b\|^2 - < a|b >^2)$$

mais on a aussi

$$[a, b, a \wedge b]^2 =< a \wedge b|a \wedge b >^2 = \|a \wedge b\|^4$$

D'où le résultat désiré.□

Corollaire 1133

$$\|a \wedge b\| = \|a\| \times \|b\| \times sin(\theta)$$

avec θ l'écart angulaire entre a et b.

Maintenant quelques propriétés un peu plus difficiles.

Proposition 1134

$$[x_1, ..., x_n]^2 = det(< x_i|x_j >_{(i,j)\in[1,n]^2})$$

Démonstration : On pose M la matrice $M_{i,j} = e_i^*(x_j)$ avec $(e_i)_{i\in[1,n]}$ une base orthonormale.

$$[x_1, ..., x_n] = det\ M = det\ {}^tM$$

$$[x_1, ..., x_n] = det\ {}^tM.M$$

$$M_{i,j} =< e_i|x_j >$$

$$({}^tM.M)_{i,j} = \sum_{k=1}^{n} < e_k|x_i > . < e_k|x_j >$$

d'où le résultat.□

Proposition 1135

$$|[x_1, ..., x_n]| \leq \Pi_{i=1}^n \|x_i\|$$

$$|[x_1, ..., x_n]| = \Pi_{i=1}^n \|x_i\| \iff (\exists i / x_i = 0 \vee (x_i)_{i \in [1,n]} \textit{ est orthogonale})$$

Démonstration : Si le système est lié, le résultat est clair (on rappelle que le déterminant d'une famille liée est nul).
Sinon, on peut construire par la méthode d'orthonormalisation de Schmidt une base $(e_1, .., e_n)$ avec $\forall (i,j) \in [1,n]^2 \ < x_j | e_i >> 0$.
La matrice de passage $P_{(e_i)_{i\in[1,n]},(x_i)_{i\in[1,n]}}$ est triangulaire (voir 1115) ; son déterminant est le produit des $< x_i | y_i >$, donc par l'inégalité de Schwartz est inférieur ou égal en valeur absolue au produit des $\|x_i\|$, avec égalité si et seulement si pour tout i x_i et y_i sont liés, donc si la famille est orthonormale.□

29.5.4 Formes quadratiques sur un espace euclidien

Pour plus d'informations on pourra consulter la partie 28.4.3.

29.6 Espaces hermitiens

Pour introduction on pourra consulter la partie 29.2, sur les espaces préhilbertiens complexes.

29.6.1 Définition et premières propriétés

Définition 1136 *On appelle* **espace hermitien** *un espace préhilbertien complexe de dimension finie, non réduit à* $\{0\}$.

On rappelle quelques propriétés, issues plus ou moins directement de la définition des espaces préhilbertiens complexes (voir la partie 29.2 pour des rappels) :

• Un espace hermitien est un espace de Hilbert ; toutes les propriétés des espaces de Hilbert sont donc valables ici (voir la partie 29.4)

• $< .|. >$ est sesquilinéaire (i.e. semi-linéaire par rapport à la première variable et linéaire par rapport à la deuxième variable)

• si x est non nul alors $< x|x >$ est un réel > 0 ($< .|. >$ est positive car pour tout $x < x|x >$ est positif et définie car x non nul $\rightarrow < x|\underline{x > \text{non}}$ nul)

• $< .|. >$ est hermitienne, c'est-à-dire $< x|y >= \overline{< y|x >}$

• toute famille orthonormale peut être prolongée en une base orthonormale

• Pour tout sous-espace F d'un espace hermitien E, on a $E = F \oplus F^\perp$ et $F = F^\perp$

• Etant donnée une base <u>orthonormale</u> $(e_i)_{i\in[1,n]}$ de E hermitien, tout vecteur x de E vérifie

$$x = \sum_{i=1..n} < e_i | x > e_i$$

• Etant donnée une base $(x_i)_{i\in[1,n]}$ d'un espace hermitien, il existe une base orthonormale $(y_i)_{i\in[1,n]}$ telle que pour tout j y_j est combinaison linéaire des x_i pour $1 \leq i \leq j$ (voir 1115)

• Etant donné E un espace hermitien l'application qui à x associe l'application $y \mapsto < x|y >$ est un semi-isomorphisme de E sur E^*.

• En corollaire de la propriété ci-dessus, étant donné E un espace hermitien de dimension n, pour toute famille $(x_1, ..., x_n)$ de $\mathbb{C}^n$ et toute base $(e_1, ..., e_n)$ de E, il existe un unique x dans E tel que $< x|e_i >= x_i$ (on aurait pu, au lieu de $< x|e_i >= x_i$, demander $< e_i|x >= x_i$, comme on s'en rend facilement compte en considérant le fait que $< .|. >$ est hermitienne).

• Etant donnée $(e_i)_{i\in[1,n]}$ une base <u>orthonormale</u> de E avec E hermitien, la famille des $(x \mapsto < e_i|x >$ est la base duale de la base des $(e_i)_{i\in[1,n]}$ (il s'agit de l'image de la famille des e_i par le semi-isomorphisme ci-dessus).

La dernière propriété nécéssite que la famille soit orthonormale !

$\mathbb{C}^n$ muni de l'application $(x, y) \mapsto \sum_{i=1}^n \overline{x}_i \times y_i$ est un espace hermitien.

De même que les espaces euclidiens sont isomorphes à $\mathbb{R}^n$ muni du produit scalaire euclidien usuel, on retiendra que les espaces hermitiens sont isomorphes à $\mathbb{C}^n$ muni du produit scalaire hermitien usuel (ce qui fait que beaucoup de propriétés intuitives se retrouvent vraies).

29.6.2 Adjoint d'un endomorphisme d'un espace hermitien

Théorème 1137 *Pour tout endomorphisme f de l'espace hermitien E il existe un et un seul endomorphisme f^* de E tel que pour tout $(x, y) \in E^2$ on ait $< f(x)|y >=< x|f^*(y) >$.*

<u>**Démonstration :**</u>

• Existence

L'application f^* $x \mapsto \sum_{i=1}^n < f(e_i)|x > .e_i$.

• Unicité

$< e_i|f^*(x) >=< f(e_i)|x >$, donc $f^*(x)$ est entièrement déterminé par ses coordonnées.□

Proposition 1138 *L'application F qui à f associe f^* est un semi-endomorphisme de $L(E)$. F est involutif, c'est à dire que $F \circ F = I$.*

Démonstration : Facile !□

> **Définition 1139** • f^* *s'appelle l'***adjoint** *de* f.
> • f *endomorphisme d'un espace hermitien* E *est dit* **unitaire** *lorsque* $f \circ f^* = f^* \circ f = I$.
> • f *endomorphisme d'un espace hermitien* E *est dit* **hermitien** *lorsque* $f = f^*$.
> • *Une matrice carrée* M *à coefficients dans* $\mathbb{C}$ *est dite* **unitaire** *si elle vérifie* $\overline{{}^tM}.M = I$.
> • *Une matrice carrée* M *à coefficients dans* $\mathbb{C}$ *est dite* **hermitienne** *si elle vérifie* $\overline{{}^tM} = M$.

On remarque donc qu'un endomorphisme unitaire est un endomorphisme inversible f tel que pour tout x et tout y de E $< f(x)|y >=< x|f^{-1}(y) >$.

Par ailleurs un endomorphisme f d'un espace hermitien est hermitien lorsque pour tout x et tout y de E on a $< f(x)|y >=< x|f(y) >$.

Sans surprise suite au cas euclidien, un endomorphisme f d'un espace hermitien est unitaire si et seulement si il conserve le produit scalaire, i.e.

$$< f(x)|f(y) >=< x|y >$$

Ou même simplement la norme, i.e.

$$\|f(x)\| = f(x)$$

De même que dans le cas des euclidiens, on note que l'adjoint de l'inverse (quand il existe) est l'inverse de l'adjoint (qui dans ce cas existe nécessairement), que l'orthogonal de l'image est le noyau de l'adjoint, et que l'image de l'adjoint est l'orthogonal du noyau ; que $(f \circ g)^* = g^* \circ f^*$.

De même qu'un endomorphisme d'un espace euclidien est orthogonal si et seulement si l'image d'une base orthonormale est orthonormale, un endomorphisme d'un espace hermitien est unitaire si et seulement si l'image d'une base orthonormale est une base orthonormale.

Alors que dans le cas euclidien et *dans une base orthonormale* la matrice de l'adjoint est la transposée de la matrice, dans le cas hermitien et *dans une base orthonormale* la matrice de l'adjoint est la conjuguée de la matrice transposée. C'est-à-dire

$$Mat_{(e_i)_{i\in[1,n]}}(f) = \overline{{}^tMat_{(e_i)_{i\in[1,n]}}(f)}$$

Conséquence logique de ce qui précède, un endomorphisme est unitaire si et seulement si sa matrice dans une base orthonormale est unitaire.

Et un endomorphisme est hermitien si et seulement si sa matrice dans une base orthonormale est hermitienne.

Les valeurs propres d'une matrice hermitienne (ou d'un endomorphisme hermitien) sont toutes réelles.

Le polynôme caractéristique de f^* est le conjugué du polynôme caractéristique de f

On pourra consulter la partie 21.10.7 pour plus d'informations sur la structure de l'ensemble des endomorphismes unitaires.

Le spectre d'une matrice unitaire est inclus dans le cercle unité ; son déterminant appartient donc aussi à ce cercle unité.

Proposition 1140 (Sur les endomorphismes hermitiens) *L'image et le noyau d'un endomorphisme hermitien sont orthogonaux.*
Les sous-espaces propres d'un endomorphisme hermitien sont en somme directe orthogonale.
Si f est un endomorphisme hermitien et si F est un sous-espace stable par f alors $F^\perp$ est stable par f qui induit sur cet espace un endomorphisme hermitien.
Si f est un endomorphisme hermitien, alors f est diagonalisable dans une certaine base orthonormale.

Démonstration : Même principe que dans le cas euclidien.□

De même, si M est une matrice complexe hermitienne, alors il existe P unitaire telle que ${}^t\overline{P}.M.P$ soit diagonale.

29.6.3 Formes quadratiques sur un espace hermitien E

Voir la partie28.4.4.

Chapitre 30

Algèbre linéaire en dimension finie

30.1 Généralités

Définition 1141 *Un espace vectoriel est dit* **de dimension finie** *lorsqu'il admet une base de cardinal finie. Dans le cas contraire il est dit* **de dimension infinie**.
Dans un espace fini le cardinal d'une base est appelé **dimension** *de l'espace (il faudra voir plus loin que toutes les bases ont alors même cardinal). Un sous-espace vectoriel d'un espace vectoriel E est dit de* **codimension finie** *si la dimension de l'espace quotient est finie. On appelle alors* **codimension** *de cet espace la dimension de l'espace quotient. Sinon il est dit de* **codimension infinie**.

Dans la suite E désigne un $\mathbb{K}$-espace vectoriel de dimension finie.

Théorème 1142 (Théorème de la base incomplète) *Si I est une famille libre et K une famille génératrice finie, avec $I \subset K$, alors il existe J avec $I \subset J \subset K$ tel que J soit une base.*

Démonstration : On considère J libre maximal au sens du cardinal vérifiant $I \subset J \subset K$. Si toute famille plus grande incluse dans K est liée, alors tout élément de K est combinaison linéaire d'éléments de J ; tout élément de E s'écrit comme combinaison d'éléments de K, qui eux-mêmes s'écrivent comme combinaisons linéaires d'éléments de J. Donc la famille J est génératrice.□

Lemme 1143 (Lemme de Steinitz) *Si E est non réduit à $\{0\}$, E admettant une famille génératrice I de cardinal n, toute famille de $n+1$ vecteurs (ou plus) est liée.*

Démonstration : • Le cas $n = 1$ est trivial. On procède alors par récurrence.

• Supposons la propriété vraie pour $n < N$ et soit $n = N$; supposons E admettant une famille génératrice $b_1, ..., b_n$ de cardinal n. Notons F le sous-espace vectoriel de E engendré par $b_1, ..., b_{n-1}$.

• Donnons-nous une famille $e_1, ..., e_{n+1}$ de E.

• Soit, pour $i \in [1, n+1]$, $f_i \in F$ et $\lambda_i \in \mathbb{K}$ tel que $e_i = f_i + \lambda_i\, b_n$.

• Si tous les λ_i sont nuls, alors l'hypothèse de récurrence appliquée à F permet de conclure.

• Supposons alors $\lambda_1 \neq 0$; dans ce cas, $b_1 = \frac{1}{\lambda_1}(e_1 - f_1)$.

• On peut alors exprimer $e_i - f_i$ en fonction de $e_1 - f_1$ pour tout i, puis $z_i = e_i - \frac{\lambda_i}{\lambda_1} e_1$ comme combinaison linéaire des y_i, donc comme éléments de F pour tout $i > 1$.

• L'hypothèse de récurrence permet alors de conclure que les z_i pour $i > 1$ sont liés.

• On écrit alors $\sum \alpha_i z_i = 0$, avec les α_i non tous nuls ; en développant $z_i = e_i - \frac{\lambda_i}{\lambda_1} e_1$, on obtient une combinaison linéaire nulle à coefficients non tous nuls des e_i. D'où le résultat.□

Théorème 1144 *Dans un espace de dimension finie, toutes les bases ont le même cardinal.*

Démonstration : Conséquence immédiate du lemme 1143.□

Théorème 1145 *Deux espaces vectoriels de dimension finie sur le même corps sont isomorphes si et seulement si ils ont même dimension.*

Démonstration : S'ils sont isomorphes, l'image d'une base de l'un par un isomorphisme est base de l'autre, donc ils ont même dimension. S'ils ont même dimension, alors avec $(e_1, ..., e_n)$ une base de l'un, et $(f_1, ..., f_n)$ une base de l'autre, l'application $\sum \lambda_i e_i \mapsto \sum \lambda_i f_i$ pour $(\lambda_1, ..., \lambda_n) \in \mathbb{K}^n$ est un isomorphisme, comme on le vérifie facilement (la fonction est bien définie car $(e_1, ..., e_n)$ est une base, linéarité évidente, injectivité immédiate car $(f_1, ..., f_n)$ est libre, surjectivité immédiate car $(f_1, ..., f_n)$ est génératrice).□

Théorème 1146 *F, sous-espace vectoriel de E, est égal à E si et seulement si $dim\ E = dim\ F$.*

Démonstration : Il est clair que si F et E sont égaux, alors ils ont même dimension. Réciproquement, s'ils ont même dimension, alors une base de F est une famille libre de même cardinal qu'une base de E, donc c'est une base de E (voir le lemme de Steinitz ci-dessus).□

Théorème 1147 *Tout sous-espace vectoriel de E (E est supposé non nul) admet un supplémentaire, et si $E = F \oplus G$, alors $dim\ E = dim\ F + dim\ G$, et une base de E s'obtient en réunissant une base de F et une base de G.*

Démonstration : Conséquence du théorème de la base incomplète.□

Maintenant quelques théorèmes faciles, sans démonstration, mais fort pratiques néanmoins :

Théorème 1148 *Etant donnés F et G des sous-espaces vectoriels de E, on a $dim\ (F+G) + dim\ (F \cap G) = dim\ F + dim\ G$.*

Théorème 1149 *F et G sont supplémentaires dans E si et seulement si leur intersection est nulle et si la somme de leurs dimensions soit la dimension de E.*

Théorème 1150 *F et G sont supplémentaires dans E si et seulement si leur intersection est nulle et si leur somme est égale à E.*

Théorème 1151 *F et G sont supplémentaires dans E si et seulement si leur somme est égale à E et si la somme de leurs dimensions est égale à la dimension de E.*

Théorème 1152 *Etant donnés F sous-espace vectoriel de E, la dimension de E/F est égale à la dimension de E moins la dimension de F.*

Théorème 1153 *Si les E_i sont de dimension finie, alors $dim\ \Pi_{i=1..n} E_i = \sum_i dim\ E_i$.*

Théorème 1154 *Si E et F sont de dimension finie, alors $\mathcal{L}(E, F)$ est de dimension finie et $dim\ \mathcal{L}(E, F) = dim\ E.dim\ F$.*

Démonstration : On considère une base de E et une base de F, notées respectivement (e_i) et (f_i) ; alors les $\delta_{i,j}$ définies par $\delta_{i,j}(e_i) = f_j$ et $\delta_{i,j}(e_l) = 0$ pour $l \neq i$ forment une base de $\mathcal{L}(E, F)$.$\square$

Définition 1155 *On appelle* **rang** *d'une famille finie de vecteurs la dimension de l'espace vectoriel que cette famille engendre.*
On appelle **rang** *d'une application linéaire la dimension de son image lorsque celle-ci est finie. C'est en fait le rang de l'image d'une base lorsque l'espace de départ est de dimension finie*

Théorème 1156 (Théorème du rang) *Si $f \in \mathcal{L}(E, F)$ et E et F de dimension finie, alors $Im\ f$ et $Ker\ f$ sont de dimension finie, et $dim\ E = dim\ Im\ f + dim\ Ker\ f$.*

Démonstration : On considère un supplémentaire du noyau, et on montre qu'il est isomorphe à l'image de f (voir théorème 970).$\square$

Corollaire 1157 *Soit f une application d'un espace vectoriel E vers un espace vectoriel F, avec E et F de dimension finie et de même dimension, alors f est un isomorphisme si et seulement si l'une des propriétés suivantes est vérifiée :*
- *f a un noyau réduit à un singleton*
- *f est injective*
- *f est surjective*
- *f est bijective*
- *Le rang de f est égal à la dimension de E*

Proposition 1158 (Quelques propriétés du rang) • *Le rang d'une famille finie de vecteurs est invariant par permutations*
• *Multiplier un vecteur d'une famille par un scalaire non nul ne change pas le rang de la famille*
• *On ne change pas le rang d'une famille de vecteurs en additionnant à un vecteur une combinaison linéaire des autres vecteurs*
• *Le rang d'une famille de vecteurs est le même si l'on suppirme un vecteur nul*

30.2 Dualité en dimension finie

30.2.1 Dualité simple

E est un espace vectoriel normé de dimension finie n.

Définition 1159 *Etant donnée une base $(e_i)_{i\in[1,n]}$ de E, on appelle* **formes coordonnées associées** *les n formes linéaires e_j définies par $e_j^*(e_i) = \delta_{i,j}$.*

On note que $x = \sum_j e_j^*(x).e_j$.

Théorème 1160 • $dim\ E^* = dim\ E$.
• *La famille des e_j^* est une base de E^* ; on l'appelle* **la base duale** *de la base (e_i).*
• *Tout f appartenant à E^* s'écrit $f = \sum_i f(e_i).e_i^*$*
• *Pour toute base (f_i) de E^*, il existe une base de E dont la base duale est la base des (f_i).*

Démonstration : La dimension de E^* est égale à la dimension de E, car $dim\ \mathcal{L}(E,K) = dim\ E.dim\ K$; il suffit donc de montrer que la famille est libre. On se donne une combinaison linéaire nulle des e_j^* ; en considérant l'image de la combinaison linéaire nulle de e_i on voit que le coefficient de e_i^* est nulle pour tout i.
On se donne une application ϕ qui à x associe $(f_1^*(x), ..., f_n^*(x))$. On montre sans trop de difficultés que ϕ est un isomorphisme de E dans K^n. Donc il existe une base (e_i) telle que $\phi(e_i) = (d_{1,i}, ..., \delta_{i,i}, ...\delta_{n,i})$; cette base convient.□

Théorème 1161 *Soit E et F des $\mathbb{K}$-espace vectoriel de dimensions finies (non nécéssairement égales). Alors l'application transposition qui à $f \in \mathcal{L}(E, F)$ associe ${}^t f \in \mathcal{L}(F^*, E^*)$ est un isomorphisme. En outre le rang de ${}^t f$ est égal au rang de f et*

$$Ker\ {}^t f = (Im\ f)^\perp$$

Démonstration : Facile !□

30.2.2 Bidual

E est un espace vectoriel normé de dimension finie n.

Théorème 1162 (Bidual) *L'application de E dans E^{**} qui à x associe la fonction qui à ϕ dans E^* associe $\phi(x)$ est un isomorphisme ; on l'appelle* **isomorphisme canonique** *de E dans E^{**}. On peut donc identifier E et E^{**}.*

Démonstration : Les dimensions de E et E^{**} étant égales (on est toujours dans le cadre de la dimension finie) il suffit de voir que cette fonction est un isomorphisme. Supposons x d'image nulle ; alors quel que soit f, $f(x) = 0$. Si x est non nul, alors on peut l'appeler e_1 et le compléter en une base e_i ; alors on constate que $e_i^*(x) = 0$ pour tout i, ce qui est contradictoire.□

30.2.3 Orthogonalité

E est un espace vectoriel normé de dimension finie n.

Théorème 1163 *Quel que soit F sous-espace vectoriel de E, on a :*
- $dim\ E = dim\ F + dim\ F^\perp$
- $F^{\perp^o} = F$

Démonstration : Si $F = \{0\}$ alors le résultat est clair. Sinon, on considère une base $(f_i)_{i\in[1,f]}$ de F, et on la complète en une base $(f_i)_{i\in[1,n]}$ de E. On considère la base duale (f_i^*) ; il est clair alors que les $(f_i)_{i>f}$ forment une famille libre génératrice de $F^\perp$, d'où la première assertion.
On constate bien que $F^{\perp^o}$ est inclus dans F ; la dimension permet de conclure.□

Théorème 1164 *Quel que soit F sous-espace vectoriel de E', on a :*
• $dim\ E = dim\ F + dim\ F^o$
• $F^{o\perp} = F$

Démonstration : La méthode est la même que précédemment ; il faut juste se souvenir que toute base du dual est la base duale d'une certaine base.□

Théorème 1165 *Avec ϕ l'isomorphisme canonique de E dans E^{**}, pour tout F sous-espace vectoriel de E, on a $F^{\perp^\perp} = \phi(F^{\perp^o}) = \phi(F)$.*

Démonstration : Il n'y a qu'à l'écrire et ça marche tout seul...□

30.3 Calcul matriciel

30.3.1 Bases sur les matrices

Définition 1166 (Définitions de base sur les matrices) *On appelle* **matrice de type** (n,p) **sur un corps** $\mathbb{K}$ *toute application de* $[1,n] \times [1,p]$ *(intervalles de* $\mathbb{N}$*) dans* $\mathbb{K}$*. On la représente généralement comme suit :*

$$\begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,p} \\ m_{2,1} & m_{2,2} & \dots & m_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,p} \end{pmatrix}$$

On note $M_{n,p}(\mathbb{K})$ *l'ensemble des matrices de type* (n,p) *sur le corps* $\mathbb{K}$*.*
On appelle **matrice ligne** *une matrice de type* $(1,p)$*, et bf matrice colonne une matrice de type* $(n,1)$*.*
On appelle **matrice extraite** *d'une matrice de type* (n,p) *la restriction de cette matrice à* $I \times J$*, avec* $I \subset [1,n]$ *et* $J \subset [1,p]$*.*
On appelle i**-ième vecteur-ligne** *de la matrice* M *de type* (n,p) *la restriction de cette matrice à* $\{i\} \times [1,p]$*. On peut identifier un vecteur-ligne à un élément de* $\mathbb{K}^p$*.*
On appelle j**-ième vecteur-colonne** *de la matrice* M *de type* (n,p) *la restriction de cette matrice à* $[1,n] \times \{j\}$*. On peut identifier un vecteur-colonne à un élément de* $\mathbb{K}^n$*.*
On appelle **matrice associée** *à un morphisme* f *de l'espace vectoriel* E *de dimension* p *dans l'espace vectoriel* F *de dimension* n *et aux bases* $B = (e_i)$ *et* $B' = (f_i)$ *de* E *et* F *respectivement la matrice* M *de type* (n,p) *telle que* $M_{i,j} = f_i^*(e_j)$*. On la note* $Mat_{B,B'}(f)$*.*
Inversement, on appelle **application linéaire canoniquement associée à la matrice** M *le morphisme de* $\mathbb{K}^p$ *dans* $\mathbb{K}^n$ *dont la matrice associée est* M*.*

Proposition 1167 *Une matrice de type* n,p *admet* $C_n^{n'}.C_p^{p'}$ *matrices extraites de type* (n',p')*.*

Proposition 1168 E *et* F *étant de dimension respectives* p *et* n *sur le corps* $\mathbb{K}$*, on a*

$$\mathcal{L}(E,F) \simeq \mathcal{M}_{n,p}(\mathbb{K})$$

via l'isomorphisme $f \mapsto Mat_{B,B'}(f)$*.*
Avec $E = \mathbb{K}^p$ *et* $F = \mathbb{K}^n$*, et* B *et* B' *les bases canoniques, on a alors un isomorphisme canonique entre* $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ *et* $\mathcal{M}_{n,p}(\mathbb{K})$*.*

Démonstration : Evident !□

Proposition 1169 (Matrice canonique d'une application linéaire en dimension finie) *Soit f une application linéaire entre E, espace vectoriel de dimension p, et F, espace vectoriel de dimension n ; soit r le rang de f. Alors il existe une base B et E et une base B' de F telles que*

$$Mat_{B,B'}(f) = M$$

avec

$$M_{i,j} = 1 \text{ si } i = j \leq r$$

$$M_{i,j} = 0 \text{ sinon}$$

On appelle cette matrice **matrice canonique de** f.

Démonstration : Considérer une base B_1 d'un supplémentaire du noyau de f ; considérer $B_1' = f(B_1)$; considérer B_2 une base du noyau de f, et compléter B_1' en une base B'. Il reste à considérer $B = B_1 \cup B_2$.□

Remarque 1170 *La matrice d'une forme linéaire est une matrice-ligne.*

Définition 1171 (produit matriciel) *On appelle* **matrice produit** *des matrices A et B de types respectifs (n,q) et (q,p) la matrice C de type (n,p) telle que*

$$C_{i,j} = \sum_{k \in [1,q]} A_{i,k}.B_{k,j}$$

On note $C = A \times B$ ou $C = A.B$.

Proposition 1172 • *Le produit matriciel défini ci-dessus est associatif et bilinéaire.*
• *Avec $A = Mat_{B,B'}(f)$ et $B = Mat_{B',B''}(g)$, $B \times A = Mat_{B,B''}(g \circ f)$*
• *Avec $x \in E$, X le vecteur des coordonnées de x dans la base B, alors les coordonnées de $y = f(x)$ dans la base B' sont celles de Y, avec $Y = M \times X$, et $M = Mat_{B,B'}(f)$.*

Démonstration : Facile (et non moins important).□

30.3.2 Espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$

E et F $\mathbb{K}$-espace vectoriel de dimensions respectives p et n.

Définition 1173 *On appelle* **matrices élémentaires de type** (n,p) *les matrices $E_{i,j}$ avec $E_{i,j}(a,b) = \delta_{a,i}.\delta_{b,j}$; c'est à dire les matrices de type (n,p) ne comportant qu'un 1 et des 0 partout ailleurs.*

Proposition 1174 *$\mathcal{M}_{n,p}(K)$ est un $\mathbb{K}$-espace vectoriel pour l'addition terme à terme et pour la multiplication terme à terme par un scalaire.*
L'application de $\mathcal{L}(E,F)$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ qui à une application f associe $Mat_{B,B'}(f)$ est un isomorphisme.
Les matrices élémentaires forment une base de cet espace vectoriel , qui est donc de dimension $n.p$.

30.3.3 Transposition de matrices

Définition 1175 (Transposée d'une matrice) *Etant donnée une matrice M on appelle* **matrice transposée** *de M la matrice N avec $N_{i,j} = M_{j,i}$. Si M est de type (n,p), alors N est de type (p,n).*
On note $N =^t M$.

Proposition 1176 • *L'application $M \mapsto^t M$ est un isomorphisme entre $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{M}_{p,n}(\mathbb{K})$ en tant que $\mathbb{K}$-espaces vectoriels .*
• $^t(A \times B) =^t B \times^t A$
• $Mat_{B'^*,B^*}(^tf) =^t Mat_{B,B'}(f)$

30.3.4 Le cas des matrices carrées : la $\mathbb{K}$-algèbre $\mathcal{M}_{n,p}(\mathbb{K})$

Définition 1177 *On appelle* **matrice carrée** *une matrice de type (n,n) pour un certain n. On note $\mathcal{M}_n(\mathbb{K}) = \mathcal{M}_{n,n}(\mathbb{K})$.*
On appelle **matrice d'un endormophisme** *f associée à une base B la matrice $Mat_{B,B'}(f)$; on la note aussi $Mat_B(f)$.*
On appelle **diagonale** *d'une matrice carrée M de type (n,n) le vecteur $(M_{1,1},...,M_{i,i},...,M_{n,n})$.*
On appelle **trace** *d'une matrice carrée M la somme $\sum_{i\in[1,n]} M_{i,i}$. On la note $tr(M)$. L'application $M \to tr(M)$ est une application linéaire.*
On appelle **matrice unité d'ordre** *n la matrice M avec $M_{i,j} = \delta_{i,j}$. C'est la matrice de l'endomorphisme identité.*
On appelle **matrice scalaire** *une matrice égale à $\lambda.I$ avec λ un scalaire et I une matrice unité.*
On appelle **matrice diagonale associée à un** *n***-uple** *m la matrice M de type (n,n) définie par $M_{i,i} = m_i$ et $M_{i,j} = 0$ si $i \neq j$. On note $M = diag(m)$.*
Une matrice est dite **symétrique** *si elle est égale à sa transposée.*
Une matrice M est dite **antisymétrique** *si elle est égale à l'opposée de sa transposée, c'est à dire si ${}^tM = -M$.*
Une matrice carrée est dite **triangulaire supérieure** *si $j < i \to M_{i,j} = 0$*
On note $\mathcal{T}_n^s$ l'ensemble des matrices carrées triangulaires supérieures d'ordre n.
Une matrice carrée est dite **triangulaire inférieure** *si $j > i \to M_{i,j} = 0$*
On note $\mathcal{T}_n^i$ l'ensemble des matrices carrées triangulaires inférieures d'ordre n.

Proposition 1178 • *$\mathcal{M}_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre, isomorphe à la $\mathbb{K}$-algèbre des endomorphismes de $\mathbb{K}^n$ (ou de E, pour tout E espace vectoriel de dimension n).*
• *Si M est inversible, alors sa transposée aussi et $({}^tM)^{-1} =^t (M^{-1})$.*
• *$tr(AB) = tr(BA)$ (que A et B soient carrées ou non, pourvu que le produit soit carré)*
• *Une matrice est scalaire si et seulement si c'est la matrice associée à un endomorphisme de la forme $x \mapsto \lambda.x$.*
• *L'ensemble des matrices diagonales de $\mathcal{M}_n(\mathbb{K})$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$.*
• *L'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{K})$ et l'ensemble des matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$; si $\mathbb{K}$ n'est pas de caractéristique 2, ils sont supplémentaires ; ils sont alors de dimensions respectives $\frac{n.(n+1)}{2}$ et $\frac{n.(n-1)}{2}$, engendrés respectivement par les $E_{i,j} + E_{j,i}$ pour $i \leq j$ et par les $E_{i,j} - E_{j,i}$ pour $i < j$. Toute matrice carrée M s'écrit $A + S$, avec A antisymétrique et S antisymétriques, avec $A = \frac{1}{2}M -^t M$ et $S = \frac{1}{2}M +^t M$*
• *Le produit de deux matrices symétriques est symétrique si et seulement si les deux matrices commutent.*
• *L'ensemble des matrices triangulaires supérieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $n.(n+1)/2$.*
• *L'ensemble des matrices triangulaires inférieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $n.(n+1)/2$*

• *L'ensemble des matrices triangulaires inférieures est isomorphe à l'ensemble des matrices triangulaires supérieures*
• *Les éléments inversibles de l'ensemble des matrices triangulaires supérieures (resp. inférieures) sont les matrices triangulaires dont tous les coefficients diagonaux sont non nuls ; ils forment un sous-groupe multiplicatif de $\mathcal{M}_n(\mathbb{K})$.*
• *Etant donnée une matrice A de $\mathcal{M}_n(\mathbb{K})$, on appelle* **commutant de** *A le sous-ensemble de $\mathcal{M}_n(K)$ des matrices commutant avec A.*
• *Etant donnée A une matrice de $\mathcal{M}_n(\mathbb{K})$, on définit $A^0 = I$ avec I la matrice*

Démonstration : Seul le dernier point pose difficulté. Il nécéssitera le théorème de Cayley-Hamilton, qui montre que $A^n \in Vect(A^0, ..., A^{n-1})$. La suite est claire.$\square$

30.3.5 Changement de bases

Soit E espace vectoriel de dimension n, et (e_i) et (f_j) des bases de E.

Définition 1179 *On appelle* **matrice de passage** *de la base (e_i) à la base (f_j) la matrice P de type (n, n) définie par $P_{i,j} = e_i^*(f_j)$; on la note $P_{(e_i),(f_j)}$. Il s'agit en fait de la matrice $Mat_{(f_j),(e_i)}(I)$.*

Proposition 1180 • *Le produit de la matrice de passage de B à B' par la matrice de passage de B' à B'' est la matrice de passage de B à B''.*
• $P_{B,B'}^{-1} = P_{B',B}$
• *Etant donné X le vecteur des coordonnées de x dans une base B, les coordonnées de x dans la base B' sont données par X' avec $X' = P_{B',B}X$.*
• $Mat_{B',C'}(f) = P_{C',C}.Mat_{B,C}(f).P_{B,B'}$
• *Dans le cas d'un endomorphisme $Mat_{B'}(f) = P_{B',B}.Mat_B(f).P_{B,B'}$*

La matrice de passage de B à B' donne les coordonnées dans B en fonction des coordonnées dans B' et pas le contraire... La terminologie vaut ce qu'elle vaut ! On peut le retenir en considérant que la matrice de passage de la base B à la base B' est la matrice dans B de l'endomorphisme dont l'image de B est B'.

30.3.6 Groupe linéaire et groupe spécial linéaire

Voir 21.10.2.

30.3.7 Groupe orthogonal réel et groupe spécial orthogonal réel

Voir 21.10.4.

30.3.8 Rang d'une matrice

Définition 1181 (Matrice associée à une famille finie de vecteurs) *Etant donnée une famille de vecteurs* $(x_1, ..., x_p)$ *d'un espace vectoriel de dimension finie* n *et une base* $(v_1, ..., v_n)$ *de* E*, on appelle* **matrice associée à la famille des** x_i **et à la base des** v_i *la matrice* M *définie par* $M_{i,j} = v_i^*(x_j)$.
On appelle **rang d'une matrice** M *le rang de la famille de ses vecteurs colonnes. On le note* $rg(M)$.
Une matrice N *extraite de* M*, avec* N *carrée inversible d'ordre le rang de* M*, est appelée* **matrice principale** *de* M.

Proposition 1182 • *Le rang d'une matrice associée à un système de vecteurs et à une base est indépendant de cette base.*
• $rg(Mat_{B,B'}(f)) = rg(f)$
• $rg({}^tM) = rg(M) \leq min(n, p)$*, avec* M *de type* (n, p).
• $rg(M) = n \iff M$ *surjective*
• $rg(M) = p \iff M$ *injective*
• $rg(M.M') \leq min(rg(M), rg(M'))$

Démonstration : Refaire cette preuve facile réveillera les neurones consacrés à l'algèbre linéaire chez ceux pour qui tout cela est un peu oublié...□

Proposition 1183 *Le rang d'une matrice* M *non nulle est égal à* n *maximal tel qu'il existe une matrice extraite inversible de type* (n, n) *de* M.

Démonstration : • Il est clair que le rang de M est supérieur au rang de toute matrice extraite inversible (considérer une combinaison linéaire nulle des vecteurs correspondants de M).
• Etant donné r le rang de M il existe une famille libre de r vecteurs parmi la famille des vecteurs colonnes de M.
On peut alors considérer la transposée de cette matrice ; son rang est le même, et donc on peut se restreindre au même nombre de colonnes.□

30.3.9 Matrices équivalentes, matrices semblables

⊡ Matrices équivalentes

Définition 1184 (Matrices équivalentes) *Deux matrices A et B de même type (n,p) sont dites* **équivalentes** *si il existe P et Q des matrices inversibles de types respectifs (p,p) et (n,n) telles que*

$$B = Q.A.P$$

Quelques propriétés immédiates :

Proposition 1185 • *Il s'agit d'une relation d'équivalence.*
• *Deux matrices de même type sont équivalentes si et seulement si elles représentent un même morphisme dans des bases différentes.*
• *Deux matrices de même type sont équivalentes si et seulement si elles ont le même rang.*

Dans la deuxième propriété, il faut bien voir qu'il faut changer éventuellement à la fois la base de départ et la base d'arrivée.

⊡ Matrices semblables

Définition 1186 (Matrices semblables) *Deux matrices carrées A et B de même type sont dites semblables s'il existe une matrice P telle que*

$$A = P.B.P^{-1}$$

Proposition 1187 • *Il s'agit d'une relation d'équivalence*
• *Deux matrices sont semblables si elles représentent un même endomorphisme dans deux bases différentes*
• *Deux matrices semblables ont même rang*
• *Deux matrices semblables ont même trace*

Démonstration : Facile pour les trois premiers points ; pour le quatrième il suffit de se rappeller que la trace de AB est égale à la trace de BA.□

Cette fois-ci, contrairement au cas des matrices équivalentes, deux matrices de même rang ne sont pas nécéssairement semblables.

⚠ La deuxième caractérisation fait appel à un endomorphisme et pas une application linéaire quelconque ; la base de départ est la même que celle d'arrivée.

30.3.10 Cofacteurs

Définition 1188 (Mineur et cofacteur) *Le déterminant de la matrice M à laquelle on ôte la j-ième colonne et la i-ième ligne est appelé* **mineur** (i, j) **de** M*. On le note généralement $\Delta_{i,j}$.*
Le déterminant de la matrice M à laquelle on ôte la j-ième colonne pour la remplacer par le i-ième vecteur de la base est appelé **cofacteur** (i, j) *de M. On le note généralement $\gamma_{i,j}(M)$. On la note généralement $com(M)$.*
La matrice γ ainsi définie est appelée **comatrice** *de M.*
La matrice ${}^t\gamma$ est appelée **matrice complémentaire** *de M. On la note généralement $\tilde{M}$.*

Pour y voir plus clair, le mineur (i, j) de M est :

$$\begin{vmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & M_{1,j+1} & \dots & M_{1,n} \\
M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & M_{2,j+1} & \dots & M_{2,n} \\
M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & M_{3,j+1} & \dots & M_{3,n} \\
M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & M_{4,j+1} & \dots & M_{4,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & M_{i-1,j+1} & \dots & M_{i-1,n} \\
M_{i,1} & M_{i,2} & M_{i,3} & \dots & M_{i,j-1} & M_{i,j+1} & \dots & M_{i,n} \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & M_{i-1,j+1} & \dots & M_{i-1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & M_{n,j+1} & \dots & M_{n,n}
\end{vmatrix}$$

et le cofacteur (i, j) de M est :

$$\begin{vmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & 0 & M_{1,j+1} & \dots & M_{1,n} \\
M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & 0 & M_{2,j+1} & \dots & M_{2,n} \\
M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & 0 & M_{3,j+1} & \dots & M_{3,n} \\
M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & 0 & M_{4,j+1} & \dots & M_{4,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\
M_{i,1} & M_{i,2} & M_{i,3} & \dots & M_{i,j-1} & 1 & M_{i,j+1} & \dots & M_{i,n} \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & 0 & M_{n,j+1} & \dots & M_{n,n}
\end{vmatrix}$$

qui est d'ailleurs égal à

$$\begin{vmatrix}
M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & 0 & M_{1,j+1} & \dots & M_{1,n} \\
M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & 0 & M_{2,j+1} & \dots & M_{2,n} \\
M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & 0 & M_{3,j+1} & \dots & M_{3,n} \\
M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & 0 & M_{4,j+1} & \dots & M_{4,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\
0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\
M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & 0 & M_{n,j+1} & \dots & M_{n,n}
\end{vmatrix}$$

Proposition 1189

$$\gamma_{i,j} = (-1)^{i+j}.\Delta_{i,j}$$

Proposition 1190 *La comatrice de la transposée de M est la transposée de la comatrice de M.*

Il est nécéssaire pour la suite d'avoir lu la partie 28.3.

Théorème 1191 *Si $\tilde{M}$ désigne la matrice complémentaire*

$$\tilde{M}.M = det\ M.Id$$

et

$$M.\tilde{M} = det\ M.Id$$

en particulier, si M est inversible, $M^{-1} = \frac{1}{det\ M}\tilde{M}$.

Démonstration : Considérons le terme (i,j) de la matrice $\tilde{M}.M$. Il s'agit de $\sum_{k=1..n} \gamma_{k,i}.M_{k,j}$.
Considérons le terme (i,j) de la matrice $det\ M.Id$.
Il s'agit de $\delta_{i,j}.det\ M$.

On cherche donc à montrer que $\delta_{i,j}.det\ M = \sum_{k=1..n} \gamma_{k,i}.M_{k,j}$.

On distingue deux cas :

• $i \neq j$

On considère alors la matrice M, sur laquelle on remplace la colonne i par la colonne j (on ne les permute pas, on supprime la colonne i, et on copie la colonne j à la place). Le déterminant de la matrice obtenue est nul, puisqu'on a deux fois la même colonne.

On développe par rapport à la colonne i. On obtient

$$\sum_{k=1..n} M_{k,j}.\gamma_{k,i} = 0$$

et donc le résultat dans ce cas est bien montré.
• $i = j$
Dans ce cas on considère le développement de M par rapport à la i-ième colonne et on obtient

$$det\ M = \sum_{k=1..n} \gamma_{k,i}.M_{k,i} = \sum_{k=1..n} \gamma_{k,i}.M_{k,j}$$

d'où le résultat souhaité.
Le second résultat se déduit du premier en considérant la transposée de chacun des deux produits.□

30.4 Opérations sur les lignes et les colonnes

On trouvera des manipulations proches de celles décrites ici dans la partie sur les déterminants28.3. La proposition 811 illustre aussi des opérations sur les lignes et les colonnes.

FLEMMARD -> determinants -> resol syst. lineaires

Définition 1192 *On appelle* **système d'équations linéaires** *une équation de la forme $MX = Y$, où M (matrice) et Y (vecteur) sont donnés et où X est l'inconnue.*
Les **opérations sur les lignes et les colonnes** *d'une matrice ou d'un système linéaire sont par définition :*
• l'addition d'une ligne (resp. colonne) i à une ligne (resp. colonne) $j \neq i$
• la multiplication d'une ligne (resp. colonne) i par un scalaire $\lambda \neq 0$
• la permutation de deux lignes (resp. colonnes) i et $j \neq i$
Ces opérations seront notées respectivement :
• $L_i \leftarrow L_i + L_j$ (resp. $C_i \leftarrow C_i + C_j$)
• $L_i \leftarrow L_i$ (resp. $C_i \leftarrow C_i$)
• $L_i \leftrightarrow L_j$ (resp. $C_i \leftrightarrow C_j$)
On pourra éventuellement ajouter à une ligne (resp. une colonne) une autre ligne (resp. colonne) multipliée par un scalaire λ ; cela se notera $L_i \leftarrow L_i + \lambda L_j$ (resp. $C_i \leftarrow C_i + \lambda C_j$).

Exemples :
• Sur un système :
Avant :

$$\begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 1 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Après $L_1 \leftarrow L_1 + 2\ L_2$:

$$\begin{pmatrix} 1 & 4 & -1 \\ 0 & 1 & 1 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix}$$

• Sur une matrice Avant :

$$\begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix}$$

Après $C_1 \leftrightarrow C_2$

$$\begin{matrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 9 \end{matrix}$$

Proposition 1193 *Les opérations sur les lignes correspondent à des multiplications à droite par des matrices inversibles ; les opérations sur les colonnes correspondent à des multiplications à gauche par des matrices inversibles. L'inverse d'une opération sur les lignes (resp. les colonnes) est une opération sur les lignes (resp. les colonnes).*

Démonstration :

Proposition 1194 • *Le déterminant d'une matrice n'est pas modifié en ajoutant à une ligne une combinaison linéaire des autres lignes.*
• *Multiplier une ligne par λ multiplie le déterminant par λ.*
• *Permuter des lignes multiplie le déterminant par -1.*
Les mêmes résultats sont valables pour les colonnes.

Démonstration : Cela découle immédiatement des propriétés du déterminant, et du fait que le déterminant d'une matrice est égal au déterminant de sa transposée. Ceux qui ont besoin de rappels peuvent consulter la partie 28.3.□

Théorème 1195 (Formules de Cramer) *Considérons le système d'équations linéaire $MX = Y$, avec M de type (n, n) :*

$$M = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,n} \end{pmatrix}$$

$$Y = (y_1, ..., y_p)$$

On suppose en outre que M est inversible.
Alors X est solution, avec

$$x_i = \frac{\begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & Y_1 & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & Y_2 & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & Y_n & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}}{det\ M}$$

Démonstration : La solution X est clairement unique, car par inversibilité de M $X = M^{-1}Y$.

On a alors $Y = \sum_k X_k C_k$ avec C_k la k-ième colonne de M.

Donc, quel que soit i, on a alors $det\ {M_k}^{(i)} = \sum X_k det\ {M'_k}^{(i)}$
avec

$$
{M_k}^{(i)} = \begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & Y_1 & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & Y_2 & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & Y_n & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}
$$

et

$$
{M'_k}^{(i)} = \begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & M_{1,k} & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & M_{2,k} & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & M_{n,k} & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}
$$

On en déduit donc, en supprimant de la somme $\sum X_k C_k$ les termes nuls, $X_i det\ M = {M'_i}^{(i)}$, ce qui est précisément le résultat désiré.□

Théorème 1196 (Méthode du pivôt de Gauss) *La* **méthode de Gauss** *consiste à :*
1) permuter les lignes pour avoir un coefficient non nul en haut à gauche de la matrice ; ce coefficient est appelé **pivôt**
2) soustraire la première ligne multipliée par un coefficient adéquat à chacune des autres lignes de manière à avoir des zéros sur toute la première colonne en dehors du premier coefficient
3) Procéder de même sur la matrice extraite, simplement dépourvue de sa première ligne et sa première colonne.
Le point $1)$ *pourra toujours être réalisé si on trouve toujours un coefficient non nul à échanger ; pour peu que la matrice soit inversible, cette condition sera toujours vérifiée. Si elle ne l'est pas, on peut travailler sur la matrice extraite par suppression de la première colonne.*
En réitérant cette méthode, on arrive à obtenir une matrice triangulaire supérieure. En fait la matrice obtenue est de la forme illustrée sur la figure 30.1, du moins après permutation des colonnes.

La matrice ainsi obtenue est donc beaucoup plus maniable : le calcul du déterminant (si la matrice est carré) se résume à la multiplication des éléments diagonaux, la résolution de systèmes linéaires est plus aisée (il convient de noter pour cela que les opérations sur les colonnes sont de simples permutations sur les inconnues), le calcul du rang est immédiat (nombre d'éléments avant le dernier élément non nul de la dernière colonne).

Afin de minimiser les pertes de précision d'un calcul numérique, il est préférable de choisir un pivôt grand en valeur absolue.

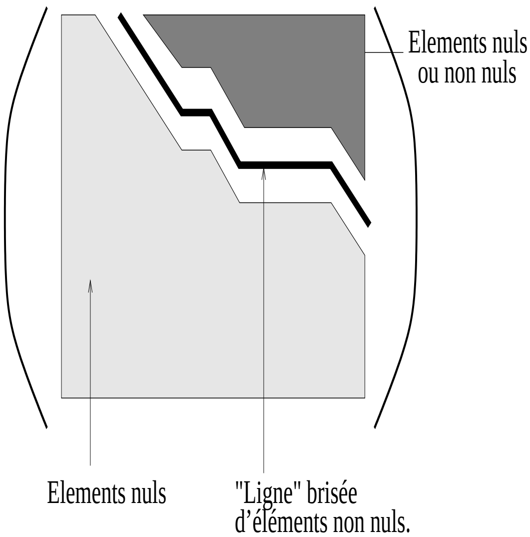


FIG. 30.1 – Matrice obtenue après pivôt de Gauss. La "ligne brisée" évoquée comporte soit des sauts à droite, soit des sauts en diagonale en bas à droite. Le premier élément de la ligne brisée se trouve quelque part sur la première ligne (éventuellement en haut à gauche).

La méthode du **pivôt total** consiste à chercher le pivôt non pas seulement sur la colonne en cours, mais d'éventuellement permuter les colonnes pour avoir un pivôt plus grand. Par opposition au pivôt total, la méthode ci-dessus est dite **pivôt partiel**.

Théorème 1197 (Décomposition $A = LU$) *Etant donnée une matrice A supposée inversible, on définit $a_k = |(A_{i,j})_{i,j\leq k}|$, déterminant de la matrice obtenue en se restreignant aux k premières lignes et k premières colonnes.*
On appelle **décomposition** *$A = LU$ un produit du type $A = LU$ avec L matrice triangulaire inférieure ne comportant que des 1 sur la diagonale, U matrice triangulaire supérieure, Alors il existe une décomposition $A = LU$ si et seulement si les a_k sont non nuls pour tout k dans $[1, n]$.*

Démonstration : Il suffit d'utiliser la méthode de Gauss, en considérant les matrices correspondant aux opérations sur les lignes et les colonnes. C'est-à-dire que l'on obtient un produit $\pi_{i=1}^{n} M_{n-i} A = U$, avec M_i la matrice correspondant à l'opération sur les lignes et les colonnes effectué à la i-ième étape. Mais l'inverse d'une opération sur les lignes ou les colonnes est une opération sur les lignes ou les colonnes (facile à trouver) ; donc on peut aussi écrire $A = \pi_{i=1}^{n} N_i U$, avec N_i l'inverse de M_i. Le produit des N_i est bien de la forme désirée, triangulaire supérieure à diagonale unité, comme on s'en convaincra en considérant la stabilité de l'ensemble des matrices triangulaires supérieures à diagonale unité par multiplication par les matrices des opérations sur les lignes et les colonnes.

Proposition 1198 *Si A est inversible, il existe une matrice de permutation P telle que $PA = LU$.*

Démonstration : Si a_k est nul, il existe nécéssairement une permutation de lignes qui arrange ça, sinon A ne pourrait pas être de rang plein - il suffit donc de multiplier les différentes permutations de lignes nécéssaires pour obtenir une matrice vérifiant les conditions demandées.

Théorème 1199 (Décomposition de Cholesky) *On appelle* **décomposition de Cholesky** *ou* **décomposition** *$A =^t RR$, un produit de la forme $A =^t RR$, avec R triangulaire inférieure inversible.*
A admet une décomposition de Cholesky si et seulement si A est symétrique définie positive.

Démonstration : On pourra consulter [7] pour cette preuve.□

30.5 Matrices par blocs

Les méthodes expliquées ci-dessous dans le cas de quatre blocs, se généralisent naturellement au cas d'un nombre quelconque de blocs.

30.5.1 Produit par blocs

Etant données les matrices A, B, C, D, A', B', C' et D', avec

$$largeur(A) = largeur(C) = hauteur(A') = hauteur(B')$$

$$largeur(B) = largeur(D) = hauteur(C') = hauteur(D')$$

on considère les matrices M et M' définies comme suit :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

$$M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

Alors

$$M.M' = \begin{pmatrix} A.A' + B.C' & A.B' + B.D' \\ C.A' + D.C' & C.B' + D.D' \end{pmatrix}$$

30.5.2 Inverse par blocs

Si $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, avec A et B inversibles, alors M est inversible et $M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}.C.B^{-1} \\ 0 & B^{-1} \end{pmatrix}$.

30.5.3 Déterminant par blocs

Si $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, alors $det\ M = det\ det\ A.det\ B$.

30.6 Exercices sur les matrices

Exercice 1200 *Une forme linéaire f sur $\mathcal{M}_n(\mathbb{K})$ telle que $f(XY) = f(YX)$ est proportionnelle à l'application $M \mapsto tr(M)$.*

Démonstration : Il suffit de considérer X et Y des matrices élémentaires.□

Exercice 1201 *Une matrice de $\mathcal{M}_n(\mathbb{K})$ commute avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$ si et seulement si elle est scalaire.*

Démonstration : Simplement vérifier qu'une telle matrice commute avec les matrices élémentaires.□

30.7 Zoologie sur les matrices et leurs déterminants

Définition 1202 (Matrice circulante) *On appelle* **matrice circulante associé au n-uple** *$(x_1, ..., x_n)$ la matrice M définie par $M_{i,j} = x_{j-i}$ (modulo n), c'est à dire*

$$M = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ x_n & x_1 & x_2 & \ddots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \ddots & x_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \dots & x_1 \end{pmatrix}$$

Proposition 1203 *L'ensemble des matrices circulantes de type (n, n) est engendré par la matrice suivante :*

$$M = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Démonstration : Il suffit de voir que la matrice circulante associée à $(x_1, ..., x_n)$ est la matrice
$$x_1.M^0 + x_2.M^1 + ... + x^n.M^{n-1}$$
et le résultat est acquis.□

30.8 Zoologie de la dualité en dimension finie

30.8.1 Polynômes de Lagrange

On considère l'espace $E = \mathbb{R}_n[X]$ des polynômes à une indéterminée et de degré au plus n. Soit a_0, ... , a_n des réels deux à deux distincts. On définit $n+1$ formes linéaires sur E par $f_i(P) = P(a_i)$.

Proposition 1204 *Les polynômes de Lagrange f_i forment une base de E.*

Démonstration : Puisque la dimension de E est égale à la dimension de E^*, il suffit de voir que la famille est de rang $n+1$, ce qui est vérifié si et seulement si l'espace vectoriel dual est de dimension 0. Supposons qu'un certain polynôme P appartienne à cet orthogonal, alors il s'annule en a_0, ... ,a_n ; donc il est nul, puisqu'il est de degré au plus n.□

Proposition 1205 *Les polynômes de lagrange (f_i) sont la base duale des P_i, avec*
$$P_i(x) = \Pi_{j\neq i}\frac{X - a_j}{a_i - a_j}$$

Démonstration : Facile, il suffit de vérifier que $P_i(a_j) = \delta_{i,j}$.□

Corollaire 1206 (Interpolation de Lagrange) *On en déduit notamment que tout polynôme de degré n s'écrit comme combinaison linéaire des P_i, les coefficients étant donnés par les f_i. C'est-à-dire que tout P de degré $\leq n$ s'écrit*
$$P = \sum_{i=1}^{n} f_i(P)P_i$$

Un exemple d'utilisation Maple :

Exemple Maple
$> \; interp([0,1,2,3],[exp(0),exp(1),exp(2),exp(3)],x);$
$\frac{1}{6}e^3x^3 - \frac{1}{2}e^3x^2 + \frac{1}{3}e^3x - \frac{1}{2}e^2x^3 + 2e^2x^2 - \frac{3}{2}e^2x + \frac{1}{2}ex^3 - \frac{5}{2}ex^2 + 3ex - \frac{1}{6}x^3 + x^2 - \frac{11}{6}x + 1$
Il est intéressant de tracer ensuite la courbe exponentielle et les graphes des interpolations à différents ordres superposées.

30.8.2 Définition d'un sous-espace vectoriel par une famille d'équations

Proposition 1207 *Soit F un sous-espace vectoriel de E de dimension p, E espace vectoriel de dimension finie n. Alors il existe $n-p$ formes linéaires linéairement indépendantes f_i telles que $F = \{x/\forall i \; f_i(x) = 0\}$ c'est à dire que F s'exprime comme intersection de $n-p$ hyperplans. Il est en outre impossible de définir F comme intersection de moins de $n-p$ hyperplans.*

Démonstration : Pour voir qu'une telle famille existe, il suffit de considérer une base de F et sa base duale. Pour vérifier qu'on ne peut faire moins, il suffit de considérer que pour tout G sous-espace vectoriel de E et tout H hyperplan de E on a $dim\,(G\cap H) \geq dim\,G - 1$ (par la formule $dimG + dimH = dim(G+H) + dim(G\cap H)$)$\square$

30.9 Approximation de fonctions holomorphes par des fractions rationnelles

Lemme 1208 *Soit K un compact de $\mathbb{C}$, inclus dans un ouvert Ω. Alors il existe $\Gamma_1, \Gamma_2, ..., \Gamma_n$ des segments orientés dans $\Omega \setminus K$ tels que pour toute fonction holomorphe f sur Ω et pour tout k dans K*

$$f(z) = \sum_{l=1}^{n} \frac{1}{2i\Pi} \int_{\Gamma_l} \frac{f(t)}{t-z} dt$$

Démonstration :

Lemme 1209 (Intuitif pour les topologistes dans l'âme !) *Il existe $\eta > 0$ tel que la distance entre un point k de K à un point du complémentaire de Ω soit toujours $> \eta$.*

Démonstration :

• En effet sinon on pourrait construire une suite de points $(k_n)_{n\in\mathbb{N}}$ de K à distance $\leq 1/n$ de $X \setminus K$.

• On pourrait alors extraire une suite convergente (par le théorème 199, puisque $\mathbb{C}$ est muni d'une topologie métrique !), et le point limite serait à une distance 0 du fermé complémentaire de Ω, et serait donc dans K sans être dans Ω ; ce qui est contradictoire.$\square$

• On construit alors une grille recouvrant le compact K, avec un maillage inférieur à $\eta/2$, comme indiqué sur la figure 30.2.

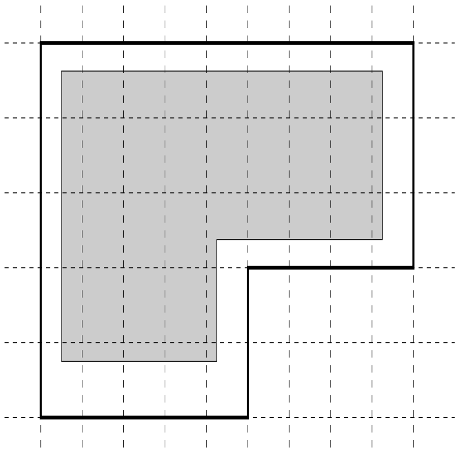


FIG. 30.2 – Un maillage.

• On considère alors les contours γ_1, ... , γ_p, orientés positivement, des carrés $C_1, ..., C_{p/4}$ intersectant K.

• On conserve alors seulement les segments des contours qui ne sont parcourus qu'une fois ; les autres étant parcourus deux fois, une fois dans chaque sens.

• On note ces segments orientés Γ_1, ... , Γ_n.

• On se donne alors f holomorphe sur Ω, et z à l'intérieur de l'un des carrés du maillage, $z \in K$. La fonction g qui à t appartenant à la réunion des γ_l associe $\frac{f(t)-f(z)}{t-z}$ est continue.

• On calcule alors

$$\sum_{l=1}^{n} \frac{1}{2i\Pi} \int_{\Gamma_l} \frac{g(t)}{t-z} dt$$

• Cette somme est égale à

$$\sum_{l=1}^{p} \frac{1}{2i\Pi} \int_{\gamma_l} \frac{g(t)}{t-z} dt$$

• Par le lemme 660, étendu au cas d'un carré, on en déduit que la somme est nulle.

• Donc

$$\sum_{l=1}^{n} \frac{1}{2i\Pi} \int_{\Gamma_l} \frac{f(t)}{t-z} dt$$

$$= \sum_{l=1}^{n} \frac{1}{2i\Pi} \int_{\Gamma_l} \frac{f(z)}{t-z} dt$$

$$= f(z) \sum_{l=1}^{n} \frac{1}{2i\Pi} \int_{\Gamma_l} \frac{1}{t-z} dt$$
$$= f(z)$$

• Il ne reste qu'à prolonger par continuité pour avoir le résultat souhaité.□

Théorème 1210 (Runge, version faible, lemme pour la version forte) *Soit K un compact de $\mathbb{C}$, inclus dans un ouvert Ω.*
Soit Z une partie de $\mathbb{C}$ contenant au moins un point dans chaque composante connexe de $\mathbb{C} \cup \{\infty\} \setminus K$.
Alors l'ensemble des fractions rationnelles dont les pôles sont inclus dans Z est dense dans l'ensemble des fonctions holomorphes sur Ω, pour la topologie de la convergence uniforme sur K.

Démonstration :

• Soit FR l'ensemble des fractions rationnelles dont les zéros sont inclus dans Z.

• D'après le corollaire 700, il suffit de montrer que toute forme linéaire continue nulle sur toutes les fractions rationnelles de FR est nulle sur toute application holomorphe sur Ω.

• D'après le théorème de représentation de Riesz, il nous suffit donc de montrer qu'étant donnée une mesure de Borel complexe μ sur K telle que l'intégrale pour μ sur K de tout élément de FR soit nulle, l'intégrale sur K pour μ de f holomorphe est nulle.

• Soit donc une telle mesure complexe μ, et f holomorphe sur Ω.

• Définissons, pour $t \in \mathbb{C} \cup \{\infty\} \setminus K$,

$$h(t) = \int_K \frac{d\mu(k)}{k-t} \tag{30.1}$$

• h est holomorphe, au vu de la proposition 657.

• Soit z dans Z et n'appartenant pas à K ; notons V_z la composante connexe de $\mathbb{C} \cup \{\infty\} \setminus K$ contenant z. On va montrer que h est nulle sur cette composante connexe ; pour cela, par le théorème 665, il sera suffisant de montrer que h est nulle sur un voisinage de z.

- Supposons tout d'abord $z = \infty$. L'objectif est donc de montrer que pour t assez grand en module, $h(t) = 0$.

Ecrivons, pour k dans K et t quelconque,

$$\frac{k-t}{=} - \frac{1}{k} \frac{1}{1-k/t} = -\sum_{l=0}^{\infty} \frac{k^n}{t^{n+1}}$$

La convergence étant uniforme en k pour t suffisamment grand, on peut alors intervertir l'intégrale et la somme dans l'équation 30.1 et on obtient bien $h(t) = 0$.

- Supposons maintenant que $z \neq \infty$

Ecrivons, pour k dans K et t quelconque,

$$\frac{1}{k-t} = \frac{1}{(k-z)-(t-z)} = \frac{1}{k-z} \frac{1}{1-\frac{t-z}{k-z}}$$

$$= \sum_{l=0}^{\infty} \frac{(t-z)^l}{(k-z)^{l+1}})$$

La convergence étant uniforme en k pour t suffisamment proche de z, on peut alors intervertir l'intégrale et la somme dans l'équation 30.1 et on obtient bien $h(t) = 0$.

• On applique alors le lemme 1208, pour pouvoir exprimer f comme une intégrale sur un contour hors de K :

$$\int_K f d\mu = \int_K \frac{1}{2i\Pi} \sum_l \int_{\Gamma_l} \frac{f(t)}{t-k} dt \, d\mu(k)$$

Et par Fubini 378,

$$= \frac{1}{2i\Pi} \sum_l \int_{\Gamma_l} f(t) \int_K \frac{1}{t-k} d\mu(k) \, dt$$

$$= -\frac{1}{2i\Pi} \sum_l \int_{\Gamma_l} f(t)h(t)dt$$

$$= 0$$

D'où le résultat tant attendu !□

On peut facilement passer à la topologie de la convergence uniforme sur tout compact.

En outre, le cas où K est simplement connexe (i.e. son complémentaire a une seule composante connexe) donne lieu à un corollaire important.

Corollaire 1211 (Corollaire important) • *Si f est une fonction holomorphe définie sur un ouvert Ω, si Z est un ensemble contenant au moins un point dans chaque composante connexe de $\hat{C} \setminus \Omega$, alors f est dans l'adhérence de l'ensemble des fractions rationelles à pôles dans Z pour la topologie de la convergence uniforme sur tout compact.*

• *Si f est une fonction holomorphe définie sur un ouvert simplement connexe Ω, alors il existe une suite de polynômes convergeant uniformément vers f sur tout compact.*

<u>Démonstration :</u>

• On définit les K_n comme dans le lemme 262, et on définit F_n comme étant une fraction rationnelle tel que $|f(z) - F_n(z)| \leq 1/n$ pour z dans K_n.

• Dans le deuxième cas, on peut simplement imposer $Z = \{\infty\}$, et on obtient bien une suite de polynômes.□

30.10 Endomorphismes semi-simples

Définition 1212 *Un endomorphisme f est dit* **semi-simple** *si tout sous-espace vectoriel stable par f a un supplémentaire stable par f.*

Théorème 1213 *On se place en dimension finie.*
On suppose u semi-simple ; on note μ_u le polynôme caractéristique de u.
$\mu_u = \Pi_{i=1}^p P_i^{n_i}$ avec les P_i irréductibles et premiers 2 à 2.
$n_i \geq 1$.
On va montrer qu'en fait $n_i = 1$.
En outre on montrera que réciproquement si tous les n_i sont égaux à 1, alors u est semi-simple.

Démonstration : $E = \oplus_{i=1}^p E_i$ avec $E_i = KerP_i(u)^{n_i}$

On va montrer que pour tout i $n_i = 1$.
Par l'absurde $n_1 \geq 2$. $KerP_1(u)$ admet un supplémentaire u-stable. Donc $KerP_1(u) \oplus F = E$ avec F u-stable.

$$F = \oplus_{i=1}^p (E_i \cap F)$$

donc

$$\begin{array}{ccccc} E = & \underbrace{KerP_1(u) \oplus (E_1 \cap F)} & \oplus & \oplus_{i=2}^p E_i \cap F \\ & \bigcap & & \bigcap \\ E = & E_1 & \oplus & \oplus_{i=2}^p E_i \end{array}$$

Les inclusions sont en fait des égalités puisque la dimension est finie. On considère $F' = E_1 \cap F$. $E_1 = KerP_1(u) \oplus F'$; si $F' = \{0\}$ alors $P_1.\Pi_{i=2}^p P_i^{n_i}$ serait dans l'idéal annulateur. Impossible car $n_1 \geq 2$, donc $F' \neq \{0\}$.

On a donc : $P_1^{n_1}(u_{|F'}) = 0$
Soit $x \in E_1 \setminus KerP_1^{n_1-1}(u)$
$x = x_1 + x_2$ avec $x_1 \in KerP_1(u)$ et $x_2 \in F'$.

$$P_1^{n_1-1}(u)(x) = \underbrace{P_1^{n_1-1}(u)(x_1)}_{0 \text{ car } n_1 \geq 2} + P_1^{n_1-1}(u)(x_2)$$

Donc $P_1^{n_1-1}(u)(x_2) \neq 0$ donc $P_1^{n_1-1}(u_{|F'}) \neq 0$.
donc (car P_1 est irréductible) $\mu_{u_{|F'}} = P_1^{n_1}$.

Donc $P_1(u_{|F'})$ est nilpotent, donc pas injectif, or $KerP_1(u) \cap F' = \{0\}$ d'où une contradiction.

Réciproquement :
On suppose $\mu_u = \Pi_{i=1}^p P_i$

Alors $E = \oplus_{i=1}^{p} E_i$ avec $E_i = Ker\ P_i(u)$

F stable $\rightarrow F = \oplus_{i=1}^{p} (F \cap E_i)$

il suffit de trouver un supplémentaire stable dans E_i de $F \cap E_i$. Or $\mu_{u_{|E_i}} = P_i$ donc on est ramené au cas ou μ_u est irréductible.

Soit alors F non réduit à 0 et différent de E, stable par u.
Soit $x \in E \setminus F$. $\{Q \in \mathbb{K}[X] | Q(u)(x) = 0\}$ est un idéal non nul engendré par P unitaire différent de 1. P divise μ_u donc $P = \mu_u$ parce que μ_u est irréductible.

Soit m le degré de μ_u. $(x, u(x), ..., u^{m-1}(x))$ est une base de $G = Vect_{i \in \mathbb{N}} u^i(x)$.
$x \notin F$ donc $F \cap G \neq G$, donc $dim\ F \cap G < m$.

Si $F \cap G \neq \{0\}$ alors $\mu_{u_{|F \cap G}} | \mu_u$ donc $\mu_{u_{|F \cap G}} = \mu_u$.

Or $deg \mu_{u_{|F \cap G}} \leq$ (par C.H.) $dim F \cap G < m = deg \mu$, d'où contradiction, donc $F \cap G = \{0\}$. On a donc $F \oplus G$ avec G stable, $G \neq \{0\}$. Si $F + G \neq E$ on recommence ; en un nombre fini d'étapes on conclut.$\square$

En résumé :
Pour montrer qu'un endomorphisme semi-simple admet pour polynôme simple un produit de polynômes irréductibles :
• On considère le polynôme minimal, avec ses facteurs P_i et leurs exposants n_i.
• On suppose $n_1 \geq 2$.
• On considère un supplémentaire F de $Ker\ P_1(u)$ u-stable.
• On considère $E_i = Ker\ P_i^{n_i}(u)$, et on note que F est la somme directe des intersections de F avec les E_i.
• On considère $F' = E_1 \cap F$; on montre que F' n'est pas réduit à 0.
• On considère x dans $E_1 \setminus Ker\ P_1^{n_1 - 1}(u)$
• On décompose x sur $Ker\ P_1(u)$ et F' ; on en déduit que $P_1^{n_1-1}(u_{|F'}) \neq 0$
• On sait alors que $\mu_{u_{|F'}} = P_1^{n_1}$
• $P_1(u_{|F'})$ est injectif, mais pourtant nilpotent, d'où contradiction.

Pour montrer l'existence d'un supplémentaire stable pour tout sous-espace stable à partir d'un endomorphisme ayant un produit de polynômes irréductibles distincts pour polynôme minimal :
• On se ramène à un polynôme minimal irréductible en considérant les restrictions aux E_i.
• On considère un espace stable F, et un élément x or de cet espace.
• On considère le polynôme qui engendre l'ensemble des polynômes Q tels que $Q(u)(x) = 0$; ce polynome divise le polynome minimal de μ et lui est donc égal puisque ce dernier est irréductible.
• On considère alors G l'espace engendré par les images successives de x par des puissances de u ; cet espace est engendré par les m premiers éléments avec m le degré de μ.
• On en déduit que $F \cap G$ est de dim $< m$.
• On considère alors le polynôme caractéristique de u restreint à $F \cap G$; si ce dernier espace est non nul, alors ce polynôme est égal à μ_u, or son degré est plus petit que m

par application de Cayley-Hamilton, donc la somme de F et G est directe. Il ne reste qu'à récurer pour avoir un supplémentaire de F.

Chapitre 31

Réduction des endomorphismes

Il est nécéssaire avant d'étudier cette partie d'avoir pris connaissance de la partie 27et de la partie 30.

La réduction d'un endomorphisme va constituer dans le cas général en la recherche des éléments propres de cet endomorphisme (définition à venir) et dans le cas de la dimension finie en la recherche d'une base dans laquelle cet endomorphisme est représenté par une matrice de forme agréable (voir plus loin).

Dans toute cette partie, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

31.1 Le cas général

Définition 1214 *Soit E un $\mathbb{K}$-espace vectoriel et f un endomorphisme de E ; c'est-à-dire $f \in L(E)$. On se donne P un polynôme appartenant à $\mathbb{K}[X]$; $P = \sum_i p_i.X^i$.*
Alors :
• *On dit que F, sous-espace vectoriel de E, est* **stable par** *f si et seulement si $f(F) \subset F$*
• *Si il existe n tel que $Ker\ f_n = Ker\ f_{n+1}$, alors le n minimal vérifiant cette propriété est appelé* **indice de** *f.*
• *On note $P(f)$ l'endomorphisme qui à x associe $\sum_i p_i.f^i(x)$.*
• *On note $\mathbb{K}[f]$ l'ensemble des $Q(f)$ pour $Q \in \mathbb{K}[X]$; c'est une algèbre commutative (sous-algèbre de l'algèbre des endomorphismes $L(E)$)*
• *On appelle* **idéal annulateur de** *f l'ensemble des $Q \in \mathbb{K}[X]$ tels que $Q(f) = 0$ (c'est un idéal, les matheux sont pas encore assez vicieux pour nous faire cette mauvaise blague). On le note $\mathcal{I}(f)$. $\mathbb{K}[X]$ étant principal, si $\mathcal{I}(f)$ est non réduit à $\{0\}$, il est engendré par un polynôme unitaire que l'on appellera* **polynôme minimal de** *f et que l'on notera P_f.*
• *On appelle* **valeur propre** *de f un scalaire λ tel que $f - \lambda.I$ ne soit pas injectif ; dans ce cas $E_f(\lambda) = Ker\ f - \lambda.I$ est appelé* **sous-espace propre** *associé à la valeur propre λ.*
• *On appelle* **spectre de** *f l'ensemble des valeurs propres de f. On le note $Sp(f)$.*
• *On appelle* **vecteur propre de** *f associé à la valeur propre λ un élément non nul du sous-espace propre associé à la valeur propre λ.*
• *On appelle* **valeur propre de** *M avec M une matrice carrée de type (n, n) une valeur propre d'endomorphisme de $\mathbb{K}^n$ canoniquement associé à M.*
• *On appelle* **sous-espace propre de** *M* **associé à** *λ avec M une matrice carrée de type (n, n) l'espace propre de l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à M pour λ valeur propre de M.*
• *On appelle* **spectre de** *M avec M une matrice carrée de type (n, n) l'ensemble des valeurs propres de M.*

⚠ En dimension infinie un endomorphisme n'a pas nécéssairement un indice

Proposition 1215 • *Avec P et Q deux polynômes, $PQ(u) = P(u) \circ Q(u) = Q(u) \circ P(u)$*
• *Si deux endomorphismes commutent, alors le noyau et l'image de l'un sont stables par l'autre.*
• *La suite F_n définie par $F_n = Ker\ f^n$ est croissante*
• *Si f a un indice fini n, alors pour tout i supérieur à n, on a $F_{i+1} = F_i$*
• *Soit P et Q deux polynômes, alors $Ker\ P(f) \cap Ker\ Q(f) = Ker\ ((P \wedge Q)(f))$*
• *Soit P et Q deux polynômes premiers entre eux (i.e. de pgcd 1), alors $Ker\ (PQ)(f) = Ker\ P(f) \oplus Ker\ Q(f)$.*
• *Soit $P_1, ..., P_p$ des polynômes premiers deux à deux tels que $\Pi P_i(f) = 0$, alors $E = \oplus\ Ker\ P_i(f)$.*
Les trois derniers • ont valeur de théorèmes ; on les regroupe souvent sous l'appelation **lemme des noyaux**.

Cela sert un peu partout, dans les résultats de réduction ; citons par exemple la partie sur les suites récurrentes linéaires, 31.3.2, ou le théorème de diagonalisation 1223.

Démonstration : Les quatre premiers • sont évidents.
Le 5ème • se montre grâce à la propriété de Bezout.
Pour le 6ème • , la somme est directe en vertu du • précédent ; et les deux inclusions s'obtiennent par Bezout et par le premier • .
Le dernier • est facile à déduire du 6ème. □

Proposition 1216 (Pratique de la réduction (sans hypothèse de dimension finie))
• *une somme finie de sous-espaces propres distincts est directe*
• *une famille de vecteurs propres associés à des valeurs propres distinctes est libre*
• *si f et g sont deux endomorphismes qui commutent, alors les sous-espaces propres de f sont stables par g*
• $\forall (\lambda, P) \in Sp(f) \times \mathbb{K}[X] \to P(\lambda) \in Sp(P(f))$
• *Si $P(f) = 0$, alors $\lambda \in Sp(f) \to P(\lambda) = 0$*
• *Si $f \in GL(E)$, alors $Sp(u^{-1}) = \{\frac{1}{\lambda} / \lambda \in Sp(u)\}$*

31.2 Le cas de la dimension finie

Dans le cas de la dimension finie, la réduction d'un endomorphisme va consister en la recherche d'une base dans laquelle l'endomorphisme est diagonale, ou, à défaut, triangulaire.

On a vu ci-dessus que pour décomposer un endomorphisme il fallait trouver ses sous-espaces propres et ses valeurs propres. Cela est très lié au polynômes annulateurs de l'endomorphisme ; mais dans le cas général, il y a assez peu de choses à dire. Nous allons voir que la dimension finie est beaucoup plus pratique.

Voyons tout d'abord une liste de propriétés élémentaires très utiles pour la suite :

Proposition 1217 • *En dimension finie, tout endomorphisme a un indice, inférieur ou égal à la dimension.*
• *En dimension finie, l'indice est aussi le plus petit n tel que $Im\ f^n = Im\ f^{n+1}$*
• *En dimension finie, avec n l'indice de f, $E = Ker\ f^n \oplus Im\ f^n$*
• *En dimension finie, tout endomorphisme admet un idéal annulateur non réduit à 0, et donc admet un polynôme minimal.*
• *Si F est un sous-espace vectoriel de E, si $(e_1, ..., e_p, e_{p+1}, ..., e_n)$ est une base de E avec $(e_1, ..., e_p)$ une base de F, alors F est stable par f si et seulement si la matrice de f dans la base $(e_1, ..., e_n)$ est de la forme suivante*

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

avec A une matrice de type (p,p), B une matrice de type $(p, n-p)$, D une matrice de type $(n-p, n-p)$.
• *Si F_1, F_2, ... et F_p sont des sous-espaces vectoriels de E tels que $E = \oplus F_i$, alors les F_i sont stables par f si et seulement si la matrice de f dans une base constituée d'une base de F_1, suivie d'une base de F_2, ..., suivie d'une base de F_p, est de la forme*

$$\begin{pmatrix} M_1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & M_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & M_{p-2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & M_{p-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & M_p \end{pmatrix}$$

avec M_i de type $(dim\ F_i, dim\ F_i)$.

On va maintenant introduire quelques définitions intéressantes dans le cadre de la dimension finie.

Définition 1218 (Définitions dans le cadre de la réduction en dimension finie)
• *On appelle* **polynôme caractéristique d'une matrice carrée** *M le polynôme $det(X.I - M)$.*
• *Le polynôme caractéristique de la matrice d'un endomorphisme en dimension finie (le polynôme caractéristique n'est défini que dans ce cadre là) est indépendant de la base choisie ; on l'appelle* **polynôme caractéristique de cet endomorphisme**.
On note χ_M le polynôme caractéristique de la matrice M et χ_f le polynôme caractéristique de l'endomorphisme f.

Voyons maintenant quelques propriétés fondamentales de ces notions, toujours dans le cadre de la réduction en dimension finie ; la preuve, très simple, n'est pas donnée.

Proposition 1219 • *Le polynôme caractéristique d'une matrice M est un polynôme (preuve en considérant la définition première du déterminant).*
• *Le polynôme caractéristique d'une matrice M de type (n, n) est de degré n, de coefficient dominant 1 (preuve là aussi en considérant la définition du déterminant, et en cherchant l'unique permutation qui donne le terme de degré n dans le polynôme caractéristique). Le second coefficient est $-tr\ M$.*
• *Le polynôme caractéristique d'une matrice M est le polynôme caractéristique de toute matrice N semblable à M.*
• *L'ensemble des valeurs propres d'un endomorphisme en dimension finie est égal à l'ensemble des 0 de son polynôme caractéristique.*
• *Le polynôme caractéristique de M est égal au polynôme caractéristique de ${}^tM$.*
• *E $\mathbb{K}$-espace vectoriel de dimension finie, F sous-espace vectoriel de E, f endomorphisme de E, F stable par f, alors $\chi_{f_{|F}} | \chi_f$*
• *Dans un $\mathbb{C}$-espace vectoriel de dimension finie ≥ 1, tout endomorphisme admet au moins une valeur propre.*
• *Le polynôme caractéristique d'une matrice M de type $(2, 2)$ est $X \mapsto X^2 - tr\ M.X + det\ M$*
• *Le polynôme caractéristique d'une matrice M de type $(3, 3)$ est $X \mapsto X^3 - tr\ M.X^2 - tr\ com(M).X - det\ M$*
• *Le polynôme caractéristique d'une matrice M triangulaire de type (n, n) est $\Pi_{i=1}^n (X - M_{i,i})$*

Définition 1220 *On appelle* **ordre** *d'une valeur propre d'un endomorphisme en dimension finie le degré de multiplicité de cette valeur propre comme racine du polynôme caractéristique.*
On appelle **sous-espace caractéristique** *associé à la valeur propre λ de l'endomorphisme f d'un $\mathbb{K}$-espace vectoriel E de dimension finie le sous-espace vectoriel $Ker\ (f - \lambda.I)^m$ avec m l'ordre de multiplicité de λ comme racine du polynôme caractéristique de f.*
Un endomorphisme en dimension finie est dit **diagonalisable** *si il existe une base dans laquelle cet endomorphisme se représente par une matrice diagonale.*
Une matrice en dimension finie est dite **diagonalisable** *si elle est semblable à une matrice diagonale, c'est-à-dire si elle représente un endomorphisme diagonalisable.*
Un endomorphisme en dimension finie est dit **trigonalisable** *si il existe une base dans laquelle cet endomorphisme se représente par une matrice triangulaire.*
Une matrice en dimension finie est dite **trigonalisable** *si elle est semblable à une matrice triangulaire, c'est-à-dire si elle représente un endomorphisme trigonalisable.*

Toute matrice triangulaire supérieure est semblable à une certaine matrice triangulaire inférieure et toute matrice triangulaire inférieure est semblable à une certaine matrice triangulaire supérieure, comme on s'en convaincra simplement en changeant l'ordre des éléments d'une base ; ainsi lorsqu'un endomorphisme est trigonalisable, on pourra en considérer indifféremment une représentation triangulaire supérieure ou inférieure.

Notons que tout sous-espace caractéristique de f, endomorphisme en dimension finie, est stable par f.

Passons maintenant à quelques "gros" résultats de réduction en dimension finie.

Théorème 1221 *Un endomorphisme f de E $\mathbb{K}$-espace vectoriel de dimension finie est diagonalisable si et seulement si l'une des conditions suivantes est vérifiée :*
• E est somme directe des sous-espaces propres
• il existe une base de E formée de vecteurs propres de f
• χ_f est scindé dans $\mathbb{K}[X]$ <u>et</u> la dimension de tout sous-espace propre est égale à l'ordre de multiplicité de la valeur propre associée.

<u>Démonstration :</u> Les deux premiers • sont évident.

Le troisième est plus intéressant :

- la condition est suffisante, car la somme des sous-espaces propres est toujours directe, et la somme des dimensions des sous-espaces propres est bien la dimension de l'espace, donc l'espace est bien somme directe des sous-espaces propres.

- la condition est nécéssaire ; on le voit de manière évidente en considérant une matrice diagonale pour laquelle la matrice de f est diagonale.□

Corollaire 1222 *Si χ_f est scindé à racines simples, alors f est diagonalisable.*

<u>Démonstration :</u> Si chaque valeur propre est d'ordre 1, forcément l'espace propre associé est de dimension 1 ; donc ce corollaire découle immédiatement du théorème précédent.□

On va voir un renforcement de ce corollaire ci-dessous.

Théorème 1223 *Un endomorphisme f de E $\mathbb{K}$-espace vectoriel de dimension finie est diagonalisable si et seulement si il existe P polynôme <u>scindé à racines simples</u> tel que $P(f) = 0$.*

<u>Démonstration :</u> • Si il existe un tel P, alors E est la somme directe des $Ker\ f - \lambda.I$ pour λ dans l'ensemble des racines de P (voir le lemme des noyaux, proposition 1215). Donc f est diagonalisable, clairement.

• Si f est diagonalisable, alors le polynôme caractéristique est scindé et la dimension de chaque espace propre est l'ordre de multiplicité de la valeur propre associée comme racine de ce polynôme (d'après le théorème précédent).

On peut alors écrire $\chi_F = \Pi(X - x_i)^{r_i}$ avec les x_i distincts deux à deux ; considérons $P = \Pi(X - x_i)$. P est scindé à racines simples. Il reste à voir que $P(f) = 0$.

E est la somme directe des $Ker\ f - x_i.I$. Soit x appartenant à $Ker\ f - x_{i_0}.I$; on a $f(x) = x_{i_0}.x$. $P(f)(x) = \Pi_{i \neq i_0}(y \mapsto f(y) - x_i.y) \circ (y \mapsto f(y) - x_{i_0}.y)^{r_{i_0}-1}\ (f(x) - x_i.x) = 0$. Donc $P(f)$ est nul sur des sous-espace vectoriel dont la somme est E ; donc $P(f)$ est nul.$\square$

Finalement, voici la méthodologie de la diagonalisation :

• On se donne f, endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie n.

• On détermine le polynôme caractéristique de f

• Si χ_F n'est pas scindé, alors on ne peut pas diagonaliser

• Si χ_F est scindé, pour chaque valeur propre λ, on cherche si $E_f(\lambda) = Ker\ (f - \lambda.I)$ est de dimension l'ordre de λ

• Si oui, on obtient une base dans laquelle l'endomorphisme est diagonal en réunissant des bases des sous-espaces propres

• Dans le cas contraire, l'endomorphisme n'est pas diagonalisable ; on essaie alors de trigonaliser (voir plus loin).

On peut ajouter quelques remarques permettant de simplifier la détermination de la diagonalisabilité d'une matrice :

• une matrice M est diagonalisable si pour tout $\lambda \in \mathbb{K}$ $M + \lambda.I$ est diagonalisable ; si $P^{-1}.(M + \lambda.I).P = D$, alors $P^{-1}.M.P = D + \lambda.I$

• La restriction d'un endomorphisme diagonalisable à un sous-espace stable est diagonalisable (en effet la diagonalisabilité d'un endomorphisme équivaut à l'existence d'un polynôme scindé à racines simples annulant cet endomorphisme)

Maintenant le théorème fondamental de la trigonalisation :

Théorème 1224 *Un endomorphisme en dimension finie est trigonalisable si et seulement si son polynôme caractéristique est scindé.*

Démonstration : Il est évident que si l'endomorphisme est trigonalisable, alors son polynôme caractéristique est scindé. La réciproque est donc le problème intéressant. Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie.

On procède par récurrence sur la dimension de l'espace. En dimension 1 le résultat est clair ; on suppose maintenant le résultat vrai jusqu'en dimension $n-1$, et on suppose E de dimension n.

Le polynôme caractéristique de f étant scindé, il possède une racine, disons λ. Soit e_1 un vecteur propre associé à λ, et complétons de manière à avoir $(e_1, ..., e_n)$ une base de E.

La matrice de f dans cette base est une matrice (n, n) de la forme

$$\begin{pmatrix} \lambda & * & \dots & * \\ 0 & M_{1,1} & \dots & M_{1,n-1} \\ 0 & M_{2,1} & \dots & M_{2,n-1} \\ 0 & \vdots & \vdots & \vdots \\ 0 & M_{n-1,1} & \dots & M_{n-1,n-1} \end{pmatrix}$$

La matrice M a un polynôme caractéristique scindé, comme on s'en convaincra facilement en considérant la définition du déterminant.

Par la propriété de récurrence, M est donc trigonalisable ; $PMP{-}1 = T$, avec T matrice triangulaire supérieure.

Alors on considère le produit suivant :

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & P_{1,1} & \dots & P_{1,n-1} \\ 0 & P_{2,1} & \dots & P_{2,n-1} \\ 0 & \vdots & \vdots & \vdots \\ 0 & P_{n-1,1} & \dots & P_{n-1,n-1} \end{pmatrix} \begin{pmatrix} \lambda & * & \dots & * \\ 0 & M_{1,1} & \dots & M_{1,n-1} \\ 0 & M_{2,1} & \dots & M_{2,n-1} \\ 0 & \vdots & \vdots & \vdots \\ 0 & M_{n-1,1} & \dots & M_{n-1,n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & P^{-1}_{1,1} & \dots & P^{-1}_{1,n-1} \\ 0 & P^{-1}_{2,1} & \dots & P^{-1}_{2,n-1} \\ 0 & \vdots & \vdots & \vdots \\ 0 & P^{-1}_{n-1,1} & \dots & P^{-1}_{n-1,n-1} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & * & \dots & * \\ 0 & T_{1,1} & \dots & T_{1,n-1} \\ 0 & T_{2,1} & \dots & T_{2,n-1} \\ 0 & \vdots & \vdots & \vdots \\ 0 & T_{n-1,1} & \dots & T_{n-1,n-1} \end{pmatrix}$$

(les $*$ représentant des scalaires quelconques)

Le résultat est ainsi établi.$\square$

Corollaire 1225 *Toute matrice à coefficients complexes est trigonalisable dans $M_n(\mathbb{K})$.*

Démonstration : Rappelons juste le théorème de D'Alembert-Gauss : un polynôme à coefficients complexes est scindé sur $\mathbb{C}$.$\square$

Ce dernier théorème va permettre de démontrer un théorème important, le théorème de Cayley-Hamilton :

Théorème 1226 (Cayley-Hamilton) *Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie. Alors $\chi_f(f) = 0$.*

Démonstration :

• On montre tout d'abord le résultat pour $\mathbb{K} = \mathbb{C}$. Ensuite, puisque le polynôme caractéristique d'une matrice à coefficients réels est le même que l'on la considère comme matrice réelle ou complexe, et puisqu'un endomorphisme est nul si et seulement si sa matrice associée est nulle, le résultat sera aussi valable pour $\mathbb{K} = \mathbb{R}$.

• χ_f est un polynôme scindé, puisque l'on travaille dans $\mathbb{C}$; donc on peut trigonaliser f.

• On se donne $(e_1, ..., e_n)$ une base dans laquelle f est représenté par une matrice triangulaire supérieure.

• On note $\chi_f = \Pi_{i=1}^n (X - x_i)$, avec x_i le i-ième terme sur la diagonale de la matrice de f dans la base $(e_1, ..., e_n)$.

• On note $P_i = \Pi_{j=1}^n (X - x_i)$.

• On considère la propriété $P(i)$ définie par $P(i) \iff \forall j \in [1, i] P_i(e_j) = 0$; on montre facilement par récurrence qu'elle est vraie pour tout i compris entre 1 et n.$\square$

Proposition 1227 *Si le polynôme caractéristique est scindé, la dimension d'un sous-espace caractéristique est égale à l'ordre de multiplicité de la valeur propre associée.*

Démonstration :

Pendant l'ensemble de la preuve, on note I_p la matrice identité à p lignes et p colonnes.

• Notons f un endomorphisme d'un espace vectoriel E de dimension n, de polynôme caractéristique scindé.

• Soit F sous-espace caractéristique associé à la valeur propre λ de multiplicité m.

• Dans une certaine base, f est représenté par une matrice triangulaire supérieure (théorème 1224). Soit M cette matrice :

$$M = \begin{pmatrix} \lambda.Id_m + N & P \\ 0 & Q \end{pmatrix}$$

où N est triangulaire supérieure à diagonale nulle, P quelconque, Q triangulaire supérieure (n'admettant pas λ pour valeur propre).

• Une vérification immédiate révèle alors que $(f - \lambda Id_n)^m$ est représenté par la matrice suivante (N est nilpotente d'indice de nilpotence $< m$) :

$$\begin{pmatrix} 0 & P' \\ 0 & (Q - \lambda.Id_{n-m})^m \end{pmatrix}$$

• On en déduit donc que le rang de $(f - \lambda Id_n)^m$ est $n - m$ (rappelons que λ n'est pas valeur propre de Q).

• On a donc bien $dim\ F = n - (n - m) = m.\square$

On peut alors passer au théorème suivant, fournissant une jolie représentation d'un endomorphisme trigonalisable :

Théorème 1228 *Soit f un endomorphisme d'un $\mathbb{K}$-espace vectoriel E de dimension finie n. Alors E est la somme directe des sous-espaces caractéristiques associés à f, et dans une certaine base f a pour matrice*

$$\begin{pmatrix} M_1 & 0 & 0 & \dots & 0 \\ 0 & M_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 0 & M_l \end{pmatrix}$$

avec M_i une matrice triangulaire supérieure de type (m, m) comportant seulement des λ_i sur la diagonale et avec m l'ordre de λ_i.

Le corollaire qui suit, et la partie 31.3.3 de calcul de l'exponentielle d'un endomorphisme donnent des applications à ce résultat.

Démonstration : Il suffit de considérer les espaces caractéristiques ; leur somme est directe et égal à E (par le théorème de Cayley-Hamilton), ils sont stables par f, on peut donc considérer les restrictions aux différents sous-espaces caractéristiques.

D'où le résultat.□

Corollaire 1229 *Si f est un endomorphisme de polynôme caractéristique scindé d'un $\mathbb{K}$-espace vectoriel E de dimension finie n, alors f s'écrit de manière unique comme somme d'un endomorphisme diagonalisable et d'un endomorphisme nilpotent, qui commutent.*

Démonstration : L'existence de cette écriture est facile ; il suffit de considérer la matrice M donnée par le théorème précédent, et d'écrire $M = D + N$, avec D une matrice diagonale, et N une matrice triangulaire supérieure à diagonale nulle.

L'unicité sera ici admise.□

Voyons maintenant la méthodologie de la trigonalisation ; elle fait suite à celle de la diagonalisation, dans le cas où la diagonalisation est impossible. En fait on ne va pas simplement chercher à trigonaliser ; on va essayer si possible d'obtenir une décomposition comme celle proposée dans le théorème 1228.

On se donne un endomorphisme f d'un $\mathbb{K}$-espace vectoriel E de dimension n.

• On cherche si le polynôme caractéristique est scindé. S'il ne l'est pas, f n'est pas trigonalisable

• On détermine les sous-espaces caractéristiques

• On détermine une base de chacun de ces sous-espaces, dans laquelle la restriction de l'endomorphisme est représentée par une matrice triangulaire supérieure ; cela se fait par récurrence, comme on peut le voir dans la démonstration du théorème 1224.

• On prend la réunion de ces bases, et on a une représentation comme souhaitée.

31.3 Applications de la réduction d'un endomorphisme

31.3.1 Application au calcul d'un polynôme d'endomorphisme

On se donne M la matrice de l'endomorphisme f d'un espace vectoriel de dimension finie n. Son polynôme caractéristique est appelé P.

Par le théorème de Cayley-Hamilton

Alors supposons que l'on cherche $Q(f)$, avec Q un polynôme de $\mathbb{K}[X]$. On détermine alors la division euclidienne $Q = P.A + B$.

$$Q(f) = P(f).A(f) + B(f) = B(f) \text{ car } P(f) = 0$$

Il suffit donc ensuite d'avoir préalablement calculé les M^k pour $k \in [0, n-1]$.

Par diagonalisation

On suppose ici que M est diagonalisable. Si $M = L.D.L^{-1}$, alors $Q(M) = L.Q(D).L^{-1}$; si D est choisie diagonale, $Q(D)$ est donc vite calculée. Cette méthode est efficace seulement si il s'agit de calculer de nombreuses fois des polynômes d'un même endomorphisme ; lorsque l'on change d' endomorphisme à chaque fois, la diagonalisation est trop laborieuse.

31.3.2 Application aux suites récurrentes linéaires

Définition 1230 *On considère une suite d'éléments d'un corps $\mathbb{K}$ de caractéristique nulle définie par récurrence par $u_n = a_0 u_{n-p} + a_1 u_{n-p+1} + a_2 u_{n-p+2} + ... + a_{p-1} u_{n-1}$ (u_i étant donné pour $i < n$). On peut en fait noter $X_n = (u_{n+p-1}, u_{n+p}, ..., u_n)$; alors $X_{n+1} = M X_n$, avec*

$$M = \begin{matrix} 0 & 1 & 0 & ... & ... & 0 \\ 0 & 0 & 1 & 0 & ... & 0 \\ 0 & ... & 0 & 1 & 0 & \vdots \\ \vdots & ... & ... & \ddots & \ddots & \vdots \\ 0 & ... & ... & ... & 0 & 1 \\ a_0 & a_1 & ... & ... & a_{p-2} & a_{p-1} \end{matrix}$$

On suppose a_0 non nul, cas auquel on peut toujours se ramener.
Une telle suite est dit **suite récurrente linéaire**.
Le polynôme caractéristique de M est $P = X^p - a_0 - a_1 X - a_2 X^2 - ... - a_{p-1} X^{p-1}$. Par définition, ce polynôme est dit **polynôme caractéristique** *de la relation de récurrence linéaire.*

L'espace vectoriel E des suites vérifiant la relation de récurrence donnée est le noyau de $P(f)$, avec f l'application qui à une suite (u_n) associe (u_{n+1}) (ceci découle immédiatement de la définition de P !).

On va faire l'hypothèse que P est scindé, ce qui dans le cas de suites réelles ou complexes est une hypothèse qui n'enlève rien à la généralité, puisqu'on peut toujours supposer la suite complexe. On peut alors appliquer D'Alembert-Gauss pour conclure que P est scindé .

Par le lemme des noyaux (proposition 1215), on constate que E est égal à la somme directe des $Ker\ (f - \lambda_i I)^{\nu_i}$ avec les λ_i les racines de P avec pour ordres de multiplicité les ν_i.

On va donc chercher à exprimer les solutions comme combinaisons linéaires d'éléments des $Ker\ (f - \lambda_i I)^{\nu}_i$. Il convient donc de déterminer une base de $Ker\ (f - \lambda I)^{\nu}$.

Pour cela on considère l'opérateur $\mathbb{D}$ de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$ qui à Q associe $Q \circ (X + 1) - Q$. On identifie $\mathbb{N}$ à $\mathbb{N}.1_{\mathbb{K}}$, de manière évidente. On considère alors Q appartenant à $\mathbb{K}[X]$, de degré $< \nu$; on définit ensuite $u_n = u_n^Q = \lambda^n Q(n)$. L'image de la suite u_n par $f - \lambda I$ est $n \mapsto \lambda^{n+1} \mathbb{D}(Q)$; son image par $(f - \lambda I)^2$ est $n \mapsto \lambda^{n+2} \mathbb{D}^2(Q)$, et ainsi de suite jusqu'à son image par $(f - \lambda I)^{\nu}$ qui est $n \mapsto \lambda^{n+\nu} \mathbb{D}^{\nu}(Q)$. Or on constate immédiatement que $deg\ \mathbb{D}(Q) = deg\ Q - 1$ si Q est de degré ≥ 1, et $\mathbb{D}(Q) = 0$ sinon, et donc vu la limite sur le degré de Q, $(f - \lambda I)^{\nu}(u_n) = 0$.

Donc on obtient des éléments de E en considérant des suites de la forme $n \mapsto \lambda_i^n n^q$, pour $0 \leq q \leq \nu_i$. Il reste à voir si l'on obtient bien une base de E. Si la famille est libre, alors son cardinal fait que l'on a bien une base de E. Il suffit donc de montrer que cette famille est libre.

Il suffit donc de montrer que l'application $Q \mapsto u^Q$ est linéaire et injective. La linéarité est évidente. Supposons maintenant que $u^Q = 0$. Alors $Q(n) = 0$ pour tout n (en effet λ_i est non nul, car nous avons imposé a_0 non nul ; donc Q est nul.

D'où le résultat désiré. Etant donnés les premiers termes u_i pour $i < n$, on peut reconstituer alors la suite en écrivant à priori la suite sous la forme $\sum Q_i(n) \lambda_i^n$ avec

Q_i de degré $< \nu_i$; les coefficients se déduisent par une résolution de système linéaire.

31.3.3 Calcul d'exponentielle de matrice

On suppose donnée une matrice M, de polynôme caractéristique scindé (hypothèse peu restrictive dans la mesure où l'on considère éventuellement la clotûre algébrique).

On cherche à calculer l'exponentielle de M.

M peut être trigonalisée, dans une base constituée d'une réunion de bases d'espaces caractéristiques ; avec P la matrice de passage correspondante, $N = PMP^{-1}$ est donné par la proposition 1228 ; une matrice diagonale par blocs, chaque bloc étant somme d'une homothétie et d'un nilpotent.

On a alors $exp(M) = P\,exp(N)P^{-1}$. Le calcul de $exp(N)$ peut se faire par blocs. Il suffit donc d'être capable de calculer $exp(\lambda\, I + Q)$, avec Q nilpotent.

Pour cela il suffit de constater que $exp(\lambda\, I + Q) = e^{\lambda} exp(Q)$ car I et Q commutent, et que

$$exp(Q) = \sum \frac{1}{k!} Q^k = I + Q + \frac{1}{2} Q^2 + ... + \frac{1}{p} Q^p$$

si $Q^{p+1} = 0$; on est ainsi ramené à une somme finie.

Chapitre 32

Géométrie affine

Dans l'ensemble de ce chapitre, les démonstrations seront succinctes et approximatives ; en effet il est souvent suffisant d'avoir de bonnes connaissances en algèbre linéaire pour comprendre cette partie sans difficulté et pour être capable de refaire les démonstrations.

Il est donc nécessaire et presque suffisant pour maîtriser tout cela d'avoir convenablement étudié la partie27, la partie30 et la partie 29.

32.1 Définitions et généralités

Définition 1231 *Etant donné E un $\mathbb{K}$-espace vectoriel , on appelle* **espace affine de direction** E *un ensemble X muni d'une application $(x, y) \mapsto \overrightarrow{xy}$ de $X \times X$ dans E telle que :*
• $\forall (x, y, z) \in X^3, \overrightarrow{xy} + \overrightarrow{yz} = \overrightarrow{xz}$
• $\forall x \in X, \forall u \in E, \exists ! y \in X \ \overrightarrow{xy} = u$
On définit alors une addition de $X \times E$ dans X par $x + u = y$ avec y tel que $\overrightarrow{xy} = u$.
On appelle **scalaires** *les éléments du corps $\mathbb{K}$.*
On appelle **vecteurs** *les éléments de l'espace vectoriel E.*
On appelle **points** *les éléments de l'espace affine.*
On note parfois surmontés d'une flèche les éléments de E, pour les distinguer des éléments de X ; ainsi au lieu de $u = \overrightarrow{xy}$ ou $x + u = y$ on peut noter $x + \overrightarrow{u} = y$.
Une convention usuelle est aussi de noter $\overrightarrow{X}$ un espace vectoriel associé à X ; il faut bien voir toutefois que la direction n'est pas unique et qu'il ne suffit pas que X soit un espace affine de direction E pour qu'on puisse définir canoniquement sa direction $\overrightarrow{X}$.
On appelle **dimension d'un espace affine** *la dimension de sa direction lorsque celle-ci est finie ; un espace affine est dit* **de dimension infinie** *lorsque sa direction est de dimension infinie.*
On appelle **translation de vecteur** *a avec $a \in E$ l'application qui à x dans X associe $x + a$.*
On appelle **variété affine d'un espace affine** X **de direction** F *ou* **sous-espace affine de** X *tout sous-ensemble de X de la forme $x + F$ avec $x \in X$ et F sous-espace vectoriel de E (***droite affine** *si ce sous-espace vectoriel est une droite,* **plan affine** *si ce sous-espace vectoriel est un plan,* **hyperplan affine** *si ce sous-espace vectoriel est un hyperplan,etc).*
On appelle **vectorialisé de** X **en** $x_0 \in X$ *l'espace vectoriel $(X, +)$ pour l'addition $x + y = z$ avec z tel que $\overrightarrow{x_0x} + \overrightarrow{x_0y} = \overrightarrow{x_0z}$. Pour l'application qui à (x, y) associe z tel que $x + z = y$, X est un espace affine de direction son vectorialisé.*

Bien noter l'unicité ($\exists !$) dans le deuxième axiome !

On peut lier facilement cette notion à celle d'espace affine définie dans le cadre des espace vectoriel ; en effet un espace affine (au sens des espaces vectoriels) est un espace affine pour l'application $(x, y) \mapsto \overrightarrow{xy} = y - x$, et un espace affine A (au sens ici défini) est un espace affine au sens des espaces vectoriels ; il suffit de choisir un point x arbitraire dans A et de considérer l'addition dans A définie par $b + c = d$, avec d tel que $\overrightarrow{xd} = \overrightarrow{xb} + \overrightarrow{xc}$.

Proposition 1232 *Les propriétés suivantes sont des propriétés élémentaires faciles à démontrer :*
• $\overrightarrow{xy} = \overrightarrow{0} \iff x = y$ *;*
• $\overrightarrow{xy} + \overrightarrow{yz} = \overrightarrow{xz}$ *(relation de Chasles) ;*
• $\overrightarrow{xy} = \overrightarrow{zt} \iff \overrightarrow{xz} = \overrightarrow{yt}$ *(propriété du parallélogramme) ;*
• *l'ensemble des translations de X est un groupe additif isomorphe au groupe additif de la direction de X ;*
• *les vectorialisés d'un espace affine sont tous isomorphes entre eux.*

32.2 Barycentre

On se donne X un espace affine.

Définition 1233 - Proposition *Etant donnés $x_1, ..., x_n$ des points de l'espace affine X, des scalaires $t_1, ..., t_n$ de somme 0, le vecteur $t_1.\overrightarrow{Ox_1} + t_2.\overrightarrow{Ox_2} + ... + t_n.\overrightarrow{Ox_n}$ est indépendant de $O \in X$.*
Etant donnés $x_1, ..., x_n$ des points de X, des scalaires $t_1, ..., t_n$ de somme 1, le point $O + t_1.\overrightarrow{Ox_1} + t_2.\overrightarrow{Ox_2} + ... + t_n.\overrightarrow{Ox_n}$ est indépendant de $O \in X$. On l'appelle **barycentre** *des (x_i, t_i) pour $i \in \{1, ..., n\}$.*
Un barycentre est ici défini pour une somme de coefficients égale à 1 ; mais on peut aussi le définir pour une somme quelconque, en remplaçant $O + t_1.\overrightarrow{Ox_1} + t_2.\overrightarrow{Ox_2} + ... + t_n.\overrightarrow{Ox_n}$ par

$$O + \frac{t_1}{t}\overrightarrow{Ox_1} + \frac{t_2}{t}\overrightarrow{Ox_2} + ... + \frac{t_n}{t}\overrightarrow{Ox_n}$$

avec $t = \sum t_i$.
On appelle **isobarycentre** *de n points le barycentre de ces points pondérés par $\frac{1}{n}$.*

Démonstration : Facile comme tout en utilisant Chasles !

Quelques remarques :

• On ne change pas le barycentre en multipliant tous les t_i par une même constante non nulle.

• Si $t_i = 0$ on peut se passer de (x_i, t_i).

• x est le barycentre des (x_i, t_i) si et seulement si $\sum_{i \in [1,n]} t_i.\overrightarrow{xx_i} = \overrightarrow{0}$.

32.3 Coordonnées cartésiennes, coordonnées barycentriques

Définition 1234 - Proposition
Etant donné X un espace affine , la famille finie $(x_i)_{i\in[1,p]}$ de points de X est dite **affinement libre** *si l'une des deux conditions* <u>*équivalentes*</u> *suivantes est vérifiée :*
• *il existe i tel que $(\overrightarrow{x_ix_j})_{j\in[1,p],j\neq i}$ soit une famille libre.*
• *pour tout i dans $[1,p]$ la famille $(\overrightarrow{x_ix_j})_{j\in[1,p],j\neq i}$ est une famille libre.*
Une famille infinie de X est dite **affinement libre** *lorsque toute sous-famille finie de cette famille est affinement libre.*
Etant donné X un espace affine de dimension finie et $\overrightarrow{X}$ sa direction, on se donne $\overrightarrow{e}_1, ..., \overrightarrow{e}_n$ une base de $\overrightarrow{X}$, et O un point de X. Alors tout point x de X s'écrit de manière unique sous la forme

$$x = O + t_1.\overrightarrow{e}_1 + ... + t_n.\overrightarrow{e}_n \qquad (*)$$

Avec x_i défini par $x_i = O + e_i$, la famille $(O, x_1, ..., x_n)$ est appelée **repère affine de** X *; c'est en particulier une famille affinement libre. Il est suffisant pour qu'une famille affinement libre soit un repère affine que son cardinal soit $(n+1)$.*
$(O, \overrightarrow{e}_1, ..., \overrightarrow{e}_n)$ *est appelé* **repère cartésien de** X.
Les t_i sont appelés **coordonnées cartésiennes** *de x.*
On note qu'avec $t_0 = 1-t_1-t_2-...-t_n$, la relation $(*)$ équivaut à la relation

$$\overrightarrow{Ox} = t_0.\overrightarrow{OO} + t_1.\overrightarrow{Ox_1} + t_2.\overrightarrow{Ox_2} + ... + t_n.\overrightarrow{Ox_n}$$

avec $\sum_{i=0}^n t_i = 1$, si bien que pour tout M

$$\overrightarrow{Mx} = t_0.\overrightarrow{MO} + t_1.\overrightarrow{Mx_1} + t_2.\overrightarrow{Mx_2} + ... + t_n.\overrightarrow{Mx_n}$$

On a (dans ce cas, c'est-à-dire avec les $\overrightarrow{e_i}$ formant une base de $\overrightarrow{X}$) existence et unicité de cette décomposition barycentrique.
Les t_i tels que $\sum_{i=0..n} t_i \neq 0$ et $(\sum_{i=0..n} t_i).\overrightarrow{Mx} = t_0.\overrightarrow{MO} + t_1.\overrightarrow{Mx_1} + t_2.\overrightarrow{Mx_2} + ... + t_n.\overrightarrow{Mx_n}$ sont appelés **des coordonées barycentriques de** x **dans le repère affine** $(O, x_1, ..., x_n)$.
Les t_i tels que $\sum_{i=0..n} t_i = 1$ et $\overrightarrow{Mx} = t_0.\overrightarrow{MO} + t_1.\overrightarrow{Mx_1} + t_2.\overrightarrow{Mx_2} + ... + t_n.\overrightarrow{Mx_n}$ sont appelés **les coordonées barycentriques normalisées de** x **dans le repère affine** $(O, x_1, ..., x_n)$.

La méthode de passage des coordonées cartésiennes aux coordonnées barycentriques est à retenir ; si $\overrightarrow{OM} = \sum t_i\overrightarrow{Ox_i}$, alors les coordonnées barycentriques normalisées de M dans $(O, x_1, x_2, ..., x_n)$ sont $(1-\sum t_i, t_1, t_2, ...t_n)$.

On remarque (preuves faciles) que :

• une famille finie est affinement libre si et seulement si aucun de ses points ne peut s'exprimer comme barycentre des autres.

• deux points distincts forment toujours une famille affinement libre

• étant donné un repère vectoriel R, une famille de $(n+1)$ points $x_1, ..., x_{n+1}$ en dimension n est affinement libre si et seulement si le déterminant suivant est non nul, avec x_i de coordonnées cartésiennes $(t_{i,1}, ..., t_{i,n})$ dans R :

$$\begin{vmatrix} t_{1,1} & t_{2,1} & \dots & t_{n,1} & t_{n+1,1} \\ t_{1,2} & t_{2,2} & \dots & t_{n,2} & t_{n+1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{1,n} & t_{2,n} & \dots & t_{n,n} & t_{n+1,n} \\ 1 & 1 & \dots & 1 & 1 \end{vmatrix}$$

(il suffit pour le voir de soustraire la première ligne à toutes les autres puis de revenir à la définition d'un famille affinement libre (proposition-définition 1234))

• étant donné un repère affine, une famille de $n+1$ points en dimension n est affinement libre si et seulement si le produit mixte de leurs vecteurs de coordonées barycentriques dans ce repère est non nul. Autrement dit les $(n+1)$ points $x_0, ..., x_n$, avec x_i de coordonnées barycentriques $(u_{i,0}, ..., u_{i,n})$ forment une famille affinement libre si et seulement si le déterminant suivant est non nul :

$$\begin{vmatrix} u_{1,0} & u_{2,0} & \dots & u_{n,0} & u_{n+1,0} \\ u_{1,1} & u_{2,1} & \dots & u_{n,1} & u_{n+1,1} \\ u_{1,2} & u_{2,2} & \dots & u_{n,2} & u_{n+1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{1,n} & u_{2,n} & \dots & u_{n,n} & u_{n+1,n} \end{vmatrix}$$

(il suffit pour le comprendre de considérer ce déterminant et de remplacer la première ligne par la somme de toutes les lignes - puis de se ramener au point précédent)

Définition 1235 - Proposition *On se place dans un espace affine de dimension finie n.*
On a vu qu'une famille de n points est liée si et seulement si le produit mixte de leurs coordonnées barycentriques dans un repère affine donné est nul. Donc, si l'on se donne n points affinement libres, l'ensemble des points appartenant à l'hyperplan engendré par ces n points $(P_i)_{i\in[1,n]}$, avec P_i de coordonnées barycentriques $(t_{i,j})_{j\in[0,n]}$, est l'ensemble des points de coordonnées barycentriques $x_0, ..., x_n$ telles que

$$\begin{vmatrix} t_{1,0} & t_{2,0} & \dots & t_{n,0} & x_0 \\ t_{1,1} & t_{2,1} & \dots & t_{n,1} & x_1 \\ t_{1,2} & t_{2,2} & \dots & t_{n,2} & x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{1,n} & t_{2,n} & \dots & t_{n,n} & x_n \end{vmatrix} = 0$$

En développant le déterminant suivant la dernière colonne, on obtient une équation linéaire en les x_i, de la forme $u_0.x_0 + u_1.x_1 + \cdots + u_n.x_n = 0$; le vecteur $(u_0, u_1, ..., u_n)$ est un vecteur de **coordonnées tangentielles** *de l'hyperplan.*
Les coordonnées tangentielles d'un hyperplan sont uniques à proportionalité près. Un n-uplet est un vecteur de coordonnées tangentielles d'un hyperplan si et seulement si ses coordonnées ne sont pas toutes nulles.

32.4 Applications affines

Définition 1236 - Proposition
Une application f de X dans X' (deux espaces affines) ; est dite **affine** *si l'une des deux conditions équivalentes suivantes est vérifiée :*
• *il existe x tel que l'application $\overrightarrow{u} \mapsto \overrightarrow{f(x)f(x+u)}$ soit une application linéaire de $\overrightarrow{X}$ dans $\overrightarrow{X}'$;*
• *pour tout x l'application $\overrightarrow{u} \mapsto \overrightarrow{f(x)f(x+u)}$ est une application linéaire de $\overrightarrow{X}$ dans $\overrightarrow{X}'$.*
L'application linéaire est alors indépendante du point x_0 ; on l'appelle **application linéaire associée à** *f ; on la note $\overrightarrow{f}$.*
Si $\overrightarrow{f}$ est une forme linéaire, c'est-à-dire si X' est de dimension 1 (est une droite), ont dit que f est une **forme affine***.*
Un **isomorphisme d'espaces affines** *est une bijection linéaire entre deux espaces affines .*

Proposition 1237 *L'application linéaire* $\overrightarrow{f}$ *associée à une application affine* f *vérifie*

$$\forall (x,y) \in X^2 \overrightarrow{f}(\overrightarrow{xy}) = \overrightarrow{f(x)f(y)}$$

ce que l'on peut aussi écrire

$$\forall (x,y) \in X^2 f(y) = f(x) + \overrightarrow{f}(\overrightarrow{xy})$$

Etant donné x *appartenant à* X*, avec* X *un espace affine , deux applications affines de même espace de départ* X *et de même espace d'arrivée* X' *sont égales si et seulement si leurs applications linéaires associées sont égales et si elles coïncident en* x*.*
Toute application affine f *d'un* $\mathbb{R}^n$ *dans un* $\mathbb{R}^m$ *est différentiable en tout point ; la différentielle de* f *(en n'importe quel point) est égale à* $\overrightarrow{f}$*.*
f *d'un espace affine dans un autre est affine si et seulement si pour toute famille* $(x_1, ..., x_n)$ *d'éléments de* X *et toute famille* $(t_1, ..., t_n)$ *de réel de somme* $\sum_{i=1}^n t_i = 1$ *on a*

$$f(\sum_{i=1}^n t_i.x_i) = \sum_{i=1}^n t_i.f(x_i)$$

Une composée d'applications affines est affine.
La réciproque d'une bijection affine est une bijection affine.
L'ensemble des bijections affines d'un espace affine E *sur lui-même est un groupe pour la composition. On l'appelle* **groupe affine** *de* E *(noté* $GA(E)$*).*
On pourra consulter21.10.5 pour plus d'informations.
Une application f *d'un espace affine dans un autre est affine si et seulement si elle conserve le barycentre, c'est-à-dire si l'image du barycentre des* (x_i, t_i) *est le barycentre des* $(f(x_i), t_i)$ *pour toute famille finie* (x_i) *et toute famille de scalaires* (t_i)*.*

Proposition 1238 (Sur les points fixes des applications affines) *Si* f *est affine de* X *dans* X *(avec* X *un espace affine) et a un point fixe* x*, alors l'application* f *induit une application linéaire du vectorialisé de* f *en* x *dans le vectorialisé de* f *en* x*.*
Si f *affine de* X *dans* X *admet un point fixe* x *alors l'ensemble des points fixes de* f *est* $x + F$ *avec* F *le noyau de* $(f - I)$*.*
Si X *est un espace affine de dimension finie, si* f *est affine de* X *dans* X *avec* $\overrightarrow{f}$ *ayant un unique point fixe, alors* f *a un unique point fixe (et l'unique point fixe de* $\overrightarrow{f}$ *est nécéssairement* 0*).*

Définition 1239 *Une application affine* f *est appelée* **homothétie affine** *si et seulement si* $\overrightarrow{f}$ *est une homothétie (i.e. de la forme* $x \mapsto \lambda.x$ *avec* $\lambda \neq 1$*).*

Proposition 1240 *Une homothétie affine admet un et un seul point fixe.*

Démonstration : Application immédiate du dernier point de la proposition ci-dessus, une homothétie linéaire de rapport différent de 1 ayant 0 pour unique point fixe.$\square$

32.5 Sous-espaces affines d'un espace affine

Définition 1241 *On appelle* **sous-espace affine** *d'un espace affine X une partie P telle que l'une des deux propriétés équivalentes suivantes soit vérifiée pour un certain sous-espace vectoriel F de $\overrightarrow{X}$:*
• il existe x dans X tel que $P = F + x$
• pour tout x dans X $P = F + x$; F est la **direction** *de X.*
Deux sous-espaces affines d'un espace affine sont dits **supplémentaires** *si et seulement si leurs directions sont supplémentaires.*
On appelle **hyperplan affine** *d'un espace affine X un sous-espace affine de X admettant un sous-espace affine supplémentaire de dimension 1.*

Un sous-espace affine est un espace affine .

Définition 1242 - Proposition *Etant donné P une partie non vide d'un espace affine X, l'ensemble des barycentres de parties finies de P est un sous-espace affine de X ; on l'appelle* **sous-espace affine engendré par** P.

Proposition 1243 *Une partie non vide P d'un espace affine X est un sous-espace affine de X si et seulement si tout barycentre d'une famille finie de points de P est dans P.*
Une partie non vide P d'un espace affine X est un sous-espace affine de X si et seulement si P contient toute droite engendrée par deux points de P.
Le sous-espace affine engendré par une partie finie non vide P est de dimension $card(P)-1$ (cardinal de P moins 1) si et seulement si la famille des éléments de P est affinement libre.
L'image d'un sous-espace affine Y par une application affine f d'un espace affine X dans un espace affine X' est un sous-espace affine de l'espace affine X' de direction $\overrightarrow{f}(\overrightarrow{Y})$.
L'image réciproque d'un sous-espace affine Y par une application affine f est soit vide soit un sous-espace affine de direction l'image réciproque par $\overrightarrow{f}$ de $\overrightarrow{Y}$.
Si Y est parallèle à Z, avec Y et Z deux sous-espaces affines de X, si f est une application affine, alors $f(Y)$ est parallèle à $f(Z)$.
Un sous-espace affine Y de dimension p d'un espace affine X de dimension n s'exprime comme intersection de $(n-p)$ images réciproques de singletons par des formes affines indépendantes ; c'est-à-dire

$$Y = \{x \in X / \forall i \in [1, n-p] f_i(x) = b_i\}$$

pour une certaine famille f_i de formes affines indépendantes et des scalaires b_i. La direction de Y est l'espace vectoriel

$$\overrightarrow{Y} = \{\overrightarrow{x} \in \overrightarrow{X} / \forall i \in [1, n-p], \overrightarrow{f_i}(\overrightarrow{x}) = 0\}$$

Etant donnés A et B deux sous-espaces affines disjoints d'un espace affine X, la dimension du sous-espace affine engendré par $A \cup B$ est égale à la dimension de $\overrightarrow{A} + \overrightarrow{B}$ plus un.
Etant donnés A et B deux sous-espaces affines de dimension finie et non disjoints d'un espace affine X, la dimension du sous-espace affine engendré par $A \cup B$ est égale à la dimension de A plus la dimension de B moins la dimension de $A \cap B$ ($\dim A + \dim B - \dim A \cap B = \dim\ Vect(A \cup B)$).

32.6 Projections dans un espace affine

Définition 1244 *On se donne Y et Z deux sous-espaces affines supplémentaires de X (un espace affine). Alors (voir plus haut) l'intersection de Y et Z est réduite à un singleton ; appelons O ce singleton. On considère alors $\overrightarrow{X}$ le vectorialisé de X et O, $\overrightarrow{Y}$ le vectorialisé de Y en O, et $\overrightarrow{Z}$ le vectorialisé de Z en O. Tout point x de X est aussi un point de $\overrightarrow{X}$; or $\overrightarrow{X} = \overrightarrow{Y} \oplus \overrightarrow{Z}$; donc $x = y + z$ avec y dans $\overrightarrow{Y}$ et z dans $\overrightarrow{Z}$. On appelle y le* **projeté de** x **sur** Y **parallèlement à** Z. *L'application qui à x associe son projeté sur Y parallèlement à Z est appelée* **projecteur sur** Y **parallèlement à** Z.

On remarque que si Z et Z' sont parallèles alors le projecteur sur Y parallèlement à Z est égal au projecteur sur Y parallèlement à Z'.

Voyons quelques propriétés des projecteurs :

Proposition 1245 (Propriétés des projecteurs) *En notant p le projecteur sur Y parallèlement à Z :*
- *pour tout x dans X $p(x)$ est l'unique point de Y tel que $\overrightarrow{xp(x)} \in \overrightarrow{Z}$, avec $\overrightarrow{Z}$ la direction de Z.*
- *$p \circ p = p$ (p est idempotent)*
- *p est affine*
- *$p(X) = Y$*
- *p induit l'identité sur Y*
- *$\overrightarrow{Z}$ est le noyau de $\overrightarrow{p}$.*

Et maintenant des caractérisations :

Proposition 1246 (Caractérisations des projecteurs)
- *Toute application affine idempotente est un projecteur.*
- *Toute application f affine ayant un point fixe et telle que $\overrightarrow{f} \circ \overrightarrow{f} = \overrightarrow{f}$ est un projecteur*

32.7 Symétries dans un espace affine

Définition 1247 *Etant donné X un espace affine , le point $x+\frac{1}{2}\overrightarrow{xy}=y+\frac{1}{2}\overrightarrow{yx}$ est appelé* **milieu de** *xy. C'est le barycentre de $(x,\frac{1}{2})$ et $(y,\frac{1}{2})$.*
On appelle **symétrie** *une application affine f d'un espace affine dans lui-même telle que $f \circ f = I$ (i.e. f est involutive).*

Quelques caractérisations :

Proposition 1248 • *Une application f d'un espace affine dans lui-même est une symétrie si et seulement si f est affine, admet un point fixe, et vérifie $\overrightarrow{f} \circ \overrightarrow{f} = \overrightarrow{f}$*
• *Une application f d'un espace affine X dans lui-même est une symétrie si et seulement si il existe deux sous-espaces affines Y et Z supplémentaires de X tels que, pour tout x, le milieu de $xf(x)$ appartienne à Y et $\overrightarrow{xf(x)} \in \overrightarrow{Z}$.*
• *Une application f d'un espace affine X dans lui-même est une symétrie si et seulement si il existe p un projecteur de X tel que $\overrightarrow{xf(x)} = 2.\overrightarrow{xp(x)}$.*
• *Une application f d'un espace affine X dans lui-même est une symétrie si et seulement si il existe p un projecteur de X tel que $p(x)$ soit le milieu de $xf(x)$.*

Définition 1249 *Le projecteur évoqué dans les deux derniers points de la proposition ci-dessus est unique ; la symétrie et le projecteur en question sont dits* **associés***.*

Définition 1250 - Proposition *Une symétrie f est entièrement caractérisée par l'ensemble Y de ses points fixes et par le sous-espace vectoriel $\overrightarrow{Z}$ de $\overrightarrow{X}$ des vecteurs de la forme $\overrightarrow{xf(x)}$; on l'appelle* **symétrie par rapport à** *Y* **parallèlement à** *$\overrightarrow{Z}$ (ou par rapport à Z, avec Z sous-espace affine quelconque de X de direction $\overrightarrow{Z}$).*
Le sous-espace vectoriel $\overrightarrow{Z}$ de $\overrightarrow{X}$ est aussi le sous-espace propre associé à la valeur propre -1 pour l'endomorphisme $\overrightarrow{f}$.

32.8 Mesure dans un espace affine

32.8.1 Abscisse le long d'une droite

Définition 1251 *On se donne X un espace affine .*
Etant donné a et b distincts dans X, l'ensemble des $a + t.\overrightarrow{ab}$ est une droite (i.e. un sous-espace affine de dimension 1). Tout point x de cette droite s'écrit de manière unique $a + t.\overrightarrow{ab}$; t est l'abscisse de x suivant $(a, \overrightarrow{ab})$. Etant donnés x et y sur la même droite, la différence $t_x - t_y$ avec t_x (resp. t_y) l 'abscisse de x (resp. y) est appelée **mesure algébrique de l'arc** (xy) **suivant** $(a, \overrightarrow{ab})$.

32.8.2 Mesure de Lebesgue sur $\mathbb{R}^n$

Cette partie ne sera pas détaillée formellement.

La définition du volume par la mesure de Lebesgue coïncide avec la définition du volume donnée par le produit mixte : le volume $\{O + t_1.x_1 + ... + t_n.x_n/(t_1, ..., t_n) \in [0, 1]^n\}$ pour la mesure de Lebesgue est égal à la valeur absolue du produit mixte $|[x_1, ..., x_n]|$.

Avec P une partie affine de X, espace affine , et f une application affine de X dans lui-même, alors $Volume(f(P)) = |det\ \overrightarrow{f}|.Volume(P)$.

32.9 Définitions supplémentaires

On donne ici quelques définitions pouvant servir, sans développer intensément...

Définition 1252 *Etant donné X un espace affine , H un hyperplan affine, D une droite affine supplémentaire de H, et λ un scalaire non nul, on appelle* **dilatation affine d'hyperplan** H**, de direction** D **et de rapport** λ *l'application dont la restriction au vectorialisé de H en $H \cap D$ est l'identité, et dont la restriction au vectorialisé de D en $H \cap D$ est une homothétie de rapport λ.*

Une dilatation est une bijection affine. Sa bijection réciproque s'obtient en remplaçant λ par $\frac{1}{\lambda}$.

L'application linéaire associée à une dilatation affine est une dilatation (au sens des espaces vectoriels).

Définition 1253 *Etant donné X un espace affine , on appelle* **transvection affine** *une application f telle qu'il existe H un hyperplan affine, h une forme affine sur X telle que $\{x \in X/h(x) = 0\} = H$, un vecteur $\overrightarrow{u}$ dans $\overrightarrow{H}$ tels que pour tout x dans X, $f(x) = x + h(x).\overrightarrow{u}$.*

⚠ Bien noter que $\overrightarrow{u} \in \overrightarrow{H} = \ker(\overrightarrow{h})$.

Une transvection est une bijection affine. Sa transvection réciproque s'obtient en remplaçant $\overrightarrow{u}$ par $-\overrightarrow{u}$.

L'application linéaire associée à une transvection affine est une transvection (au sens des espaces vectoriels).

32.10 Pour se ramener à l'algèbre linéaire

Il est bien évident au vu de tout ceci que géométrie affine et algèbre linéaire sont très fortement liées. On va préciser ces liens dans cette partie.

On se donne X un espace affine .

Théorème 1254 *Etant donné X un espace affine de direction $\overrightarrow{X}$, il existe un espace vectoriel $\hat{X}$ tel que X est un hyperplan affine de $\hat{X}$, $\overrightarrow{X}$ est un hyperplan (vectoriel) de $\hat{X}$;*
en outre :
• chaque élément de $\hat{X}$ qui n'est pas dans $\overrightarrow{X}$ s'écrit sous la forme $\lambda.x$ avec $x \in X$ et $\lambda \in \mathbb{K}$;
• pour tout élément y dans X, $\hat{X} = X \oplus \mathbb{K}.y$;
• toute famille affinement libre de X est une famille libre de $\hat{X}$;
• les repères affines de X sont les bases de $\hat{X}$ formées d'éléments de X ;
• les coordonnées barycentrique (normalisées) dans un repère affine de X sont exactement les coordonnées dans la même base en tant que base de $\hat{X}$.
• Etant donné $(R, e_1, ..., e_n)$ un repère cartésien de X, si $x \in X$ a pour coordonnées cartésiennes $(x_1, ..., x_n)$ dans ce repère, x a pour coordonnées $(1, x_1, ..., x_n)$ dans la base de $\hat{X}$ égale à $(R, e_1, ..., e_n)$ (bien voir que R est un élément de X, donc un élément de $\hat{X}$, et que e_i est un élément de $\overrightarrow{X}$, donc aussi un élément de X).

Il n'y a pas unicité de $\hat{X}$, mais on peut le construire explicitement en considérant l'ensemble réunion de X, de $\overrightarrow{X}$, et de l'ensemble des couples $(\lambda, \overrightarrow{x})$ appartenant à $(\mathbb{K} \setminus \{0, 1\}) \times \overrightarrow{X}$. Les lois sont celles que l'on attend en identifiant $\overrightarrow{X}$ à $\{0\} \times \overrightarrow{X}$ et X à $\{1\} \times \overrightarrow{X}$.

Définition 1255 *Cet espace vectoriel est noté par la suite $\hat{X}$.*

Enfin un petit théorème immédiat :

Théorème 1256 *Pour toute application affine f de X dans X', il existe une et une seule application linéaire $\hat{f}$ de $\hat{X}$ dans $\hat{X}'$ telle que $\hat{f}$ induise f. En outre $\overrightarrow{f}$ est la restriction de f à $\overrightarrow{X}$.*

32.11 Formes 2-affines et quadriques affines

Définition 1257 *On appelle* **forme 2-affine** *une application d'un espace affine X de dimension finie dans $\mathbb{K}$ telle qu'étant donné un certain repère cartésien R, $f(x) = \sum_{(i,j)\in[1,n]^2} A_{i,j}.x_i.x_j + 2\sum_{i\in[1,n]} b_i.x_i + c$, avec x_i les coordonnées cartésiennes de x dans R et b un vecteur à n composantes, c un réel, A une matrice de type (n,n) qui ne soit pas antisymétrique.*
On appelle **quadrique affine** *d'un espace affine de dimension finie une équation du type $f(x) = 0$ avec f une forme 2-affine sur cet espace affine .*

Proposition 1258 • *La définition ci-dessus stipule seulement qu'il existe un certain repère cartésien dans lequel on peut exprimer ainsi l'application ; en fait s'il en est ainsi, l'application s'exprime de même dans tout référentiel cartésien.*
• *On peut aussi imposer, sans changer la notion de forme 2-affine, que la matrice A soit symétrique non nulle.*

Proposition 1259 *L'application $f \mapsto f_{|X}$ de l'ensemble des formes quadratiques sur $\hat{X}$ dont la restriction à $\overrightarrow{X}$ n'est pas la forme quadratique nulle dans l'ensemble des formes 2-affines sur X est une bijection.*

Ce résultat est important ; il permet donc de ramener l'étude d'une forme 2-affine sur un espace affine de dimension n à l'étude de la restriction à un hyperplan d'une forme quadratique sur un espace vectoriel de dimension $(n+1)$.

32.12 Zoologie de la géométrie affine

32.12.1 Etude de quadriques en dimension 3

Les figures données ici sont obtenues avec Maple. Les commandes ne sont pas toutes données, car elles sont très similaires. On se contentera de donner la commande fournissant la figure 32.3 : il s'agit de

with(plots) ; implicitplod3d(x2̂-y2̂=z,x=-3..3,y=-3..3,z=-3..3,axes=BOXED,style=PATCHCONTOUR,numpoints=10000) ;

Pour les autres figures, la modification de "numpoints" n'est pas utile.

⊡ Quadriques de rang 3 (alias quadriques à centre)

Ces quadriques sont celles de signature $(3,0)$ (ellipsoïde éventuellement réduit à un point, figure 32.1 à gauche), $(2,1)$ (hyperboloïde à 1 nappe 32.1 à droite, à 2 nappes figure 32.2 à gauche).

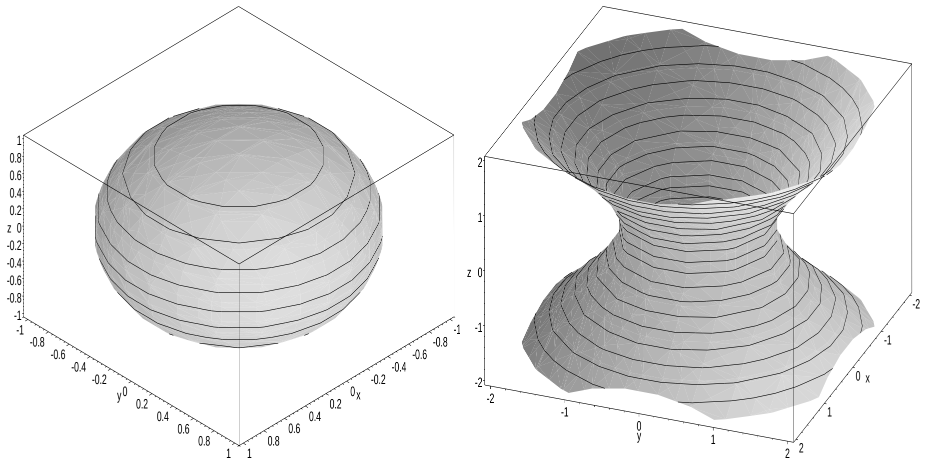


FIG. 32.1 – $x^2 + y^2 + z^2 = 1$ $x^2 + y^2 - z^2 = 1$

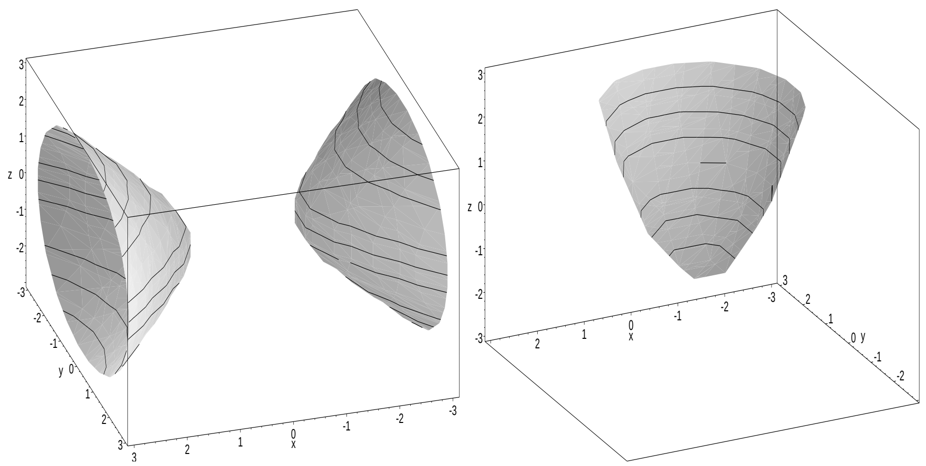


FIG. 32.2 – $x^2 - y^2 - z^2 = 1$ $x^2 + y^2 = 1 + z$

⊡ Quadriques de rang 2

Ces quadriques sont le paraboloïde elliptique (figure 32.2, à droite) et le paraboloïde hyperbolique (figure 32.3).

32.12.2 Exemples très banals d'espaces affines

⊡ Espace vectoriel

On se donne E un espace vectoriel, par exemple $\mathbb{R}^n$; alors $X = E$ muni de l'application $(x, y) \mapsto \overrightarrow{xy} = y - x$ est un espace affine.

⊡ Avec une bijection sur un espace vectoriel

Etant donné X un ensemble et E un espace vectoriel , avec f une bijection de X sur E, alors l'application $(x, y) \mapsto \overrightarrow{xy} = f(y) - f(x)$ de $X \times X$ dans E définit une

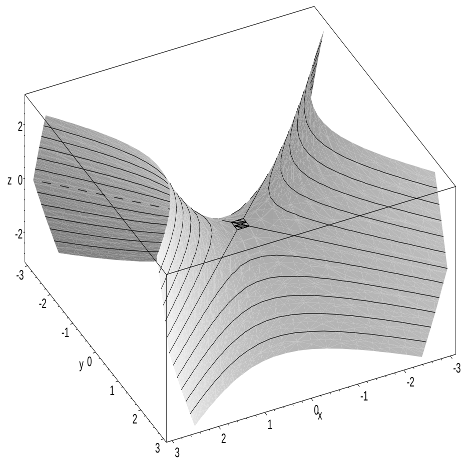


FIG. 32.3 – $x^2 - y^2 = 1 + z$

structure d'espace affine sur X. En fait, tout espace affine peut se définir de cette façon ; mais il n'y a pas unicité de la fonction f (unicité à composition par une translation près toutefois).

32.12.3 Solutions d'un système linéaire

Etant donné un système linéaire $A.X = b$ avec A une matrice de type (n, p), X l'inconnue (vecteur colonne de dimension p), et b un vecteur colonne de taille n, l'ensemble des solutions est un espace affine, de direction l'espace vectoriel des solutions de $A.X = 0$.

S'il existe une solution, la dimension de l'espace affine des solutions est égal à $p-r$ avec r le rang de A.

Donc pour trouver toutes les solutions, il suffit de trouver une solution X_0, et l'espace des solutions est alors $X_0 + Ker(X \mapsto A.X)$.

32.12.4 Solutions d'une équation différentielle

Du premier ordre

Etant donné I un intervalle de $\mathbb{R}$, u et v des applications continues de I dans $\mathbb{R}$, les solutions de l'équation différentielle du premier ordre

$$y'(t) = u(t).y(t) + v(t)$$

forment un espace affine de direction l'ensemble des $y(t)$ tels que $y'(t) = u(t).y(t)$.

Cet espace affine est de dimension 1, sa direction est une droite vectorielle (de l'espace vectoriel des fonctions continues de I dans $\mathbb{R}$) engendrée par l'application $x \mapsto exp(U(x))$ (de I dans $\mathbb{R}$) avec $U' = u$ sur I.

Du second ordre

Etant donné I un intervalle de $\mathbb{R}$, u, v et w des applications continues de I dans $\mathbb{R}$, les solutions de l'équation différentielle du deuxième ordre

$$y''(t) = u(t).y'(t) + v(t).y(t) + w(t)$$

forment un espace affine de direction l'ensemble des $y(t)$ tels que $y''(t) = u(t).y'(t) + v(t)y(t)$.

Cet espace affine est de dimension 2, sa direction est un plan vectoriel (de l'espace vectoriel des fonctions continues de I dans $\mathbb{R}$), dont on trouvera une base en utilisant la transformation donnée en partie 13.3 pour se ramener à une équation d'ordre 1.

32.12.5 Pour les algébristes, action du groupe additif d'un espace vectoriel sur un ensemble

On se donne X un ensemble quelconque, et E un espace vectoriel , et on considère une action du groupe $(E, +)$ sur X ; on impose que cette action soit fidèle et transitive, c'est-à-dire que

$$e.(f.x) = (ef).x \text{ et } 0.x = x \text{ (définition d'une action)}$$

et

$$\forall (x, y) \in X\ \exists! u \in E/u.x = y$$

(existence = transitivité, unicité = fidélité). Alors X est un espace affine de direction E pour l'application qui à (x, y) associe l'élément u tel que $u.x = y$.

La notation multiplicative, usuelle pour les actions de groupe, est peut-être ici assez malvenue, du fait que l'on se retrouve avec $0.x = x$... On peut remplacer $e.x$ pour (e, x) dans $E \times X$ par $T_e(x)$, et employer le vocabulaire plus imagé des translations.

Chapitre 33

Géométrie projective

33.1 Homographies, birapport et droite projective

33.1.1 Théorie

Définition 1260 - Proposition *On appelle* **droite projective associée au corps** $\mathbb{K}$ *et on note* $P^1(\mathbb{K})$ *l'ensemble* $\mathbb{K} \cup \{\infty\}$*, sur lequel on prolonge les lois d'addition et de multiplication en posant* $x + \infty = \infty$ *pour tout* $x \in \mathbb{K} \cup \{\infty\}$ *et* $\{x.\infty\} = \infty$ *pour tout* $x \in \mathbb{K}^* \cup \{\infty\}$*, le produit* $0.\infty$ *n'étant pas défini. On prolonge la division en posant* $x/\infty = 0$ *et* $x/0 = \infty$ *pour tout* $x \in \mathbb{K}^* \cup \{\infty\}$.
Toute application affine f *de* $\mathbb{K}$ *dans* $\mathbb{K}$ *peut être prolongée en une application de* $P^1(\mathbb{K})$ *dans* $P^1(\mathbb{K})$*, en posant* $f(\infty) = \infty$.
Soit M *une matrice de* $GL_2(\mathbb{K})$*, noté* $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *; alors* **l'homographie de** $P^1(\mathbb{K})$ **associée à** M *est par définition l'application qui à* x *dans* $P^1(\mathbb{K})$ *associe*
• $\frac{a.x+b}{c.x+d}$ *si* $c.x + d \neq 0$
• ∞ *si* $c.x + d = 0$
• $\frac{a}{c}$ *si* $x = \infty$
On note cette application $H(M)$.
Si $c \neq 0$, $H(M)$ *induit une bijection de* $\mathbb{K} \setminus \{-\frac{d}{c}\}$ *sur* $\mathbb{K} \setminus \{\frac{a}{c}\}$*. On en déduit que* $H(M)$ *est une bijection de* $P^1(\mathbb{K})$ *sur* $P^1(\mathbb{K})$*. Un calcul simple montre que sa réciproque est l'homographie associée à l'inverse de la matrice* M.
L'ensemble des homographies de $P^1(\mathbb{K})$ *est un groupe pour la composition. On l'appelle* **groupe projectif de** $\mathbb{K}$ *; on le note* $GP^1(\mathbb{K})$.
L'application qui à M *associe* $H(M)$ *est un morphisme du groupe* $GL_2(\mathbb{K})$ *dans* $GP^1(\mathbb{K})$*. Son noyau est l'ensemble* $(\mathbb{K}^*).I$*, c'est-à-dire le groupe des matrices non nulles proportionnelles à la matrice identité (ce noyau, comme tout noyau d'un morphisme de groupe, est distingué).*
On en déduit donc $GP^1(\mathbb{K}) \simeq GL_2(\mathbb{K})/\mathbb{K}^*.I = PGL_2(\mathbb{K})$.

Définition 1261 - Proposition *Etant donné une droite affine D (c'est-à-dire un espace affine de dimension 1), on peut considérer sa* **droite projective complétée** $\widetilde{D}$*, consistant en $D \cup \{\infty\}$. Une abscisse m sur D se prolonge en une bijection de $\widetilde{D}$ sur $P^1(\mathbb{K})$, par $m(\infty) = \infty$; cette application sera encore appelée* **abscisse***, sur $\widetilde{D}$.*
On appelle **homographie de $\widetilde{D}$ sur $\widetilde{D}'$** *une application h de D sur D' telle qu'il existe des abscisses m et m' respectivement sur $\widetilde{D}$ et $\widetilde{D}'$ telles que $h = m'^{-1} \circ H \circ m$, avec H une homographie de $P^1(\mathbb{K})$.*
Un composée d'homographies (quel que soit le contexte) est encore une homographie.
Une homographie est toujours bijective.
L'ensemble des homographies de $\widetilde{D}$ sur lui-même forme un groupe pour $\circ$, isomorphe à $GP^1(\mathbb{K})$; on le note $GP(\widetilde{D})$, et on l'appelle **groupe projectif de** $\widetilde{D}$*.*
Etant donnée E une droite affine munie d'un repère affine, donc d'une abscisse, on se donne 4 points A, B, C et D <u>distincts</u> d'abscisses respectives a, b, c et d.
On appelle **birapport de A, B, C et D** *et on note $[A,B,C,D] = \frac{a-c}{a-d}.\frac{b-d}{b-c}$.*
On a les propriétés suivantes du birapport :
• *le birapport est indépendant de l'abscisse choisie.*
• *le birapport n'est <u>pas</u> indépendant de l'ordre.*
• $[A,B,C,D] = [C,D,A,B]$
• $[A,B,C,D] = \frac{1}{[A,B,D,C]} = \frac{1}{[B,A,C,D]}$
• $[A,B,C,D] + [A,C,B,D] = 1$
En constatant que le birapport avec 3 points fixés est une homographie, on va la prolonger.
Ainsi, on prolonge le birapport au cas où l'un des points est ∞, de la manière logique intuitivement, c'est-à-dire que les termes incluant ∞ "se compensent". Ainsi on a par exemple

$$[A,B,C,\infty] = \frac{a-c}{b-c}$$

$$[A,B,\infty,D] = \frac{b-d}{a-d}$$

$$[A,\infty,C,D] = \frac{a-c}{a-d}$$

$$[\infty,B,C,D] = \frac{b-d}{b-c}$$

Ensuite, on prolonge le birapport au cas où deux points sont égaux :

$$[A,A,C,D] = [A,B,C,C] = 1$$

$$[A,B,A,D] = [A,B,C,B] = 0$$

$$[A,B,C,A] = [A,B,B,D] = \infty$$

Le birapport est invariant par homographie : $\forall A,B,C,D \in P^1(\mathbb{K}), \forall f \in PGL_2(\mathbb{K}), [f(A),f(B),f(C),f(D)] = [A,B,C,D]$.

Définition 1262 - Proposition
Toute homographie de $\widetilde{D}$, avec D une droite affine, s'exprime de manière unique sous la forme $M \mapsto [M, A, B, C]$ pour un certain triplet (A, B, C) de points distincts de $\widetilde{D}$. On appelle **repère projectif de** $\widetilde{D}$ *un triplet A, B, C de points distincts, et $[M, A, B, C]$ est appelé* **coordonnée de** M **dans le repère projectif** (A, B, C).
L'application qui à un point associe sa coordonnée dans un repère projectif est égale à composition par une homographie près à l'application qui à un point associe sa coordonnée dans un autre repère projectif.
Etant donnés deux triplets de points distincts (A, B, C) et (A', B', C') respectivement sur $\widetilde{D}$ et $\widetilde{D}'$ (deux droites projectives), il existe une et une seule homographie h de $\widetilde{D}$ sur $\widetilde{D}'$ telle que $h(A) = A'$, $h(B) = B'$ et $h(C) = C'$.
Etant donnés deux quadruplets de points distincts (A, B, C, D) et (A', B', C', D') respectivement sur $\tilde{E}$ et $\tilde{E}'$ (deux droites projectives), il existe une homographie h de $\tilde{E}$ sur $\tilde{E}'$ telle que $h(A) = A'$, $h(B) = B'$, $h(C) = C'$ et $h(D) = D'$ si et seulement si $[A, B, C, D] = [A', B', C', D']$.
Une bijection entre deux droites projectives est une homographie si et seulement si elle conserve le birapport.
Etant donné $\widetilde{D}$ une droite projective, et $d \in \widetilde{D}$, $\widetilde{D}$ induit sur $E = \widetilde{D} \setminus \{d\}$ une structure de droite affine ; $GA(E)$ est l'ensemble des applications induites sur E par des homographies de $\widetilde{D}$ dont d est un point fixe (on en déduit donc que tous les points de $\widetilde{D}$ jouent le même rôle, même ∞).

33.1.2 Visualisation

On considère le plan vectoriel $\mathbb{K}^2$ muni de sa base canonique, et l'ensemble $\widetilde{D}$ des droites vectorielles de $\mathbb{K}^2$ (donc seulement les droites affines passant par l'origine).

On note $\mathbb{D}$ la droite $\mathbb{K} \times \{0\}$ et D l'ensemble $\widetilde{D} \setminus \{\mathbb{D}\}$.

Etant donné d appartenant à $\widetilde{D}$, on considère $x(d)$ l'abscisse de son intersection avec la droite $\mathbb{K} \times \{1\}$, et on pose $x(\mathbb{D}) = \infty$.

D est une droite affine pour l'application qui à deux droites d et d' associe $\overrightarrow{dd'} = x(d') - x(d)$. $\widetilde{D}$ en est une droite projective complétée. Une abscisse x' sur D, qui est affine par rapport à $d \mapsto x(d)$, se complète en une abscisse sur $\widetilde{D}$ par $x'(\Delta) = \Delta$.

On peut donc visualiser la droite projective $P^1(\mathbb{K})$ comme l'ensemble des droites vectorielles du plan euclidien $\mathbb{K}^2$.

On peut aussi considérer la droite projective complétée comme l'ensemble des points non nuls du plan $\mathbb{K}^2$, en considérant des classes d'équivalence égales aux droites vectorielles privées de 0. L'image est sans doute d'autant plus visualisable, car l'homographie $h(M)$ pour $M \in GL_2(\mathbb{K})$ est alors la multiplication par M dans $\mathbb{K}^2$ quotienté par la relation d'équivalence «être dans le même alignement avec 0».

En identifiant un point d'abscisse t dans un repère affine à la classe d'équivalence $\mathbb{K}.(t, 1)$ si $t \neq \infty$ et à la classe d'équivalence $\mathbb{K}.(1, 0)$ sinon, on peut identifier toute droite projective complétée à l'ensemble des classes d'équivalence précédemment définies de $\mathbb{K}^2 \setminus \{0\}$.

Ainsi une homographie dans une droite projective complétée est l'application quotiente d'une transformation linéaire dans $\mathbb{K}^2$.

Il reste maintenant à passer aux espaces projectifs de dimension supérieure.

33.2 Espaces projectifs

33.2.1 La théorie

Définition 1263 *Etant donné E un espace vectoriel de dimension finie, on appelle* **espace projectif associé à** *E l'ensemble $P(E)$ des droites vectorielles de E. On note $P^{n-1}(\mathbb{K})$ pour $P(\mathbb{K}^n)$.*
On appelle **dimension de** *$P(E)$ la dimension de E moins un.*
Ne pas confondre avec l'ensemble des parties de E !
On appelle **droite de** *$P(E)$ l'image d'un plan vectoriel de E,* **plan vectoriel de** *$P(E)$ l'image d'un sous-espace vectoriel de E de dimension 3,* **sous-espace projectif de** *$P(E)$* **de dimension** *q l'image d'un sous-espace vectoriel de E de dimension $(q+1)$,* **hyperplan projectif de** *$P(E)$ un hyperplan vectoriel de E, sous-espace projectif de $P(E)$ engendré par q points de $P(E)$ (i.e. q droites de E) l'image du sous-espace vectoriel de E engendré par les q droites correspondantes.*
On note $< P >$ le sous-espace projectif de $P(E)$ engendré par une partie P incluse dans $P(E)$.
On dit de q points de $P(E)$ qu'ils sont **projectivement indépendants** *si les droites correspondantes sont en somme directe dans E, c'est-à-dire s'ils engendrent un sous-espace projectif de $P(E)$ de dimension $(q-1)$.*
On dit de n points $x_1, ..., x_n$ de $P(E)$ qu'ils forment un **repère projectif de** *$P(E)$ si la dimension de $P(E)$ est $n-2$ et si toute sous-famille de $n-1$ points des x_i est projectivement indépendante.*
On dit de x dans $P(E)$ qu'il a pour **coordonnées homogènes** *$(t_1, ..., t_n)$* **dans un certain repère projectif** *$(d_0, d_1, ..., d_n)$ si un certain y appartenant à x a pour coordonnées $(t_1, ..., t_n)$ dans une base $(e_0, ..., e_{n-1})$ avec $e_n = \sum_{i=[1,n-1]} e_i$ et $d_i =< e_i >$ pour tout $i \in [1, n]$.*

Cette définition des espaces projectives coïncide avec celle des droites projectives donnée plus tôt.

Les coordonnées homogènes ne sont pas uniques, même dans un repère donné ! Par contre elles sont unique à multiplication par un scalaire non nul près. La preuve en sera plus facile après certaines autres propositions $\rightarrow$ voir plus loin.

Proposition 1264 *Etant donnés F et G deux sous-espaces projectifs d'un espace projectif $P(E)$, on a*

$$dim\ F + dim\ G = dim\ < F \cup G > + dim\ F \cap G$$

Définition 1265 (Sur les repères projectifs) - Proposition • *Si* $(y_1, ..., y_{n-1})$ *est une base de* E*, et si les* x_i *pour* $i \in [1, n]$ *sont des éléments de* $P(E)$ *avec* $y_i \in x_i$*, et* $\sum_{i=1..n-1} y_i \in x_n$ *(intuitivement* x_n *est au milieu des* x_i *pour* $i \in [1, n-1]$ *- la division par* n *est superflue puisque l'on travaille dans* $P(E)$ *à une constante multiplicative près), alors* $(x_1, ..., x_n)$ *est un repère projectif de* $P(E)$*.*

• *Réciproquement, si* $(x_1, ..., x_n)$ *est un repère projectif de* $P(E)$*, alors il existe* $(y_1, ..., y_{n-1})$ *une base de* E*, avec* $y_i \in x_i$*, et* $\sum_{i=1..n-1} y_i \in x_n$*.*

Pour bien voir ce que signifie ce résultat, il faut comprendre que n'importe quel repère projectif s'exprime donc comme l'image d'une base par la surjection canonique de $E \setminus \{0\}$ *sur son quotient PLUS l'image de la somme des éléments de cette base ; et que n'importe lequel des éléments peut s'exprimer comme étant l'image de l'élément somme.*

Une famille $(y_i)_{i\in[1,n-1]}$ *et une famille* $(x_i)_{i\in[1,n]}$ *étant définies comme dans le deuxième point, la base* $(y_i)_{i\in[1,n-1]}$ *et le repère projectif* $(x_i)_{i\in[1,n]}$ *sont dits* **associés***.*

Démonstration :

Remarquons tout d'abord que pour que des droites x_i soient en somme directe, il faut et il suffit qu'une famille quelconque de y_i avec y_i engendrant x_i soit libre ; dans ce cas, toute famille de y_i avec y_i engendrant x_i est libre.

Dans la suite j'assimile « y engendre x » à « y engendre la droite vectorielle engendrée par x », lorsque y est un point de E et x une droite vectorielle privée de 0.

• La famille $x_1, ..., x_{n-1}$ est clairement projectivement indépendante. Il reste à voir que n'importe quelle autre famille de $(n-1)$ éléments est projectivement indépendante. Pour cela, on peut se contenter de la famille $x_2, ..., x_n$, puisque la situation est invariante par permutation des termes du vecteur[1]. On se donne alors $y_1, ..., y_n$ dans E engendrant respectivement $x_1, ..., x_n$. Si $(y_1, ..., y_{n-1})$ est une base de E et $y_n = y_1 + ... + y_{n-1}$, alors la famille obtenue à partir de $(y_1, ..., y_{n-1})$ en remplaçant n'importe quel vecteur par y_n est de nouveau une base de E, si bien que n'importe quelle famille de $(n-1)$ éléments parmi $x_1, ..., x_n$ est projectivement indépendante : $(x_1, ..., x_n)$ est un rep'ere projectif de $P(E)$.

• On se donne y'_i engendrant x_i dans E. Par définition, les y'_i pour $i \in [1, n-1]$ forment une famille libre de E, donc une base de E. Donc y'_n est combinaison linéaire des y'_i pour $i \in [1, n-1]$. On écrit alors

$$y'_n = \sum_{i\in[1,n-1]} \lambda_i.y'_i$$

puis $y_n = y'_n$ et $y_i = \lambda_i.y'_i$ pour $i \in [1, n-1]$, et on a le résultat souhaité sous réserve que les λ_i soient non nuls.

• Les λ_i sont tous non nuls ; en effet, dans le cas contraire, une sous-famille d'au

[1] σ_n agit transitivement sur les paires de $\{1, ..., n\}$, donc, par passage au complémentaire, sur les paires de $(n-1)$-uplets de $\{1, ..., n\}$

plus $n-1$ vecteurs parmi $y'_1, ..., y'_n$. $\square$

> **Proposition 1266 (Coordonnées homogènes)** *Les coordonnées homogènes dans un repère donné sont uniques à multiplication par un scalaire non nul près.*

Démonstration : • Les y_i donnés par le deuxième point de la proposition ci-dessus sont uniques à multiplication par un scalaire non nul près (comme on s'en convaincra en consultant la preuve ci-dessus).

• Les coordonnées homogènes sont donc uniques à multiplication par un scalaire non nul près.$\square$

> **Proposition 1267 (Propriétés des espaces projectifs)** • *Dans un espace projectif, par deux points distincts passe une droite et une seule.*
> • *Dans un espace projectif, l'intersection d'une droite et d'un hyperplan qui ne la contient pas est constituée d'un point et d'un seul.*

> **Proposition 1268 (Propriétés des plans projectifs)** *Dans un plan projectif :*
> • *Deux droites distinctes se coupent en un point et un seul $\rightarrow$ pas de droite sans point commun !*

On a défini laborieusement les homographies dans une droite projective ; on va maintenant les définir dans un espace projectif de dimension quelconque.

> **Définition 1269 (Homographies)** *Soit E et F des espaces vectoriels de même dimension finie, et $f \in Isom(E, F)$. Alors f conservant l'alignement, on peut restreindre f à E privé de 0 et F privé de 0, on a encore une bijection ; on peut alors considérer l'application quotient de f ; on obtient une bijection de $P(E)$ sur $P(F)$. Cette application est appelée* **homographie de** $P(E)$ **sur** $P(F)$.
> *Une* **matrice associée à une homographie de** $P(E)$ **dans** $P(F)$ **dans des repères projectifs** R **et** R' *(de $P(E)$ et $P(F)$ respectivement) est la matrice dans des bases associées à R et R' d'un certain isomorphisme de E dans E' engendrant cette homographie. Un endomorphisme ayant pour matrice cette même matrice dans les mêmes bases est dit lui aussi* **associé** *à cette homographie.*

Il n'y a pas unicité des matrices associées à une homographie ! Ni des endomorphismes ! Il y a par contre unicité à multiplication par un scalaire près, comme on le vérifie dans la partie 21.10.6.

> **Proposition 1270** *Soit R un repère projectif de $P(E)$ et R' un repère projectif de $P(E)$. Avec $x \in P(E)$ et X un vecteur de coordonnées homogènes de x dans R et M une matrice associée à l'homographie H de $P(E)$ dans $P(F)$ pour les repères R et R', $M.X$ est un vecteur de coordonnées homogènes de $H(x)$ dans R'.*

Proposition 1271 *La restriction d'une homographie à un sous-espace projectif est une homographie.*
Etant donnés $P(E)$ et $P(F)$ deux espaces projectifs de même dimension et R_E et R_F des repères projectifs de $P(E)$ et $P(F)$ respectivement ; alors il existe une unique homographie de $P(E)$ sur $P(F)$ par laquelle R_F soit l'image de R_E.

Définition 1272 *L'ensemble des homographies d'un espace projectif $P(E)$ sur lui-même forme un groupe pour $\circ$; on l'appelle* **groupe projectif de** E*, et on le note $PGL(E)$.*

D'après la proposition précédente, $PGL(E)$ agit simplement transitivement sur l'ensemble des repères projectifs de $P(E)$. On trouvera plus d'informationsà 21.10.6.

33.2.2 La visualisation

On a introduit les espaces projectifs en considérant les classes d'équivalence d'un espace vectoriel de dimension finie privé de 0 pour la relation d'équivalence "appartenir à la même droite vectorielle". Certes cette représentation est elle-même bien imagée et intuitive. Toutefois il reste à fournir une représentation rappelant plus la géométrie affine.

J'assimile abusivement une droite vectorielle et la même droite privée de 0 dans la suite de ce paragraphe.

On se donne f une forme linéaire non nulle sur un espace vectoriel E de dimension finie. On considère l'hyperplan affine $F = \{x/f(x) = 1\}$, de direction $\overrightarrow{F}$ avec $\overrightarrow{F} = \{x/f(x) = 0\}$. On considère alors l'ensemble X des éléments de $P(E)$ qui intersectent F ; il est clair qu'il s'agit du complémentaire dans $P(E)$ de $P(\overrightarrow{F})$. On peut en outre l'identifier à F, par l'application qui à un élément de X associe son intersection avec F.

L'ensemble des droites vectorielles contenues dans $\overrightarrow{F}$ est un hyperplan projectif de $P(E)$. L'application qui à une homographie h de $P(E)$ laissant $\overrightarrow{F}$ invariant associe l'application g de F dans F définie par $g(x) = h(x)$ (rappelons que l'on a identifié X et F) est un morphisme injectif de $PGL(E)$ dans $GA(F)$.

En outre pour toute droite affine D de F, il existe un unique point D_∞ de $\overrightarrow{F}$ tel que $D \cup \{D_\infty\}$ soit une droite projective. On a $D_\infty = D'_\infty$ si et seulement si D est parallèle à D'. Enfin l'application qui à une droite de $P(E)$ non contenue dans $\overrightarrow{F}$ associe son intersection avec F est une bijection de l'ensemble des droites projectives de $P(E)$ non contenues dans $\overrightarrow{F}$ dans l'ensemble des droites affines de F.

On peut donc voir $P(E)$ comme un hyperplan de E, muni en outre de points à l'infini complétant les droites, correspondant à la direction d'une droite (i.e. deux droites parallèles ont même direction) ; les droites contenues dans $\overrightarrow{F}$ sont en fait les droites constituées uniquement de points à l'infini.

Enfin il faut bien noter que la notion de point à l'infini est relative ; n'importe quel point de $P(E)$ pourrait être un point à l'infini en choisissant F convenablement.

Pour le redire autrement, un espace projectif est un espace vectoriel, muni de points à l'infini pour compléter les droites ; deux droites parallèles ont alors un point d'intersection à l'infini.

Définition 1273 *Un hyperplan affine H de E ne passant pas par 0, identifié à l'ensemble des éléments de $P(E)$ qui ne sont pas inclus dans $\overrightarrow{H}$, comme précédemment, et complété par les points à l'infini correspondant à ses droites, est appelé* **complété projectif de** H.
*Etant donné H un hyperplan affine de E ne passant pas par 0, $\overrightarrow{H}$ est appelé l'***hyperplan à l'infini** *du complété projectif de E.*

Pour y voir clair, une visualisation en petite dimension sera pratique.

⊡ **En dimension** 1

Un espace projectif de dimension 1 est une droite projective, c'est-à-dire une droite affine, avec en outre un point à l'infini. On peut aussi la voir comme un cercle où les points diamétralement opposés sont identifiés.

⊡ **En dimension** 2

Un espace projectif de dimension 2 est un plan projectif, c'est-à-dire un plan affine, avec en outre des points à l'infini. On peut se représenter cela par un plan, avec un cercle à l'infini ; une droite peut être soit une droite du plan, avec en bonus les deux points du cercle correspondant à sa direction, soit l'unique droite constituée des points à l'infini.

On peut aussi le voir comme la sphère unité de $\mathbb{R}^3$, en identifiant les points diamétralement opposés.

⊡ **En dimension** 3

Un espace projectif de dimension 3 peut se représenter comme un espace tridimensionnel classique, muni de points à l'infini (qu'on peut se représenter comme une sphère loin loin loin). Les droites en sont les droites usuelles (munies des deux points correspondant sur la sphère), plus les grands cercles (i.e. de diamètre maximal) de la loin-loin-lointaine sphère.

⊡ **Le cas général**

Un espace projectif de dimension n peut se représenter comme un espace affine de dimension n, muni de points à l'infini correspondant aux directions des droites. C'est à dire que les points à l'infini sont un espace projectif de dimension $(n-1)$.

Il est très important de noter que, comme le laisse comprendre la méthode imagée de voir tout ça, n'importe quel hyperplan peut être considéré comme l'hyperplan à l'infini.

33.2.3 Liste de résultats de géométrie projective

Pour prouver les résultats qui suivent il suffit généralement de traduire la question en termes d'espaces vectoriels.

Proposition 1274 • *Soit H un hyperplan projectif de E espace projectif. Soit x un point de $E \setminus H$; alors toute droite projective passant par x intersecte H.*
• *Si F et G sont deux sous-espaces projectifs de E de dimensions f et g de somme $\geq n$, alors F et G ont une intersection non vide.*
• *Deux droites d'un plan projectif se coupent toujours.*

33.2.4 Topologie des espaces projectifs réels ou complexes

Théorème 1275 (Compacité des espaces projectifs) *Soit E un espace vectoriel normé, muni de la topologie liée à sa norme (toutes les normes étant de toute façon équivalentes en dimension finie) ; alors l'espace projectif $P(E)$ est compact pour la topologie quotient.*

Démonstration :
• On considère la sphère unité de E. Cette sphère contient exactement deux points de chaque droite vectorielle de E ; l'image de la sphère unité par la projection est donc exactement $P(E)$. $P(E)$ est donc l'image d'un compact par une application continue, et donc est compact sous réserve que $P(E)$ soit séparé (voir théorème 171). Il suffit donc pour conclure de vérifier que $P(E)$ est séparé.

• On se donne deux droites vectorielles D et D' distinctes de E ; on considère deux couples de points x, y et x',y' de E, avec $\{x, y\} = D \cap S$ et $\{x', y'\} = D' \cap S$ (S désigne la sphère unité).

• Il existe quatre ouverts X, X', Y et Y' de S disjoints, avec $x \in X$, $y \in Y$, $x' \in X'$ et $y' \in Y'$.

• On considère alors l'intersection des projections de X et Y d'une part, et l'intersection des projections de X' et Y' d'autre part (rappelons que la projection canonique sur un ensemble quotient est ouverte lorsque les classes d'équivalence sont les orbites pour une action d'un groupe sur un espace topologique par homéomorphismes, voir proposition144 ; on obtient ainsi deux ouverts distincts séparant nos deux droites.$\square$

Théorème 1276 (Connexité par arcs des espaces projectifs de dimension ≥ 1)
Un espace projectif de dimension ≥ 1 est connexe par arcs.

Démonstration : Rappelons juste que l'image d'un espace connexe par arcs par une application continue est connexe par arcs...$\square$

On peut noter que le résultat est vrai aussi pour l'espace projectif de dimension 0 (réduit à un singleton) bien que $\mathbb{R} \setminus \{0\}$ ne soit pas connexe.

Chapitre 34

Zoologie de la géométrie

Cette partie concerne tant la géométrie affineque la géométrie projective. En fait souvent la géométrie projective et la géométrie affine sont liées ; et certains théorèmes de géométrie affine seront prouvés par la géométrie projective (et réciproquement).

34.1 La dualité

⊡ Quelques rappels, pour ceux qui ont oublié l'algèbre linéaire

On considère un espace vectoriel E de dimension finie. On note, classiquement, E^* son dual (i.e. l'espace vectoriel des formes linéaires continues sur E, c'est-à-dire, puisque E est de dimension finie, l'espace vectoriel des formes linéaires sur E).

Etant donné F un sous-espace vectoriel de E, on note F' l'ensemble des formes linéaires nulles sur tout F. Il est clair que F' un sous-espace vectoriel de E^* (voir la partie sur l'algèbre linéaire). On a $dim\ F + dim\ F' = dim\ E$, et $dim\ E^* = dim\ E$.

Il est fondamental de rappeler que $F \subset G$ dans E équivaut à $G' \subset F'$ dans E^*.

⊡ Le rapport avec la géométrie projective

Notons maintenant bien ce qu'il se passe, lorsque l'on identifie E à E^* :

On suppose que E est de dimension n, et on note p la surjection canonique de E dans $P(E)$.

L'important dans le tableau ci-dessous est la seconde et la quatrième colonne ; les autres colonnes sont là pour la compréhension.

dans E $dim\ E = n$	dans $P(E)$ $dim\ P(E) = n-1$	dans E^* $dim\ E^* = n$	dans $P(E^*)$ $dim\ P(E^*) = n-1$
F $dim\ F = f$	$p(F)$ $dim\ p(F) = f-1$	F' $dim\ F' = n-f$	$p(F')$ $dim\ p(F') = n-f-1$
droite	point	hyperplan (vectoriel)	hyperplan (projectif)
hyperplan (vectoriel)	hyperplan (projectif)	droite	point

On peut aller plus loin, en traduisant cette fois-ci les relations d'inclusion et d'intersection. Je note par la suite par des lettres les éléments de $P(E)$, et par les mêmes lettres munies d'un $'$ leurs images dans $P(E^*)$.

dans $P(E)$	dans $P(E^*)$
$A \subset B$	$B' \subset A'$
$A \in B$	$B' \in A'$

On peut ensuite spécialiser à la dimension 2 : les lettres majuscules désignent des droites, les lettres minuscules des points ; la même lettre désigne un élément et son dual ; on ne fait que passer de minuscule à majuscule, et réciproquement.

dans $P(E)$	dans $P(E^*)$
a (un point)	A une droite
$a = B \cap C$	$A = (bc)$
A, B, C concourantes	a, b, c alignés
Triangle de côtés A, B et C	Triangle de sommets a, b et c

Pourquoi utiliser cela en géométrie projective au lieu de l'utiliser simplement en géométrie affine ?

$\rightarrow$ par deux points passe toujours une droite... il est donc souhaitable que deux droites se croisent toujours en un point pour avoir une analogie complète !

On verra une application fort sympathique avec le théorème de Desargues 1279.

34.2 Géométrie dans le plan et dans le plan projectif

34.2.1 Théorème de Pappus

En géométrie affine

Théorème 1277 (Pappus) *Soient D et D' deux droites du plan, distinctes.*
Soient A, B et C trois points de D, et A', B' et C' trois points de D'.
Si AB' est parallèle à BA',
et si CB' est parallèle à BC',
alors AC' est parallèle à CA',

Démonstration :

• Supposons que D et D' ne soient PAS parallèles. Alors on considère O leur point d'intersection.

• On oriente les deux droites D et D', et on constate que $\frac{\overline{OB}}{\overline{OA}} = \frac{\overline{OA'}}{\overline{OB'}}$ et que $\frac{\overline{OC}}{\overline{OB}} = \frac{\overline{OB'}}{\overline{OC'}}$ (par la réciproque du théorème de Thalès).

• On en déduit que $\frac{\overline{OA}}{\overline{OC}} = \frac{\overline{OC'}}{\overline{OA'}}$, ce qui par Thalès nous donne bien le résultat souhaité.

NB : il est possible de raisonner sur des homothéties pour se débarasser des cas particuliers OC=0, etc. Il suffit de remplacer «$\frac{\overline{OB}}{\overline{OA}} = \frac{\overline{OA'}}{\overline{OB'}}$» par «il existe une homothétie h centrée sur O telle que $h(A) = B$ et $h(B') = A'$».

• Le cas D et D' parallèles n'est pas plus difficile, on utilise des translations pour exprimer les parallélismes...$\square$

⊡ **En géométrie projective**

34.2.2 Théorème de Desargues

⊡ **En géométrie affine**

Théorème 1278 (Théorème de Desargues) *On se donne ABC et $A'B'C'$ deux triangles sans sommet commun et à côtés respectivement parallèles* [a]. *Alors AA', BB' et CC' sont concourantes ou parallèles.*

[a] Ces deux triangles sont translatés l'un de l'autre ou homothétiques car directement semblables.

Démonstration :

• Premier cas : AA' et BB' ne sont pas parallèles.

Alors elles se coupent en un point, disons O.

$$\frac{\overline{OA'}}{\overline{OA}} = \frac{\overline{OB'}}{\overline{OB}}$$

On considère alors le point M de OC tel que

$$\frac{\overline{OA'}}{\overline{OA}} = \frac{\overline{OM}}{\overline{OC}}$$

Par la réciproque du théorème de Thalès, AC et $A'M$ sont parallèles, et BC et $B'M$ sont parallèles.

M est donc à la fois sur $B'C'$ et sur $A'C'$, et donc $M = C'$. On en déduit le résultat demandé.

NB : là aussi on peut raisonner via des homothéties.

• Second cas : AA' et BB' sont parallèles.

Alors $\overrightarrow{AA'} = \overrightarrow{BB'}$, et on construit M tel que $\overrightarrow{CM} = \overrightarrow{AA'}$; on a alors magiquement $\overrightarrow{CM} = \overrightarrow{BB'}$, et donc $A'M$ est parallèle à AC et $B'M$ est parallèle à BC ; donc $M = C'$.□

Il y a une réciproque au théorème de Desargues ; voir paragraphe suivant...

⊡ **En géométrie projective**

Théorème 1279 (Théorème de Desargues) *On se donne ABC et $A'B'C'$ deux triangles d'un plan projectif. On note a, b et c les points d'intersection respectifs de BC et $B'C'$, AC et $A'C'$, AB et $A'B'$. Alors a, b et c sont alignés si et seulement si AA', BB' et CC' sont concourantes.*

NB : Notez bien qu'il s'agit d'une généralisation du théorème précédent ! Deux droites parallèles en géométrie affine se croisent à l'infini dans leur complété projectif.

Démonstration :

• Supposons tout d'abord a, b et c alignés. Ils déterminent donc une droite. On peut supposer que cette droite est la droite à l'infini, puisque, comme on l'a déjà dit, n'importe quelle droite (ou hyperplan dans le cas général) peut être considérée comme droite (ou, donc, hyperplan dans le cas général) à l'infini.

On peut alors appliquer le théorème 1278.

Il nous reste maintenant à montrer la réciproque.

Pour cela on va utiliser la dualité... En fait, il n'y a tout simplement rien à faire, car l'énoncé dual du théorème de Desargues, dans le sens où on l'a montré est précisément sa réciproque.□

34.3 $O_2(\mathbb{R})$, $O_3(\mathbb{R})$, les polygones réguliers, les polyèdres réguliers

34.3.1 Dimension 2

⊡ Définition de la notion d'angle

Définition 1280 (Angle) *L'***angle orienté** *entre deux vecteurs unitaires u et v de $\mathbb{R}^2$ (pris dans cet ordre) est par définition l'unique rotation de $\mathbb{R}^2$ dont l'image de u est v.*
*L'***angle orienté** *entre deux vecteurs non nuls quelconques u et v de $\mathbb{R}^2$ (pris dans cet ordre) est par définition l'angle orienté entre $\frac{1}{\|u\|}u$ et $\frac{1}{\|v\|}v$.*
On appelle **angle nul** *l'angle entre u et u pour u vecteur non nul quelconque.*
On appelle **angle plat** *l'angle entre u et $-u$ pour u vecteur non nul quelconque.*
On appelle **angle orienté de deux demi-droites** *$\mathbb{R}^+u$ et $\mathbb{R}^+v$ l'angle orienté entre u et v.*
Pour tous ces angles, l'angle non orienté correspondant est la paire $\{r, r^{-1}\}$ avec r l'angle orienté correspondant.
L'angle orienté de deux droites $\mathbb{R}u$ et $\mathbb{R}v$ est la paire des angles entre $\mathbb{R}^+u$ et $\mathbb{R}^+v$ et entre $\mathbb{R}^+u$ et $\mathbb{R}^-v$.
L'angle non-orienté correspondant est l'ensemble à 4 éléments constitué des angles entre $\mathbb{R}^+u$ et $\mathbb{R}^+v$, entre $\mathbb{R}^+u$ et $\mathbb{R}^-v$, et leurs inverses.
Etant donnée une base orthonormée directe de $\mathbb{R}^2$ et un angle orienté r, on appelle **mesure de cet angle** *l'unique $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ tel que la matrice de r dans cette base soit*

$$\begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix}$$

Notons que la valeur de θ est indépendante du choix de la base orthonormée directe.
On appelle **mesure principale d'un angle** *la mesure de cet angle comprise dans $]-\pi, \pi]$. On notera $\widehat{X, Y}$ l'angle orienté entre X et Y, quelle que soit la nature de X et Y.*

Proposition 1281

$$\widehat{a,b} - \widehat{c,d} = \widehat{a,c} - \widehat{b,d}$$
$$\widehat{a,b} + \widehat{b,c} = \widehat{a,c}$$

⊡ Sous-groupes finis de $O_2(\mathbb{R})$

Proposition 1282 *Tout sous-groupe fini de $O_2^+(\mathbb{R})$ s'identifie à un sous-groupe fini de $(\mathbb{U}, \times)$* [a].
Un tel sous-groupe est donc de la forme $\{e^{\frac{2ik\pi}{n}} / k \in [0, n-1]\}$, et est isomorphe à $\mathbb{Z}/n\mathbb{Z}$, avec n l'ordre du groupe. Il est ainsi monogène et cyclique.
Tout sous-groupe fini de $O_2(\mathbb{R})$ s'identifie à $\mathbb{Z}/n\mathbb{Z}$ ou à $D_n \simeq \mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$, groupe diédral d'ordre $2n$.

[a] Ensemble des nombres complexes de module 1

Démonstration :

Soit G un sous-groupe fini de $O_2^+(\mathbb{R})$. Il est engendré par un nombre fini de rotations (puisqu'il n'y a que des rotations dans $O_2^+(\mathbb{R})$. Donc il est engendré par des rotations d'ordre fini, d'angle $e^{\frac{2\pi}{n}}$. En posant m le $ppcm$ des n en question, on constate que G est engendré par la rotation d'angle $e^{\frac{2\pi}{n}}$.□

Considérons maintenant G un sous-groupe fini de $O_2(\mathbb{R})$ et montrons qu'il est de l'une des deux formes annoncées. S'il est inclus dans $O_2^+(\mathbb{R})$ c'est terminé. En cas contraire, soit s appartenant à $G \cap O_2^-(\mathbb{R})$. s est une symétrie. L'intersection H de G et de $O_2^+(\mathbb{R})$ est un sous-groupe fini de $O_2^+(\mathbb{R})$; il est donc engendré par la rotation r d'angle $2\pi/n$ pour un certain n. Le groupe G est alors engendré par r et s ; $G = H \cup sH = \{Id, s\} \times H$ (au sens du produit ensembliste et non du produit direct). H est distingué, $\{Id, s\} \cap H = \{Id\}$: donc $G = \mathbb{Z}/2\mathbb{Z} \rtimes \mathbb{Z}/n\mathbb{Z} = D_n$.□

⊡ Polygones réguliers

Définition 1283 *On appelle polygone régulier de $\mathbb{R}^2$ l'orbite d'un vecteur non nul sous l'action d'un sous-groupe fini de $O_2^+(\mathbb{R})$.*

On consultera la partie 21.10.12 pour constater que D_n est le groupe des isométries d'un polygone régulier.

On pourra consulter le livre [10] pour la constructibilité des polygones réguliers ; la caractérisation des n tels que le polygone régulier à n côtés est très surprenante, faisant intervenir les nombres de Fermat : un tel polygone est constructible si n s'exprime comme le produit d'une puissance de 2 par un produit de nombres premiers de Fermat distincts[1].

[1] Un nombre premier de Fermat est un nombre premier de la forme $2^{2^n} + 1$. Les seuls nombres premiers

34.3.2 Dimension 3 : Sous-groupes finis de $O_3^+(\mathbb{R})$

Théorème 1284 (Sous-groupes finis de $O_3^+(\mathbb{R})$) *Les sous-groupes finis de $O_3^+(\mathbb{R})$ sont de l'une des 5 formes suivantes :*
• Groupes constitués exclusivement de rotations, isomorphes à $\mathbb{Z}/n\mathbb{Z}$
• Groupes isomorphes aux groupes diédraux (ordre pair)
• Groupe des isométries positives laissant invariant un tétraèdre régulier (ordre 12)
• Groupe des isométries positives du cube = groupe des isométries positives de l'octaèdre (ordre 24)
• Groupe des isométries positives de l'icosaèdre (ordre 60)
Il existe des exemples pour chacun de ces cas.

Démonstration : Donner une preuve complète est très laborieux, et être vraiment rigoureux est très difficile... On pourra voir le livre [10].□

⊡ **Polyèdres réguliers**

On n'étudiera pas ici l'existence des 5 polyèdres réguliers et les relations de dualité. Le lecteur intéressé pourra se référer à [8] ; la preuve qui y est donnée n'est pas complètement rigoureuse, car le formalisme nécessaire pour manier de tels objets est difficile, et il est sans doute préférable de se limiter à des notions intuitives.

Les 5 polyèdres réguliers sont le tétraèdre (4 faces triangulaires), le cube (6 faces carrées), l'octaèdre (8 faces triangulaires), l'icosaèdre (20 faces triangulaires), le dodécaèdre (12 faces pentagonales).

34.4 La droite de Simson et quelques suites

Ce théorème, à quelques modifications mineures près, est extrait du livre de Hahn, "Complex numbers and geometry".

Théorème 1285 *Soit A, B, C un triangle, D un point.*
Soient P, Q et R les projections orthogonales de D sur (BC), (CA) et (AB).
Alors P,Q et R sont colinéaires $\iff$ D appartient au cercle circonscrit à ABC.

Démonstration : On considère le triangle ABC inscrit dans un cercle de centre 0 et de rayon 1 (sans perte de généralité).
On note α, β, γ, δ, λ, μ et ν les abscisses respectives de A,B,C,D,P,Q,R.

Une équation de (BC) est donnée par le calcul suivant :

$$z \in (BC) \iff \frac{z-\beta}{\gamma-\beta} \in \mathbb{R}$$

de Fermat connus sont 3, 5, 17, 257, 65537.

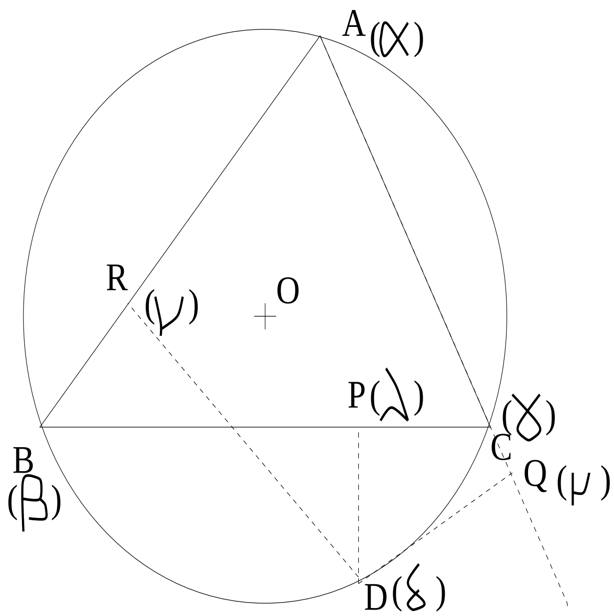


FIG. 34.1 – Figure préparatoire à l'étude de la droite de Simson

$$\frac{z-\beta}{\gamma-\beta} = \frac{\overline{z}-\overline{\beta}}{\overline{\gamma}-\overline{\beta}}$$

c'est-à-dire[2] $z+\gamma.\beta.\overline{z} = \beta+\overline{\gamma}$

Une équation de (PD) est :

$$z \in (BC) \iff \frac{z-\gamma}{\gamma-\beta} \in i\mathbb{R}$$

$$\frac{z-\delta}{\gamma-\beta} = -\frac{\overline{z}-\overline{\delta}}{\overline{\gamma}-\overline{\beta}}$$

Donc l'intersection de ces deux droites doit vérifier :

$$z-\gamma.\beta.\overline{z} = \delta-\gamma.\beta.\overline{\delta}$$

On en déduit que

$$\lambda = \frac{1}{2}(\beta+\gamma+\delta-\beta.\gamma.\overline{\delta})$$

de même :

$$\mu = \frac{1}{2}(\alpha+\gamma+\delta-\alpha.\gamma.\overline{\delta})$$

$$\nu = \frac{1}{2}(\alpha+\beta+\delta-\alpha.\beta.\overline{\delta})$$

[2]En constatant que $\frac{\alpha-\beta}{\overline{\alpha}-\overline{\beta}} = -\beta\gamma$ et $\frac{\beta\overline{\gamma}-\overline{\beta}\gamma}{\overline{\gamma}-\beta} = \beta+\gamma$.

P, Q, R colinéaires $\iff \frac{\lambda-\nu}{\mu-\nu} \in \mathbb{R}$

$$\iff [\gamma, \frac{\delta}{|\delta|^2}, \beta, \alpha] \in \mathbb{R}$$

$\iff \alpha, \beta, \gamma, \frac{\delta}{|\delta|^2}$ cocycliques ou alignés

$$\iff |\delta| = 1$$

Ce qui est précisément le résultat souhaité.□

On suppose maintenant $|\delta| = 1$.

λ vérifie $z = \frac{1}{2}(\beta + \gamma + \delta - \frac{\beta.\gamma}{\delta})$

Soit $\sigma_1 = \alpha + \beta + \gamma$
$\sigma_2 = \alpha.\beta + \alpha.\gamma + \beta.\gamma$
$\sigma_3 = \alpha.\beta.\gamma$

Il s'agit des trois polynômes symétriques élémentaires en α, β, γ. Notons que $\overline{\sigma}_1 = \overline{\alpha} + \overline{\beta} + \overline{\gamma} = \frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = \frac{\sigma_2}{\sigma_3}$.

$z = \frac{1}{2}(\sigma_1 + \delta - \alpha - \frac{\sigma_3}{\delta.\alpha})$ et $\overline{z} = \frac{1}{2}(\sigma_2/\sigma_3 + 1/\delta - 1/\alpha - \frac{\delta.\alpha}{\sigma_3})$

En éliminant α :

$\delta.z - \sigma_3.\overline{z} = \frac{1}{2}(\delta\sigma_1 + \delta^2 - \sigma_2 - \sigma_3/\delta)$

Cette expression, étant symétrique en α, β, γ, est vérifiée aussi pour μ et ν. C'est l'équation d'un sous espace affine, qui est :

- non vide
- non réduit à un point car P, Q, R ne sont jamais confondus.
- non égal au plan car σ_3 est non nul

Définition 1286 *Cette équation est donc l'équation d'une droite, appelée* **droite de Simson**.

Théorème 1287 *Soient L, M, N trois points sur le cercle circonscrit à ABC. Alors les trois droites de Simson correspondantes sont concourantes si et seulement si $\widehat{AL} + \widehat{BM} + \widehat{CN}$ est nul modulo 2π.*

Démonstration :

Soit l, m, n les abscisses respectives de L, M, N.

$$l.z - \sigma_3.\overline{z} = \frac{1}{2}(l^2 + \sigma_1.l - \sigma_3 - \sigma_3/l)$$

$$m.z - \sigma_3.\overline{z} = \frac{1}{2}(m^2 + \sigma_1.m - \sigma_3 - \sigma_3/m)$$

$$n.z - \sigma_3.\overline{z} = \frac{1}{2}(n^2 + \sigma_1.n - \sigma_3 - \sigma_3/n)$$

Les deux premières se coupent en $z = \frac{1}{2}(l + m + \sigma_1 + \sigma_3/(l.m))$.
Les deux dernières se coupent en $z = \frac{1}{2}(m + n + \sigma_1 + \sigma_3/(m.n))$.

La condition nécéssaire et suffisante est donc $l - n + \sigma_3.(n - l)/(l.m.n) = 0$, c'est-à-dire $\sigma_3 = m.l.n$ ou encore $\alpha.\beta.\gamma = l.m.n$
D'où le résultat par passage aux arguments.□

34.5 Le cercle des neuf points, et une suite par Coolidge

Cette partie est, elle aussi, fortement inspirée du livre de Hahn "Complex Numbers and Geometry".

Le cercle des neuf points

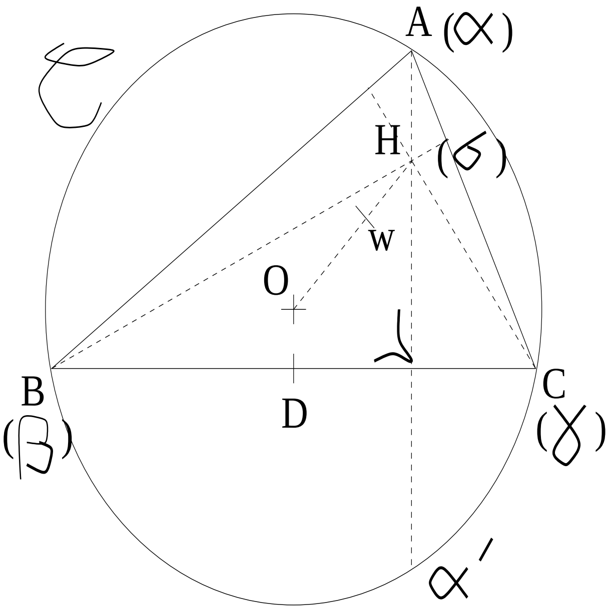


FIG. 34.2 – Schéma préparatoire au cercle des neuf points

Théorème 1288 *Soit (ABC) un triangle, O le centre du cercle circonscrit, G le barycentre de A, B et C, et H l'intersection des hauteurs.*
Alors :
• (O, G, H) *sont alignés.*
• *Les pieds des hauteurs, les milieux des côtés et les milieux des segments joignant l'orthocentre aux sommets A, B et C sont tous sur un même cercle.*

Démonstration : On choisit O pour origine du plan complexe.
On note α, β, γ les affixes respectives de A, B, C. On note $\mathcal{C}$ le cercle circonscrit.

Soit H le point d'affixe $\theta = \alpha + \beta + \gamma$. Comme $\frac{\sigma-\alpha}{2} = \frac{\beta+\gamma}{2}$, on a $\overrightarrow{AH} = 2\overrightarrow{OD}$, où D est le milieu de BC. En particulier, comme OA' est la médiatrice de $[BC]$, on a (AH) et (BC) orthogonales.

On peut raisonner de même pour obtenir $BH \perp AC$ et $CH \perp AB$; on retrouve ainsi le fait que les hauteurs sont concourantes en H. Comme G a pour affixe $\frac{1}{3}(\alpha + \beta + \gamma)$, les points O, G et H sont alignés.

Soit W d'affixe $\sigma/2$; W est le milieu de $[OH]$.
$WD = |\frac{\beta+\gamma}{2} - \frac{\sigma}{2}| = |\alpha|/2 = \frac{1}{2}$
De même $WE = WF = \frac{1}{2}$ avec E et F les milieux respectifs de $[CA]$ et $[AB]$.
La distance de W au milieu de AH est elle aussi $\frac{1}{2}$, et de même pour les autres.
On note Λ l'affixe du pied de la hauteur issue de A.
Pour situer Λ on va d'abord s'intéresser à α', l'intersection de $\alpha\Lambda$ avec $\mathcal{C}$. On va montrer que $B\alpha'\sigma$ est isocèle.

$\overrightarrow{A\alpha'} \perp \overrightarrow{BC}$ donne $\frac{\alpha'-\alpha}{\gamma-\beta} \in i.\mathbb{R}$.
Lemme : On a $\alpha' = -\frac{\beta\gamma}{\alpha}$.

Démonstration : Comme le quotient $\frac{\beta\gamma}{\alpha\alpha'}$ est invariant par rotation, on peut supposer que $\overrightarrow{A\alpha'}$ est vertical et $\overrightarrow{BC}$ horizontal, de sorte que $\alpha\alpha' = 1$ et $\beta\gamma = -1$, si bien que ce quotient vaut -1 et $\alpha' = -\frac{\beta\gamma}{\alpha}$ dans tous les cas. □

On obtient donc par le lemme ci-dessus $\alpha' = -\frac{\beta\gamma}{\alpha}$. On montre ensuite facilement que $|B - \alpha'| = |B - \sigma|$, donc le triangle $B\alpha'\sigma$ est isocèle. Le pied λ de la hauteur issue de A est en fait aussi le milieu de $[\alpha'\sigma]$.

On en déduit alors que $|\Lambda - \sigma/2| = 1/2$. On ferait de même pour les pieds des autres hauteurs. Finalement $\sigma/2$ est le centre d'un cercle de rayon $1/2$ comportant les pieds des trois hauteurs, les milieux des trois côtés, et les milieux des segments joignants l'orthocentre aux sommets A,B et C sont tous sur un même cercle.□

Enfin, voici un résultat dû à Coolidge :

• Soit z_1,...,z_4 sur le cercle unité. Le centre du cercle d'Euler de z_2, z_3, z_4 est $\tau_1 = \frac{1}{2}(z_2 + z_3 + z_4)$; τ_2, τ_3, τ_4 sont définis de manière analogue. On définit en outre $\tau = \frac{1}{4}(z_1 + z_2 + z_3 + z_4)$; alors pour tout i, τ_i est sur le cercle de centre τ et de rayon $\frac{1}{2}$; ce cercle est appelé cercle d'Euler du quadrilatère z_1,z_2,z_3,z_4.

• On peut faire la même chose avec 5 points ; on définit 5 cercles d'Euler de 5 quadrilatères, et un cercle d'Euler pour le pentagone.
• On peut procéder de même avec un polygone quelconque.

Pour plus de précisions, on pourra consulter le livre de Hahn cité ci-dessus.

Chapitre 35

Combinatoire et dénombrements

On consultera avec grand profit le chapitre 1 de [10], très bien fait et fournissant une bonne quantité de résultats originaux, complétant utilement les résultats très classiques suivants. On pourra aussi aller consulter le paragraphe4.2 pour les cardinaux d'ensembles infinis.

35.1 Cardinaux d'ensembles finis

Proposition 1289 *Avec A un ensemble fini, on a la* **Formule d'inclusion exclusion** *; avec $F_i \in \mathcal{A} = P(A)$, on a* $|\cup_{i \leq n} F_i| = \sum_{i \leq n} |F_i| - \sum_{1 \leq i < j \leq n} |F_i \cap F_j| + \sum_{1 \leq i < j < k \leq n} |F_i \cap F_j \cap F_k| ... + (-1)^{n-1} | \cap_{1 \leq i \leq n} F_i|$

Le nombre de parties à p éléments d'un ensemble à n éléments est C_n^p ; voir à 35.4.

35.2 Dénombrement de fonctions

On considère E et F deux ensembles finis, de cardinaux e et f.

35.2.1 Ensemble des applications de E dans F

Proposition 1290 *L'ensemble des applications de E dans F, noté F^E, a pour cardinal f^e.*

35.2.2 Ensemble des injections de E dans F

Proposition 1291 *L'ensemble des injections de E dans F a pour cardinal $A_f^e = \frac{f!}{(f-e)!}$ si $f \geq e$; 0 sinon.*

La preuve se fait facilement par récurrence. Voir 35.3 pour les premières valeurs.

35.2.3 Ensemble des surjections (et bijections) de E dans F

Proposition 1292 *En notant S_e^f le cardinal de l'ensemble des surjections de E dans F (dans le cas $e \geq f$) divisé par $f!$ (c'est-à-dire, dans le cas de E et F totalement ordonnés, le nombre de surjections croissantes), on a les formules :*

$$S_e^1 = S_e^e = 1$$
$$\forall (e,f)\ S_{e+1}^f = S_e^{f-1} + f.S_e^f$$

On obtient ainsi les valeurs suivantes de S_e^f :

	$f=1$	$f=2$	$f=3$	$f=4$	$f=5$
$e=1$	1	0	0	0	0
$e=2$	1	1	0	0	0
$e=3$	1	3	1	0	0
$e=4$	1	7	6	1	0
$e=5$	1	15	25	10	1

Le nombre de bijections de E vers F vaut $e!$ si $e = f$, 0 sinon.

35.2.4 Ensemble des applications croissantes de E vers F

E et F sont maintenant munis d'un ordre total.

L'ensemble des applications croissantes de E dans F a même cardinal que l'ensemble des applications strictement croissantes de E dans $[1, f+e-1]$; une fois que l'on s'est convaincu de cela (en voyant que 1] si u de $\{1, 2, 3, ..., e\}$ dans $\{1, 2, 3, ..., f\}$ est croissante alors $x \mapsto v_u(x) = u(x) + x$ de $\{1, 2, 3, ..., e\}$ dans $\{1, 2, 3, ..., e+f\}$ est strictement croissante 2] $u \mapsto v_u$ est bijective de l'ensemble des applications croissantes de E dans F vers l'ensemble des applications strictement croissantes de E dans $[1, f+e-1]$) il est facile de montrer que cet ensemble a pour cardinal C_{f+e-1}^e (les coefficients binomiaux C_n^p sont présentés plus bas).

35.3 Arrangements

Définition 1293 *On appelle p-**arrangement de** E une application injective de $\mathbb{N}_p$ dans E.*
On note A_n^p le cardinal de l'ensemble des p-arrangements d'un ensemble à n éléments.

Au vu des résultats précédents, on peut dire que $A_n^p = \frac{n!}{(n-p)!}$.
Les premières valeurs sont :

	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$n=1$	1	0	0	0	0
$n=2$	2	2	0	0	0
$n=3$	3	6	6	0	0
$n=4$	4	12	24	24	0
$n=5$	5	20	60	120	120

35.4 Combinaisons

Définition 1294 *On appelle p-**combinaison de** E tout sous-ensemble de E de cardinal p. On note C_n^p ou $\begin{pmatrix} n \\ p \end{pmatrix}$ le cardinal de l'ensemble des p-combinaisons de E.*

On montre facilement, par récurrence, que $C_n^p = \frac{n!}{p!(n-p)!}$.
Cette formule peut aussi se déduire sans récurrence en voyant que $p!$ p-arrangements donnent la même p-combinaison donc $C_n^p = \frac{1}{p!}A_n^p = \frac{n!}{p!(n-p)!}$.
Elle peut aussi se déduire du fait que le groupe $\sigma(E)$ des permutations de $E = \{1, 2, ..., e\}$ agit transitivement sur l'ensemble des p-combinaisons de E, que le stabilisateur S de $F = \{1, 2, ..., f\} \subset E$ est le produit $A \times B$ avec A et B respectivement les groupes de permutations de $\{1, 2, ..., f\}$ et $\{f+1, f+2, ..., e\}$; S a donc pour cardinal $S = f!(e-f)!$, d'où $C_e^f = \frac{\sigma(E)}{|S|} = \frac{e!}{f!(e-f)!}$. Un argument similaire permet d'ailleurs de montrer la formule $A_n^p = \frac{n!}{(n-p)!}$ sans récurrence.

En outre on a les formules suivantes :

Proposition 1295

$$C_n^p = C_n^{n-p}$$

$$C_{n+1}^p = C_n^p + C_n^{p+1}$$

Formule de Newton, valable dans un anneau : si $a.b = b.a$ et $n > 0$, alors :

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k.b^{n-k}$$

$$\sum_{k=0}^{n} C_n^k = 2^n$$

$$\sum_{k=0}^{n} (-1)^k C_n^k = 0$$

(ces deux formules sont obtenues en spécialisant la formule de Newton)

$$1 \leq p \leq n \to p.C_n^p = n.C_{n-1}^{p-1}$$

$$\sum_{k=0}^{n} k.C_n^k = n.2^{n-1}$$

$$\sum_{k=0}^{n} (-1)^k C_n^k = 1$$

$$\sum_{k=0}^{n} (C_n^k)^2 = C_{2n}^n$$

	$p=0$	$p=1$	$p=2$	$p=3$	$p=4$
$n=0$	1	0	0	0	0
$n=1$	1	1	0	0	0
$n=2$	1	2	1	0	0
$n=3$	1	3	3	1	0
$n=4$	1	4	6	4	1

35.5 Quelques applications

FLEMMARD dim espaces polynomes homogenes, anagrammes, tirages du loto, dates de naissance

35.5.1 Une formule utile

Proposition 1296 *Soit n un entier naturel. On a alors*

$$\sum_{k=0}^{n} C_n^k (-1)^{n-k} k^p = \begin{cases} 0 & \text{si} \quad p < n \\ n! & \text{si} \quad p = n \end{cases} .$$

Ce résultat sera utile pour la proposition 585.

Démonstration : Comme $C_n^k = C_n^{n-k}$, la somme présentée est le coefficient c_n du produit de Cauchy suivant :

$$f_p(x) = \sum_{m \geq 0} c_m x^m = \underbrace{\left(\sum_{k=0}^{n} C_n^k (-1)^k x^k \right)}_{(1-x)^n} \underbrace{\left(\sum_{l \geq 0} l^p x^l \right)}_{g_p(x)},$$

où toutes les séries entières présentées ont un rayon de convergence d'au moins 1. La famille de fonctions $(g_p)_{p \geq 0}$ est également définie par récurrence comme suit :

$$g_0(x) = \frac{1}{1-x} \text{ et } g_{p+1}(x) = x g_p'(x).$$

Lemme 1297 *Pour tout entier naturel p, il existe un polynôme h_p de degré p tel que $g_p(x) = h_p(x)/(1-x)^{p+1}$.*

Démonstration : Récurrence évidente.□

Si $p < n$, alors $f_p(x) = h_p(x)(1-x)^{n-p-1}$ est un polynôme de degré $(n-1)$, si bien que $c_n = 0$.

Supposons maintenant que $p = n$. On a alors :

$$f_n(x) = \frac{h_n(x)}{1-x} = h_n(x)(1 + x + x^2 + \ldots).$$

Comme h_n est de degré n, le coefficient c_n vaut la somme des coefficients de h_n. Autrement dit, $c_n = h_n(1)$. On déduit de la relation $g_{p+1}(x) = x g_p'(x)$ que

$$\frac{h_{p+1}(x)}{(1-x)^{p+2}} = \frac{x h_p'(x)}{(1-x)^{p+1}} + \frac{(p+1) x h_p(x)}{(1-x)^{p+2}}.$$

En multipliant par $(1-x)^{p+2}$ puis en prenant $x = 1$, il vient $h_{p+1}(1) = (p+1) h_p(1)$. Comme $h_0(1) = 1$, on obtient $h_p(1) = p!$ pour tout p. En particulier, $c_n = h_n(1) = n!$ et la proposition est démontrée.□

35.5.2 Généralisation du binôme de Newton

Proposition 1298 (Généralisation de la formule de Newton) *Dans un anneau commutatif, pour n non nul,*

$$(x_1 + ... + x_p) = \sum_{i_1+i_2+...+i_p=n, i_j \geq 0} \left(\frac{n!}{i_1!i_2!...i_p!} \right) x_1^{i_1} x_2^{i_2} ... x_p^{i_p}$$

Le coefficient $\frac{n!}{i_1!i_2!...i_p!}$ pouvant se noter C_n^i avec $i = (i_1, ..., i_p)$.

35.5.3 Familles sommables infinies

Définition 1299 *Une famille $(x_i)_{i\in I}$ de nombres complexes est* **sommable de somme** *x si pour tout ϵ il existe $J \subset I$ finie telle que, pour tout K fini, $J \subset K \subset I$ implique $|x - \sum_{i\in K x_i}| \leq \epsilon$.*

Lemme 1300 *Si une famille (x_i) de nombres réels est sommable, alors la famille de ses termes positifs est sommable, et la famille de ses termes qui sont négatifs est sommable.*

Démonstration : Supposons que la famille des (x_i) soit sommable, et supposons que la famille des $(max(x_i, 0))$ ne le soit pas. Alors la somme des $(max(x_i, 0))$ pour $i \in J$ avec J fini peut être arbitrairement grande. Or la somme des (x_i) pour i dans $J' = \{j \in J / x_j > 0\}$ est tout aussi grande, et peut donc être arbitrairement grande elle aussi. D'où le résultat pour la famille des réels positifs. Le raisonnement pour la famille négative est le même.

Proposition 1301 *Toute famille sommable de nombres réels est de support dénombrable (ie seule une quantité au plus dénombrable de ces réels est non nulle). Il en va de même des familles de nombres complexes.*

Démonstration :

En vertu du lemme précédent, je me contente de démontrer ce résultat pour une famille de nombres réels positifs. Le résultat dans le cas général se démontre de manière similaire.

Il existe un nombre fini de réels plus grands que $1/n$, pour tout n. En notant A_n la famille des réels $> 1/n$, on voit que la réunion des A_n est le support de la famille ; une réunion dénombrable d'ensembles finis étant dénombrable, la famille est dénombrable.□

Chapitre 36

Probabilités

36.1 Espaces mesurés

On trouvera en 5.1 les fondements de la topologie, et en 6.1 les fondements de la théorie de la mesure. Je ne rappelle ci-dessous que quelques définitions qui sont données dans les paragraphes que je viens de citer.

Une **topologie** sur X est un sous-ensemble de $P(X)$ contenant $\emptyset$, X, et stable par réunion quelconque et par intersection finie. Les éléments d'une topologie sont appelés **ouverts**, leurs complémentaires sont appelés **fermés**.
Une **algèbre** sur X est un sous-ensemble de $P(X)$ contenant X, stable par passage au complémentaire et stable par union finie.
Une **tribu** sur X est un sous-ensemble de $P(X)$ contenant X, stable par passage au complémentaire et par union dénombrable. Une tribu est aussi appelée σ**-algèbre** .
Un **espace mesurable** est un couple $(X, \mathcal{A})$ avec $\mathcal{A}$ tribu sur X.
Dans $\mathbb{R}$ ou dans $\mathbb{R}^n$, ou en général dans un espace topologique, la tribu usuelle est la tribu engendrée par les ouverts. Dans le cas de $\mathbb{R}^n$ et $\mathbb{R}$, cette tribu est aussi la tribu engendrée par les boules ouvertes. Une tribu engendrée par une topologie s'appelle **tribu des boréliens** ; ses éléments s'appellent les **boréliens**.
Une mesure positive sur un espace mesurable $(X, \mathcal{A})$ est une fonction μ de $\mathcal{A}$ dans $\mathbb{R}^+$ telle que
• $\mu(\emptyset) = 0$
• Si les A_i sont deux à deux disjoints et I dénombrable alors $\mu(\cup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$
Un **espace mesuré** est un triplet $(X, \mathcal{A}, \mu)$ avec $\mathcal{A}$ une tribu sur X, μ une mesure positive sur $(X, \mathcal{A})$.
Une fonction de $(X_1, \mathcal{A}_1)$ dans $(X_2, \mathcal{A}_2)$ est dite **mesurable** si et seulement si l'image réciproque de tout ensemble mesurable est mesurable.
La σ-algèbre engendrée par une base d'ouverts d'une topologie est égale à la σ-algèbre engendrée par cette topologie.
La partie 6.2 est indispensable pour s'attaquer aux probabilités.
Une mesure est dite **finie** si et seulement si la mesure de l'espace tout entier X est finie et alors $\forall A$ mesurable $\mu(A) \leq \mu(X)$.

Définition 1302 *Une mesure est une* **mesure de probabilité** *si la mesure de l'espace tout entier est* 1.

36.2 Evènements

36.2.1 Définitions de base

Définition 1303 (Définitions de base) *On appelle* **triplet de probabilité** *un triplet* $(\Omega, \mathcal{F}, P)$*, avec* P *une mesure de probabilité sur* $(\Omega, \mathcal{F})$.
Ω *est appelé* **l'univers**.
Un élément de Ω *est appelé* **possible**.
On appelle **évènement** *une partie mesurable de* Ω*, c'est à dire un élément de* $\mathcal{F}$*, c'est à dire une partie* $\mathcal{F}$*-mesurable.*

Proposition 1304 *Si chaque* $(F_n)_{n\in\mathbb{N}}$ *est de mesure* 1*, c'est à dire que chaque évènement* F_n *est réalisé presque sûrement*[a]*, alors* $\cap_{n\in\mathbb{N}}F_n$ *se réalise avec une probabilité* 1*. Si* E_n *est une suite d'évènements tels que* $\sum_i P(E_i) < +\infty$*, alors* $P(limsup\ E_n) = 0$*. Ce résultat est connu sous le nom de premier lemme de* **Borel-Cantelli**.

[a]Notez que **presque sûrement** est l'analogue, pour une mesure de probabilité, de **presque partout**, en théorie de la mesure.

On notera une nouvelle façon de voir $limsup$ des E_n, avec les E_n des évènements ; en l'écrivant $\cap_n \cup_{k\leq n} E_n$, on voit maintenant cette limite comme l'évènement qui arrive "infiniment souvent" ; c'est l'ensemble de ω qui appartiennent à une infinité de E_n. De même on peut voir différemment $liminf$ des E_n, avec les E_n des évènements ; en l'écrivant $\cup_n \cap k \leq nE_n$, on voit $liminf\ E_n$ comme l'ensemble des ω tels que ω est dans tout E_n pour n assez grand ($\geq N_\omega$ avec N_ω dépendant de ω).

On peut trouver ici des corollaires du lemme de Fatou ; notamment les deux propriétés suivantes :
• $P(liminf\ E_n) \leq liminf\ P(E_n)$
• $P(limsup\ E_n) \geq limsup\ P(E_n)$
La première propriété est vraie dans le cas d'un espace mesuré quelconque ; la seconde demande que la mesure soit finie, ce qui est donc le cas dans le cadre d'un espace de probabilité.

36.2.2 Quelques mesures de probabilité

⊡ **Ensemble fini**

Lorsque l'univers est fini, on peut prendre (et on prend usuellement) pour tribu l'ensembles de toutes les parties de l'univers. Par exemple, si l'on lance 3 fois une pièce, et que l'on peut obtenir pile ou face, l'univers est :
$\{PPP, PPF, PFP, PFF, FPP, FPF, FFP, FFF\}$
On peut dans ce cas prendre pour mesure l'application qui à un ensemble E associe égale à $\frac{card\ E}{card \Omega}$. L'univers contient des éléments correspondant à chaque manière dont peut se réaliseer le phénomène aléatoire étudié (dépend de la finesse de la description). La structure de Ω lui-même est secondaire, ce sont les variables aléatoires définies dessus qui importent (pourvu que Ω soit suffisamment vaste pour les définir suffisamment fines).

⊡ **Distribution sur** $[0, 1]^n$

On peut utiliser comme mesure sur $[0, 1]^n$ la restriction d'une mesure sur les boréliens à $[0, 1]^n$ telle que la mesure de l'espace soit 1 ; c'est en particulier le cas de la mesure de Lebesgue. On généralise facilement la méthode à une partie mesurable (de mesure > 0) de $\mathbb{R}^n$, en divisant par la mesure de la partie. Dans le cas d'une distribution uniforme, on choisit bien sûr la mesure de Lebesgue.

36.3 Variables aléatoires

36.3.1 Définitions : variable aléatoire, loi, fonction de répartition

Définition 1305 (Variable aléatoire) *Une* **variable aléatoire** *est simplement une fonction mesurable d'un univers vers* $\mathbb{R}$ *(muni de sa tribu borélienne pour la topologie usuelle).*

On peut voir alors de nombreux outils qui seront des variables aléatoires ; par définition il suffit que $f^{-1}(B)$ soit mesurable pour tout B borélien pour que f soit une variable aléatoire (il faut bien entendu que l'espace de départ soit un univers, donc que la mesure de l'espace soit 1). Toutes les opérations sur des fonctions à variables réelles qui conservent la mesurabilité sont alors possibles pour construire des variables aléatoires ; la somme, le produit, la valeur absolue... Ce qui est cohérent par rapport à la notion intuitive de variable aléatoire ; si un tirage aléatoire donne un réel et si l'on répète dix fois ce tirage aléatoire, alors la somme des résultats de ces dix tirages est une variable aléatoire, le produit aussi ; ils ont eux aussi leurs *distributions de probabilité* *(notion à définir plus tard).*

Définition 1306 *Supposons que X soit une variable aléatoire sur un triplet de probabilité $(\Omega, \mathcal{F}, P)$. X est alors une application de Ω dans $\mathbb{R}$, et P est une application de $\mathcal{F}$ dans $[0,1]$.Alors on définit la* **loi de probabilité** *$\mathcal{L}_X$ de X par $\mathcal{L}_X = P \circ (X^{-1})$ (X^{-1} n'est pas l'application réciproque - non nécessairement bien définie - mais l'application qui à une partie associe son image réciproque) ; $\mathcal{L}_X$ est ainsi définie sur l'ensemble des boréliens de $\mathbb{R}$.*

$\mathcal{L}_X(A)$ est la probabilité pour que X soit dans A.

Proposition 1307 *$\mathcal{L}_X$ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B})$.*

Démonstration : Pas dur !$\square$

Définition 1308 *Soit $F_X(t) = L_X(]-\infty, t])$; alors F_X est appelée* **fonction de répartition** *de X.*

Proposition 1309 *La donnée de F_X détermine $\mathcal{L}_X$ de manière unique.*

Démonstration : $\{]-\infty, t]/t \in \mathbb{R}\}$ est un Π-système qui engendre l'ensemble des boréliens. Donc par le lemme 336 $\mathcal{L}_X$ est entièrement défini par la fonction $F_X(t) = L_X(]-\infty, t]) = P(\{\omega \in \Omega / X(\omega) \leq t\})$.$\square$

Proposition 1310 (Quelques propriétés des distributions) *Soit F est une fonction de répartition d'une certaine variable aléatoire si et seulement si les quatre propriétés suivantes sont vérifiées :*
- *F est croissante de $\mathbb{R}$ dans $[0,1]$*
- *$F(x) \to 1$ quand $x \to +\infty$*
- *$F(x) \to 0$ quand $x \to -\infty$*
- *F est continue à droite en tout point.*

Démonstration : Le sens "seulement si" ne pose pas trop de problème. Les trois premiers points sont clairs, le quatrième utilise le fait que l'ensemble des ω inférieurs à $x+\frac{1}{n}$ tend vers l'ensemble des ω inférieurs à x, en décroissant (en effet la mesure d'une suite décroissante d'ensembles mesurables est égale à la mesure de l'intersection).
Le sens "si" est plus délicat ; la loi X correspondant à F est définie par $X(\omega) = inf\ \{z/F(z) \geq \omega\}$. Il convient alors de vérifier que X ainsi défini est FLEMMARD $\square$

36.3.2 Variables aléatoires indépendantes

Intuitivement des variables aléatoires sont indépendantes lorsqu'aucune d'elles ne dépend, d'aucune façon, des autres. Par exemple dans un sondage en vue d'un référendum, il est souhaitable que la variable "être sondée" soit indépendente de la variable "être partisan du oui" (chose très difficile à réaliser en pratique), sans quoi le sondage risque d'être biaisé. La notion d'indépendance est formalisée ci-dessous.

Définition 1311 (σ-algèbres indépendantes) *Soit S un σ-algèbre , et $(S_i)_{i\in I}$ une famille de sous-σ-algèbres de S ; alors les S_i sont dites σ-algèbres indépendantes si pour toute famille $(s_i)_{i\in J} \in \Pi_{i\in J} S_i$ avec J partie finie de I, on a $P(\cap_{i\in J} s_i) = \Pi_{i\in J} P(s_i)$.*
Soit $(X_i)_{i\in I}$ une famille de variables aléatoires ; alors les X_i sont dites **variables aléatoires indépendantes** *si les S_i sont des σ-algèbres indépendantes avec $S_i = X_i^{-1}(\mathcal{B})$ ($\mathcal{B}$ la σ-algèbre borélienne de $\mathbb{R}$).*
Des évènements E_i sont dits **évènements indépendants** *si les σ-algèbres $\{E_i, \Omega \setminus E_i, \emptyset, \Omega\}$ sont indépendantes ; ce qui équivaut au fait que les fonctions caractéristiques des E_i, en tant que variables aléatoires, sont indépendantes.*
On appelle suite de **variables aléatoires identiquement distribuées** *une suite de variables aléatoires indépendantes et ayant la même fonction de répartition.*

Rappel de vocabulaire :
• Un triplet de probabilité est un triplet $(\Omega, \mathcal{F}, P)$, avec $\mathcal{F}$ une tribu sur X, et P une mesure de probabilité sur $(\Omega, \mathcal{F})$.
• Ω est appelé univers.
• Un évènement est une partie mesurable de Ω, c'est à dire un élément de $\mathcal{F}$.
• Une variable aléatoire sur Ω est une application mesurable de Ω dans $\mathbb{R}$ (muni des boréliens usuels).
• La loi de probabilité d'une variable aléatoire X est l'application $\mathcal{L}_X$ qui à un borélien B inclus dans $\mathbb{R}$ associe $P(X^{-1}(B))$.
• La fonction de répartition d'une variable aléatoire X est l'application qui à un réel t associe $P(\{\omega \in \Omega / X(\omega) \leq t\})$.

Proposition 1312 *Pour que des évènements E_i pour $i \in I$ soient indépendants il suffit de vérifier que $P(\cap_{i\in J} E_i) = \Pi_{i\in J} P(E_i)$ pour tout J fini dans I.*

Démonstration : Il suffit de considérer les propriétés d'additivité de P, pour voir que cette formule permet de déduire les cas où des E_i sont remplacés par leurs complémentaires (cas de E_i remplacé par l'ensemble vide ou Ω trivial).$\square$

Définition 1313 *Deux Π-système P_1 et P_2 sur un même ensemble sont dits* **indépendants** *si pour tout $p_1 \in P_1$ et tout $p_2 \in P_2$ on a*

$$P(p_1 \cap p_2) = P(p_1).P(p_2)$$

Lemme 1314 *Supposons S_1 et S_2 deux σ-algèbres sur X engendrées respectivement par P_1 et P_2 des Π-systèmes. Alors S_1 et S_2 sont indépendantes si et seulement si P_1 et P_2 sont des Π-systèmes indépendants.*

Démonstration : Le sens "seulement si" étant trivial on se préoccupe de l'autre sens.
Fixons p_1 dans P_1, et considérons les mesures m_1 et m_2 définies sur S_2 par

$$m_1(E) = P(E \cap p_1)$$

et

$$m_2(E) = P(E).P(p_1)$$

Ces deux mesures coïncident sur P_2 et donnent une mesure finie à X ; donc par le lemme 336 elles sont égales.
Fixons maintenant p_2 dans P_2.
On définit maintenant les deux mesures m_3 et m_4 sur S_1 par

$$m_3(E) = P(E \cap p_2)$$

et

$$m_4(E) = P(E).P(p_2)$$

Elles coïncident sur P_1 et donnent une mesure finie à X ; donc par le lemme 336 elles sont égales.

On a donc montré le résultat souhaité pour p_1 et p_2 dans les cas suivants :
- $p_1 \in P_1$ et $p_2 \in P_2$, clair par hypothèse.
- $p_1 \in P_1$ et p_2 quelconque, par le lemme 336.
- p_1 quelconque et p_2 quelconque, en fixant p_2 et en utilisant le lemme 336.

Le résultat est donc prouvé.□

Lemme 1315 (Second lemme de Borel-Cantelli) *Si $(E_n)_{n\in\mathbb{N}}$ est une suite d'évènements indépendants, alors*

$$\sum_n P(E_n) = +\infty \to P(limsup\ E_n) = 1$$

Intuitivement, cela signifie que si la somme des probabilités pour qu'un évènement arrive à l'instant n pour $n \in N$ tend vers $+\infty$, alors l'évènement a une probabilité 1 d'avoir lieu une infinité de fois.

<u>Démonstration :</u> Notons que

$$(limsup\, E_n)^c = (\cap_m \cup_{n \geq m} E_n)^c$$

$$= \cup_m((\cup_{n \geq m} E_n)^c)$$

$$= \cup_m \cap_{n \geq m} (E_n^c)$$

$$= liminf\ (E_n^c)$$

Avec $p_n = P(E_n)$, pour tout p, par définition de l'indépendance, on a

$$P(\cap_{p \geq n \geq m} E_n^c) = \Pi_{p \geq n \geq m}(1 - p_n)$$

Donc en passant à la limite, par monotonie de l'intersection des E_n^c, on a

$$P(\cap_{n \geq m} E_n^c) = \Pi_{n \geq m}(1 - p_n)$$

$1-x \leq exp(-x)$, donc $\Pi_{n \geq m}(1-p_n) \leq \Pi_{n \geq m} exp(-p_n) \leq exp(-\sum_{n \geq m} p_n) \leq 0$, d'où le résultat.□

Le premier lemme de Borel-Cantelli, évoqué en partie **??**, fournit un complément à ce lemme.

Exemple 1316 (Application des deux lemmes de Borel-Cantelli) *On définit les $(E_n)_{n \geq 1}$ comme des évènements aléatoires, indépendants, par*

$$E_n = [\alpha.log(n), +\infty[$$

avec X variable aléatoire définie par sa fonction de répartition $F_X = e^{-x}$.
$P(E_n) = n^{-\alpha}$, donc $\sum_{n \geq 1} P(E_n) = +\infty$ si et seulement si $\alpha \leq 1$, et donc par les deux lemmes de Borel-Cantelli E_n a lieu infiniment souvent (c'est à dire que $P(limsup\, E_n) = 1$) si et seulement si $\alpha > 1$.

Définition 1317 (σ-algèbre asymptotique) *Etant donnée une suite de variables aléatoires X_1, ... , X_n, ... , on appelle σ-**algèbre asymptotique de la suite** $(X_n)_{n \in \mathbb{N}}$ la σ-algèbre $\cap_n \tau_n$, avec τ_n la σ-algèbre engendrée par $(X_{n+1}, X_{n+2}, ...)$.*

Pour bien comprendre cette définition il faut voir que :
• τ_n est la σ-algèbre qui rendre toutes les variables aléatoires X_i mesurables pour $i > n$.
• τ est l'intersection des τ_n.

Intuitivement la σ-algèbre asymptotique contient les évènements qui ne dépendent que du comportement à la limite.

Proposition 1318 *Les évènements suivants sont par exemples mesurables pour la σ-algèbre asymptotique des X_i (on les appelle* **évènements asymptotiques***) :*
• $\{\omega/lim_{n\to+\infty}X_i(\omega)$ *converge*$\}$
• $\{\omega/\sum_{n\to+\infty}X_i(\omega)$ *converge*$\}$
• $\{\omega/\lim_{n\to+\infty}sum_{i\in[0,1]}X_i(\omega))/n$ *converge*$\}$
Les variables aléatoires suivantes sont dans τ (on les appelle **variables asymptotiques***) :*
• $\limsup_{n\to+\infty}sum_{i\in[0,1]}X_i(\omega)/n)$
• $\liminf_{n\to+\infty}sum_{i\in[0,1]}X_i(\omega)/n)$

Pour contre-exemples on peut citer par exemple X_10 (variable aléatoire non-asymptotique dans le cas général), ou la somme des X_i pour $0\leq i\leq 10$.

Démonstration : pour les trois premiers points il faut donc montrer que l'ensemble E donné est inclus dans chaque τ_n.
• Les X_i sont mesurables, donc $liminf\ X_i$ et $limsup\ X_i$ sont mesurables, donc $(liminf\ X_i - limsup\ X_i)^{-1}(\{0\})$ est une partie mesurable. Donc $E\in\tau_1$; de la même manière $E\in\tau_n$, pour tout n, donc $E\in\tau$.
• Les X_i sont mesurables, donc une somme finie de X_i est mesurables, donc

$$limsup_{n\to+\infty}\sum_{i=1..n}X_i$$

est mesurable, pareil avec $liminf$, d'où le résultat en considérant

$$(limsup_{n\to+\infty}\sum_{i=1..n}X_i - limsup_{n\to+\infty}\sum_{i=1..n}X_i)^{-1}(\{0\}).$$

• Par une méthode similaire aux deux cas précédents on montre que E appartient à τ_1, il suffit de voir ensuite que $\lim_{n\to+\infty}sum_{i\in[0,1]}X_i(\omega))/n$ converge si et seulement si $\lim_{n\to+\infty}sum_{i\in[0,1]}X_i(\omega))/(n-K+1)$ converge pour conclure que E appartient aussi à τ_K.
Pour les variables aléatoires qui suivent les façons de raisonner sont les mêmes.$\square$

Théorème 1319 (Loi $0-1$ de Kolmogorov) *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes, et soit τ la σ-algèbre asymptotique des X_n ; alors :*
• *Tout évènement asymptotique a une probabilité 0 ou 1.*
• *Pour toute variable asymptotique Y, il existe un unique $z\in[-\infty,+\infty]$ tel que $P(Y=z)=1$.*

Démonstration : On procède par étapes :
• On montre tout d'abord que les σ-algèbres suivantes sont indépendantes pour tout n :

- la σ-algèbre engendrée par $X_1, ..., X_n$, notée par la suite Y_n.
- la σ-algèbre engendrée par $X_{n+1}, X_{n+2}, ...$, notée comme d'habitude τ_n.

En effet :
- la première de ces deux σ-algèbres est engendrée par le Π-système des ensembles de la forme $\{\omega/\forall i \in [1, n] X_i(\omega) \leq x_i\}$ avec $x_i \in]-\infty, +\infty]$.
- la seconde de ces deux σ-algèbres est engendrée par le Π-système des ensembles de la forme $\{\omega/\forall j \in [n+1, n+1+K] X_i(\omega) \leq x_j\}$ avec $x_j \in]-\infty, +\infty]$.

Par le lemme 1314, nos deux σ-algèbres sont donc indépendantes.

• Y_n et τ sont indépendantes

Il suffit de voir que $\tau \subset \tau_n$.

• τ_1 et τ sont indépendantes.

L'union des Y_n est un Π-système (facile), qui engendre τ_1 (facile aussi). D'après l'étape précédente, l'union des Y_n et τ sont indépendantes au sens des Π-systèmes ; donc les σ-algèbre engendrées sont indépendantes.

• τ étant inclus dans τ_1, τ est indépendante de τ, et donc pour tout $E \in \tau$, on a $P(E) = P(E \cap E) = P(E).P(E)$.

Le premier des deux points est alors prouvé. Pour trouver z du second point il suffit de considérer le plus grand z tel que $P(Y \leq z) = 0$.□

36.3.3 Espérance

⊡ Définitions

Définition 1320 (Espérance d'une variable aléatoire dans L^1) *Etant donnée X une variable aléatoire de $L^1(X, \mathbb{R})$, on définit son espérance par*

$$E(X) = \int_\Omega X.dP$$

Cette définition peut éventuellement être étendue aux variables aléatoires intégrables positives.
On définit en outre $E(X; F)$, avec X une variable aléatoire $\mathcal{L}^1$ ou bien une variable aléatoire intégrable positive, et F une partie mesurable, par

$$E(X; F) = \int_F X.d\mu = E(X.\chi_F)$$

avec χ_F la fonction caractéristique de F.

⊡ **Théorèmes et inégalités**

Théorème 1321 (Théorèmes de passage à la limite en probabilités) *Soit X_n une suite de variables aléatoires et X une variable aléatoire telles que*

$$P(X_n \to X) = 1$$

c'est à dire

$$P(\{\omega / X_n(\omega) \to X(\omega)\}) = 1$$

Alors les résultats de convergence monotone, de Fatou, de convergence dominée et de Scheffé, que l'on peut trouver dans la partie6, se reformulent comme suit :

• Convergence monotone :
Si les X_n sont ≥ 0 et $X_n(\omega)$ croit vers $X(\omega)$ pour $n \to +\infty$, alors $E(X_n) \to E(X)$.

• Lemme de Fatou :
Si $X_n \geq 0$ alors $E(X) \leq liminf\ E(X_n)$

• Théorème de convergence dominée de Lebesgue :
Si pour tout n et tout ω on a $|X_n(\omega)| \leq |Y(\omega)|$, avec $E(Y) \leq +\infty$, alors $E(|X_n - X|) \to 0$, et en particulier $E(X_n) \to E(X)$.

• Lemme de Scheffé :
Si $E(|X_n|) \to E(|X|)$, alors $E(|X_n - X|) \to 0$.

Démonstration : Voir le chapître6 pour les preuves correspondantes, qui s'appliquent directement.□

Théorème 1322 (Inégalité de Markov) *Supposons X variable aléatoire , et f mesurable de $\mathbb{R}$ dans $[0, +\infty]$, avec f croissante. Alors*

$$E(f \circ X) \geq E(f \circ X; X \geq c) \geq f(c). \int \chi_{\{\omega / X(\omega) \geq c\}}$$

que l'on peut aussi noter

$$E(f \circ X) \geq E(f \circ X; X \geq c) \geq f(c).P(X \geq c)$$

Démonstration : Il n'y a rien à prouver, il suffit de bien voir que f est positive.□

Corollaire 1323 *Avec X une variable aléatoire positive,*

$$E(X) \geq c.P(X \geq c)$$

Démonstration : C'est l'inégalité de Markov avec f la fonction identité.□

Corollaire 1324 *Pour X variable aléatoire positive, $P(X \geq z) \leq E(X)/z$.*

Démonstration : Il s'agit simplement de l'inégalité ci-dessus, reformulée.□

Théorème 1325 (Inégalité de Jensen) *On se donne f une application de U dans $\mathbb{R}$, avec U intervalle ouvert de $\mathbb{R}$, et X une variable aléatoire, avec les hypothèses suivantes :*

$$f \text{ convexe}$$

$$P(X \in U) = 1$$

$$E(|X|) < +\infty \text{ (c'est à dire que } X \text{ est intégrable)}$$

$$E(|f(X)|) < +\infty \text{ (c'est à dire que } f \circ X \text{ est intégrable)}$$

Alors :

$$E(f(X)) \geq f(E(X))$$

Voir par exemple les propriétés des fonctions caractéristiques en probabilités, 1343.

Démonstration : • Les dérivées à gauche et à droite de f, notée d^- et d^+, existent et sont croissantes ; on a en outre $d^-(x) \leq d^+(x)$.
• Considérons maintenant $z \in U$, et $x \in U$.
Soit $x < u < z$, alors la pente entre x et u est inférieure à la pente entre u et z ; en faisant tendre u vers z on constate alors que la pente entre x et z est inférieure à $d^-(z)$.
Donc :

$$f(x) \geq f(z) + d^-(z)(x - z)$$

De même pour $x > z$ on montrerait

$$f(x) \geq f(z) + d^+(z)(x - z)$$

• Comme $d^-(z) \leq d^+(z)$, on peut résumer ces assertions en

$$f(x) \geq f(z) + d^-(z)(x - z)$$

valable pour tout x.
• Il est facile de voir que $E(X) \in G$

• On peut donc spécialiser l'affirmation de l'avant-dernier point en $f(x) \geq f(E(X)) + d^-(E(X))(x - E(X))$ • En intégrant l'inégalité ci-dessus il vient

$$E(f(X)) \geq f(E(X))$$

La preuve est ainsi complète□

⊡ Espaces L^p

⚠ ⚠ ⚠ Dans le contexte des probabilités, L^p désignera, étant donné un univers Ω, $L^p(\Omega)$, Ω étant muni d'une mesure de probabilité (L^p est en fait dépendant de l'univers Ω, de la tribu définie sur Ω, et de la mesure définie sur cette tribu). Ne pas généraliser les résultats qui suivent au cas général de $L^p(X)$ pour X espace mesuré quelconque !

Proposition 1326 *Pour $p \in [1, +\infty]$, alors $L^{p'} \subset L^p$ pour tout $p' \geq p$ (éventuellement p' infini). En outre pour tout X dans $L^{p'}$, on a $N_{p'}(X) \leq N_p(X)$.*

Démonstration : Pour l'inclusion, il suffit de voir la proposition 405.
Pour l'inégalité, on peut clairement se ramener au problème des variables aléatoires positives.
Etant donné X à valeurs positives dans $L^{p'}(X)$, on définit $X_n(\omega) = min(X(\omega), n)$.
Alors X_n est bornée, et donc $X_n^{p'}$ et X_n^p aussi, donc $X_n^{p'}$ et X_n^p sont dans L^1 (on utilise le fait que la mesure est finie). On peut donc appliquer l'inégalité de Jensen (voir théorème 1325) avec la variable aléatoire X_n^p et la fonction convexe $x \mapsto x^{p'/p}$, et écrire

$$E(X_n^p)^{p'/p} \leq E(X_n^{pp'/p}) = E(X_n^{p'}) \leq E(X^{p'})$$

On applique alors le théorème de convergence monotone à X_n^p et

$$E(X^p)^{p'/p} \leq E(X^{p'})$$

En élevant à la puissance $1/p'$ on a alors

$$N_p(X) \leq N_{p'}(X)$$

La preuve est ainsi complète.□

Les résultats usuels dans L^p sont valables, notamment l'inégalité de Schwartz, de Hölder, de Minkowski, pour lesquels on consultera la partie 8.
Pour rappeler l'essentiel :
• Si X et Y sont des variables aléatoires de L^2, alors le produit $X.Y$ appartient à L^1, et

$$|E(X.Y)| \leq E(|X.Y|) \leq N_2(X).N_2(Y)$$

• Si X et Y sont des variables aléatoires de L^2, alors la somme $X + Y$ appartient à L^2, et

$$N_2(X + Y) \leq N_2(X) + N_2(Y)$$

Une proposition est nécéssaire pour bien comprendre ce qu'il se passe :

Proposition 1327 *Soit X une variable aléatoire , et soit f une fonction mesurable de $\mathbb{R}$ dans $\mathbb{R}$, alors $f \circ X$ est une variable aléatoire de L^1 (au sens donné ici, c'est à dire $L^1(\Omega)$, avec Ω muni d'une mesure de probabilité) si et seulement si f est dans $L^1(\mathbb{R}, L_X)$ avec L_X la loi de X.*
On a alors

$$E(f \circ X) = \int f(x).dL_X$$

Voir la proposition 1307 et la définition qui la précède pour bien cerner ce qu'est une loi de probabilité.

Démonstration : Si vous n'arrivez pas à le faire vous-mêmes, mieux vaut relire tout le chapître. La méthode est la suivante :
• Si f est une fonction caractéristique d'un borélien, il s'agit simplement de la définition de la loi de probabilité.
• Si f est simple, alors par linéarité la propriété est aussi vraie.
• Si f est positive, alors f est limite croissante de fonctions simples, donc on peut appliquer le théorème de convergence monotone.
• Enfin dans le cas général, f s'écrit comme différence de deux fonctions mesurables positives.□

Définition 1328 (Mesure image) *Etant donnée f une application mesurable d'un espace Ω mesuré par une mesure μ dans $\mathbb{R}$ muni des boréliens, on note μ^f la mesure appelée* **mesure image de** μ **par** f *définie sur l'ensemble des boréliens de $\mathbb{R}$ par*

$$\mu^f(E) = \mu(f^{-1}(E)$$

Il s'agit bien d'une mesure.

Théorème 1329 (Théorème de transport) *Pour toute fonction mesurable ϕ de $\mathbb{R}$ dans $\mathbb{R}$,*

$$\int_{\mathbb{R}} \phi d\mu^f = \int_{\Omega} \phi \circ f \, d\mu$$

On ramène ainsi les intégrales du type $\int_\Omega dP$ à des intégrales sur $\mathbb{R}$ pour la mesure de Lebesgue ; on n'a pas besoin de connaître la structure de Ω, mais seulement les lois.

Démonstration : Le chapître sur l'intégration permet de comprendre clairement les notions en jeu. Il s'agit en fait simplement de vérifier la formule dans le cas d'une fonction caractéristique d'un borélien, puis d'un le cas d'une fonction simple (i.e. étagée[1] et mesurable) grâce à la linéarité de l'intégrale, puis pour une fonction positive

[1]Etagée = ne prenant qu'un nombre fini de valeurs

par passage au sup, puis dans le cas général en exprimant une fonction comme différence de deux fonctions l'une positive et l'autre négative (utilisation du théorème de convergence monotone à la fois pour les fonctions simples tendant vers ϕ et pour les fonctions simples tendant vers $\phi \circ f$).□

Corollaire 1330 *On peut écrire le même théorème avec une fonction ϕ de $\mathbb{R}^n$ dans $\mathbb{R}$ et f de Ω dans $\mathbb{R}^d$.*

Démonstration : Même principe que ci-dessus.□

⊡ Densité de probabilité

Définition 1331 *Etant donné X une variable aléatoire , une application f_X mesurable est appelée une* **densité de probabilité de** *X si et seulement si pour tout borélien E de $\mathbb{R}$ $P(X^{-1}(E)) = \int_E f_X$.*

Notons que bien sûr $\int_{\mathbb{R}} f_X = 1$

⊡ Variance, covariance, lois jointes, densités jointes, fonctions de répartition jointes

Définition 1332 (Covariance et variance) *Etant donnée X une variable aléatoire , on définit la* **déviation** *de X par $\tilde{X} = X - E(X)$.*
Etant données X et Y des variables aléatoires dans L^2, on définit la **covariance** *de X et Y par*

$$Cov(X, Y) = E(\tilde{X}.\tilde{Y})$$

Etant donnée X une variable aléatoire dans L^2, on définit la **variance** *de X par*

$$Var(X) = Cov(X, X)$$

Le **produit scalaire de deux variables aléatoires** *X et Y de L^2 est l'espérance de $X.Y$ (comme dans le cadre d'un espace L^2 quelconque), noté $< X|Y >$.*
On appelle **corrélation** *entre deux variables aléatoires X et Y de norme 2 non nulles le réel de $[0, 1]$ $cor(X, Y) = \frac{<\tilde{X}|\tilde{Y}>}{N_2(\tilde{X}).N_2(\tilde{Y})}$. On appelle* **angle** *entre deux variables aléatoires X et Y de norme 2 non nulles le réel θ appartenant à $[0, \Pi]$ tel que $cos(\theta) = \frac{<X|Y>}{N_2(X).N_2(Y)}$.*
Deux variables aléatoires sont dites **non corrélées** *si leur covariance est nulle.*
On appelle **matrice de covariance** *d'un suite fini de variables aléatoires $(X_1, ..., X_d)$ la matrice M définie par $M_{i,j} = cov(X_i, X_j)$.*

Corollaire 1333 (Inégalité de Tchébitchev) *Pour X variable aléatoire, $P(|X - E(X)| > \epsilon) \leq Var(X)/\epsilon^2$.*

Voir le théorème 417 sur les polynomes de Bernstein.

Démonstration : Il suffit d'appliquer le corollaire 1324 de l'inégalité de Markov à $(X - E(X))^2$.$\square$

La définition de la covariance et de la variance se justifie par le fait que si X est dans L^2, alors $X - E(X)$ aussi, et donc $(X - E(X)).(Y - E(Y))$ est dans L^1 par l'inégalité de Schwartz.
La définition de la corrélation se justifie par l'inégalité de Schwartz.
La corrélation entre deux variables aléatoires est le cosinus de l'angle entre les déviations de ces variables aléatoires.
On a $cov(X,Y) = E(X.Y) - E(X).E(Y) =< \tilde{X}|\tilde{Y} >$ et $var(X) = E(X^2) - E(X)^2$.
Si $X_1, ..., X_n$ sont des variables aléatoires , alors

$$var(\sum_{i\in[1,n]} X_i) = \sum_{(i,j)\in[1,n]^2} cov(x_i, x_j)$$

$$var(\sum_{i\in[1,n]} X_i) = \sum_{i\in[1,n]} var(x_i) + \sum_{(i,j)\in[1,n]^2, i\neq j} cov(x_i, x_j)$$

$$var(\sum_{i\in[1,n]} X_i) = \sum_{i\in[1,n]} var(x_i) + 2.\sum_{(i,j)\in[1,n]^2, i<j} cov(x_i, x_j)$$

Pour plus d'informations voir8 et plus spécialement8.4.2.

Théorème 1334 (Une propriété fondamentale des variables aléatoires indépendantes)
Soient X et Y des variables aléatoires indépendantes appartenant à L^1. Alors $X.Y$ est L^1 et

$$E(X.Y) = E(X).E(Y)$$

Soient X et Y des variables aléatoires indépendantes appartenant à L^2. Alors :

$$cov(X,Y) = 0$$

$$var(X+Y) = var(X) + var(y)$$

Démonstration : On se préoccupe tout d'abord du premier résultat :
• Si X et Y sont des fonctions caractéristiques, alors $X = \chi_E$ et $Y = \chi_F$, et $E(X.Y) = P(\chi_{E\cap F}) = P(E).P(F)$ par indépendance.
• Si X et Y sont des fonctions simples alors ce sont des combinaisons linéaires de fonctions caractéristiques, donc le résultat est aussi valable.
• Si X et Y sont positives, alors ce sont des limites de fonctions simples, donc le résultat est aussi valable par le théorème de convergence monotone.
• Dans le cas général, X et Y s'écrivent comme différences de deux fonctions positives.
La suite se déduit facilement, au vu des définitions de la covariance et de la variance.$\square$

⚠ Notez bien qu'il n'y a PAS d'erreur dans l'énoncé, X et Y sont supposées dans le premier cas appartenant à L^1, et pas nécéssairement à L^2.

Pour cerner plus précisément l'intérêt de l'indépendance des variables aléatoires , on a besoin de définitions supplémentaires utilisant les mesures produits (voir la partie6.8 pour connaître les bases requises).

Définition 1335 *Etant données $X_1, ..., X_n$ des variables aléatoires , on appelle*
• **loi jointe** *de $X_1, ..., X_n$ ou simplement* **loi** *de $X_1, ..., X_n$ l'application $L_{X_1,...,X_n}$ qui à un borélien E de $\mathbb{R}^n$ associe $P(F)$ avec $F = \{\omega \in \Omega/(X_1(\omega), ..., X_n(\omega)) \in E\}$.*
• **fonction de répartition** *de $X_1, ..., X_n$ l'application qui à $(x1, ..., x_n)$ dans $\mathbb{R}^n$ associe $L_{X_1,...,X_n}(]-\infty, x_1], ...,]-\infty, x_n])$.*
• **densité de probabilité** *ou simplement* **densité de probabilité** *de $X_1, ..., X_n$ une application f de $\mathbb{R}^n$ dans $\mathbb{R}$ telle que pour tout borélien E de $\mathbb{R}^n$ on ait $L_{X_1,...,X_n}(E) = \int_E f$.*

On note que le théorème de Fubini permet d'affirmer qu'étant donnée f densité de probabilité jointe de $X_1, ..., X_n$ l'application

$$x \mapsto \int_{(x_1,...,x_{i-1},x_{i+1},...,x_n)\in\mathbb{R}^{n-1}} f(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_n)$$

est une densité de probabilité de X_i.

Théorème 1336 *Soient $X_1, ..., X_n$ des variables aléatoires . On note L_{X_i} la loi de probabilité de X_i, F_{X_i} la fonction de répartition de X_i, $L_{X_1,...,X_n}$ la loi de probabilité jointe de $X_1, ..., X_n$, $F_{X_1,...,X_n}$ la fonction de répartition de $X_1, ...X_n$, f_{X_i} une densité de probabilité de X_i, $f_{X_1,...,X_n}$ une densité de probabilité de $X_1, ...X_n$.*
Alors

$$X_1, ..., X_n \text{ sont indépendantes} \iff L_{X_1,...,X_n} = L_{X_1} \otimes ... \otimes L_{X_n}$$

$$X_1, ..., X_n \text{ sont indépendantes} \iff$$
$$F_{X_1,...,X_n}(x_1, ..., x_n) = F_{X_1}(x_1) \times ... \times F_{X_n}(x_n)$$

$$X_1, ..., X_n \text{ sont indépendantes} \iff$$
$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_{X_1}(x_1) \times ... \times f_{X_n}(x_n) \text{presque partout}$$

Démonstration : Admise.□

Proposition 1337 (Egalité de Bienaymé) *Si les X_i sont deux à deux non corrélées (par exemples indépendantes), alors*

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

Démonstration :

$$Var(\sum_i X_i)$$

$$= E((\sum X_i - E(\sum X_i))^2)$$

$$= E((\sum_i X_i - E(X_i))^2)$$

$$= \sum_{(i,j)\in[1,n]^2} E((X_i - E(X_i)).(X_j - E(X_j)))$$

$$= \sum_{i\in[1,n]} Var(X_i)$$

La preuve est complète...□

Corollaire 1338 (Inégalité de Bienaymé-Tchébitchev) *Si les $(X_i)_{i\in[1,n]}$ sont deux à deux indépendantes, pour $t > 0$,*

$$P(|\sum_i X_i - E(X_i)| \geq t) \leq \frac{\sum_i Var(X_i)}{t^2}$$

On peut par exemple voir 36.12.5.

Démonstration : Il suffit de combiner l'inégalité de Tchébitchev et l'égalité de Bienaymé.□

36.4 Somme de variables aléatoires et transformée de Fourier

Définition 1339 (Produit de convolution) *On appelle* **produit de convolution** *de deux lois de probabilités indépendantes P^X et P^Y sur $\mathbb{R}$ la loi $P^X * P^Y$ sur $\mathbb{R}$ définie par*

$$(P^X * P^Y)(E) = \int_E (\int_E dP^Y(x-y))dP^X(x)$$

Proposition 1340 (Propriétés fondamentales du produit de convolution) *Si X et Y sont deux variables aléatoires réelles indépendantes de lois P^X et P^Y :*
• La loi de $X+Y$ est $P^X * P^Y$
• $P^X * P^Y = P^Y * P^X$
• Pour toute fonction mesurable f de $\mathbb{R}$ dans $\mathbb{R}$,

$$\int_{\mathbb{R}} f(x).d(P^X * P^Y)(x)$$

$$= \int_{\mathbb{R}} (\int_{\mathbb{R}} f(x+y)dP^Y(y))dP^X(x)$$

Démonstration :

$$P^{X+Y}(E) = P^{X\times Y}(som(X,Y) \in E)$$

avec $som : (x,y) \mapsto x+y$

$$= \int_{\Omega}\int_{\Omega} \chi_E(Y(\omega_1) + X(\omega_2))d\omega_2 d\omega_1$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \chi_E(Y(\omega_1) + X(\omega_2))dP^X dP^Y$$

Et le premier point en découle ; le deuxième point découle de la commutativité de l'addition, et le troisième point est une application immédiate du théorème de transport.□

Proposition 1341 (Liste de propriétés du produit de convolution) *On se donne X, Y et Z des variables aléatoires réelles et P^X, P^Y et P^Z leurs lois.*
• Le produit de convolution de la loi P^X par une masse de Dirac située[a] en 0 est la loi P^X elle-même.
• Le produit de convolution de P^X par une masse de Dirac située en x est la loi de $X+x$.
• Le produit de convolution est commutatif, associatif.
• Le produit de convolution est distributif, en un sens bien précis ; pour t dans $[0,1]$, on a :

$$P^X * (t.P^Y + (1-t).P^Z) = t.P^X * P^y + (1-t).P^X * P^Y$$

[a] Une masse de Dirac en X est une mesure δ_x telle que $\delta_x(E) = 1$ si $x \in E$ et $\delta_x(E) = 0$ sinon.

Démonstration : Les trois premiers • sont évidents, au vu de la proposition précédente. Le quatrième vient simplement du fait que $t.P^Y + (1-t).P^Z$ est bien une loi de probabilité.□

Définition 1342 (Fonction caractéristique) *Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$. On appelle* **fonction caractéristique de** X *la fonction*

$$\phi^X : \mathbb{R}^d \to \mathbb{R}^d$$

définie par

$$\phi^X(t) = E(e^{i<t,X>})$$

Les adeptes auront reconnu une transformée de Fourier.

Proposition 1343 • $\phi^X(t) = \int_{\mathbb{R}^d} cos(< t,x >)dP^X(x) + i.\int_{\mathbb{R}^d} sin(< t,x >)dP^Y(y)$
• $\phi^X(0) = 1$
• ϕ^X *est à valeurs dans le disque unité fermé de* $\mathbb{C}$
• $\phi^X = \phi^Y$ *implique* $P^X = P^Y$.

Démonstration :
Point par point :
• Par définition !
• Clair !
• Grâce à l'inégalité de Jensen (voir 1325).
• Ce point, délicat, sera ici admis.□
Quelques exemples de fonctions caractéristiques :
- Si P^X est un dirac en x, alors $\phi^X(t) = e^{i<t.x>}$.
- Etant donné X une variable aléatoire à valeurs dans $\mathbb{R}^d$, M une matrice de type (d,d), et C un vecteur dans $\mathbb{R}^d$, avec $Y = M.X + C$, on a

$$\phi^Y(t) = E(e^{i<t,Y>})$$
$$= E(e^{i<t,MX>+i<t,C>})$$
$$= e^{i<t,C>}.E(e^{i<t,MX>})$$
$$= e^{i<t,C>}.E(e^{i<X,{}^tMt>})$$
$$= e^{i<t,C>}.\phi^X({}^tMt)$$

- On trouvera d'autres exemples dans la partie 36.8.

Théorème 1344 (Formule d'inversion de Fourier) *On suppose que ϕ^X, fonction caractéristique de la variable aléatoire X, est intégrable. Alors X admet une densité continue bornée f^X, et on a*

$$f^X(x) = \frac{1}{(2\Pi)^d}\int_{\mathbb{R}^d} e^{-i<t,x>}\phi^X(t).dt$$

Démonstration : On se réfère à la partie consacrée aux séries de Fourier, où l'on trouvera d'ailleurs de nombreux résultats complémentaires...□

Définition 1345 *On appelle* **moment d'ordre** k *de la variable aléatoire* X *à valeurs dans* $\mathbb{R}$ *l'espérance de* X^k. *On appelle* **moment centré d'ordre** k *de la variable aléatoire* X *à valeurs dans* $\mathbb{R}$ *l'espérance de* $(X - E(X))^k$.

Le résultat suivant est donné sans preuve.

Proposition 1346 *Deux variables aléatoires bornées ayant les mêmes moments à tous ordres sont égales.*

36.5 Probabilités conditionnelles

Cette partie sera indispensable pour bien comprendre la partie sur les martingales (36.6). Les démonstrations, souvent abstraites et difficiles, seront laissés de côté. FLEMMARD

36.6 Martingales

Définition 1347 *On appelle* **espace filtré** *un quadruplet* $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, P)$ *avec* $(\Omega, \mathcal{F}, P)$ *triplet de probabilité, et* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *une* **filtration**, *c'est à dire une suite croissantes de* σ*-algèbres incluses dans* $\mathcal{F}$.
On appelle **processus adapté à un espace filtré** *une suite* $(X_n)_{n\in\mathbb{N}}$ *de variables aléatoires à valeurs dans* $\mathbb{R}$ *telles que*

$$\forall n \in \mathbb{N} X_n \text{ est } \mathcal{F}_n\text{-mesurable}$$

On appelle **processus prévisible** *(relativement à un espace filtré) une suite* $(X_n)_{n>0}$ *de variables aléatoires à valeurs dans* $\mathbb{R}$ *telles que pour tout* $n > 0$ C_n *est* $\mathcal{F}_{n-1}$*-mesurable.*
On appelle **temps d'arrêt** *une application de* Ω *dans* $\mathbb{N}$ *telle que pour tout* n $\{\omega / T(\omega) \leq n\}$ *appartient à* $\mathcal{F}_n$.
On appelle **processus prévisible** *associé à un temps d'arrêt le processus prévisible* C *tel que* $C_n(\omega)$ *est égal à* 1 *si* $n \leq T(\omega)$ *et égal à* 0 *sinon.*
Etant donnés X *un processus et* T *un temps d'arrêt, on note* X^T *le* **processus** X **stoppé à l'instant** T *défini par* $X^T_n(\omega) = X_{min(T(\omega),n)}(\omega)$.
Etant donnés un processus prévisible C *et une martingale* X, *on note* $(C \bullet X)_n = \sum_{i=1}^n C_i(X_i - X_{i-1})$ *pour* $n > 1$.
Un processus C *est dit* **borné** *si il existe* K *tel que pour tout* n *et tout* ω, $|C_n(\omega)|$ *est majoré par* K.

En fait l'espace filtré représente les connaissance disponibles à l'instant $n \in \mathbb{N}$, dans un espace à temps discret ; c'est-à-dire qu'une fonction est $\mathcal{F}_n$-mesurable à condition qu'elle puisse être connue à l'instant n. Ensuite le fait qu'un processus soit adapté, signifie simplement que la valeur de $X_n(\omega)$ est connue à l'instant n. Un processus prévisible est en fait un processus déterminé à l'avance, ie le processus à l'instant n est connu dès l'instant $n-1$. Un processus prévisible sera notamment usuellement une stratégie élaborée par un joueur, qui peut donc agir en fonction de ce qui a déjà eu lieu, la stratégie étant supposée déterministe. Un temps d'arrêt est en fait une façon de décider un instant, sachant que la décision d'un instant ne peut être faite qu'en fonction des évènements antérieurs. $(C \bullet X)_n$ représente le total des gains à l'instant n, C_n représentant la mise, et $X_n - X_{n-1}$ le gain avant multiplication par la mise. Le processus prévisible associé à un temps d'arrêt est en fait une façon de jouer où l'on ne choisit pas la mise, mais pour laquelle on peut choisir le moment où le jeu s'arrête.

On verra un temps d'arrêt sympathique et un processus stoppé sympathiques en partie 36.12.10.

Définition 1348 *Un processus adapté X est une* **martingale** *si pour tout n X_n est L^1 ET si pour tout n l'espérance conditionnelle $E(X_n|\mathcal{F}_{n-1})$ est égale à X^{n-1}.*
Un processus adapté X est une **surmartingale** *si pour tout n X_n est L^1 ET si pour tout n l'espérance conditionnelle $E(X_n|\mathcal{F}_{n-1})$ est $\leq$ à X^{n-1}.*
Un processus adapté X est une **sous-martingale** *si pour tout n X_n est L^1 ET si pour tout n l'espérance conditionnelle $E(X_n|\mathcal{F}_{n-1})$ est $\leq$ à X^{n-1}.*

On comprend bien ce que signifie le fait que X_n soit L^1 ; la condition sur l'espérance conditionnelle, signifie, elle, simplement que la moyenne de X_n, toutes les informations étant connues jusqu'à l'étape $n-1$, est égale à X_{n-1}. C'est à dire que si l'on fixe les $n-1$ premières étapes, la n-ième est centrée (a sa moyenne) sur l'étape $n-1$.

En voyant X_n comme le gain à un jeu jusqu'à l'instant n inclus, une surmartingale est un jeu où en moyenne on perd, une sous-martingale un jeu où en moyenne on gagne.

Proposition 1349 *Si X est une surmartingale, $-X$ est une sous-martingale. X est une surmartingale si et seulement si X est une surmartingale et une sous-martingale.*

Exemple 1350 *On aura souvent comme filtration $\mathcal{F}_n = \sigma(W_0, \dots, W_n)$ (σ-algèbre engendrée par $W_0, \dots, W_n$), et $X_n = f_n(W_0, \dots, W_n)$ avec f_n mesurable de $\mathbb{R}^n$ dans $\mathbb{R}$ comme processus adapté.*

Exemple 1351 *Soit $(X_n)_{n \in \mathbb{N}}$ des variables aléatoires indépendantes L^1 d'espérance nulle.*

On définit $S_n = \sum_{i\in[0,n]} Y_i$. La filtration choisie est définie par : $\mathcal{F}_n$ est la σ-algèbre engendrée par $(X_0,\dots,X_n)$. Alors S_n est une martingale.

Avec $X_i = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$, variable aléatoire à valeurs dans $\{-1,1\}$ (équirépartie sur ces deux valeurs), on a une **marche aléatoire sur** $\mathbb{Z}$.

On peut aussi prendre des variables aléatoires X_i positives, d'espérance 1, indépendantes, pour X, et définir Π_n le produit des X_i pour $i \leq n$. La filtration se définit comme dans le cas ci-dessus.

Théorème 1352 (On peut pas gagner si on a un porte-monnaie fini) • *Si C est un processus prévisible et <u>borné</u> et <u>positif</u> et si X est une surmartingale, une sous-martingale, une martingale (respectivement), alors $(C \bullet X)$ est une surmartingale, une sous-martingale, une martingale.*
• Si C est un processus prévisible et <u>borné</u> et X une martingale, alors $C \bullet X$ est une martingale.

<u>Démonstration :</u> En utilisant les propriétés de l'espérance conditionnelle,

$$E((C \bullet X)_n - (C \bullet X)_{n-1}|\mathcal{F}_{n-1} = C_n E(X_n - X_{n-1}|\mathcal{F}_{n-1}) =$$

$$C_n(E(X_n|\mathcal{F}_{n-1}) - E(X_{n-1}|\mathcal{F}_{n-1})) = C_n(E(X_n|\mathcal{F}_{n-1}) - X_{n-1},$$

d'où les résultats en appliquant les définitions des martingales, des surmartingales,d es sousmartingales.□

Corollaire 1353 *Si X est une surmartingale, et T un temps d'arrêt, alors le processus stoppé X^T est une surmartingale et $E(X_{T^n}) \leq E(X_0)$. Si X est une martingale, et T un temps d'arrêt, alors le processus stoppé X^T est une martingale et $E(X_{T^n}) = E(X_0)$.*

Le résultat suivant provient de [21] :

Théorème 1354 (Théorème d'arrêt éventuel de Doob) *Soit T un temps d'arrêt et X une surmartingale, alors si l'une des conditions suivantes est vérifiée : • $\exists N / \forall \omega T(\omega) < N$*
• $\exists K / \forall(\omega, n) |X_n(\omega)| < K$ et pour presque tout ω T est fini.
• $E(T) < \infty$ et $\exists K / \forall(n,\omega) |X_n(\omega) - X_{n-1}(\omega)| \leq K$
on peut conclure que $E(X_T) \leq E(X_0)$

<u>Démonstration :</u> Application facile des résultats ci-dessus, en utilisant la convergence dominée de Lebesgue 366 dans le troisième cas. □

36.7 Processus stochastique. Processus de Markov

Définition 1355 *On dit qu'une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires à valeurs dans un ensemble E au plus dénombrable muni de la σ-algèbre $P(E)$ est une* **chaîne de Markov** *dans E* **espace des états** *si pour tout $(i_0,\dots,i_n)$ suite finie d'élements de E telle que $P(X_0 = i_0 \wedge X_1 = i_1 \wedge \dots \wedge X_{n-1} = i_{n-1}) > 0$, $P(X_n = i_n | X_0 = i_0 \wedge X_1 = i_1 \wedge \dots \wedge X_{n-1} = i_{n-1}) = P(X_n = i_n | X_{n-1} = i_{n-1})$.*
La chaîne de Markov est dite **homogène** *si pour tout n $P(X_{n+1} = j | X_n = i)$ est indépendant de n tel que $P(X_n = i) > 0$.*
On appelle **matrice stochastique** *une application M de E^2 dans $[0,1]$ telle que pour tout i $\sum_{j=0}^{n} M_{i,j} = 1$. La* **matrice stochastique***, dite aussi* **matrice de transition***, associée à une chaîne de Markov homogène telle que pour tout i dans E il existe n tel que $P(X_n = i) > 0$* [a] *est la matrice M définie par $M_{i,j} = P(X_{n+1} = j | X_n = i)$.*

[a] Cas auquel on peut toujours se ramener, en restreignant E.

Cela signifie simplement que l'état à l'instant n (ie X_n) ne dépend que de l'état à l'instant x_{n-1} et pas des états aux instants antérieurs. La chaîne est homogène si les changements d'états ne dépendent que de l'état, et pas de la date. Dans beaucoup de modélisations, la chaîne est homogène.

Les marches aléatoires, définies dans la partie 36.6, sont des exemples de chaînes de Markov.

Remarquons l'égalité de Chapman-Kolmogorov : $P(X_{m+n} = j | X_0 = i) = \sum_{k\in E} P(X_m = j | X_0 = k) P(X_n = k | X_0 = i)$

Notons que les produits de matrices stochastiques, définis comme généralisation du produit usuel de matrice par $MN = P$ avec $P_{i,j} = \sum_{k\in E} M_{i,k} N_{k,j}$, sont bien définis et sont encore des matrices stochastiques. On remarque aussi que :

Proposition 1356 *Si X est un processus de Markov de matrice de transition M* [a]
- $P(X_0 = i_0 \wedge X_1 = i_1 \wedge X_n = i_n) = P(X_0 = i_0) M_{i_0,i_1} M_{i_1,i_2} \dots M_{i_{n-1},i_n}$.
- $P(X_n = i) = \sum_{j\in E} (P^n)_{j,i}$

[a] Ceci impliquant que X est une chaîne de Markov homogène et que pour tout i dans E il existe n tel que $P(X_n = i) > 0$.

FLEMMARD

36.8 Zoologie des lois de probabilité

36.8.1 Lois normales

FLEMMARD

36.8.2 Loi de Bernoulli

Proposition 1357 (Loi de Bernouilli) • *Paramètre :* $B(p)$ *a pour paramètre* $p \in [0,1]$
• *A valeurs dans* $\{0,1\}$
• *Loi :* $P(X=1)=p$ *&* $P(X=0)=1-p$
• *Espérance :* p
• *Variance :* $p.(1-p)$
• *Fonction caractéristique :* $\phi(t)=1-p+p.e^{it}$
• *Intuition : pile ou face si* $p=\frac{1}{2}$*, pile ou face "biaisé" sinon*

36.8.3 Loi binomiale et multinomiale

⊡ Loi binomiale

Proposition 1358 (Loi binomiale) • *Paramètres :* $B(n,p)$ *a pour paramètres* n *dans* $\mathbb{N}$ *et* $p \in [0,1]$
• *A valeurs dans* $\{0,1,2,...,n\}$
• *Loi :* $P(X=k)=C_n^k p^k(1-p)^{n-k}$ *si* $k \in [0,n]$ *et* $P(X=k)=0$ *sinon*
• *Espérance :* $n.p$
• *Variance :* $n.p.(1-p)$
• *Fonction caractéristique :* $\phi(t)=(1-p+p.e^{it})^n$
• *Intuition : somme de* n *lois de Bernoullis de même paramètre* p*.*
• *Signe particulier : la somme de deux variables aléatoires lois binomiales* $B(n_1,p)$ *et* $B(n_2,p)$ *est une variable aléatoire de loi* $B(n_1+n_2,p)$ *(les deux lois binomiales en question etant supposées indépendantes !). On peut de la même manière sommer un nombre quelconque de lois binomiales (conformément à l'intuition ci-dessus d'ailleurs).*
• *Cas particulier :* $B(1,p)=B(p)$*, loi de Bernoulli.*
• *Cas limite : Si* $lim_{n\to\infty} n.p_n=\lambda$*, alors* $B(n,p_n)$ *converge en loi vers une variable aléatoire de loi de Poisson* $P(\lambda)$*. Noter que seul le produit* $n.p_n$ *compte pour ce passage à la limite ; d'où l'additivité des lois de Poisson quel que soient leurs paramètres.*

⊡ Loi géométrique

Proposition 1359 (Loi géométrique) • *Paramètre : $G(p)$ a un paramètre $p \in]0,1]$*
• *A valeurs dans i*
• *Loi : $P(X = k) = p.(1-p)^k$*
• *Espérance : $\frac{1-p}{p}$*
• *Variance : $\frac{1-p}{p^2}$*
• *Fonction caractéristique : $(\frac{p}{1-q.e^{it}})^n$ $FLEMMARD$*
• *Intuition : on tire au sort jusqu'à ce que l'on gagne, sachant qu'à chaque étape on a une probabilité p de gagner. Le nombre d'échecs avant la première victoire suit une loi géométrie $G(p)$.*

⊡ loi binomiale négative

Proposition 1360 (Loi binomiale négative) • *Paramètres : $B^-(n,p)$ a deux paramètres $n \in \mathbb{N}$ et $p \in]0,1]$*
• *A valeurs dans $\mathbb{N}$*
• *Loi : $P(X = k) = C^{n-1}_{n+k-1}p^n.(1-p)^k$ pour tout $k \in \mathbb{N}$*
• *Espérance : $n.\frac{1-p}{p}$*
• *Variance : $n.\frac{1-p}{1^2}$*
• *Fonction caractéristique : $\phi(t) = (\frac{p}{1-q.e^{it}})^n$ $FLEMMARD$*
• *Intuition : c'est un peu comme une série géométrique, à part que l'on attend d'avoir gagné n fois ; on compte le nombre d'échecs.*
• *Cas limite : $B^-(1,p) = G(p)$*

⊡ Loi multinomiale

Proposition 1361 (Loi multinomiale) • *Paramètre : $\mathcal{M}(n, p_1, p_2, ..., p_d)$ a pour paramètres $n \in \mathbb{N}$ et $(p_1, ..., p_d) \in [0,1]^d$ avec $\sum_{i=1}^d p_i = 1$*
• *A valeurs $(n_1, ..., n_d) \in [0,n]^d$, avec $\sum_{i=1}^d n_i = n$*
• *Loi : $P(X = (n_1, ..., n_d)) = \frac{n!}{n_1!.n_2!...n_3!}$ si $\sum_{i=1}^d n_i = n$ et 0 sinon*
• *Espérance : $(n.p_1, n.p_2, ..., n.p_d)$*
• *Matrice de covariance : $M_{i,j} = -n.p_i.p_j$ si $i \neq j$, $M_{i,i} = n.p_i.(1-p_i)$*
• *Fonction caractéristique : $\phi(t) = FLEMMARD$*
• *Intuition : on tire au sort n fois un nombre entier entre 1 et d, et la i-ième composante représente le nombre de fois que l'on a tiré l'entier i.*

36.8.4 Loi de Poisson

Proposition 1362 (Loi de Poisson) • *Paramètre :* $\mathcal{P}(\lambda)$ *a pour paramètre* $\lambda \in \mathbb{R}$
• *A valeurs dans* $\mathbb{N}$
• *Loi :* $P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}$
• *Espérance :* λ
• *Variance :* λ
• *Fonction caractéristique :* $\phi(t) = e^{\lambda.(e^{it}-1)}$
• *Intuition : cas limite de la loi binomiale (voir partie 36.8.3)*
• *Signe particulier : la somme de deux variables aléatoires de lois de Poisson* $\mathcal{P}(\lambda)$ *et* $\mathcal{P}(\mu)$ *est une variable aléatoire de loi de Poisson* $\mathcal{P}(\lambda+\mu)$*.*
• *Cas limite : Si* X *suit une loi* $\mathcal{P}(\lambda)$*, alors* $\frac{X-\lambda}{\sqrt{\lambda}}$ *converge en loi vers une loi normale* $\mathcal{N}(0,1)$ *quand* λ *tend vers* $+\infty$*.*

36.8.5 Loi hypergéométrique

Proposition 1363 (Loi hypergéométrique) • *Paramètre :* $H(N,n,p)$ *a pour paramètres* N *un entier,* n *un entier* $\leq N$*, et* p *de la forme* q/N *avec* $q \in \{0,1,...,N\}$*.*
• *A valeurs dans* $\{0,1,...,n\}$
• *Loi :* $P(X=k) = \frac{C^k_{N.p}.C^{n-k}_{N.(1-p)}}{C^n_N}$ *si* k *est supérieur ou égal à* 0 *et à* $n - N.(1-p)$ *et inférieur ou égal à* n *et à* $N.p$*.*
• *Espérance :* $n.p$ *(indépendante de* N *!)*
• *Variance :* $\frac{N-n}{N-1}.n.p.(1-p)$
• *Fonction caractéristique : FLEMMARD*
• *Intuition : une urne contient* N *boules, dont une proportion* p *de boules noires. On tire* n *boules ; la loi hypergéométrique* $H(N,n,p)$ *décrit le comportement du nombre de boules noires tirées.*
• *Cas limite : une suite de variables aléatoires suivant une loi hypergéométrique* $H(N,n,p)$ *converge en loi quand* $N \to \infty$ *vers une loi binomiale* $B(n,p)$ *(logique intuitivement !).*

FLEMMARD + illustration en matlab d'une au moins de ces fonctions (toutes les caracs.)

36.9 Loi des grands nombres

FLEMMARD

Exemple Matlab : texte du programme lgn.m

```
function f = lgn(i,j,n,p)

x=rbeta([p,n],i,j);

m=cumsum(x')'.*(ones([1,p])'*(1./(1:n)));
m=[m;((i/(i+j))*ones([1,n]))];
plot(m');
xlabel(sprintf('Convergence de %d moyennes de k vas
beta(%d,%d) vers l''esperance pour k dans
                                    1,%d',p,i,j,n));
```

Le résultat se trouve en figure 36.1.

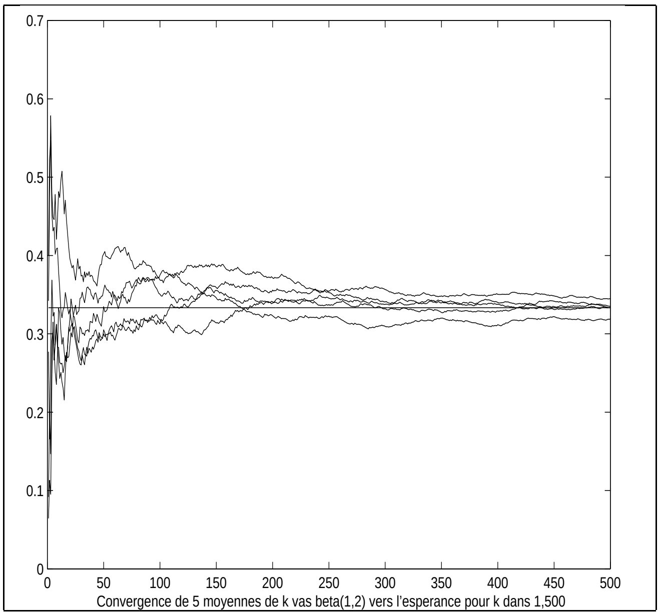


FIG. 36.1 – Loi des grands nombres

36.10 Théorème central limite

FLEMMARD citer comme application les grandes déviations et d'autres trucs. On donne ici deux exemples d'illustration, l'un en matlab, l'autre en Maple.

Exemple Matlab : texte du programme tcl.m

```
function f = tcl(i,j,n,p,nb,step,A)
N=n/step;
x=sum(rbeta([p,N,step],i,j),3);
E=step*i/(i+j);
sigma=(i*j/((i+j+1)*(i+j)^2))*step;
m=(cumsum(x')'.*(ones([1,p])'*(1./(1:N)))-E)/sqrt(sigma);
rf=pnorm(-A:2*A/nb:A,0,1);
index=cumsum(ones([1,nb])/(nb+2));
for k=1:N,
vecteur=(m(:,k))*sqrt(k);
e=quantile(vecteur,index);
plot(-A:2*A/nb:A,rf,e,index);
xlabel(sprintf('Fonction de repartition de la moyenne de
                        %d vas beta(%d,%d)',k*step,i,j));
if (k ~= N)
             sprintf('Appuyez sur une touche pour la
                                              suite'), pause;
             end
end;
```

L'intérêt de cet exemple (notamment par rapport à l'illustration suivante en Maple) est le fait qu'ici on montre la convergence simple de la fonction de répartition, alors que la convergence illustrée en examinant des histogrammes est moins directement liée au théorème central limite. Je n'ai pas représenté la figure obtenue car il s'agit d'une séquence, qu'on ne peut rendre sur papier sans occuper une place abusive...

Exemple Maple

```
> restart;with(stats);
```

$$[anova, describe, fit, importdata, random, statevalf, statplots, transform]$$

```
> unif:= i -> random[uniform[0,1]](i);
```

$$unif := random_{uniform_{0,\,1}}$$

```
> normale:= i -> random[normald[0,1]](i);
```

$$normale := random_{normald_{0,\,1}}$$

```
> fit[leastsquare[[x,y,z],z=d*x*y+e*x^2+f*y^2+a*x+b*y+c,
{a,b,c,d,e,f }]]([[1,1,1,2,2,2,3,3,3],[1,2,3,1,2,3,1,2,3],
[2,4,6,9,12,16,12,13,14]]);
```

$$z = -\frac{1}{2}\,x\,y - \frac{23}{6}\,x^2 + \frac{1}{6}\,y^2 + \frac{125}{6}\,x + \frac{5}{2}\,y - \frac{160}{9}$$

```
> with(statplots);histogram([normale(50)],[normale(200)],
[normale(800)],[normale(6400)],numbars=40,area=1);
```

$$[boxplot, histogram, scatterplot, xscale, xshift, xyexchange, xzexchange, yscale, yshift, yzexchange, zscale, zshift]$$

36.11 Inégalité de Cramer-Chernoff, grandes déviations

On pourra consulter le livre [18], lecture 16, avec profit. Une application possible des grandes déviations est illustrée en 36.12.5.

FLEMMARD

Le programme suivant illustre le résultat prédit par le résultat FLEMMARD ci-

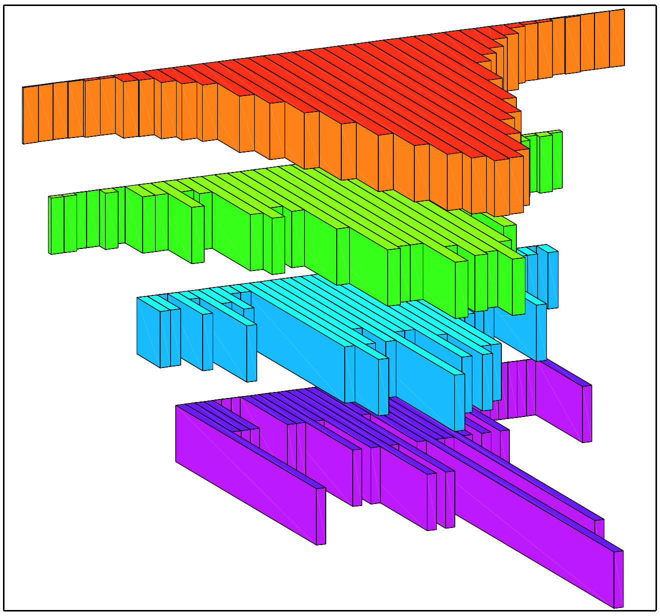


FIG. 36.2 – Théorème central limite : convergence vers la loi normale.

dessus.

Exemple Matlab : texte du programme gd.m

```
function f = gd(i,j,n,p,c,astuce)

x=mean(rbeta([p,n,astuce],i,j),3);
M=(abs(cumsum(x')'.*(ones([1,p])'*(1./(1:n)))-i/(i+j)));
m(1,:)=mean(M>c);
m(2,:)=mean(M>2*c);
m(3,:)=mean(M>3*c);
m(4,:)=mean(M>4*c);
m=(log(m))/astuce;
plot(m');
title(sprintf('1/k log de la proportion des %d moyennes
\ \ \ de k vas beta(%d,%d) a distance > %g x 1:4 de
\ \ \ l''esperance pour k dans 1,%d',p,i,j,c,n*astuce));
text(n/2,m(1,floor(n/2)),sprintf('%g',c));
text(n/2,m(2,floor(n/2)),sprintf('%g',2*c));
text(n/2,m(3,floor(n/2)),sprintf('%g',3*c));
text(n/2,m(4,floor(n/2)),sprintf('%g',4*c));
```

Le résultat est illustré en figure 36.3 ; il faut noter le fait que la courbe est bien linéaire.

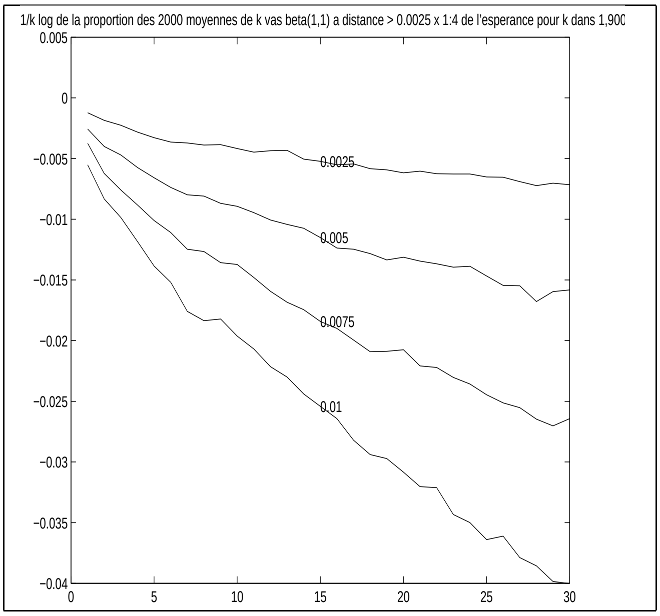


FIG. 36.3 – Grandes déviations

36.12 Applications des probabilités

36.12.1 Liste d'exemples simples

⊡ **Les singes tapant Shakespeare**

On note SK la concaténation de toutes les œuvres de Shakespeare. Un singe est placé devant une machine à écrire ; il tape un caractère par seconde, chaque caractère ayant une probabilité > 0 d'être tapé à la seconde t.
$\rightarrow$ la probabilité pour que le singe tape SK une infinité de fois est 1

Démonstration : Par le deuxième lemme de Borel-Cantelli.□

Si la suite de caractères de SK est désigné par $(u_n)_{n\in[1,N}$, alors la probabilité pour que SK soit tapé au moins k fois en temps P est FLEMMARD.□

Démonstration : FLEMMARD□

36.12.2 Application des probabilités au calcul d'intégrales

On se donne une fonction f L^1 de $[0,1]$ dans $\mathbb{R}$. On va chercher à calculer l'intégrale I de f sur $[0,1]$.

I est l'espérance de f, vue comme variable aléatoire. Donc par l'inégalité de Tchebitchev, on peut écrire que

$$P(|\frac{1}{n}\sum_{i=1}^{n} f(X_i) - I| \leq \epsilon) \leq FLEMMARD$$

avec les X_i des variables aléatoires identiquement distribuées uniformes sur $[0, 1]$.

36.12.3 ABRACADABRA : application des martingales

Le résultat suivant est extrait de [21] (en tant qu'exercice non corrigé) :

Proposition 1364 *Soit T la variable aléatoire égale au temps (en secondes) pendant lequel un singe tape sur une machine à écrire une lettre au hasard par seconde jusqu'à obtenir la suite "ABRACADABRA", chaque lettre ayant une probabilité $1/26$ d'être tirée au sort à chaque seconde.*
Alors $E(T) = 26^{11} + 26^4 + 1$.

Démonstration : On imagine qu'un nouveau joueur intervient à chaque seconde $n \in [0, T]$, et parie sur la prochaine lettre fournie par le singe. Il mise un franc sur A, et gagne 26 francs s'il a raison. A la seconde suivante (alors même qu'un autre joueur arrive), s'il a perdu il arrête de jouer, et s'il a gagné il remise tout ce qu'il a gagné sur B, gagnant 26^2 francs s'il gagne, et perdant 26 francs s'il perd. Il continue ainsi jusqu'à gagner 26^{11} francs avec le mot complet ABRACADABRA ou jusqu'à perdre.

Il est clair que le gain total de l'ensemble des joueurs est une martingale (par le théorème 1352), et que donc par le théorème d'arrêt éventuel de Doob (1354) permet de conclure que l'espérance est nulle. On en déduit le résultat démandé en constatant que les pertes sont égales à $T - 26^{11} - 26^4 - 26$ (examiner les joueurs encore en jeu à l'instant T).$\square$

36.12.4 Application au calcul de la longueur d'une courbe

FLEMMARD

36.12.5 Application à l'évaluation de la perte de précision dans un algorithme

On peut utiliser ici des résultats divers concernant la limite d'une somme de variables aléatoires identiquement distribuées ; grandes déviations, théorème central limite, inégalité de Bienaymé-Tchébytchev 1338 FLEMMARD...

On suppose qu'un programme effectue un calcul et qu'à chaque étape on ajoute un nombre, et que la perte de précision est la valeur absolue de la somme de variables aléatoires indépendantes identiquement distribuées.

A chaque étape, on a un u_n représentant la valeur obtenue après le n-ième calcul, et X_n la différence entre u_{n+1} et $f(u_n)$ la valeur que l'on obtiendrait si l'ordinateur avait une précision infinie à cette étape.

C'est-à-dire que, S_n représentant la somme des X_i pour i variant de 1 à n, les X_i étant des variables aléatoires identiquement distribuées , en supposant que l'espérance de X_i est nulle, on a les résultats suivants :

- $E(S_n) = 0$
- Par l'inégalité de Bienaymé-Tchébitchev, si la variance est $\sigma^2 < \infty$: $P(|S_n| \geq c) \leq \frac{n\sigma^2}{c^2}$
- Par le théorème central limite,

$$P(|S_n| > c\sqrt{n}\sigma) \to_{n\to\infty} 2P(N > c)$$

avec N variable aléatoire de loi normale $N(0,1)$ (ceci permet donc la construction d'intervalles de confiance).

• Par les résultats sur les grandes déviations,

$$P(|S_n| > cn)$$

avec c positif décroit exponentiellement en n.

36.12.6 Application des probabilités à la géométrie euclidienne

Soient $(x_i)_{i\in[1,d]}$ une famille de d vecteurs de $\mathbb{R}^d$.

On se donne $(p_i)_{i\in[1,d]}$ une famille de réels dans $[0,1]$.

Définissons $x = \sum_i p_i.x_i$.

Montrons qu'il existe $(\epsilon_i)_{i\in[1,d]}$ une famille de d éléments de $\{0,1\}$ tels que $\|w - \sum_i \epsilon_i.x_i\| \leq \sqrt{\sum_{i\in[1,d]} \|u_i\|^2}/2$.

On définit $(X_1,...,X_d)$ un vecteur aléatoire de loi

$$P(X = (x_1,...,x_d)) = \Pi_{i\in[1,d]} p_i^{x_i}.(1-p_i)^{1-x_i}$$

pour $(x_1,...,x_d) \in \{0,1\}^d$

(c'est à dire, intuitivement, que la probabilité pour X_i d'être égal à 1 est p_i, et que les composantes sont indépendantes)

et on définit $X = \sum X_i.x_i$; X est un vecteur aléatoire.

Calculons l'espérance de $\|X - w\|^2$.

$$E(\|X-w\|^2)$$

$$= E(\sum_{(i,j)\in[1,d]^2} (X_i - p_i).(X_j - p_j).< u_i, u_j >)$$

$$= \sum_{i=1}^{d} \|u_i\|^2 p_i.(1-p_i)$$

$$= \sum_{i=1}^{d} \|u_i\|^2/4$$

D'où le résultat.□

36.12.7 Probabilité pour que le rapport de #piles/(#piles+#faces) tende vers $\frac{1}{2}$

FLEMMARD

36.12.8 Proportion de diviseurs de n dans $[1,i]$

Ce résultat est extrait de [1].

Définition 1365 *On note $\mu(n)$ le nombre de diviseurs premiers de n.*

Soit u_n une suite tendant vers $+\infty$. On définit E_n^i l'ensemble des $i \in [1, n]$ tels que

$|\mu(i) - log(log(n))|$ est supérieur ou égal à $a(n)\sqrt{ln(ln(n))}$ Alors l'objectif va être de prouver que :

$lim_{n\to\infty} \frac{|E_n^i|}{n} = 0$

L'intérêt sera de prendre alors $a(n) = \sqrt{ln(ln(n))}$, et on arrivera à la conclusion que la proportion d'entiers i tels que $\mu(i)/ln(ln(i))$ soit plus loin de 1 que ϵ tendra vers 0 pour $n \to \infty$.

Montrons donc notre résultat.

• Pour cela on considère les lois de probabilités $(prob_n)$ à valeurs dans $\mathbb{N}^*$ pour $n \in \mathbb{N}$, définie par $prob_n(i) = 1/n$ si $i \leq n$ et 0 sinon.

• On définit ensuite $div(p, n)$, application de $P \times \mathbb{N}^*$ dans $\{0, 1\}$, avec P l'ensemble des nombres premiers, avec $div(p, n) = 1$ si $p|n$ et 0 sinon.

• On constate que $\mu(n) = \sum_{p\in P}(div(p, n))$.

• On constate aussi que $\frac{|E_n^i|}{n} = prob_n(|\mu - ln(ln(n))| \geq a(n).ln(ln(n)))$

• Déterminons maintenant l'espérance de $div(p, .)$, pour la probabilité $prob_n$.

$$E(div(p, .)) = \lfloor n/p \rfloor/n = 1/p + O(1/n)$$

• On détermine maintenant l'espérance de μ ; elle est somme des espérances des $div(p, .)$, or le théorème des nombres premiers FLEMMARD affirme que

$$\sum_{p\in P \wedge p\leq n} 1/p = ln(ln(n)) + o(1)$$

pour $n \to \infty$

Donc toujours pour $prob_n$

$$E(\mu) = \sum_{p\in P\wedge p\leq n} 1/p + O(1/n)$$

$$= ln(ln(n)) + o(1)$$

• Déterminons maintenant la variance de $div(p, .)$.

$$Var(div(p, .)) = E((div(p, .) - E(div(p, .))^2)$$

$$= E((div(p, .) - 1/p + O(1/n))^2$$

$$= \frac{1}{p}.\frac{p-1}{p} + \frac{p-1}{p}\frac{1}{p} + O(1/n)$$

$$= \frac{2.(p-1)}{p^2}$$

• Déterminons maintenant la covariance de $div(p, .)$ et $div(q, .)$

$$Cov(div(p, .), div(q, .)) = E(div(p, .).div(q, .)) - E(div(p, .))E(div(q, .))$$

$$= \frac{\lfloor n/pq \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n}$$

$$\leq \frac{1}{pq} - (p - 1/n).(q - 1/n)$$

$$\leq \frac{1}{n}(1/p + 1/q)$$

• On peut donc maintenant calculer la variance de μ.

$$Var(\mu) = \sum_{p \in P \wedge p \leq n} Var(div(p,.))$$
$$+2 \sum_{(p,q) \in P^2 \wedge p,q \leq n} Cov(div(p,.), div(q,.))$$
$$\leq ln(ln(n)) + O(1)$$

• On peut alors appliquer l'inégalité de Tchebitchev :

$$prob_n((X - E(x))^2 \geq t^2) \leq \frac{Var(\mu)}{t^2}$$

donc

$$prob_n(|X - ln(ln(n))| \geq t\sqrt{ln(ln(n))}) \leq 1/t^2$$

La preuve est ainsi complète !□

36.12.9 Processus de branchement

Je suis ici la démarche du chapitre 0 de [21], qui m'a semblé la plus intuitive ; on trouvera des résultats similaires dans [18], et des prolongements et illustrations dans d'autres livres cités par Williams dans [21] ; Feller, Ross.

On se donne une variable aléatoire X à valeurs dans $\mathbb{N}$. Intuitivement, X correspond au nombre d'enfants d'un animal donné.

On suppose que $P(X = 0) > 0$.

On note f la **fonction génératrice** de X, c'est-à-dire

$$f(\theta) = E(\theta^X) = \sum_{k \geq 0} P(X = k)\theta^k$$

On considère une suite double de variables aléatoires identiquement distribuées $X_{n,m}$ toutes distribuées comme X.

On définit $Z_0 = 1$, et $Z_{n+1} = \sum_{i \in [1,Z_n]} X_{n,m}$.

On note f_n la **fonction génératrice** de Z_n, c'est-à-dire $f_n(\theta) = E(\theta^{Z_n}) = \sum_{k \geq 0} \mathcal{P}(Z_n = k)\theta^k$.

Intuitivement, Z_n est le nombre d'individus d'une espèce à l'instant n, descendant d'un même individu, qui est seul à l'instant $n = 0$ (puisque $Z_0 = 1$).

On définit aussi Π la probabilité d'extinction, définie par $\Pi = P(\exists n/Z_n = 0)$, et Π_n la probabilité d'extinction avant l'instant n, définie par $\Pi_n = P(Z_n = 0)$.

Lemme 1366 *Pour θ dans* $[0, 1[$

$$f_n(theta) = f^n(\theta)$$

c'est-à-dire, par définition de $f^n(\theta)$,

$$= \underbrace{f \circ f \circ \cdots \circ f}_{n \text{ fois}}$$

Démonstration :

Le cas $n = 0$ est clair, $n = 1$ aussi, on procède ensuite par récurrence ; il suffit donc de montrer que $f_{n+1}(\theta) = f_n(f(\theta))$.

Or,

$$f_{n+1}(\theta) = E(\theta^{Z_{n+1}})$$
$$= E(E(\theta^{Z_{n+1}}|Z_n))$$

par les propriétés de l'espérance conditionnelle,

Or

$$E(\theta^{Z_{n+1}}|Z_n = k) = E(\theta^{X_{n,1}+X_{n,2}+\cdots+X_{n,k}})$$
$$= \Pi_{l=1}^{k} E(\theta_{n,l}^{X}) = E(\theta^X)^k$$

vu l'indépendance des $\theta^{X_{n,m}}$

$$= f(\theta)^k$$

$$f^{n+1}(\theta) = E(\theta^{Z_{n+1}}|Z_n) = \sum_k E(\theta^{Z_{n+1}}|Z_n = k)P(Z_n = k) = f_n(f(\theta))$$

D'où le résultat.□

Théorème 1367 *Si $E(X) > 1$ alors la probabilité d'extinction Π est l'unique racine de l'équation $\Pi = f(\Pi)$ située entre 0 et 1 strictement, et sinon, $\Pi = 1$.*

Démonstration :

$\pi = \limsup_n \pi_n$ par le théorème de convergence monotone 355.

Par le lemme 1366, $\pi_n = f(\pi_{n-1})$. Par continuité de f, $\pi = f(\pi)$.

Il ne reste plus qu'à considérer le graphe de f, convexe, croissante, vérifiant $f(0) > 0$, pour conclure...□FLEMMARD faire un dessin

FLEMMARD on peut dire plus de choses sur les processus (i ?) de branchement

36.12.10 Calcul de surface minimale

On se donne un compact K de $\mathbb{R}^2$, et δK son contour. On suppose donnée une fonction g définie sur δK. Soit E l'ensemble des applications de K dans $\mathbb{R}$. Chaque f appartenant à E définit une surface, l'ensemble des $\{(x, y, f(x, y)/(x, y) \in K\}$. On cherche parmi E la fonction définissant la surface minimale. On admet le fait que la fonction vérifiant cette propriété est celle dont le laplacien est nulle, et qu'elle est unique. On va s'intéresser ici à une méthode probabiliste résolvant le problème discrétisé. On pourrait bien sûr s'attaquer à un problème plus général, mais par simplicité de notations on considèrera que $K = [0, n]^2$, et on s'intéressera seulement aux points de coordonnées entières de K. La fonction g peut-être quelconque ; on s'intéressera pour nos représentations schématiques à la fonction définie ci-dessous, front.m :

Exemple Matlab : front.m

```
function f=front(x,y)


a=abs(x-round(x));
b=abs(y-round(y));

if (a<b) f=1; else f=0; end;
```

Pour résoudre le problème, on calcule séparément les valeurs de f en les différents points de coordonnées entières de K. Considérons par exemple $(i,j) \in [0,n]^2$. On considère alors le processus de Markov $(X^{(i,j)})_n$ ayant K pour espace des états, partant de (i,j), et effectuant une marche aléatoire simple sur K (ie les 4 directions sont équiprobables). On définit un temps d'arrêt T égal au nombre d'étapes avant que la marche aléatoire atteigne δK, ie une abscisse ou une ordonnée égale à 0 ou n, ce qui a une probabilité 1 d'arriver. On considère alors $f \in E$ définie par $f(i,j) = E(g((X^{(i,j)})_T))$.

Il est clair que l'application f ainsi définie vérifie bien $\Delta f = 0$. Le programme matlab correspondant est le suivant :

Exemple Matlab : lapla.m

```
function v=lapla(n,e)
u=zeros(n+1,n+1);
for i=0:n,
for j=0:n,
disp(sprintf('%g %%',(i*(n+1)+j)*100/((n+1)*(n+1))))
nb=0;
t=0;
err=[];
while ((2*t > e )|(nb<30))
a=i;
b=j;
while((a<n)&(a>0)&(b>0)&(b<n))
switch(floor(rand*4))
case 0
a=a+1;
case 1
a=a-1;
case 2
b=b+1;
case 3
b=b-1;
end;
end;
a=a/n; b=b/n; nb=nb+1; err=[err,front(a,b)];
if (nb>1) t=std(err)/sqrt(nb-1); end;
end;
u(i+1,j+1)=mean(err);
end;
end;
surfl(u)
shading interp
colormap autumn
v=u;
```

On pourra regretter que les pourcentages affichés pendant le calcul ne sont pas les pourcentages du temps de calcul, mais les pourcentages du nombre de points calculés. La figure obtenue par "lapla(10,0.05)" est 36.5, à gauche. En remplaçant la fonction "front.m" par "cos(4*atan((x-0.5)/(y-0.5)))", on obtient la figure 36.5, à droite.

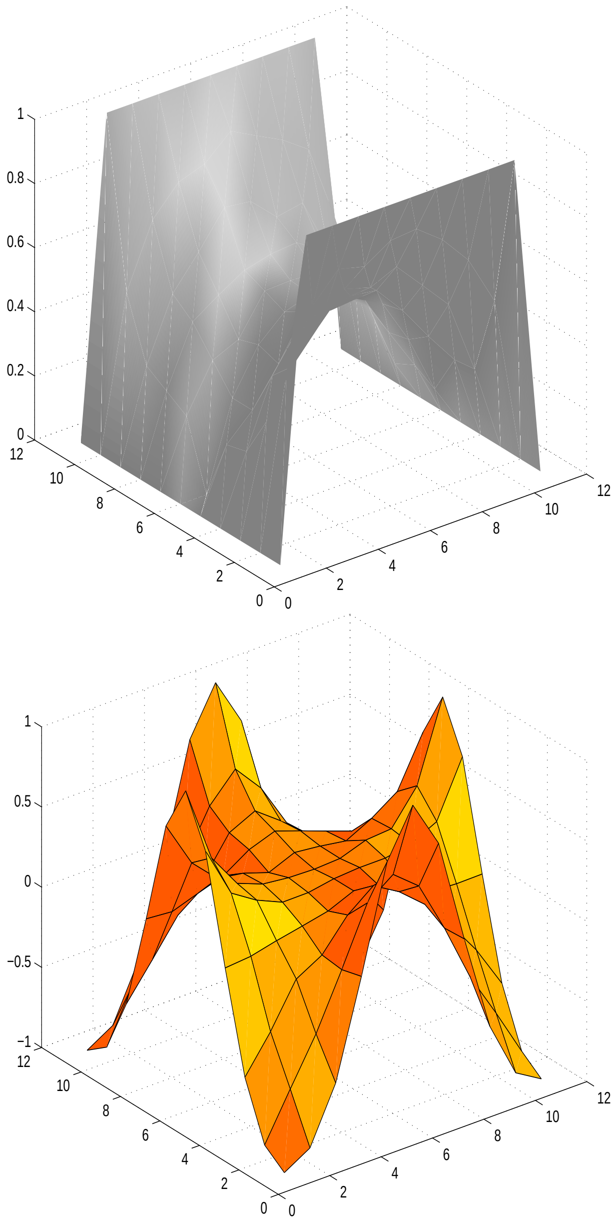


FIG. 36.4 – Approximation de surface minimale obtenue par "lapla.m".

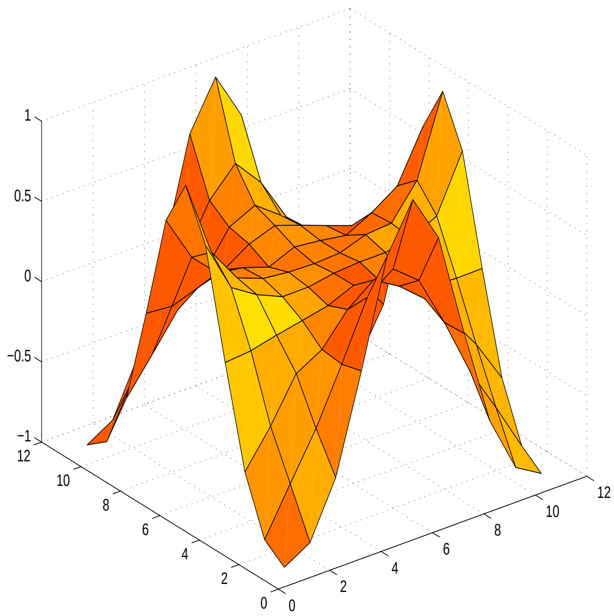


FIG. 36.5 – Approximation de surface minimale obtenue par "lapla.m".

Chapitre 37

Statistique

37.1 Quelques notions élémentaires

37.1.1 Définitions

Définition 1368 *On appelle* **moyenne arithmétique** *de n entiers x_1,...,x_n la quantité $\frac{\sum_{i=1\dots n} x_i}{n}$. On l'appelle aussi* **moyenne** *tout court lorsqu'il n'y a pas de risque de confusion, et on la note $\overline{x}$.*

Définition 1369 *On appelle* **moyenne géométrique** *de n entiers x_1,...,x_n la quantité $\sqrt[n]{\Pi_i x_i}$.*

Définition 1370 *On appelle* **moyenne harmonique** *des x_i l'inverse de la moyenne arithmétique des inverses des x_i.*

Définition 1371 *On appelle* **moyenne quadratique** *des x_i la racine carrée de la moyenne arithmétique des carrés des x_i.*

Définition 1372 *On appelle* **médiane** *d'une mesure finie sur un espace ordonné un élément x tel que la mesure de $\{y/y > x\}$ est égale à la mesure de $\{y/y < x\}$. Cette notion est définie lorsque la mesure est finie.*

Définition 1373 *On appelle* **effectif cumulé croissant** *d'une distribution sur un espace ordonné la fonction qui à x associe la mesure de $\{y/y < x\}$, et* **effectif cumulé décroissant** *la fonction qui à x associe la mesure de $\{y/y > x\}$. Les effectifs cumulés croissants sont aussi appelés* **effectifs cumulés** *tout simplement. Ces notions sont définies lorsque les mesures correspondantes sont bien finies.*

Définition 1374 *On appelle* k**-ième percentile** *d'une distribution sur $\mathbb{R}$ une valeur x telle que les effectifs cumulés en x représentent $k\%$ de la mesure de tout l'espace. Bien évidemment il est nécessaire que la mesure soit finie pour cela. On définit de même des* **quartiles**, *des* **déciles**. *On appelle* **interquartile** *la différence entre le troisième et le premier quartile.*

Définition 1375 *On appelle* **mode** *ou* **dominante** *d'une distribution la valeur x telle que la densité de probabilité en x soit maximale. S'il y a plusieurs modes la distribution est dite* **plurimodale**.

Définition 1376 *On appelle* **déviation** *de x_i la valeur $x_i - \overline{x}$.*

Définition 1377 *On appelle* **écart moyen** *la moyenne des $|x_i - \overline{x}|$; c'est donc $\overline{|x_i - \overline{x}|}$.*

Définition 1378 *On appelle* **variance** *la moyenne des $(x_i - \overline{x})^2$; on la note souvent V.*

Définition 1379 *On appelle* **écart type** *ou* **écart quadratique moyen** *la racine carrée de la variance. On le note souvent σ ; $\sigma = \sqrt{V}$.*

Définition 1380 *On procède à un* **changement d'origine** *lorsque l'on remplace les données x_i par les y_i définis par $y_i = x_i - C$, avec C une constante.*

Définition 1381 *On procède à un* **changement d'échelle** *lorsque l'on remplace les données x_i par les y_i définis par $y_i = C.x_i$, avec C une constante.*

Définition 1382 *On appelle* **moment d'ordre** p *des* x_i *par rapport à* y *la moyenne des* $(x_i - y)^p$. *Pour* $p = 1$ *il s'agit donc de la moyenne (arithmétique), pour* $p = 2$ *il s'agit de la variance.*

37.1.2 Propriétés

• Le logarithme de la moyenne géométrique est la moyenne arithmétique des $log(x_i)$.

• Moyenne harmonique $\leq$ moyenne géométrique $\leq$ moyenne arithmétique $\leq$ moyenne quadratique

• La moyenne arithmétique est peu sensible aux fluctuations d'échantillonnage.

• La médiane est peu sensible aux valeurs aberrantes.

• La somme des déviations est nulle.

• La variance V est aussi égale à $V = \overline{x^2} - \overline{x}^2$, avec $\overline{x^2}$ la moyenne arithmétique des x_i^2, et $\overline{x}^2$ le carré de la moyenne des x_i. On le prouve facilement en développant $\sum(x_i - \overline{x})^2$.

• Multiplier les données par C multiplie la moyenne arithmétique par C, la variance par C^2, et l'écart-type par C.

• Translater les données de C ajoute C à la moyenne arithmétique, et ne change ni la variance ni l'écart-type.

37.2 Applications des probabilités à l'échantillonnage

Cette partie ne se veut qu'une très brève introduction aux statistiques. Il est bien évident que dans le cadre de l'option probabilités de l'agrégation, il est indispensable de se référer à un livre plus complet. Les livres bien faits d'introduction abondent FLEMMARD.

Nous nous contenterons ici de donner un cas d'applications, et de citer d'autres cadres d'applications.

Soit X_1, ... , X_n indépendantes identiquement distribués. On suppose le théorème central limite **??** vérifié. Intuitivement, les X_i sont des mesures ; par exemple, on mesure la taille de 50 français pour évaluer la taille moyenne des français. L'intérêt des probabilités va être de fournir des bornes sur l'erreur commise par une telle évaluation.

On se donne donc $m = \frac{1}{n}(X_1 + X_2 + ... + X_m)$. On cherche $[a, b]$ tel que $M = E(X)$ soit compris dans $[a, b]$. Il faut alors noter que bien entendu, on ne peut être certain que M soit dans l'intervalle $[a, b]$, quel que soit l'intervalle que l'on donne, simplement au vu des X_i. Il est toujours possible que l'on ait été particulièrement malchanceux dans les tirages des X_i et que la moyenne soit très différente de ce que l'on suppose au vu du tirage. On doit donc plutôt donner α un réel (petit de préférence) et z tel que avec probabilité $1-\alpha$, pour toute loi de X_1, $|m-M| \leq z$ soit vrai. a et b seront alors $m-z$ et $m+z$ respectivement.

Concrètement on procède comme suit :

• On évalue (empiriquement) l'écart type σ de X_i.

• On repère t_α tel que $P(|N| \leq t_\alpha) = 1-\alpha$, avec N loi normale centrée réduite (espérance nulle et écart-type 1). Les valeurs de t_α sont tabulées (il s'agit simplement de la fonction de répartition de la loi normale). Le plus courant est de choisir $\alpha = 0.05$, t_α étant alors environ égal à 2.

• On détermine $a = m - t_a\sigma/\sqrt{n}$ et $b = m + t_\alpha\sigma/\sqrt{n}$.

• On peut alors écrire que, au **seuil de confiance** α, M est compris entre a et b. Ceci constitue un **intervalle de confiance**.

On peut citer les développements suivants :

– le cas des petits échantillons ($n < 30$). Il faut alors utiliser la loi de Student.
– le cas où l'on ne s'intéresse pas à la probabilité pour que la moyenne soit mal évaluée, mais à la probabilité pour que la moyenne soit sur-évaluée. Il suffit de voir pour cela que $P(N > t) = \frac{1}{2}P(|N| > t)$ pour toute variable aléatoire N symétrique, et en particulier donc la loi normale.
– le cas de X_i à valeur dans $\{0,1\}$.
– le cas où l'on n'étudie pas la moyenne des X_i mais leur $\max$.
– le cas de X_i non réellement indépendants.
– le cas de X_i non identiquement distribués.
– l'évaluation de la loi de m par la fameuse technique du bootstrap, très ingénieuse.
– on a fait l'approximation que pour $n > 30$, la distribution était environ celle de la loi normale. On peut se passer de cette approximation, et donner une borne absolue à l'écart à la loi normale.

Ces études et d'autres encore constituent la théorie des tests et font appel à des variantes difficiles du théorème central limite. Les trois premiers points sont utiles pour faire bonne impression à un jury d'agrégation...

Chapitre 38

Formulaires

38.1 Espaces topologiques

Catégorie	Exemples
Ouvert	$Isom(E,F)$ avec E et F Banach, dans $\mathcal{L}(E,F)$
Fermé	compact d'un séparé
Borné	compact d'un métrique
Espaces compacts	Espaces projectifs Complété d'Alexandrov d'un espace séparé non compact, localement compact Produit de compacts boule fermée d'un espace vectoriel normé de dim finie Cantor Tapis de Sierpinski
Espaces connexes par arc	Espaces projectifs Cube de Hilbert Ouvert connexe d'un espace vectoriel normé Boule Image d'un connexe par arcs par une application C^0 Tapis de Sierpinski, eponge de Menger Ensembles de Julia $SO(n)$, $SU(n)$
Espaces connexes	Produit de connexes Image d'un connexe par une application C^0
Espaces de Montel. (métrisables, compact équivaut à fermé borné[a]) [a] Borné au sens des espace vectoriel topologiques.	$\mathbb{R}^n$, $\mathbb{C}^n$ $H(\Omega)$ avec Ω ouvert de $\mathbb{C}$
Espace complet	$\mathbb{R}^n$, $\mathbb{C}^n$, $L^p(\Omega)$
Espace métrisable	Boule unité fermée du dual d'un espace séparable pour la topologie faible Cube de Hilbert
Localement connexe par arcs	
Localement compact	

38.2 Equivalents en l'infini

38.2.1 Séries

u_n et v_n suites réelles, $U_n = \sum_{k=0}^{n} u_k$, $V_n = \sum_{k=0}^{n} v_k$. Si $U_n \to U$, $R_n = U - U_n$, si $V_n \to V$, $R'_n = V - V_n$.

Hypothèse	Conclusion
$u_n > 0$ $u_n \simeq v_n$ U_n converge	V_n converge $R_n \simeq R'_n$
$v_n > 0, u_n > 0$ $v_n = o(u_n)$ V_n diverge	U_n diverge $U_n \simeq V_n$

38.2.2 Intégrales

f et g définies sur $]a, +\infty[$ intégrables sur $]a, x[$ pour tout $x > a$.

ON SUPPOSE QUE f EST POSITIVE

On définit $F(x) = \int_a^x f(t)dt$ et $G(x) = \int_a^x g(t)dt$. Si F a une limite en $+\infty$ on définit $R_f(x) = \int_x^\infty f(t)dt$, et si G a une limite en $+\infty$ on définit $R_g(x) = \int_x^\infty f(t)dt$.

Alors on a les résultats suivant au voisinage de $+\infty$:

$$\left\{\begin{array}{ll}
\int_a^x f(t)dt = +\infty & \left\{\begin{array}{c} g = o(f) \Rightarrow G = o(F) \\ g = O(f) \Rightarrow G = O(F) \\ g \simeq f \Rightarrow F \simeq G \end{array}\right. \\
\int_a^x f(t)dt < +\infty & \left\{\begin{array}{ll} g = o(f) \Rightarrow & \left\{\begin{array}{c} \int_a^\infty g(t)dt \text{ existe} \\ R_g = o(R_f) \end{array}\right. \\ g = O(f) \Rightarrow & \left\{\begin{array}{c} \int_a^\infty g(t)dt \text{ existe} \\ R_g = O(R_f) \end{array}\right. \\ g \simeq f \Rightarrow & \left\{\begin{array}{c} \int_a^\infty g(t)dt \text{ existe} \\ R_g \simeq R_f) \end{array}\right. \end{array}\right.
\end{array}\right.$$

38.3 Dérivation de limites

E et F des espaces vectoriels normés , U ouvert de E, (f_n) suite d'applications de U dans F différentiables.

Hypothèses	Conclusions
f_n convergeant simplement vers f, les Df_n convergeant <u>uniformément</u> vers une certaine application g de U dans $\mathcal{L}(E,F)$,	• f différentiable et $Df = g$ • Pour tout C convexe borné $\subset U$ la convergence de $f_{n\|C}$ vers $f_{\|C}$ est uniforme • Si les f_n sont C^1 alors f est C^1.
U connexe, F Banach. $\exists x_0$ / $f_n(x_0)$ converge, $\forall x\ \exists V_x$ voisinage de x tel que la suite des $Df_{n\|V_x}$ soit de Cauchy pour la métrique d définie par $d(f,g) = sup_{z \in V_x} \|f(z) - g(z)\|$ (ie la suite des Df_n converge normalement sur un certain voisinage de tout point)	Alors il existe f de U dans F tel que : • f est dérivable en tout point • la suite des f_n converge vers f (simplement) • tout x possède un voisinage V_x tel que les convergences de f_n et Df_n restreints à V_x soient uniformes. • Si les f_n sont C^1, f l'est aussi.

38.4 L'indispensable sous le signe intégral

X muni d'une mesure μ, (E,d) un espace métrique. $f : E \times X \to \mathbb{C}$, $F(t) = \int_X f(t,x)d\mu(x)$

Hypothèses	Conclusion
Pour tout t l'application $x \mapsto f(t,x)$ est mesurable Pour presque tout x la fonction $t \mapsto f(t,x)$ est continue en T Il existe g L^1 telle que pour tout t et presque tout x $\|f(t,x)\| \leq g(x)$	F est continue en T.
Pour tout t l'application $x \mapsto f(t,x)$ est mesurable Pour presque tout x la fonction $t \mapsto f(t,x)$ est continue sur E Pour tout compact K de E il existe g L^1 telle que pour tout t dans K et presque tout x $\|f(t,x)\| \leq g(x)$.	F est continue sur E.
E est un ouvert de $\mathbb{R}$ ou de $\mathbb{C}$ [a] Pour presque tout t $x \mapsto f(t,x)$ est L^1 Il existe N négligeable tel que pour tout $x \notin N$ la fonction $t \mapsto f(t,x)$ est dérivable (resp. C^1) [b]. Pour tout compact K de E il existe une fonction g L^1 telle que pour tout t dans K et tout $x \notin N$ $\|\frac{\delta f}{\delta t}(t,x)\| \leq g(x)$. [a] Hypothèse facile à retenir ; il s'agit de pouvoir définir une dérivée au sens le plus commun, ie dérivée d'une fonction d'une variable réelle ou complexe ! [b] Attention ! Dans le cas d'un ouvert de $\mathbb{C}$ on parle de dérivabilité au sens complexe, et pas de différentiabilité en voyant $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel !	Pour tout t la fonction $x \mapsto \frac{\delta f}{\delta x}(t,x)$ est L^1 F est dérivable (resp. C^1), de dérivée $\int_X \frac{\delta f}{\delta t}(t,x)dx$.
E est un ouvert de $\mathbb{R}$ ou un ouvert de $\mathbb{C}$. Pour presque tout t $x \mapsto f(t,x)$ est L^1 Il existe N négligeable tel que pour tout $x \notin N$ la fonction $t \mapsto f(t,x)$ est C^k. Pour tout compact K de E et tout $j \in [1,k]$ il existe une fonction g L^1 telle que pour tout t dans K et tout $x \notin N$ $\|\frac{\delta^j f}{\delta t^j}(t,x)\| \leq g(x)$.	Pour tout t la fonction $x \mapsto \frac{\delta^j f}{\delta t^j}(t,x)$ est L^1 F est C^k, et pour $j \in [0,k]$ $\frac{\delta^j F}{\delta t^j} = \int_X \frac{\delta^j f}{\delta t^j}(t,x)dx$.

38.5 Convergence d'une série à termes positifs

Hypothèse	Conclusion
Critère de D'Alembert	
$u_{n+1}/u_n \to L < 1$	Série converge
$u_{n+1}/u_n \to L > 1$	Série diverge
Critère de Cauchy	
$limsup\sqrt{u_n} = L < 1$	Série converge
$limsup\sqrt{u_n} = L > 1$	Série diverge
Critère de Raabe-Duhamel	
$\frac{u_{n+1}}{u_n} = 1 - a/n + O(1/n^2)$, $a > 1$	Série converge
$\frac{u_{n+1}}{u_n} = 1 - a/n + O(1/n^2)$, $a < 1$	Série diverge

38.6 Convergence d'une série semi-alternée

$U_N = \sum_{n=0}^{N} u_n$, $U = \lim U_n$, $R_n = U - U_n$

Hypothèse	Conclusion
Critère de Leibnitz	
(ϵ_n) suite décroissante, $\epsilon_n \to 0$, $u_n = (-1)^n \epsilon_n$	$\sum u_n$ converge $R_n \leq \lvert a_{n+1} \rvert$
Méthode d'Abel, dans un Banach	
(ϵ_n) suite décroissante, (a_n) / $\sum_{n=0}^{N} a_n$ borné $\epsilon_n \to 0$, $u_n = \epsilon_n a_n$	$\sum u_n$ converge $R_n \leq \lvert a_{n+1} \rvert$

38.7 Les séries entières

Soit $\sum a_n z^n$ une série entière, R son rayon, A sa fonction somme.

Hypothèse	Conclusion
Formule d'Hadamard	
	$R = \frac{1}{limsup \lvert a_n \rvert^{1/n}}$ (avec $1/0 = \infty$ et $1/\infty = 0$)
Critère de D'Alembert	
a_n non nul à partir d'un certain rang a_{n+1}/a_n tend vers L	$R = 1/L$
Théorème de Césaro	
(b_n) suite réelle > 0, (a_n) suite réelle, $\sum b_n$ divergente, $a_n = o(b_n)$ ou $a_n \simeq b_n$, R rayon de convergence de $\sum b_n z^n$	$A(x) = \sum a_n x^n$ sur $]-1, 1[$ bien défini, $B(x) = \sum b_n x^n$ sur $]-1, 1[$ bien défini, $a = o(b) \Rightarrow A = o(B)$ en R, $a \simeq b \Rightarrow A \simeq B$ en R

38.8 Densité, approximation

Théorème de Stone K compact, A sous-algèbre unitaire de l'algèbre $C^0(K, \mathbb{R})$. Si A sépare les points, alors A est dense dans $C^0(K, \mathbb{R})$ pour la norme infinie.
Théorème de Stone, version complexe K compact, A sous-algèbre unitaire de l'algèbre $C^0(K, \mathbb{C})$ stable par passage au conjugué. Si A sépare les points, alors A est dense dans $C^0(K, \mathbb{C})$ pour la norme infinie.
Théorème de Weierstrass K compact de $\mathbb{R}$, l'ensemble des polynômes de K dans $\mathbb{R}$ est dense dans l'ensemble des fonctions continues de K dans $\mathbb{R}$ pour la convergence uniforme. pas valable pour les polynômes d'un compact de $\mathbb{C}$ dans $\mathbb{C}$!
Théorème de Lusin f mesurable de $\mathbb{R}^n$ dans $\mathbb{C}$, dont le support est inclus dans un ensemble de mesure finie. Alors pour tout ϵ il existe g continue sur $\mathbb{R}^n$ égale à f sauf sur un ensemble de mesure $< \epsilon$, bornée en module par $\sup \lvert f \rvert$.
f mesurable bornée de $\mathbb{R}^n$ dans $\mathbb{C}$, dont le support est inclus dans un ensemble de mesure finie. Alors f est limite simple presque partout d'une suite de fonctions continues et bornées (par la même borne).
$C_c^k(\mathbb{R}^n, \mathbb{R})$ [a] est dense dans $C^k(\mathbb{R}^n, \mathbb{R})$. [a] Ensemble des fonctions C^k à support compact de $\mathbb{R}^n$ dans $\mathbb{R}$.
Si $p \neq \infty$, $C_c^\infty(\mathbb{R}^n, \mathbb{R})$ est dense dans $L^p(\mathbb{R}^n)$.
Si $p \neq \infty$, les classes des fonctions en escalier à support compact sont denses dans $L^p(\mathbb{R}^n)$.
Les classes des fonctions C^∞ à support compact sont denses dans l'ensemble des fonctions de L^∞ de $\mathbb{R}$ dans $\mathbb{R}$ tendant vers 0 en $+\infty$ et en $-\infty$.

38.9 Trigonométrie

Formule	Preuve
$cos(x) = Re(e^{ix}) = \frac{e^{ix}+e^{-ix}}{2}$ $sin(x) = Im(e^{ix}) = \frac{e^{ix}-e^{-ix}}{2}$	Définition.
$cos' = -sin$ $sin' = cos$	$(x \mapsto e^{ix})' = (x \mapsto ie^{ix})$
$cos^2 + sin^2 = 1$	$e^{ix}e^{-ix} = 1$
$\exists \pi > 0$minimal/ $cos(\pi/2) = 0$	$cos(0) = 1,$ donc si $cos(\mathbb{R}^+) > 0$ $sin \uparrow, sin(x) > sin(a),$ donc $cos(x) + xsin(a)$ décroit, or $cos(x) + xsin(a) \to \infty$ Par déf,$\pi = 2infcos^{-1}(0) \cap \mathbb{R}^{+*}$
$\pi > 0$	continuité de cos
$e^{i\pi/2} = i, e^{i\Pi} = -1,$	corollaire de ci-dessus.
$e^{i\pi} = 1,$	
$e\ 2i\pi -$ pério	
cos et sin 2π pério	
$Arccos(x) + Arcsin(x) = \Pi/2$	Dérivée nulle + valeur en 0
$FormuledeMoivre$	
$(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$	$e^{inx} = e^{ix^n}$
$Formulesdelinarisation$	
$cos(a)cos(b) = \frac{1}{2}(cos(a+b) + cos(a-b))$	$cos(x) = \frac{1}{2}(e^{ix} + e^{-ix})$(développer)
$sin(a)sin(b) = \frac{1}{2}(cos(a-b) - cos(a+b))$ $sin(a)cos(b) = \frac{1}{2}(sin(a+b) + sin(a-b))$	$sin(x) = cos(x + \pi/2)$ (apprendre la méthode, non le résultat)
$\cos(a+b) = \cos(a)\cos(b) - \sin(b)\sin(a)$	combinaison linéaire des formules de linéarisation
$\sin(a+b) = \sin(a)\cos(b) + \sin(b)\cos(a)$ $\sin(a-b) = \sin(a)\cos(b) - \sin(b)\cos(a)$ $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$	$sin(x) = cos(x + \pi/2)$ b devient $-b$
$\cos(c) + \cos(d) = 2\cos(\frac{c+d}{2})\cos(\frac{c-d}{2})$ $\cos(c) - \cos(d) = -2\sin(\frac{c+d}{2})\sin(\frac{c-d}{2})$ $\sin(c) + \sin(d) = 2\sin(\frac{c+d}{2})\cos(\frac{c-d}{2})$ $\sin(c) - \sin(d) = 2\sin(\frac{c+d}{2})\cos(\frac{c-d}{2})$	Formules ci-dessus, plus $c = a + b, d = a - b$ puis $\sin(x) = \cos(x + \pi/2)$

Retenir $\cos(a)\cos(b)$, $\cos(a+b)$, $\cos(a)+\cos(b)$ devrait suffire, modulo l'entrainement pour retrouver le reste.

38.10 Trigonométrie hyperbolique

$cosh(x) = \frac{e^x+e^{-x}}{2}$
$sinh(x) = \frac{e^{-x}-e^{-x}}{2}$
$tanh(x) = \frac{sinh(x)}{cosh(x)}$
$cotanh(x) = \frac{cosh(x)}{sinh(x)} = \frac{1}{tanh(x)}$
$cosh(x+y) = cosh(x)cosh(y) + sh(x)sh(y)$
$sinh(x+y) = sinh(x)sinh(y) + cosh(x)sinh(y)$
$cosh(x)^2 - sinh(x)^2 = 1$
$cosh(2x) = cosh(x)^2 + sinh(x)^2$
$cosh(2x) = 2cosh(x)^2 - 1$
$sinh(2x) = 2sinh(x)cosh(x)$
$cosh(x) = \frac{1+u^2}{1-u^2}$ avec $u = tanh(x/2)$
$sinh(x) = \frac{2u}{1-u^2}$

38.11 Les changements de variable magiques dans le calcul de primitive

$f(x) = cos(x)^n sin(x)^m$	$2\|n+1 \Rightarrow u = sin(x)$ $2\|m+1 \Rightarrow u = cos(x)$ Linéarisation.
$f(x) = F(cos(x), sin(x))$ $F \in \mathbb{R}(X)$	f paire $\Rightarrow u = sin(x)$ f impaire $\Rightarrow u = cos(x)$ $f\ T-$ périodique $\Rightarrow u = tan(\Pi x/T)$ sinon $u = tan(x/2),\ cos(x) = \frac{1-u^2}{1+u^2}$ $sin(x) = \frac{2u}{1+u^2},\ dx = \frac{du}{1-u^2}$
$F(x, \sqrt{ax+b})$	$u = \sqrt{ax+b}$
$F(x, \sqrt{ax^2+bx+c})$ $F(x, \sqrt{ax^2+bx+c})$	$a > 0 \Rightarrow u = \sqrt{ax^2+bx+c} - ax$ $a < 0 \Rightarrow (x, \sqrt{ax^2+bx+c})$ parcourt un bon d'ellipse on paramètre pour que x soit un cos de u et $\sqrt{ax^2+bx+c}$ un sin de u
$F(x, \sqrt[n]{\frac{ax+b}{cx+d}})$	$u = \sqrt[n]{\frac{ax+b}{cx+d}}$

38.12 Primitives usuelles

Sauf mention contraire, les primitives sont valables sur l'ensemble de définition.

$f(x)$	$\int^x f(u)du$	Précisions
x^a	$\frac{x^{a+1}}{a+1}$	$a \in]-\infty,-1[\cup]1,\infty[$
$1/x$	$ln\|x\|$	
e^{ax}	$\frac{e^{ax}}{a}$	$a \in \mathbb{C}^*$
cos	sin	
sin	$-cos$	
tan	$-ln\|cos(x)\|$	
$cotan(Remarque$:cotan(x)=1/tan(x).)	$ln\|sin(x)\|$	
$1/cos$	$ln\|tan(x/2+\pi/4)\|$	
$1/sin$	$ln\|tan(x/2)\|$	
$1/cos^2$	tan	
$1/sin^2$	$-1/tan$	
$1/(sin(x)cos(x))$	$ln\|tan(x)\|$	
tan^2	$tanx-x$	
ch	sh	
sh	ch	
$th(x)$	$ln\ ch(x)$	
$coth(x)$	$ln\|sh(x)\|$	
$1/sh(x)$	$ln\|th(x/2)\|$	
$1/ch(x)$	$2Arctan(e^x)$	
$th(x)^2$	$x-th(x)$	
$1/(sh(x)ch(x))$	$ln\|th(x)\|$	
$1/ch^2$	th	
$1/sh^2$	$-coth$	
$1/(x^2+a^2)$	$\frac{1}{a}arctan(x/a)$	$a\neq 0$
$1/(x^2-a^2)$	$\frac{1}{2a}ln\|\frac{x+a}{x-a}\|$	$a\neq 0$
$1/(a^2-x^2)$	$\frac{1}{a}argth(x/a)$	$a\neq 0, x<a$
$1/\sqrt{x^2+a}$	$ln\|x+\sqrt{x^2+a}\|$	
	ou $argsh(x/\sqrt{a})$	$a>0$
	ou $argch(-x/\sqrt{-a})$ si $x>\sqrt{-a}$	$a<0$
	ou $-argch(x/\sqrt{-a})$ si $x<\sqrt{-a}$	$a<0$
$1/\sqrt{a^2-x^2}$	$arcsin(x/\|a\|)$	$a\neq 0$
$\frac{1}{(x^2+a)^{3/2}}$	$\frac{x}{a\sqrt{x^2+a}}$	$a\neq 0$
$\frac{1}{(a-x^2)^{3/2}}$	$\frac{x}{a\sqrt{a-x^2}}$	$a\neq 0$
$I_n=1/(1+x^2)^n$	$2nI_{n+1}=\frac{x}{(1+x^2)^n}+(2n-1)I_n$	
$I_n=1/(1-x^2)^n$	$2nI_{n+1}=\frac{x}{(1-x^2)^n}+(2n-1)I_n$	

Pour une primitive de fraction rationnelle réelle décomposée en éléments simples, on a besoin d'intégrer des termes en (i) $a/(cx+d)^n$ (facile), des éléments en (ii) $\frac{b}{(x^2+dx+e)^n}$ et des éléments en (iii) $\frac{x+b}{(x^2+dx+e)^n}$. Pour (i), c'est facile. Pour (ii), il suffit de penser à reformuler x^2+dx+e en $(x+d/2)^2+f$, et d'appliquer le formulaire ci-dessus avec un changement de variable $u=\frac{x+d/2}{\sqrt{f}}$[1] et par l'une des formules du tableau ci-dessus. Pour (iii), il suffit de réécrire $x+b$ en $\frac{1}{2}(x+b+(p-b)\ -\ (p-b))$, chacun des deux quotients obtenus s'intégrant sans peine par (ii) ou parce que de la forme u'/u^n (cf $\log$ ou u^{-n+1}).

[1] f étant positif du fait que le discréminant du polynôme est négatif.

38.13 Dérivées

La plupart des dérivées s'obtiennent en utilisant le fait que la dérivée de $f \circ f^{-1}$ est 1.

$f(x)$	$f'(x)$
$ln(ax)$	$frac1t$
e^{ax}	$a\ e^{ax}$
x^a	$a\ x^{a-1}$
$cos(ax)$	$-a\ sin(ax)$
$sin(ax)$	$a\ cos(ax)$
$tan(ax)$	$\frac{a}{cos(ax)^2}$
$Arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$Arccos(x)$	$\frac{-1}{\sqrt{1-x^2}}$
$Arctan(x)$	$\frac{1}{1+x^2}$
$cosh(x)$	$sinh(x)$
$sinh(x)$	$cosh(x)$
$tanh(x)$	$\frac{1}{tanh(x)^2}$
$cotanh(x)$	$\frac{-1}{sinh(x)^2}$
$argsinh(x)$	$\frac{1}{\sqrt{1+x^2}}$
$argcosh(x)$	$\frac{1}{\sqrt{x^2-1}}$
$argtanh(x)$	$\frac{1}{1-x^2}$

FLEMMARD

38.14 Différentielles

Application	Différentielle
E, F Banach $Isom(E,F) \to Isom(F,E)$ $\phi : f \mapsto f^{-1}$	$\phi C^\infty\ \phi'(u)(h) = -u^{-1}hu^{-1}$ $\phi^{(n)}(u)(h_1, \ldots, h_n)$ $= (-1)^n \sum_{\sigma \in \sigma_n} u^{-1} \circ h_{\sigma(1)} \circ u^{-1} \circ h_{\sigma(2)} \circ \ldots h_{\sigma(n)} \circ u^{(-1)}$

38.15 développements limités

Le développement limité de e, de sin et de cos découlent directement de leur définition. Pareil pour $cosh$, $sinh$.

La plupart des autres développements limités suivant s'obtiennent grâce à la formule de Taylor-Young 583. Toutefois, le développement limité de $Arctan$ s'obtient par intégration, ainsi que ceux de $Arcsin$, $Arctanh$, $Argsinh$. Le développement limité de $Arccos$ s'obtient à partir de celui de $Arcsin$ par $Arccos(x) + Arcsin(x) = \Pi/2$.

$e^x = 1 + \frac{x}{1!} + \frac{x}{2!} + \frac{x}{3!} + ... + \frac{x^n}{n!} + O(x^{n+1})$
$cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + ... + \frac{x^{2n}}{(2n)!} + O(x^{2n+2})$
$sinh(x) = \frac{x}{1!} + \frac{x^3}{3!} + \frac{x^5}{5!} + ... \frac{x^{2n+1}}{(2n+1)!} + O(x^{2n+3})$
$sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + ... + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + O(x^{2n+3})$
$cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - ... + (-1)^n \frac{x^{2n}}{2n!} + O(x^{2n+2})$
$ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ... + (-1)^{n+1} x^n/n + O(x^{n+1})$
$ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - ... - \frac{x^n}{n} + O(x^{n+1})$
$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + ... + \frac{\alpha(\alpha-1)(\alpha-2)...(\alpha-n+1)}{n!} x^n + O(x^{n+1})$
$Arctan(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \frac{1}{9}x^9 + ... + \frac{1}{2n+1}x^{2n+1} + O(x^{2n+3})$
$Arcsin(x) = x + \frac{x^3}{2.3} + \frac{1.3.x^5}{2.4.5} + \frac{1.3.5.x^7}{2.4.6.7} + \frac{1.3.5.7.x^9}{2.4.6.8.9} + ... + \frac{1.3.5.....(2n-1)x^{2n+1}}{2.4.6.....2n.(2n+1)} + O(x^{2n+3})$
$Arccos(x) = \Pi/2 - Arcsin(x)$
$Argth(x) = x + x^3/3 + x^5/5 + x^7/7 + ... + x^{2n+1}/(2n+1) + O(x^{2n+3})$
$Argsh(x) = x - \frac{x^3}{2.3} + \frac{1.3.x^5}{2.4.5} + ... + (-1)^n \frac{1.3.5.7.....(2n-1).x^{2n+1}}{2.4.6.....2n.(2n+1)} + O(x^{2n+3})$

Attention, le développement limité d'$Arcsin(x)$ est "irrégulier", au sens où le terme en x ne suit pas la même formule que les termes en x^{2i+1} !

38.16 Espaces $\mathcal{L}^p(X)$ et $L^p(X)$, ou $\mathcal{L}^p_{\mathbb{C}}(X)$ et $L^p_{\mathbb{C}}(X)$

L^p est égal à $\mathcal{L}^p$ quotienté par la relation "être égales presque partout". Les L^p sont des espaces de Banach, $L^2(X)$ est un espace hilbertien réel pour $(f|g) = \int fg$, $L^2_{\mathbb{C}}(X)$ est un espace hilbertien complexe pour $(f|g) = \int \overline{f}g$.

Hypothèse	Conclusion
$f_i \in L^{p_i}$ $p_i \in [1, \infty]$ $\sum_{i \in [1,n]} 1/p_i = 1$	$\|\Pi_{i=1}^n f_i\|_1 \leq \Pi_i \|f_i\|_{p_i}$ (inégalité de Hölder)
$\forall n |f_n(x)| < |g(x)|$ presque partout $g \in L^p$ $f_n(x) \to f(x)$ presque partout	$f \in L^p$ $f_n \to f$ dans L^p (convergence dominée de Lebesgue)
$p < p'$ et X de mesure finie	$L^{p'} \subset L^p$

38.17 Transformée de Fourier

hypothèse	conclusion
$g(x) = f(x)e^{i\alpha x}$	$\hat{g}(t) = \hat{f}(t-\alpha)$
$g(x) = f(x-\alpha)$	$\hat{g}(t) = \hat{f}(t)e^{-i\alpha t}$
$g(x) = -ixf(x)$ $g\ L^1$	$\hat{f}$ différentiable $\hat{f}'(t) = \hat{g}(t)$
$f\ L^1$	$\hat{f}(t) = \int_{\mathbb{R}} f(x)e^{-ixt}d\mu(x)$ $\hat{f}\ C^0$ $\|f\|_\infty \leq \|f\|_1$
$f\ L^1$ et $\hat{f}\ L^1$	$x \mapsto \int_{\mathbb{R}} \hat{f}(t)e^{ixt}d\mu(t)$ est continue et égale à f presque partout
$f\ L^2$ Théorème de Plancherel	$\hat{f}\ L^2$: $\phi_A(f) = \int_{-A}^{A} \hat{f}(t)e^{ixt}d\mu(t)$ $\lim_{A\to+\infty} \phi_A(f) = f$ (dans L^2) ($f \mapsto \hat{f}$ isomorphisme de L^2 dans L^2, défini sur $L^1 \cap L^2$ par la formule ci-dessus et prolongé sur L^2 par densité FLEMMARD)

38.18 Série de Fourier - cas f 2π-périodique

hypothèse	conclusion
$n \in \mathbb{Z}$	$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}$ $\hat{f}(n) =< f\|u_n >$ ($u_n(t) = e^{int}$)
$n \in \mathbb{N}$ $n \in \mathbb{N}$	$D_n = \sum_{i=-n}^{n} u_i$ (noyau de Dirichlet) $K_n = \frac{1}{n}\sum_{i=0}^{n-1} D_i$ (noyau de Féjer)
$n \in \mathbb{N}$	$s_n(f) = \sum_{i=-n}^{n} \hat{f}(i)u_i$ (somme de Fourier) $s_n(f) = f * D_n$ $\sigma_n(f) = \sum_{i=0}^{n} s_i(f)$ (somme de Fejer) $\sigma_n(f) = f * K_n$
$f\ L^1$ f admet une dérivée à droite et à gauche en x	Théorème de Dirichlet $\sigma_n(f)(x)$ $\to \frac{1}{2}(lim_{t\to x, t<x} f(t) + lim_{t\to x, t>x} f(t))$
$f \in L^2(T)$	$\hat{f} \in l^2 = L^2(\mathbb{Z})$ ($f \mapsto \hat{f}$ isomorphisme de $L^2(T)$ dans l^2) $s_n(f) \to f$ dans $L^2(T)$
$f\ C^0$ (Th. de Fejer)	$\forall n\ \|\sigma_n(f)\|_\infty < \|f\|_\infty$ $\sigma_n(f) \to f$ uniformément
$f\ L^p$ $p < \infty$	$\forall n\ \|\sigma_n(f)\|_p < \|f\|_p$ $\sigma_n(f) \to f$ dans L^p (Th. de Fejer)

38.19 Les 1001 formules dont vous rêvez

Formule	Preuve
$\sum_{k=1}^{+}\infty\frac{1}{k^2}=\frac{\Pi^2}{6}$	Formule de Parseval (théorème 1108) sur la série de Fourier de l'identité sur $[-\Pi,\Pi[$
$\sum_{k=1}^{+}\infty 1/p_k=\infty$ p_k k-ième nombre premier	voir partie 15.12.2
$(1+\frac{x}{n})^n\to e^x$ pour x réel.	$(1+\frac{x}{n})^n=e^{n\ ln(x/n)}=e^{n(x/n+o(1/n))}\to e^x$
Formule de Stirling $n!\simeq(\frac{n}{e})^n\sqrt{2\Pi n}$	$$n!=\int_0^{\infty}e^{-t}t^n dt$$ (preuve par récurrence sur n, par une intégration par parties (théorème 555) sur $[0,x]$ et en faisant tendre x vers $+\infty$) changement de variable $u=t/n$ $$n!=n^{n+1}\int_0^{+\infty}e^{n(ln(u)-u)}du$$ Il suffit alors d'appliquer le théorème 594.
$\sum_{n\geq 1}\frac{(-1)^{n-1}}{n}=ln(2)$	$U_N=\sum_{n=1}^{N}\frac{(-1)^{n-1}}{n}$ puis écrire $1/n=\int_0^1 t^{n-1}$ puis $\sum_{n=1}^{N}(-t)^{n-1}=\frac{1-(-t)^n}{1+t}$ puis passer à la limite, par exemple par convergence uniforme sur un ensemble de mesure finie.

38.20 Espaces préhilbertiens

<table>
<tr><th>Espace préhilbertien réel</th><th>Espace préhilbertien complexe</th></tr>
<tr><td>$\|x+y\|^2=\|x\|^2+\|y\|^2$</td><td>$\|x+y\|^2=\|x\|^2+\|y\|^2+2Re(<x|y>)$</td></tr>
<tr><td colspan="2">Inégalité de Schwartz</td></tr>
<tr><td>$|<x|y>|\leq\|x\|\ \|y\|$</td><td>$|<x|y>|\leq\|x\|\ \|y\|$</td></tr>
<tr><td colspan="2">Inégalité de Minkowski</td></tr>
<tr><td>$\|x+y\|\leq\|x\|+\|y\|$
Cas d'égalité : x et y positivement liés</td><td>$\|x+y\|\leq\|x\|+\|y\|$
Cas d'égalité : x et y positivement liés</td></tr>
<tr><td colspan="2">Formule de polarisation</td></tr>
<tr><td>$<x|y>=\frac{1}{2}(q(x+y)-q(x)-q(y))$
$<x|y>=\frac{1}{4}(q(x+y)-q(x-y))$</td><td>$<x|y>=\frac{1}{4}\sum_{j=1}^{4}\overline{i^j}q(x+i^jy)$
$<x|y>$
$=\frac{1}{4}(q(x+y)-q(x-y)+iq(x-iy)-iq(x+iy)$</td></tr>
<tr><td colspan="2">En dimension finie, il existe une base orthonormale</td></tr>
</table>

38.21 Espaces de Hilbert

On suppose que $(x_i)_{i\in I}$ est une base hilbertienne.

Relation de Parseval
$\|x\|^2 = \sum_i (x\|x_i)^2$
$< x\|y >= \sum_i < x\|x_i >< y\|x_i >$
$x = \sum_{i \in I} < x\|x_i > x_i$

Relation de Parseval
$\|x\|^2 = \sum_i \|(x\|x_i)\|^2$
$< x\|y >= \sum_i < x\|x_i > \overline{< y\|x_i >}$
$x = \sum_{i \in I} < x\|x_i > x_i$

38.22 Espaces euclidiens

Espace euclidien	Espace hermitien
f^* adjoint de f $< f(x)\|y >=< x\|f^*(y) >$ $Mat_B(f^*) =^t Mat_B(f)$ $det\ f^* = det\ f$	f^* adjoint de f $< f(x)\|y >=< x\|f^*(y) >$ $Mat_B(f^*) = \overline{{}^t Mat_B(f)}$ $det\ f^* = \overline{det\ f}$
f symétrique si $f^* = f$	f hermitien si $f^* = f$
f orthogonal (ie $\in O(E)$) si $f^* f = f f^* = Id$ $\iff < f(x)\|f(y) >=< x\|y >$ $\iff \|f(x)\| = \|x\|$	f unitaire (ie $\in U(E)$) si $f^* f = f f^* = Id$ $\iff < f(x)\|f(y) >=< x\|y >$ $\iff \|f(x)\| = \|x\|$
f orthogonal est direct (ie $\in O(E)$) si $det\ f > 0$ $\iff$ conserve l'orientation d'une base	

38.23 Réduction en dimension finie - propriétés de matrices particulières

Matrice M de type (n, n), associée à l'endomorphisme f, $E = \mathbb{K}^n$, polynôme caractéristique P, $multip(\lambda)$ ordre de multiplicité de λ dans P. Dans beaucoup de cas, la réduction dans une base orthonormale est basée sur le fait que l'orthogonal d'un espace stable est stable.

diagonalisable	$\iff$ $\exists$ base de vecteurs propres $\iff E=\otimes_{\lambda\in Sp_{\mathbb{K}}(f)}\ker(f-\lambda I)$ $\iff P$ scindé et $\dim\ker(f-\lambda I)=multip(\lambda)$ $\iff \exists Q\in GL_n(\mathbb{K})/QMQ^{-1}$ diagonale
trigonalisable (Jordan)	$\iff P$ scindé $\iff \exists Q\in GL_n(\mathbb{K})\ Q^{-1}MQ$ triang. sup. $\iff \exists Q\in GL_n(\mathbb{K})\ Q^{-1}MQ$ triang. inf. $\iff \exists Q\in GL_n(\mathbb{K})\ Q^{-1}MQ=\otimes M_i$ avec $M_i=\begin{pmatrix}\lambda_i & 1 & 0 & 0 & \dots & 0\\ 0 & \lambda_i & 1 & 0 & \ddots & \vdots\\ \vdots & 0 & \ddots & \ddots & \ddots & 0\\ 0 & \dots & 0 & \lambda_i & 1 & 0\\ 0 & \dots & \dots & 0 & \lambda_i & 1\\ 0 & \dots & \dots & \dots & 0 & \lambda_i\end{pmatrix}$ $\Rightarrow \exists!(D,N)/\begin{cases} D\text{diagonalisable}\\ N\text{nilpotente}\end{cases}$ $M=D+H,\ DH=HD$
M symétrique réelle	diagonalisable <u>dans une base orthonormale</u> i.e. $\exists Q\in O_n(\mathbb{R})/QMQ^{-1}$ diagonale
M,N symétriques réelles $Sp_{\mathbb{R}}(M)>0$	$\exists Q\in GL_n(\mathbb{R})/\begin{cases} QMQ^{-1}\text{ diagonale}\\ QNQ^{-1}\text{ diagonale}\end{cases}$
M antisymétrique réeele	$Sp_{\mathbb{R}}(M)\subset\{0\}$, $Sp_{\mathbb{C}}(M)\subset i\mathbb{R}$ $rang(M)$ pair, M^2 symétrique $\exists Q\in O_n(\mathbb{R})/Q^{-1}MQ=(0)\otimes M_k$ (0) matrice nulle de la taille du noyau, M_k antisymétrique $(2,2)$
M réelle commutant avec ${}^tM$	$\exists Q\in O_n(\mathbb{R})\ QMQ^{-1}=\otimes M_i\otimes N_i$ avec $M_i=(\lambda_i)$ (matrice $(1,1)$) et $N_i=\begin{pmatrix}a_i & -b_i\\ b_i & a_i\end{pmatrix}$
M orthogonale (réelle)	$det\ M=\pm1$, $Sp_{\mathbb{R}}(M)\subset\{-1,1\}$
M unitaire (complexe)	$\lvert det\ M\rvert=1$, $\lvert Sp_{\mathbb{C}}(M)\rvert=1$
M hermitienne	$Sp_{\mathbb{C}}(M)\subset\mathbb{R}$, $\exists Q\in U_n(\mathbb{C})/Q^{-1}MQ$ diagonale
M,N hermitiennes, $Sp_{\mathbb{C}}(M)>0$	$\exists Q\in GL_n(\mathbb{C})/QMQ^{-1}$ diagonale et QNQ^{-1} diagonale
M complexe commute avec ${}^t\overline{M}$	M et ${}^t\overline{M}$ simultanément diagonalisables dans une base orthonormale
M réelle $\forall r\in[1,n]$ $detM_{[1,r]^2}\neq0$	Décomposition ALU $\exists!(L,U)/\ L$ triang. inf., $Sp(L)=1$ U triangulaire supérieure, $A=LU$
A symétrique réelle $Sp(A)>0$	$A=R\,{}^tR$ avec R triangulaire inférieure inversible, décomposition de Cholesky
M_i diagonalisables commutant deux à deux	Les M_i sont simultanément diagonalisables.
$\mathbb{K}$ alg. clos, M_i commutant 2 à 2	Les M_i sont simultanément trigonalisables

38.24 Fonctions holomorphes

courbe=application continue de $[a,b]$ dans $\mathbb{C}$
chemin=application C^1 pm de $[a,b]$ dans $\mathbb{C}$

Hypothèses	Conclusion
μ mesure complexe (donc finie) sur X, Ω ouvert de $\mathbb{C}$, $\phi : X \to \mathbb{C}$ mesurable, $\phi(X) \cap \Omega = \emptyset$	$z \mapsto \int_X \frac{d\mu(x)}{\phi(x)-z}$ est holomorphe
Γ courbe, Ω complémentaire de $\Gamma^* = \Gamma([a,b])$	$z \in \Omega \to Ind_\Gamma(z) = \frac{1}{2i\pi}\int_\Gamma \frac{dt}{t-z}$, Ind_Γ constant sur les composantes connexes de Ω, nul sur la composante connexe non bornée ($Ind_\Gamma(z)$ nombre de tours de Γ autour de z)
Γ_1 et Γ_2 courbes, $\|\Gamma_1(t) - \Gamma_2(t)\| < \|\Gamma_2(t)\|$.	$Ind_{\Gamma_1}(0) = Ind_{\Gamma_2}(0)$
Γ chemin fermé dans Ω ouvert, $f \in H(\Omega)$ $z \in \Omega$, $z \notin \Gamma^*$, Ind_Γ nulle hors de Ω	$f(z)Ind_\Gamma(z) = \frac{1}{2i\pi}\int_\Gamma \frac{f(t)}{t-z}dt$
$f \in H(\Omega)$ $D(x,r)$ disque inclus dans Ω	Théorème de Cauchy : f développable en série entière sur $D(x,r)$ $f = \sum a_n(z-x)^n$ Estimateurs de Cauchy : $\|f^{(n)}(x)\| \leq \frac{n! \sup \|f\|}{r^n}$
f C^0 de Ω ouvert dans $\mathbb{C}$ $\int_P f$ nulle pour tout P triangle (ou sur tout P carré de côtés perpendiculaires aux axes)	f holomorphe sur Ω (théorème de Morera)
f non constante holomorphe sur un ouvert connexe Ω	$Z(f) = f^{-1}(0)$ n'a pas de point d'accumulation, $\forall z \in \Omega$ on peut écrire $\forall t \in \Omega f(t) = (t-z)^m g(t)$ avec g holomorphe non nulle en z
Ω ouvert connexe $f \in H(\Omega \setminus \{a\})$	Trois cas : - singularité artificielle en a - pôle d'ordre $m \in \mathbb{N}$ en a, ie $f - \sum_{i=1}^n \frac{c_i}{(z-a)^i}$ holomorphe sur Ω ; par définition $Res(f;a) = c_1$. - singularité essentielle en a
f holomorphe sur Ω ouvert privé des z_i en nombre fini où f admet des pôles, f chemin fermé ne passant pas par les z_i, $Ind_\Gamma(z) = 0$ pour tout z dans Ω^c	$\frac{1}{2i\pi}\int_\Gamma f(t)dt = \sum_i Res(f;z_i)Ind_\Gamma(z_i)$
$f \in H(\Omega)$,$\overline{D}(a,r) \subset \Omega$, Γ cercle orienté positivement centré sur a de rayon r	Le nombre de zéros de $f - w$ dans $D(a,r)$, comptés avec multiplicités, est l'indice de w par rapport à la courbe $f \circ \Gamma$
Ω ouvert connexe $f \in H(\Omega)$ f non constante $x \in \Omega$ m ordre du zéro de $f - f(x)$ en x	f induit sur un certain ouvert V contenant x une application surjective de V sur un ouvert W telle que chaque point de W autre que $f(x)$ ait exactement m antécédents dans V
f et g dans $H(\Omega)$ $\overline{D}(a,r) \subset \Omega$ $\|f(z) - g(z)\| < \|g(z)\|$ sur $\delta D(a,r)$	f et g ont même nombre de zéros avec multiplicité dans $D(a,r)$.

38.25 Propriétés de G groupe fini de cardinal n

Hypothèses	Conclusion
G cyclique	tout s-g est cyclique
G cyclique, $d \| n$	il existe un unique s-g de cardinal d, c'est $\{x \in G / x^d = 1\} = \{x^{n/d} / x \in G\}$
n premier	$G = \mathbb{Z}/n\mathbb{Z}$
H s-g de G	$\|G\| = \|G/H\|\|H\|$
$H \lhd G$, $K \lhd G$, $K \subset H$	$H/K \lhd G/K$, $G/H \simeq (G/K)/(H/K)$
p premier, p divise n	Th. de Cauchy : p divise n, $\exists x \in G / \ o(x) = p$

FLEMMARD

38.26 Reconnaître un groupe G d'ordre n

On notera que les critères ci-dessous couvrent immédiatement les cas $n \leq 15$, sauf $n = 12$.

Hypothèses	Conclusion
n premier	$G \simeq \mathbb{Z}/n\mathbb{Z}$
tout élément est d'ordre ≤ 2	$G \simeq (\mathbb{Z}/2\mathbb{Z})^q$
$n = 2p$ avec p premier	$\exists H \subset G$ $G \simeq \mathbb{Z}/2\mathbb{Z} \times H$ ou $G \simeq \mathbb{Z}/2\mathbb{Z} \rtimes H$ (groupe diédral D_p)
Sous-groupe d'indice 2 S_m	A_m groupe alterné
$n = pq$ p et q premiers $p < q$	si $q \not\equiv 1(p)$ alors $G \simeq \mathbb{Z}/n\mathbb{Z}$ sinon deux cas : - $G \simeq \mathbb{Z}/n\mathbb{Z}$ - $G \simeq \mathbb{Z}/q\mathbb{Z} \rtimes_\phi \mathbb{Z}/p\mathbb{Z}$ avec $o(\phi(1)) = p$ dans $\mathbb{Z}/q\mathbb{Z}$ (un seul produit semi-direct non trivial possible, à isomorphisme près)
$n = 8$ abélien	$(Z/2\mathbb{Z})^3$ ou $\mathbb{Z}/8\mathbb{Z}$ $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$
$n = 8$ non abélien	diédral D_4 ou quaternions
$n = p^2$, p premier	$G \simeq \mathbb{Z}/p^2\mathbb{Z}$ ou $G \simeq (\mathbb{Z}/p\mathbb{Z})^2$

38.27 Etude d'un groupe abélien fini G

$Card(G) = \Pi p_i^{n_i}$ Existence de $H_i \subset G$ de card $p_i^{n_i}$ garantie, unique	Décomposition primaire : G isomorphe au produit des H_i
FLEMMARD	Décomposition cyclique : G isomorphe au produit des $\mathbb{Z}/n_i\mathbb{Z}$ avec $1 < n_1 \vert n_2 \vert n_3 \vert \ldots \vert n_l$ (n_i) suite des invariants de G (existence et unicité)

38.28 Etude d'un groupe fini G

<table>
<tr><td>G simple</td><td colspan="3">Existence de $N \lhd G$</td></tr>
<tr><td rowspan="5">$G = \mathbb{Z}/p\mathbb{Z}$
avec p
premier ?

Groupe
alterné
A_n avec
$n > 4$?</td><td colspan="3">$\exists H / H \cap N = \{1\} \wedge HN = G$
pour l'un des (N, H) possibles avec
$N \lhd G$</td></tr>
<tr><td>Non</td><td colspan="2">Oui</td></tr>
<tr><td rowspan="3">Etude
du
quotient</td><td colspan="2">Peut-on imposer $H \lhd G$?
Alias : $\forall (h, n) \in H \times N$ $hn = nh$?</td></tr>
<tr><td>Oui</td><td>Non</td></tr>
<tr><td>Produit
direct
$N \times H$</td><td>Produit
semi-direct
$N \rtimes H$
par conjugaison</td></tr>
</table>

38.29 Anneaux

Type d'anneau	Propriétés	Exemples
	Formule de Newton seulement pour des éléments qui commutent	$L(E, F)$
Commutatif	a associé à b implique $(a) = (b)$	$C^\infty(\Omega, \mathbb{R})$ avec Ω ouvert de $\mathbb{R}^n$
Noethérien	Commutatif, tout idéal est de type fini, toute suite croissante d'idéaux est ultimement stationnaire, tout quotient est noethérien, $A[X_1, ..., X_n]$ noethérien.	
Intègre	Commutatif, (0) est un idéal premier, a associé à b $\iff$ $(a) = (b)$, le quotient par la relation d'association est isomorphe pour l'ordre à l'ensemble des idéaux, $A[X_1, \ldots, X_n]$ est intègre	$H(\Omega)$ avec Ω ouvert connexe, quotient d'un anneau par un idéal premier
Factoriel	Intègre, irréductible $\iff$ premier, lemme d'Euclide, th. de Gauss	
Principal	Factoriel, noethérien, les idéaux sont (0) et les (p) avec p irréductible	$\mathbb{K}[X]$ avec $\mathbb{K}$ corps
Euclidien	Principal	$\mathbb{Z}$, $\mathbb{K}[X]$, $\mathbb{Z}[i]$, $\mathbb{Z}[\sqrt{2}]$
Corps $\mathbb{K}$	Seuls idéaux : $\{0\}$ et $\mathbb{K}$	$\mathbb{Z}/p\mathbb{Z}$ avec p premier, $\mathbb{Q}$,$\mathbb{R}$,$\mathbb{C}$, $\mathbb{Q}(i)$, $\mathbb{K}(X)$ avec K corps, corps des fractions d'un anneau intègre

38.30 Calcul différentiel

Rappelons qu'une dérivée ou une fonction continue (sur un intervalle de $\mathbb{R}$!) satisfont la propriété des valeurs intermédiaires.

Hypothèse	Conclusion
$f : V \to W$ V ouvert d'un Banach E W ouvert d'un Banach F Homéomorphisme C^1	f C^1-difféomorphisme si et seulement si $\forall x \in V$ $f'(x) \in Isom(E,F)$ (isomorphisme de Banach)
Théorème d'inversion locale	
f C^1 de U (ouvert d'un Banach E) dans F (Banach), $f'(a) \in Isom(E,F)$	f induit un C^1 difféomorphisme entre un voisinage ouvert de a et un voisinage ouvert de $f(a)$.
Corollaire	
f C^1 de U (ouvert d'un Banach E) dans F (Banach)	f C^1-difféo de U sur son image (ouverte) si et seulement si f injective et $\forall x$ $f'(x) \in Isom(E,F)$
$[a,b]$ segment de $\mathbb{R}$, $b \neq a$, f de $[a,b]$ dans un espace vectoriel normé . Polynôme de Taylor à l'ordre n de f en a : $P_{f,a,n}(x) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$	
Formule de Taylor-Lagrange	
f C^n $n+1$ fois dérivable sur $]a,b[$	$\exists c \in]a,b[/ f(b) = P_{f,a,n}(b) + \frac{f^{n+1}(c)}{(n+1)!}(b-a)^{n+1}$
Inégalité de Taylor-Lagrange	
f C^n $n+1$ fois dérivable sur $]a,b[$, f^{n+1} bornée par M sur $]a,b[$	$\|f(b) - P_{f,a,n}(b)\| \leq M\frac{(b-a)^{n+1}}{(n+1)!}$
Formule de Taylor avec reste intégral	
f C^{n+1} sur $[a,b]$	$f(b) = f(a) + P_{f,a,n}(b) + \frac{1}{n!}\int_a^b f(b-t)^n.f^{n+1}(t)dt$
Formule de Taylor-Young	
f n fois dérivable en a	$f(x) - P_{f,a,n}(x) = o((x-a)^n)$

La première proposition est certes plus faible que le corollaire du théorème d'inversion locale, mais elle fait appel à de moins gros résultats, d'où son intérêt.

38.31 Equations différentielles

$\frac{\delta x}{\delta t} = f(t,x)$, f application de $I \times E$ dans E, avec E un espace de Banach . CI : $f(t_0) = x_0$.

Il faut savoir ramener l'étude d'une équadif d'ordre n à l'étude d'une équadif d'ordre 1.

$f\ C^0$	Existence d'une solution
Th. de Cauchy-Lipschitz	
f localement lipschitzienne en x	$\exists!$ solution maximale x_{t_0,x_0}
Dépendance aux conditions initiales	
f lipschitzienne en x	$(t,x_0) \mapsto x_{t_0,x_0}(t)$ continue
Th. de Cauchy-Lipschitz, ordre n	
$x^{(p)} = f(t,x,x^{(1)},\ldots,x^{(n)})$	$f\ C^0 \Rightarrow$ Existence d'une solution maximale f localement lipschitzienne en $(x^{(i)})_{i\in[1,n]} \Rightarrow$ Unicité
Dépendance à un paramètre	
$\frac{\delta x}{\delta t} = f(t,x,\underbrace{\lambda}_{\in L})$ avec L espace topologique et $f\ C^0$ et k-lipschitzienne en x, avec k indépendant de t et λ	$\exists! x_\lambda / x_\lambda(t_0) = x_0$ $(t,\lambda) \mapsto x_\lambda(t)\ C^0$
Equadifs linéaires d'ordre 1	
$x' = a(t)x + b(t)$ $a\ C^0$ de I dans $\mathcal{L}(E,E)$, $b\ C^0$ de I dans E	$x = exp(\int_{t_0}^t a(u)du)x_0 + x_p(t)$ avec x_p solution particulière, cherchée sous la forme $x_p(t) = \lambda(t)exp(\int_{t_0}^t a(u)du)$,
Divers exemples	
Equation à variables séparées $x'(t) = f(t)g(x(t))$	$(dx/g(x))' = f(t)dt$
Equation de Bernouilli $x'(t) + p(t)x(t) + q(t)x^\alpha = 0$ $\alpha \in \mathbb{R} \setminus \{1\}$	pour se ramener à du linéaire : diviser par x^α et poser $z(x) = x(x)^{1-\alpha}$
Equation homogène $x'(t) = f(\frac{x(t)}{t})$	poser $y(t) = t\ x(t)$ ou passer en coord. polaires
Equation de Lagrange $x(t) = a(x'(t))t + b(x'(t))$ a et $b\ C^1$	Solutions affines : $y = cx + b(c)$, avec $a(c) = c$. Autres solutions : poser $y = x'$, $t = g(x)$. On obtient alors l'équation linéaire $\frac{dt}{dy} = \frac{a'(y)t+b'(y)}{y-a(y)}$
Equation de Riccati $x'(t) = a(t)x(t)^2 + b(t)x(t) + c(t)$	Trouver une solution particulière x_p puis poser $x = x_p + 1/y$, l'équation obtenue est linéaire en y.

Index

Index

Bibliographie

[1] P. BARBE, M. LEDOUX, *Probabilité*, BELIN, 1998.

[2] H. BRÉZIS, *Analyse fonctionnelle*, MASSON, 1983.

[3] H. CARTAN, *Calcul différentiel*, FLEMMARD.

[4] A. CHAMBERT-LOIR, S. FERMIGIER, V. MAILLOT, *Exercices de mathématiques pour l'agrégation, Analyse 1*, MASSON, 1997.

[5] A. CHAMBERT-LOIR, S. FERMIGIER, *Exercices de mathématiques pour l'agrégation, Analyse 2*, MASSON, 1995.

[6] A. CHAMBERT-LOIR, S. FERMIGIER, *Exercices de mathématiques pour l'agrégation, Analyse 3*, MASSON, 1996.

[7] P.G. CIARLET, *Introduction à l'analyse numérique matricielle et à l'optimisation*, DUNOD, 1998.

[8] F. COMBES *Algèbre et géométrie*, BRÉAL, 1998. <

[9] J.-P. DEMAILLY, *Analyse numérique et équations différentielles*, PRESSES UNIVERSITAIRES DE GRENOBLE, 1996.

[10] W. GIORGI, *Thèmes mathématiques pour l'agrégation*, MASSON, 1998.

[11] A. GRAMAIN, *Intégration*, HERMANN 1988, PARIS.

[12] J.-L. KRIVINE, *Introduction to axiomatic set theory*, D. REIDEL PUBLISHING COMPANY, DORDRECHT-HOLLAND.

[13] S. LANG, *Real analysis*, ADDISON-WESLEY PUBLISHING COMPANY, 1969.

[14] D. PERRIN, *Cours d'algèbre*, ELLIPSES 1996.

[15] A. POMMELLET, *Cours d'analyse*, ELLIPSES 1994.

[16] W. RUDIN, *Analyse réelle et complexe*, MASSON 1992.

[17] R. SMULLYAN, *Théorie de la récursion pour la métamathématique*, FLEMMARD.

[18] Y.G. SINAI *Probability theory - An introduction course*, SPRINGER TEXTBOOK, 1992.

[19] P. TAUVEL, *Mathématiques générales pour l'agrégation*, MASSON, 1997.

[20] J. VAUTHIER, J.J. PRAT, *Cours d'analyse mathématiques de l'intégration*, MASSON, 1994.

[21] D. WILLIAMS, *Probability with martingales*, CAMBRIDGE UNIVERSITY PRESS, 1991.

[22] C. ZUILY, H. QUEFFÉLEC, *Eléments d'analyse pour l'intégration*, MASSON, 1995.