Cours Mathématiques MP

david Delaunay

16 octobre 2015

Première partie

Algèbre

Chapitre 1

Groupes

1.1 L'ensemble $\mathbb{Z}/n\mathbb{Z}$

1.1.1 Relation d'équivalence

Définition

On appelle relation d'équivalence sur un ensemble E toute relation binaire $\mathcal{R}$ vérifiant
1) $\mathcal{R}$ est réflexive i.e. $\forall x \in E, x\mathcal{R}x$;
2) $\mathcal{R}$ est symétrique i.e. $\forall x, y \in E, x\mathcal{R}y \Rightarrow y\mathcal{R}x$:
3) $\mathcal{R}$ est transitive i.e. $\forall x, y, z \in E, x\mathcal{R}y$ et $y\mathcal{R}z \Rightarrow x\mathcal{R}z$;

Exemple L'égalité est une relation d'équivalence sur E.

Exemple L'équivalence des suites (ou de fonctions au voisinage de $a \in \bar{\mathbb{R}}$) est une relation d'équivalence.

Exemple L'équivalence des matrices de $\mathcal{M}_{n,p}(\mathbb{K})$.

Remarque Plus généralement, pour une application $f : E \to F$, la relation $\mathcal{R}$ donnée par

$$x\mathcal{R}y \Leftrightarrow f(x) = f(y)$$

définit une relation d'équivalence sur E.

Remarque En fait, une relation d'équivalence se comprend comme « une égalité modulo certains critères » .

1.1.2 Classe d'équivalence

Soit $\mathcal{R}$ une relation d'équivalence sur E.

Définition

On appelle classe d'équivalence d'un élément x de E pour la relation $\mathcal{R}$, le sous-ensemble noté $Cl(x)$ formé des éléments qui sont en relation avec x

$$Cl(x) \underset{\text{déf}}{=} \{y \in E / x\mathcal{R}y\}$$

La classe d'équivalence de x est encore souvent notée $x, \bar{x}, \hat{x}$,...

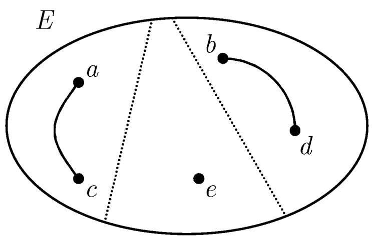
Exemple Considérons $E = \{a, b, c, d, e\}$ et $f : E \to \{0, 1, 2\}$ définie par

$$f(a) = 0, f(b) = 1, f(c) = 0, f(d) = 1 \text{ et } f(e) = 2$$

La relation $\mathcal{R}$ définie par

$$x\mathcal{R}y \Leftrightarrow f(x) = f(y)$$

est une relation d'équivalence que l'on peut visualiser ainsi



Pour celle-ci $Cl(a) = Cl(c) = \{a, c\}$, $Cl(b) = Cl(d) = \{b, d\}$ et $Cl(e) = \{e\}$.

Remarque $Cl(x)$ réunit les éléments de E qui sont « égaux modulo la relation $\mathcal{R}$ » .

Théorème

a) $\forall x \in E, x \in Cl(x)$;
b) $\forall x, y \in E, x\mathcal{R}y \Rightarrow Cl(x) = Cl(y)$;
c) $\forall x, y \in E, x \not\mathcal{R} y \Rightarrow Cl(x) \cap Cl(y) = \emptyset$
Ainsi une classe d'équivalence n'est jamais vide et deux classes d'équivalence distinctes sont disjointes.

dém. :
$x \in Cl(x)$ car la relation $\mathcal{R}$ est réflexive.
Si $x\mathcal{R}y$ alors pour tout $z \in Cl(y)$ on a $y\mathcal{R}z$ et donc $x\mathcal{R}z$ par transitivité. Ainsi $Cl(y) \subset Cl(x)$ et par symétrie on a l'autre inclusion et donc l'égalité.
Enfin, par contraposée, si $Cl(x) \cap Cl(y) \neq \emptyset$ alors pour un certain $z \in Cl(x) \cap Cl(y)$, on a $x\mathcal{R}z$ et $y\mathcal{R}z$ donc par symétrie et transitivité, on obtient $x\mathcal{R}y$.
□

Remarque Si y est élément d'une classe d'équivalence $Cl(x)$ alors $x\mathcal{R}y$ et donc $Cl(x) = Cl(y)$. Ainsi, tout élément d'une classe d'équivalence détermine celle-ci.

Définition

Tout élément y d'une classe d'équivalence est appelé représentant de celle-ci.

1.1.3 Ensemble quotient

Soit $\mathcal{R}$ une relation d'équivalence sur E. Les classes d'équivalence réalisent une partition de E ; cette partition est obtenue en regroupant entre eux les éléments qui sont « égaux modulo la relation $\mathcal{R}$ » .

Exemple Considérons la relation d'équivalence précédente sur $E = \{a, b, c, d, e\}$.
Celle-ci réalise une partition de E en 3 classes d'équivalence.

Définition

On appelle ensemble quotient de E par $\mathcal{R}$ l'ensemble des classes d'équivalence pour relation $\mathcal{R}$.
On le note $E/\mathcal{R}$.

Remarque $E/\mathcal{R}$ se comprend comme l'ensemble obtenu lorsqu'on « identifie entre eux les éléments qui sont égaux modulo $\mathcal{R}$ » .

Exemple L'ensemble $\mathbb{Q}$ des nombres rationnels se construit comme l'ensemble quotient de $\mathbb{Z} \times \mathbb{Z}^\star$ pour la relation

$$(a, b)\mathcal{R}(c, d) \Leftrightarrow ad = bc$$

La classe d'équivalence d'un couple (a, b) est alors notée a/b.

1.1.4 L'ensemble $\mathbb{Z}/n\mathbb{Z}$

Soit $n \in \mathbb{N}^\star$.

Définition

On définit sur $\mathbb{Z}$ la relation de congruence modulo n par

$$a \equiv b \quad [n] \Leftrightarrow n \mid (b - a)$$

Proposition

La relation de congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$.

dém. :
La relation est réflexive car $a \equiv a \quad [n]$ puisque $n \mid (a - a)$.
La relation est symétrique car $a \equiv b \quad [n] \Rightarrow b \equiv a \quad [n]$ puisque $n \mid (b - a) \Rightarrow n \mid (a - b)$.
Enfin, la relation est transitive car $a \equiv b \quad [n]$ et $b \equiv c \quad [n] \Rightarrow a \equiv c \quad [n]$ puisque $n \mid (b - a)$ et $n \mid (c - b) \Rightarrow n \mid (c - a)$.
□

Définition

Pour $a \in \mathbb{Z}$, on note $\bar{a}$ la classe d'équivalence de $a \in \mathbb{Z}$ pour la relation de congruence modulo n.
Ainsi

$$\bar{a} = \{a + kn/k \in \mathbb{Z}\} = a + n\mathbb{Z}$$

Définition

On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble quotient de $\mathbb{Z}$ pour la relation de congruence modulo n.

Théorème

$\mathbb{Z}/n\mathbb{Z}$ est un ensemble fini à n éléments qui sont

$$\bar{0}, \bar{1}, \ldots, \overline{(n-1)}$$

dém. :
$\bar{0}, \bar{1}, \ldots, \overline{(n-1)}$ sont des éléments de $\mathbb{Z}/n\mathbb{Z}$.
Pour $a, b \in \{0, \ldots, n-1\}$,

$$\bar{a} = \bar{b} \Rightarrow n \mid (b-a) \Rightarrow a = b$$

Par suite, les classes $\bar{0}, \bar{1}, \ldots, \overline{(n-1)}$ sont deux à deux distinctes.
Pour tout $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$, en considérant le reste $r \in \{0, 1, \ldots, n-1\}$ de la division euclidienne de a par n, on obtient $\bar{a} = \bar{r}$. Ainsi toutes les classes d'équivalence figurent parmi $\bar{0}, \bar{1}, \ldots, \overline{(n-1)}$.
□

Exemple $\mathbb{Z}/2\mathbb{Z} = \{\bar{0}, \bar{1}\}$, $\mathbb{Z}/3\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}\}$, $\mathbb{Z}/4\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$, etc.

Proposition

Pour tout $a, b, a', b' \in \mathbb{Z}$,

$$a \equiv a' \quad [n] \text{ et } b \equiv b' \quad [n] \Rightarrow a + b \equiv a' + b' \quad [n] \text{ et } ab \equiv a'b' \quad [n]$$

dém. :
$n \mid a' - a$ et $n \mid b' - b$ entraînent $n \mid (a' + b') - (a + b) = (a' - a) + (b' - b)$ et $n \mid (a'b') - (ab) = (a' - a)b' + a(b' - b)$
□

Définition

On définit deux opérations $+$ et $\times$ sur $\mathbb{Z}/n\mathbb{Z}$ en posant

$$\bar{a} + \bar{b} \underset{\text{déf}}{=} \overline{a + b} \text{ et } \bar{a} \times \bar{b} \underset{\text{déf}}{=} \overline{ab}$$

Remarque La définition ci-dessus est consistante puisque le résultat de ces opérations ne dépend pas des représentants a, b choisis pour chaque classe.

Exemple Dans $\mathbb{Z}/6\mathbb{Z}$,
$\bar{3}+\bar{5}=\bar{8}=\bar{2}$ ou encore $\bar{3}+\bar{5}=\bar{3}+\overline{-1}=\bar{2}$.
$\bar{3}\times\bar{5}=\overline{15}=\bar{3}$ ou encore $\bar{3}\times\bar{5}=\bar{3}\times\overline{-1}=\overline{-3}=\bar{3}$.

1.2 Structure de groupe

1.2.1 Définition

Définition

On appelle groupe tout couple $(G,\star)$ formé d'un ensemble G et d'une loi de composition interne $\star$ sur G vérifiant :
1) $\star$ est associative i.e.

$$\forall a,b,c\in G,(a\star b)\star c=a\star(b\star c)\,;$$

2) $\star$ possède un neutre i.e.

$$\exists e\in G,\forall a\in G,a\star e=a=e\star a$$

cet élément e est alors unique ;
3) tout élément de G est symétrisable $\star$ i.e.

$$\forall a\in G,\exists b\in G,a\star b=e=b\star a$$

cet élément b est alors unique et appelé symétrique de a, noté a^{-1} .
Si de plus la loi $\star$ est commutative, on parle de groupe abélien.
Lorsque la loi est notée $\times$ ou., on dit que le groupe est noté multiplicativement ($e\to 1$, $a\star b\to ab$)
Lorsque la loi est notée +, on dit que le groupe est noté additivement ($e\to 0$, $a\star b\to a+b$, $a^{-1}\to -a$). Cette dernière notation est réservée au groupe commutatif.

Attention : Lorsque la loi $\star$ n'est pas commutative :
- la neutralité de e se vérifie par deux compositions ;
- l'inversibilité d'un élément se vérifie par deux compositions ;
- on a $(a\star b)^{-1}=b^{-1}\star a^{-1}$.

Exemple $(\mathbb{C},+),(\mathbb{R},+),(\mathbb{Z},+)$ sont des groupes abéliens de neutre 0.

Exemple $(\mathbb{C}^\star,\times),(\mathbb{R}^\star,\times),(\mathbb{R}^{+\star},\times)$ sont des groupes abéliens de neutre 1.

Exemple $(\mathrm{GL}_n(\mathbb{K}),\times)$ est un groupe non commutatif de neutre I_n.

1.2.2 Itéré d'un élément

Soit $(G,\star)$ un groupe de neutre e.

Définition

Pour $a \in G$ et $k \in \mathbb{Z}$, on note a^k l'itéré d'ordre k de l'élément a :
- pour $k > 0$, $a^k \underset{\text{déf}}{=} a \star \cdots \star a$ (k termes) ;
- pour $k = 0$, $a^0 \underset{\text{déf}}{=} e$;
- pour $k < 0$, $a^k \underset{\text{déf}}{=} a^{-1} \star \cdots \star a^{-1}$ ($|k|$ termes).

Proposition

On a

$$\forall k, \ell \in \mathbb{Z},\, a^k \star a^\ell = a^{k+\ell} \text{ et } (a^k)^\ell = a^{k\ell}$$

dém. :
Il suffit de discuter selon les signes des exposants d'itérations considérés, c'est un peu lourd…
□

Remarque Si le groupe est noté additivement, on note $k.a$ l'itéré d'ordre k de a. On a alors

$$k.a + \ell.a = (k+\ell).a \text{ et } \ell.(k.a) = (k\ell).a$$

Attention : En général

$$(a \star b)^p \neq a^p \star b^p$$

En effet

$$(a \star b)^p = (a \star b) \star (a \star b) \star \ldots \star (a \star b)$$

et

$$a^p \star b^p = (a \star a \star \ldots \star a) \star (b \star b \star \ldots \star b)$$

Cependant, si a et b commutent alors $(a \star b)^p = a^p \star b^p$

1.2.3 Le groupe symétrique

Définition

On note $\mathcal{S}_E$ l'ensemble des permutations de E i.e. des bijections de E vers E.

Théorème

$(\mathcal{S}_E, \circ)$ est un groupe de neutre Id_E.
Ce groupe est non commutatif dès que $\mathrm{Card}E \geqslant 3$.

Exemple $\mathcal{S}_n = \mathcal{S}(\{1,\ldots,n\})$ est un groupe de cardinal $n!$.
Parmi ses éléments signalons :
- les transpositions $\tau = \begin{pmatrix} i & j \end{pmatrix}$ vérifiant $\tau^2 = \mathrm{Id}$;
- les p-cycles $c = \begin{pmatrix} a_1 & a_2 & \ldots & a_p \end{pmatrix}$ vérifiant $c^p = \mathrm{Id}$.

1.2.4 Le groupe $(\mathbb{Z}/n\mathbb{Z},+)$

Théorème

$(\mathbb{Z}/n\mathbb{Z},+)$ est un groupe abélien à n éléments de neutre $\bar{0}$.
De plus
$$\forall \bar{a}\in\mathbb{Z}/n\mathbb{Z},\ -\bar{a}=\overline{(-a)}$$

dém. :
$\bar{a}+\bar{b}=\overline{(a+b)}=\overline{(b+a)}=\bar{b}+\bar{a}$ donc $+$ est commutative sur $\mathbb{Z}/n\mathbb{Z}$.
$(\bar{a}+\bar{b})+\bar{c}=\overline{a+b}+\bar{c}=\overline{(a+b)+c}=\overline{a+(b+c)}=\bar{a}+(\bar{b}+\bar{c})$ donc $+$ est associative sur $\mathbb{Z}/n\mathbb{Z}$.
$\bar{a}+\bar{0}=\overline{a+0}=\bar{a}=\bar{0}+\bar{a}$ donc $\bar{0}$ est élément neutre de $(\mathbb{Z}/n\mathbb{Z},+)$.
$\bar{a}+\overline{(-a)}=\overline{a-a}=\bar{0}=\overline{(-a)}+\bar{a}$ donc $\bar{a}$ est symétrisable et $-\bar{a}=\overline{(-a)}$.
□

Exemple $n=2$, $\mathbb{Z}/2\mathbb{Z}=\{\bar{0},\bar{1}\}$.

$+$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

Exemple $n=3$, $\mathbb{Z}/3\mathbb{Z}=\{\bar{0},\bar{1},\bar{2}\}$.

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

Remarque Dans une table d'opérations, sur chaque ligne figure chaque élément de groupe ; cela provient de la bijectivité de l'application $x\mapsto a\star x$ sur G. On a la même propriété sur les colonnes.

Théorème

Pour tout $\bar{a}\in\mathbb{Z}/n\mathbb{Z}$ et $k\in\mathbb{Z}$
$$k.\bar{a}=\overline{k\times a}$$

dém. :
Par récurrence pour $k\in\mathbb{N}$.
Cas $k=0$: $0.\bar{a}=\bar{0}=\overline{0.a}$.
Supposons la propriété vraie au rang $k\geqslant 0$.
$$(k+1).\bar{a}=k.\bar{a}+\bar{a}\underset{HR}{=}\overline{ka}+\bar{a}=\overline{ka+a}=\overline{(k+1)a}$$
Récurrence établie.
Pour $k\in\mathbb{Z}^-$, on peut écrire $k=-p$ avec $p\in\mathbb{N}$.
On a alors
$$k.\bar{a}=-(p.\bar{a})=-\overline{pa}=\overline{-pa}=\overline{ka}$$

□

1.2.5 Produit fini de groupes

Définition

Soit $\star_1, \dots, \star_n$ des lois de composition interne sur des ensembles $E_1, \dots, E_n$. On appelle loi produit sur $E = E_1 \times \cdots \times E_n$ la loi $\star$ définie par

$$(x_1, \dots, x_n) \star (y_1, \dots, y_n) \underset{\text{déf}}{=} (x_1 \star_1 y_1, \dots, x_n \star_n y_n)$$

Proposition

Si $(G_1, \star_1)$,..., $(G_n, \star_n)$ sont des groupes de neutres $e_1, \dots, e_n$ alors $G = G_1 \times \ldots \times G_n$ muni de la loi produit $\star$ est un groupe de neutre $e = (e_1, \dots, e_n)$.
De plus :
- l'inverse d'un élément $(x_1, \dots, x_n) \in G$ est $(x_1^{-1}, \dots, x_n^{-1})$;
- si tous les groupes $(G_1, \star_1)$,..., $(G_n, \star_n)$ sont commutatifs, le groupe $(G, \star)$ l'est aussi.

dém. :
Soit $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ et $z = (z_1, \dots, z_n)$ éléments de $G_1 \times \ldots \times G_n$.
On a

$$x \star (y \star z) = (\dots, x_i \star_i (y_i \star_i z_i), \dots)$$

et

$$(x \star y) \star z = (\dots, (x_i \star_i y_i) \star_i z_i, \dots)$$

Puisque les lois $\star_i$ sont associatives, on obtient

$$x \star (y \star z) = (x \star y) \star z$$

L'élément e est neutre car

$$x \star e = (\dots, x_i \star_i e_i, \dots) = x \text{ et } e \star x = (\dots, e_i \star_i x_i, \dots) = x$$

L'élément x est symétrisable de symétrique $x' = (x_1^{-1}, \dots, x_n^{-1})$ car

$$x \star x' = (\dots, x_i \star_i x_i^{-1}, \dots) = e \text{ et } x' \star x = (\dots, x_i^{-1} \star_i x_i, \dots) = e$$

Ainsi $(G, \star)$ est bien un groupe.
Si de plus les lois $\star_i$ sont toutes commutatives

$$x \star y = (\dots, x_i \star y_i, \dots) = (\dots, y_i \star x_i, \dots) = y \star x$$

□

Exemple Si $(G, \star)$ est un groupe de neutre e alors $(G^n, \star)$ est un groupe de neutre $(e, \dots, e)$.

Exemple Pour $(G_1, \star_1) = (G_2, \star_2) = (\mathbb{Z}, +)$, la loi produit sur $\mathbb{Z}^2$ que nous notons $+$ est définie par :

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$

$(\mathbb{Z}^2, +)$ est un groupe abélien de neutre $0_{\mathbb{Z}^2} = (0, 0)$.

Exemple Pour $(G_1, \star_1) = (\mathbb{R}^{+\star}, \times)$ et $(G_2, \star_2) = (\mathbb{R}, +)$, la loi produit sur $\mathbb{R}^{+\star} \times \mathbb{R}$ que nous notons $\star$ est définie par :

$$(r, \theta) \star (r', \theta') = (rr', \theta + \theta')$$

$(\mathbb{R}^{+\star} \times \mathbb{R}, \star)$ est alors un groupe abélien de neutre $e = (1, 0)$.
De plus

$$(r, \theta)^{-1} = (1/r, -\theta)$$

1.3 Sous-groupes

$(G, \star)$ désigne un groupe de neutre e.

1.3.1 Définition

Définition

On appelle sous-groupe d'un groupe $(G, \star)$ toute partie H de G vérifiant :
1) $e \in H$;
2) $\forall x, y \in H, x \star y^{-1} \in H$.

Exemple $\{e\}$ et G des sont sous-groupes de $(G, \star)$.

Remarque Le point 1) peut aussi être transposé en $H \neq \emptyset$ car alors $H \neq \emptyset$ et 2) entraîne $e \in H$.
Le point 2) peut aussi être transposé en 2a) $\forall x, y \in H, x \star y \in H$ et 2.b) $\forall x \in H, x^{-1} \in H$.

Remarque Si le groupe est noté additivement 1) et 2) se relisent $0 \in H$ et $\forall x, y \in H, x - y \in H$.

Théorème

Si H est un sous-groupe d'un groupe $(G, \star)$ alors $(H, \star)$ est un groupe de même neutre.

Exemple L'ensemble des racines n-ième de l'unité est

$$\mathbb{U}_n = \{z \in \mathbb{C} / z^n = 1\}$$

C'est un sous-groupe de $(\mathbb{C}^\star, \times)$.
$(\mathbb{U}_n, \times)$ est le groupe des racines n-ième de l'unité.
Rappelons

$$\mathbb{U}_n = \left\{ \mathrm{e}^{2ik\pi/n} / k \in [\![0, n-1]\!] \right\} = \left\{ \omega^k / k \in [\![0, n-1]\!] \right\}$$

avec $\omega = \mathrm{e}^{2i\pi/n}$.

Exemple L'ensemble des matrices orthogonale est

$$\mathcal{O}_n(\mathbb{R}) = \left\{ A \in \mathcal{M}_n(\mathbb{R}) / {}^t AA = I_n \right\}$$

C'est un sous-groupe de $(\mathrm{GL}_n(\mathbb{R}), \times)$.
$(\mathcal{O}_n(\mathbb{R}), \times)$ est un groupe, c'est le groupe orthogonal d'ordre n.

1.3.2 Intersection d'une famille de sous-groupes

Théorème

Si $(H_i)_{i\in I}$ est une famille de sous-groupes de $(G,\star)$ alors leur intersection $H=\bigcap_{i\in I} H_i$ est un sous-groupe de $(G,\star)$.

dém. :
$H\subset G$ et $e\in H$ car e est élément de chaque H_i.
Soit $x,y\in H$. Pour tout $i\in I$, $x,y\in H_i$ donc $x\star y^{-1}\in H_i$ puis $x\star y^{-1}\in H$.
□

Remarque La réunion de deux sous-groupes n'est pas un sous-groupe sauf cas d'inclusion de l'un dans l'autre.

1.3.3 Sous-groupe engendré par un élément

Définition

On appelle sous-groupe engendré par un élément $a\in G$ l'ensemble

$$\langle a\rangle \underset{\text{déf}}{=} \left\{a^k/k\in\mathbb{Z}\right\}$$

Remarque En notation additive,

$$\langle a\rangle=\{k.a/k\in\mathbb{Z}\}$$

Théorème

$\langle a\rangle$ est un sous-groupe de $(G,\star)$ contenant a.
De plus, pour tout sous-groupe H de G

$$a\in H\Rightarrow\langle a\rangle\subset H$$

Ainsi $\langle a\rangle$ apparaît comme le plus petit sous-groupe contenant a.

dém. :
$\langle a\rangle\subset G$, $e=a^0\in\langle a\rangle$ et pour tout $x,y\in\langle a\rangle$, on peut écrire $x=a^k$, $y=a^\ell$ avec $k,\ell\in\mathbb{Z}$ et alors

$$x\star y^{-1}=a^{k-\ell}\in\langle a\rangle$$

$\langle a\rangle$ est donc un sous-groupe de $(G,\star)$ et $a=a^1\in\langle a\rangle$.
De plus, si H est un sous-groupe de $(G,\star)$ contenant a alors

$$a^0=e\in H,\ a^1=a\in H,\ a^2=a\star a\in H,\ a^3=a^2\star a\in H,\ldots$$

Par une récurrence facile,

$$\forall k\in\mathbb{N},\ a^k\in H$$

Pour $k\in\mathbb{Z}^-$, $k=-p$ avec $p\in\mathbb{N}$, $a^k=a^{-p}=(a^p)^{-1}\in H$ car $a^p\in H$.
Ainsi

$$\forall k\in\mathbb{Z}, a^k\in H$$

ce qui signifie $\langle a\rangle \subset H$.
□

Remarque Même si la loi $\star$ n'est pas commutative, le sous-groupe $\langle a\rangle$ est commutatif car

$$a^k \star a^\ell = a^{k+\ell} = a^{\ell+k} = a^\ell \star a^k$$

Exemple Dans $(\mathbb{C}, +)$,

$$\langle a\rangle = \{ak/k \in \mathbb{Z}\} = a\mathbb{Z}$$

Exemple Dans $(\mathbb{C}^\star, \times)$,

$$\langle a\rangle = \left\{a^k/k \in \mathbb{Z}\right\}$$

En particulier

$$\langle 2\rangle = \left\{2^k/k \in \mathbb{Z}\right\} = \{\ldots, 1/8, 1/4, 1/2, 1, 2, 4, 8, \ldots\}$$

et pour $\omega = \mathrm{e}^{2i\pi/n}$

$$\langle \omega\rangle = \left\{\omega^k/k \in \mathbb{Z}\right\} = \left\{1, \omega, \ldots, \omega^{n-1}\right\} = \mathbb{U}_n$$

car $\omega^n = 1$.

Exemple Dans $(\mathcal{S}_4, \circ)$ considérons le cycle $c = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}$.

$$\langle c\rangle = \left\{\mathrm{Id}, \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 2 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 3 & 2 & 1 \end{pmatrix}\right\}$$

1.3.4 Sous-groupe engendré par une partie

Définition

On appelle groupe engendré par une partie A de G l'intersection de tous les sous-groupes de $(G, \star)$ qui contiennent A. On le note $\langle A\rangle$

Théorème

$\langle A\rangle$ est un sous-groupe de $(G, \star)$ qui contient A.
De plus, pour tout sous-groupe H de $(G, \star)$,

$$A \subset H \Rightarrow \langle A\rangle \subset H$$

Ainsi $\langle A\rangle$ apparaît comme le plus petit sous-groupe contenant A.

dém. :
Posons $\mathcal{S} = \{H \text{ sous - groupe de } (G, \star)/A \subset H\}$. Par définition

$$\langle A\rangle = \bigcap_{H \in \mathcal{S}} H$$

$\langle A\rangle$ est un sous-groupe car intersection d'une famille de sous-groupes.
Puisque A est inclus dans chaque $H \in \mathcal{S}$, on a $A \subset \langle A\rangle$.
Enfin, si H est un sous-groupe de $(G, \star)$

$$A \subset H \Rightarrow H \in \mathcal{S} \Rightarrow \langle A\rangle \subset H$$

□

Exemple Pour $a \in G$,

$$\langle\{a\}\rangle = \left\{a^k/k \in \mathbb{Z}\right\} = \langle a\rangle$$

Exemple Pour $a, b \in G$,

$$\langle\{a, b\}\rangle = \left\{a^{k_1}b^{\ell_1}\dots a^{k_n}b^{\ell_n}/n \in \mathbb{N}^\star, k_1, \dots, k_n, \ell_1, \dots, \ell_n \in \mathbb{Z}\right\}$$

En fait

$$\langle\{a, b\}\rangle = \{\text{produits finis d'itérés de } a \text{ et } b\}$$

Si a et b commutent, on peut simplifier

$$\langle\{a, b\}\rangle = \left\{a^k b^\ell/k, \ell \in \mathbb{Z}\right\}$$

Exemple Dans $(\mathbb{Z}^2, +)$

$$\langle\{(a, b), (c, d)\}\rangle = \{(ka + \ell c, kb + \ell d)/k, \ell \in \mathbb{Z}\}$$

On peut montrer que ce groupe se confond avec $\mathbb{Z}^2$ si, et seulement si, $ad - bc = \pm 1$.

Exemple Dans $\mathcal{S}_n$, considérons $\mathcal{T}$ l'ensemble des transpositions éléments de $\mathcal{S}_n$. On a

$$\langle\mathcal{T}\rangle = \mathcal{S}_n$$

car il est connu que toute permutation peut s'écrire comme un produit de transpositions.

1.3.5 Les sous-groupes de $(\mathbb{Z}, +)$

Théorème

Les sous-groupes de $(\mathbb{Z}, +)$ sont les $n\mathbb{Z}$ avec $n \in \mathbb{N}$.

dém. :
$n\mathbb{Z}$ est un sous-groupe de $(\mathbb{Z}, +)$ car

$$n\mathbb{Z} = \{kn/k \in \mathbb{Z}\} = \langle n\rangle$$

Inversement, soit H un sous-groupe de $(\mathbb{Z}, +)$.
Cas $H = \{0\}$: on a $H = n\mathbb{Z}$ avec $n = 0$.
Cas $H \neq \{0\}$: on introduit $H^+ = \{x \in H/x > 0\}$.

Il existe $x_0 \in H$ tel que $x_0 \neq 0$. Si $x_0 > 0$ alors $x_0 \in H^+$, sinon $-x_0 \in H^+$. Dans les deux cas $H^+ \neq \emptyset$.
Rappelons : Toute partie non vide de $\mathbb{N}$ admet un plus petit élément.
Ici H^+ est une partie non vide de $\mathbb{N}$, on peut donc introduire $n = \min H^+$.
On a $n \in H$ donc $n\mathbb{Z} = \langle n \rangle \subset H$.
Inversement, soit $x \in H$. Par division euclidienne, $x = qn + r$ avec $0 \leqslant r < n$.
On a alors $r = x - qn \in H$ car $qn \in n\mathbb{Z} \subset H$.
Si $r > 0$ alors $r \in H^+$ ce qui est impossible car $r < n = \min H^+$.
Il reste $r = 0$ et donc $x = qn \in n\mathbb{Z}$.
Ainsi $H \subset n\mathbb{Z}$ puis par double inclusion $H = n\mathbb{Z}$.
□

Remarque Le naturel n tel que $H = n\mathbb{Z}$ est unique car
Si $H = \{0\}$ alors $n = 0$ et si $H \neq \{0\}$ alors $n = \min\{x \in H / x > 0\}$.

1.4 Morphisme de groupes

Soit $(G, \star)$, $(G', \top)$ et $(G'', \perp)$ des groupes.

1.4.1 Définition

Définition

On appelle morphisme du groupe $(G, \star)$ vers le groupe $(G', \top)$ toute application $\varphi : G \to G'$ vérifiant

$$\forall x, y \in G, \varphi(x \star y) = \varphi(x) \top \varphi(y)$$

Exemple L'application constante $\varphi : G \to G$ définie par $\varphi(x) = e$ est un morphisme du groupe $(G, \star)$ vers lui-même.

Exemple L'identité Id_G est un morphisme du groupe $(G, \star)$ vers lui-même.

Remarque Un morphisme d'un groupe vers lui-même est souvent appelé endomorphisme.

Exemple $\ln$ est un morphisme de $(\mathbb{R}^{+\star}, \times)$ vers $(\mathbb{R}, +)$.
En effet, pour tout $a, b > 0$,

$$\ln(ab) = \ln(a) + \ln(b)$$

Exemple $\exp$ est un morphisme de $(\mathbb{C}, +)$ vers $(\mathbb{C}^\star, \times)$.
En effet, pour tout $z, z' \in \mathbb{C}$,

$$\exp(z + z') = \exp(z)\exp(z')$$

Exemple Le déterminant définit par restriction un morphisme de $(\mathrm{GL}_n(\mathbb{K}), \times)$ vers $(\mathbb{K}^\star, \times)$

Exemple La signature $\varepsilon : \mathcal{S}_n \to \{1,-1\}$ avec

$$\varepsilon(\sigma) = \prod_{1\leqslant i<j\leqslant n} \frac{\sigma(j)-\sigma(i)}{j-i}$$

est un morphisme du groupe $(\mathfrak{S}_n, \circ)$ vers $(\{1,-1\}, \times)$.
En effet,

$$\forall \sigma, \sigma' \in \mathcal{S}_n, \varepsilon(\sigma\circ\sigma') = \varepsilon(\sigma)\times\varepsilon(\sigma')$$

Rappelons que si τ est une transposition alors $\varepsilon(\tau) = -1$.
En conséquence, si c est un cycle de longueur p alors $\varepsilon(c) = (-1)^{p-1}$ car c est un produit de $p-1$ transpositions

$$\begin{pmatrix} a_1 & a_2 & \dots & a_p \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \end{pmatrix} \circ \begin{pmatrix} a_2 & a_3 \end{pmatrix} \circ \ldots \circ \begin{pmatrix} a_{p-1} & a_p \end{pmatrix}$$

Exemple Soit a un élément d'un groupe $(G,\star)$.
L'application $\varphi : \mathbb{Z} \to G$ définie par $\varphi(k) = a^k$ est un morphisme de groupes.
En effet

$$\varphi(n+p) = a^{\star(n+p)} = a^{\star n} \star a^{\star p} = \varphi(n)\star\varphi(p)$$

1.4.2 Propriétés

Proposition

> Si $\varphi : G \to G'$ et $\psi : G' \to G''$ sont des morphismes de groupes alors $\psi\circ\varphi : G \to G''$ en est un aussi.

dém. :
Soit $x, y \in G$. On a

$$\psi\circ\varphi(x\star y) = \psi(\varphi(x)\top\varphi(y)) = (\psi\circ\varphi(x))\perp(\psi\circ\varphi(y))$$

□

Remarque La composée de deux endomorphismes d'un groupe $(G,\star)$ est un endomorphisme du groupe $(G,\star)$.

Proposition

> Si φ est un morphisme d'un groupe $(G,\star)$ vers un groupe $(H,\top)$ alors
>
> $$\varphi(e) = e' \text{ et } \forall x \in G, \varphi(x^{-1}) = \varphi(x)^{-1}$$
>
> Plus généralement
>
> $$\forall x \in G, \forall n \in \mathbb{Z}, \varphi(x^n) = \varphi(x)^n$$

dém. :
$\varphi(e) = \varphi(e\star e) = \varphi(e)\top\varphi(e)$ et en composant par $\varphi(e)^{-1}$ on obtient $e' = \varphi(e)$.
Aussi $\varphi(x)\top\varphi(x^{-1}) = \varphi(x\star x^{-1}) = \varphi(e) = e'$ donc en composant par $\varphi(x)^{-1}$ à gauche on obtient

$$\varphi(x^{-1}) = \varphi(x)^{-1}$$

Par récurrence, on vérifie aisément
$$\forall n \in \mathbb{N}, \varphi(x^n) = \varphi(x)^n$$
puis par passage au symétrique, on étend cette propriété à $n \in \mathbb{Z}$.
□

Remarque On peut aussi établir
$$\forall x_1, \ldots, x_n \in G, f\left(\mathop{\star}_{i=1}^{n} x_i\right) = \mathop{\top}_{i=1}^{n} f(x_i)$$

Théorème

> L'image directe (resp. réciproque) d'un sous-groupe par un morphisme de groupes est un sous-groupe.

dém. :
Soit $\varphi : G \to G'$ morphisme de groupes.
Soit H un sous-groupe de $(G, \star)$. Montrons que
$$\varphi(H) = \{\varphi(x)/x \in H\}$$
est un sous-groupe de $(G', \top)$.
D'une part $e' \in \varphi(H)$ car $e' = \varphi(e)$ avec $e \in H$.
D'autre part, pour $x', y' \in \varphi(H)$, on peut écrire $x' = \varphi(x)$ et $y' = \varphi(y)$ avec $x, y \in H$ et alors
$$x' \top y'^{-1} = \varphi(x \star y^{-1}) \in \varphi(H)$$
car $x \star y^{-1} \in H$.
Ainsi $\varphi(H)$ est un sous-groupe de $(G', \top)$.
Soit H' un sous-groupe de $(G, \top)$. Montrons que
$$\varphi^{-1}(H') = \{x \in G/\varphi(x) \in H'\}$$
est un sous-groupe de $(G, \star)$.
D'une part $e \in \varphi^{-1}(H')$ car $\varphi(e) = e' \in H'$.
D'autre part, pour $x, y \in \varphi^{-1}(H')$, on a $\varphi(x \star y^{-1}) = \varphi(x) \top \varphi(y)^{-1} \in H'$ car $\varphi(x), \varphi(y) \in H'$.
Ainsi $\varphi^{-1}(H')$ est un sous-groupe de $(G, \star)$.
□

1.4.3 Noyau et image

Définition

> Si φ est un morphisme du groupe $(G, \star)$ vers le groupe $(G', \top)$, on introduit
> - son noyau $\ker \varphi = \varphi^{-1}(\{e'\})$ qui est un sous-groupe de $(G, \star)$;
> - son image $\mathrm{Im}\varphi = \varphi(G)$ qui est un sous-groupe de $(G', \top)$.

Exemple Déterminons image et noyau du morphisme de $\varphi : \mathbb{C}^\star \to \mathbb{C}^\star$ défini par $\varphi(z) = |z|$.
$\mathrm{Im}\varphi = \mathbb{R}^{+\star}$ et $\ker \varphi = \mathbb{U}$

Exemple Déterminons image et noyau du morphisme $\exp : \mathbb{C} \to \mathbb{C}^\star$.
Pour $z = a + ib$, on a $\exp(z) = \mathrm{e}^a \mathrm{e}^{ib}$.
Pour $Z \in \mathbb{C}^\star$, on peut écrire $Z = r\mathrm{e}^{i\theta}$.
En posant $z = \ln r + i\theta$, on a $\exp(z) = Z$. Ainsi

$$\mathrm{Im}(\exp) = \mathbb{C}^\star$$

Aussi, pour $z = a + ib$

$$\exp(z) = 1 \Leftrightarrow \mathrm{e}^a = 1 \text{ et } \mathrm{e}^{ib} = 1$$

Par suite

$$\ker(\exp) = 2i\pi\mathbb{Z}$$

Exemple Déterminons image et noyau de $\det : \mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}^\star$.
On a $\mathrm{Im}\det = \mathbb{K}^\star$ car avec une matrice diagonale il est facile de construire une matrice inversible de déterminant tel que voulu. Aussi

$$\ker\det = \{M \in \mathrm{GL}_n(\mathbb{K}) / \det M = 1\} = \mathrm{SL}_n(\mathbb{K})$$

appelé groupe spécial linéaire d'ordre n.

Exemple Déterminons image et noyau de $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$ pour $n \geqslant 2$.
On a $\mathrm{Im}\varepsilon = \{1, -1\}$ et

$$\ker\varepsilon = \mathfrak{A}_n$$

appelé groupe alterné (ou groupe des permutations paires).

Théorème

Soit φ un morphisme du groupe $(G, \star)$ vers le groupe $(G', \top)$.
a) φ est injectif si, et seulement si, $\ker\varphi = \{e\}$.
b) φ est surjectif si, et seulement si, $\mathrm{Im}\varphi = G'$.

dém. :
a) Si φ est injectif, e' possède au plus un antécédent par φ. Puisque $\varphi(e) = e'$, on obtient

$$\ker\varphi = \{e\}$$

Inversement, supposons $\ker\varphi = \{e\}$. Soit $x, y \in G$ tels que $\varphi(x) = \varphi(y)$.
On a $\varphi(x \star y^{-1}) = \varphi(x) \top \varphi(y)^{-1} = e'$ et donc $x \star y^{-1} \in \ker\varphi$. Ainsi $x \star y^{-1} = e$ puis $x = y$.
b) C'est une évidence et ne dépend du fait que φ soit un morphisme.
□

1.4.4 Isomorphisme de groupes

Définition

On appelle isomorphisme de groupes tout morphisme de groupes bijectif.

Exemple $\ln : \mathbb{R}^{+\star} \to \mathbb{R}$ est un isomorphisme de $\left(\mathbb{R}^{+\star}, \times\right)$ vers $(\mathbb{R}, +)$.

Proposition

Si $\varphi : G \to G'$ et $\psi : G' \to G''$ sont des isomorphismes de groupes alors $\psi \circ \varphi : G \to G''$ en est un aussi.

Théorème

Si $\varphi : G \to G'$ est un isomorphisme de groupes alors $\varphi^{-1} : G' \to G$ est un isomorphisme de groupes.

dém. :
dém. :
Pour tout $x', y' \in G'$, il existe $x, y \in G$ tel que $\varphi(x) = x'$ et $\varphi(y) = y'$.
On a alors

$$\varphi^{-1}(x' \top y') = \varphi^{-1}\left(\varphi(x) \top \varphi(y)\right) = \varphi^{-1}\left(\varphi(x \star y)\right) = x \star y = \varphi^{-1}(x') \star \varphi^{-1}(y')$$

Ainsi φ^{-1} est un morphisme de groupes et il est de plus bien connu que φ^{-1} est bijective.
□

Définition

On appelle automorphisme du groupe $(G, \star)$ tout isomorphisme du groupe $(G, \star)$ dans lui-même.

Exemple Si a est un élément du groupe $(G, \star)$ alors l'application $\tau_a : G \to G$ définie par

$$\tau_a(x) = axa^{-1}$$

est un automorphisme de groupe.

Proposition

L'ensemble Aut(G) des automorphismes d'un groupe $(G, \star)$ est un sous-groupe de $(\mathcal{S}_G, \circ)$.

dém. :
Aut(G) est bien une partie de $\mathcal{S}_G$.
L'identité est automorphisme de groupe, la composée de deux automorphismes de groupe est un automorphisme de groupe et, enfin, l'application réciproque d'un automorphisme de groupe est encore un automorphisme de groupe.
□

1.4.5 Groupes isomorphes

Définition

S'il existe un isomorphisme entre deux groupes, ceux-ci sont dits isomorphes.
Ceux-ci se comportent alors de façon identique d'un point de vue calculatoire.

Exemple Les groupes $\left(\mathbb{R}^{+\star}, \times\right)$ et $(\mathbb{R}, +)$ sont isomorphes (via le logarithme népérien).
La multiplication sur $\mathbb{R}^{+\star}$ et l'addition sur $\mathbb{R}$ ont les mêmes propriétés.
En revanche les groupes $(\mathbb{R}^\star, \times)$ et $(\mathbb{R}, +)$ ne sont pas isomorphes.
En effet, l'équation $x^2 = 1$ possède deux solutions dans $(\mathbb{R}^\star, \times)$ alors que l'équation analogue $2x = 0$ n'en possède qu'une dans $(\mathbb{R}, +)$.

Exemple Comparons les tables d'opérations dans $(\mathbb{Z}/4\mathbb{Z},+)$ et $(\mathbb{U}_4,\times)$:

$$\begin{array}{c|cccc} + & \bar 0 & \bar 1 & \bar 2 & \bar 3 \\ \hline \bar 0 & \bar 0 & \bar 1 & \bar 2 & \bar 3 \\ \bar 1 & \bar 1 & \bar 2 & \bar 3 & \bar 0 \\ \bar 2 & \bar 2 & \bar 3 & \bar 0 & \bar 1 \\ \bar 3 & \bar 3 & \bar 0 & \bar 1 & \bar 2 \end{array} \text{ et } \begin{array}{c|cccc} \times & 1 & i & -1 & -i \\ \hline 1 & 1 & i & -1 & -i \\ i & i & -1 & -i & 1 \\ -1 & -1 & -i & 1 & i \\ -i & -i & 1 & i & -1 \end{array}$$

Les deux groupes $(\mathbb{Z}/4\mathbb{Z},+)$ et $(\mathbb{U}_4,\times)$ se comportent de façon semblables ; ils sont isomorphes via l'application φ qui envoie $\bar k$ sur i^k.

Exemple Considérons en revanche la table d'opérations dans $\left((\mathbb{Z}/2\mathbb{Z})^2,+\right)$:

$$\begin{array}{c|cccc} + & e & a & b & c \\ \hline e & e & a & b & c \\ a & a & e & c & b \\ b & b & c & e & a \\ c & c & b & a & e \end{array} \text{ en notant } \begin{cases} e = (\bar 0,\bar 0) \\ a = (\bar 1,\bar 0) \\ b = (\bar 0,\bar 1) \\ c = (\bar 1,\bar 1) \end{cases}$$

$\left((\mathbb{Z}/2\mathbb{Z})^2,+\right)$ se comporte d'une façon différente ; il n'est pas isomorphe aux groupes précédents.

1.5 Groupes engendré par un élément

1.5.1 Groupes monogènes

Définition

Un groupe $(G,\star)$ est dit monogène s'il existe $a\in G$ tel que $G=\langle a\rangle$.
Cet élément a est alors appelé générateur du groupe.

Remarque Un groupe monogène est nécessairement commutatif car

$$a^k\star a^\ell = a^{k+\ell} = a^\ell \star a^k$$

Exemple $(\mathbb{Z},+)$ est monogène car $\mathbb{Z}=\langle 1\rangle$.

Exemple $(\mathbb{U}_n,\times)$ est monogène car $\mathbb{U}_n=\langle\omega\rangle$ avec $\omega=\mathrm{e}^{2i\pi/n}$.

Exemple $(\mathbb{C},+)$ et $(\mathbb{C}^\star,\times)$ ne sont pas des groupes monogènes.

Exemple Pour $n\geqslant 3$, le groupe $(\mathcal{S}_n,\circ)$ n'est pas monogène car non commutatif.

1.5.2 Groupes cycliques

Définition

Un groupe est dit cyclique s'il est monogène et fini.

Exemple $(\mathbb{U}_n, \times)$ est un groupe cyclique.

Théorème

$(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe cyclique dont les générateurs sont les $\bar{m}$ pour $m \in \mathbb{Z}$ avec $m \wedge n = 1$.

dém. :
$\mathbb{Z}/n\mathbb{Z} = \langle \bar{1} \rangle$ car

$$\langle \bar{1} \rangle = \{k.\bar{1}/k \in \mathbb{Z}\} = \left\{\bar{k}/k \in \mathbb{Z}\right\} = \mathbb{Z}/n\mathbb{Z}$$

Si $\bar{m}$ est générateur de $\mathbb{Z}/n\mathbb{Z}$ alors il existe $k \in \mathbb{Z}$ tel que $k.\bar{m} = \bar{1}$ et donc $km \equiv 1 \quad [n]$. Il existe alors $\ell \in \mathbb{Z}$ tel que

$$km + n\ell = 1$$

et ainsi $m \wedge n = 1$ en vertu du théorème de Bézout.
Inversement, si $m \wedge n = 1$ alors il existe $k, \ell \in \mathbb{Z}$ tels que $km + \ell n = 1$ et donc

$$km \equiv 1 \quad [n]$$

d'où $k.\bar{m} = \bar{1}$. Ainsi $\bar{1} \in \langle \bar{m} \rangle$ or $\langle \bar{1} \rangle = \mathbb{Z}/n\mathbb{Z}$ donc

$$\langle \bar{m} \rangle = \mathbb{Z}/n\mathbb{Z}$$

□

1.5.3 Description des groupes monogènes

Théorème

Soit $(G, \star)$ un groupe monogène.
Si Card$G = +\infty$ alors $(G, \star)$ est isomorphe à $(\mathbb{Z}, +)$.
Si Card$G = n \in \mathbb{N}^\star$ alors $(G, \star)$ est isomorphisme à $(\mathbb{Z}/n\mathbb{Z}, +)$.

dém. :
Soit a un générateur de G. L'application $\varphi : \mathbb{Z} \to G$ définie par $\varphi(k) = a^k$ est un morphisme de groupes car

$$\varphi(k + \ell) = a^{k+\ell} = a^k \star a^\ell = \varphi(k) \star \varphi(\ell)$$

Il est de plus surjectif car a est générateur de G et donc

$$G = \left\{a^k/k \in \mathbb{Z}\right\}$$

Le noyau de φ est un sous-groupe de $(\mathbb{Z}, +)$. Il existe donc $n \in \mathbb{N}$ tel que $\ker \varphi = n\mathbb{Z}$.
Cas $n = 0$: φ est injectif, c'est un isomorphisme de groupes. $(G, \star)$ est alors isomorphe à $(\mathbb{Z}, +)$ et G est de cardinal infini.
Cas $n \neq 0$: On a

$$\varphi(k) = \varphi(\ell) \Leftrightarrow k - \ell \in \ker \varphi$$

donc

$$a^k = a^\ell \Leftrightarrow k \equiv \ell \quad [n]$$

On peut alors considérer l'application $\bar{\varphi} : \mathbb{Z}/n\mathbb{Z} \to G$ déterminée par $\bar{\varphi}(\bar{k}) = a^k$.
$\bar{\varphi}$ est un morphisme de groupes car

$$\bar{\varphi}(\bar{k} + \bar{\ell}) = \bar{\varphi}(\overline{k + \ell}) = a^{k+\ell} = a^k \star a^\ell = \bar{\varphi}(\bar{k}) \star \bar{\varphi}(\bar{\ell})$$

D'une part Im$\bar{\varphi} = \left\{a^k / k \in \mathbb{Z}\right\} = G$ et d'autre part

$$\bar{k} \in \ker \bar{\varphi} \Leftrightarrow a^k = a^0 \Leftrightarrow \bar{k} = \bar{0}$$

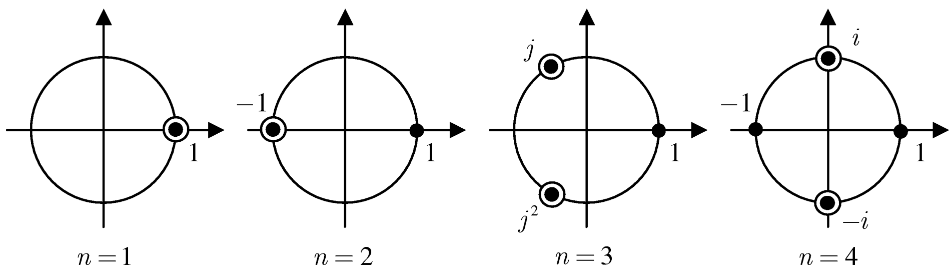
donc $\ker \bar{\varphi} = \{\bar{0}\}$. On en déduit que $\bar{\varphi}$ définit un isomorphisme.
Le groupe $(G, \star)$ est alors isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$ et en particulier G est de cardinal n.
□

Corollaire

$(\mathbb{Z}/n\mathbb{Z}, +)$ et $(\mathbb{U}_n, \times)$ sont isomorphes via l'application $\bar{k} \mapsto \omega^k = \mathrm{e}^{2ik\pi/n}$.
Les générateurs de $(\mathbb{U}_n, \times)$ sont donc les $\omega^m = \mathrm{e}^{2im\pi/n}$ avec $m \wedge n = 1$
Ces éléments sont appelés racines primitives n-ième de l'unité.

dém. :
Puisque ω est générateur de $(\mathbb{U}_n, \times)$, l'application $\bar{\varphi} : \bar{k} \mapsto \omega^k$ est un isomorphisme de groupes. Celui-ci échange les générateurs de $(\mathbb{Z}/n\mathbb{Z}, +)$ avec ceux de $(\mathbb{U}_n, \times)$.
□

Exemple Déterminons les générateurs des groupes $(\mathbb{U}_1, \times), (\mathbb{U}_2, \times), (\mathbb{U}_3, \times), (\mathbb{U}_4, \times)$.



1.5.4 Ordre d'un élément dans un groupe

Définition

On dit qu'un élément a d'un groupe $(G, \star)$ est d'ordre fini s'il existe $n \in \mathbb{N}^\star$ vérifiant $a^n = e$
On appelle alors ordre de a le plus petit $n \in \mathbb{N}^\star$ vérifiant $a^n = e$.

Exemple Dans $(\mathbb{C}^\star, \times)$, l'élément 2 n'est pas d'ordre fini.
En revanche, l'élément $\omega = \mathrm{e}^{2i\pi/n}$ est d'ordre fini égal à n.

Exemple Le neutre e est l'unique élément d'ordre fini égal à 1 du groupe $(G, \star)$.

Théorème

Si a est d'ordre fini égal à n alors

$$\forall m \in \mathbb{Z}, a^m = e \Leftrightarrow n \mid m$$

dém. :
($\Leftarrow$) immédiat.
($\Rightarrow$) Supposons $a^m = e$ et introduisons le reste r de la division euclidienne de m par n.

$$m = qn + r \text{ avec } 0 \leqslant r < n$$

On a

$$a^r = a^{m-qn} = a^m \star (a^n)^{-q} = e$$

Or n est le plus petit naturel non nul vérifiant $a^n = e$ donc $r = 0$ puis n divise m.
□

Exemple Si a est d'ordre n alors a^k est d'ordre $n/\text{pgcd}(n,k)$.

Corollaire

On a alors

$$\forall k, \ell \in \mathbb{Z}, a^k = a^\ell \Leftrightarrow k \equiv \ell \quad [n]$$

dém. :
Car

$$a^k = a^\ell \Leftrightarrow a^{k-\ell} = e$$

□

Théorème

Si a est un élément d'ordre fini d'un groupe $(G, \star)$ alors son ordre n est le cardinal du sous-groupe $\langle a \rangle$ qu'il engendre et ce dernier est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$

dém. :

$$\langle a \rangle = \left\{a^k / k \in \mathbb{Z}\right\} = \left\{e, a, \ldots, a^{n-1}\right\}$$

avec $e, a, \ldots, a^{n-1}$ deux à deux distincts.
$\langle a \rangle$ est un groupe cyclique à n éléments donc isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$ via $\bar{\varphi} : \bar{k} \mapsto a^k$.
□

1.5.5 Elément d'un groupe fini

Théorème

Si $(G, \star)$ est un groupe fini de cardinal n alors

$$\forall a \in G, a^n = e$$

dém. :
Cas $(G, \star)$ commutatif
Soit $a \in G$. L'application $\tau : x \mapsto a \star x$ est une permutation de G. On en déduit

$$\prod_{x\in G} \tau(x) = \prod_{x\in G} x$$

Or

$$\prod_{x\in G} \tau(x) = \prod_{x\in G} (a \star x) = a^{\text{Card}G} \star \prod_{x\in G} x$$

Et par conséquent

$$a^{\text{Card}G} = e$$

Cas général
On définit sur G une relation binaire $\mathcal{R}$ en posant

$$x\mathcal{R}y \Leftrightarrow \exists k \in \mathbb{Z}, y = a^k \star x$$

On vérifie aisément que $\mathcal{R}$ est une relation d'équivalence et que pour tout $x \in G$

$$Cl(x) = \{b \star x / b \in \langle a\rangle\}$$

En particulier

$$\forall x \in G, \text{Card}Cl(x) = \text{Card } \langle a\rangle$$

En notant p le nombre de classe d'équivalence de la relation $\mathcal{R}$, on obtient

$$\text{Card}G = np$$

□

Corollaire

Si $(G, \star)$ est un groupe fini alors tous ses éléments sont d'ordre fini et leur ordre divise le cardinal du groupe.

Exemple Dans $(\mathbb{Z}/6\mathbb{Z}, +)$, $\bar{0}$ est d'ordre 1, $\bar{3}$ est d'ordre 2, $\bar{2}, \bar{4}$ sont d'ordre 3 et $\bar{1}, \bar{5}$ sont d'ordre 6.

Exemple Dans un groupe à 6 éléments, il peut y a avoir des éléments d'ordre 2 et 3, mais pas d'éléments d'ordre 4.

1.5.6 Musculation : sous-groupes de $(\mathbb{Z}/n\mathbb{Z}, +)$

Exemple Montrer que les sous-groupes de $(\mathbb{Z}/n\mathbb{Z}, +)$ sont cycliques. Soit H un sous-groupe de $(\mathbb{Z}/n\mathbb{Z}, +)$.
Posons $A = \{x \in \mathbb{Z}/\bar{x} \in H\}$. On vérifie aisément que A est un sous-groupe de $(\mathbb{Z}, +)$ et donc il existe $c \in \mathbb{N}$ tel que $A = c\mathbb{Z}$. Pour $x \in \mathbb{Z}$, on a

$$\bar{x} \in H \Leftrightarrow \exists k \in \mathbb{Z}, x = kc \Leftrightarrow \exists k \in \mathbb{Z}, \bar{x} = k.\bar{c}$$

On en déduit

$$H = \langle\bar{c}\rangle$$

Exemple Montrons que $(\mathbb{Z}/n\mathbb{Z},+)$ possède un unique sous-groupe de cardinal d pour chaque d divisant n.Soit d un diviseur de n.
Posons $c = n/d$ et $H = \langle \bar{c} \rangle$. On a

$$H = \{\bar{0}, \bar{c}, 2\bar{c}, \dots, (d-1)\bar{c}\}$$

et H est un sous-groupe a exactement d éléments.
Inversement, soit H un sous-groupe à d éléments de $(\mathbb{Z}/n\mathbb{Z},+)$.
Tout élément de H d'ordre divisant d et donc

$$\forall \bar{x} \in H, d.\bar{x} = \bar{0}$$

i.e.

$$\forall \bar{x} \in H, n \mid dx$$

puis

$$\forall \bar{x} \in H, c \mid x$$

Ainsi

$$H \subset \{\bar{0}, \bar{c}, 2\bar{c}, \dots, (d-1)\bar{c}\}$$

et l'égalité est acquise par cardinalité.

Chapitre 2

Anneaux

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

2.1 Structure d'anneau

2.1.1 Définition

Définition

On appelle anneau tout triplet $(A, +, \times)$ formé d'un ensemble A et de deux lois de composition internes usuellement notées $+$ et $\times$ sur A vérifiant :
1) $(A, +)$ est un groupe abélien de neutre 0_A ;
2) $\times$ est associative et possède un neutre 1_A ;
3) $\times$ est distributive sur $+$ i.e.

$$\forall a, b, c \in A, a(b + c) = ab + ac \text{ et } (b + c)a = ba + ca$$

Si de plus la loi $\times$ est commutative, on dit que l'anneau $(A, +, \times)$ est commutatif.

Exemple $(\mathbb{Z}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$ sont des anneaux commutatifs de neutres 0 et 1.

Exemple Soit X un ensemble et $\mathcal{F}(X, \mathbb{K})$ l'ensemble des fonctions de X vers $\mathbb{K}$.
$(\mathcal{F}(X, \mathbb{K}), +, \times)$ est un anneau de neutres $\tilde{0}$ et $\tilde{1}$ (fonctions constantes).
En particulier, si $X = \mathbb{N}$, l'ensemble $\mathbb{K}^{\mathbb{N}}$ des suites d'éléments de $\mathbb{K}$ est un anneau.

Exemple $(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau de neutres O_n et I_n.

Exemple Si E est un $\mathbb{K}$-espace vectoriel, $(\mathcal{L}(E), +, \circ)$ est un anneau de neutres $\tilde{0}$ et Id_E.

Exemple $A = \{0_A\}$ est un anneau (c'est le seul pour lequel $1_A = 0_A$).

2.1.2 Calculs dans un anneau

Proposition

On a
$$\forall a, b \in A,\, 0_A \times a = a \times 0_A = 0_A,\, (-a) \times b = -(ab) = a \times (-b)$$
Plus généralement
$$\forall n \in \mathbb{Z},\, (n.a) \times b = n.(ab) = a \times (n.b)$$

Théorème

Si a et b sont deux éléments commutant (i.e. $ab = ba$) d'un anneau A on a pour tout $n \in \mathbb{N}$
$$(ab)^n = a^n b^n,\, (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$
et
$$a^n - b^n = (a-b) \sum_{k=0}^{n-1} a^k b^{n-1-k}$$

2.1.3 Groupe des inversibles

Définition

Un élément a d'un anneau $(A, +, \times)$ est dit inversible s'il existe $b \in A$ tel que
$$ab = ba = 1$$
Cet élément b est alors unique, on l'appelle inverse de a et il est noté a^{-1}.

Exemple 1_A est inversible et $1_A^{-1} = 1_A$.

Exemple Si A n'est pas l'anneau nul, 0_A n'est pas inversible.

Exemple Si $x \in A$ est inversible alors x^{-1} aussi et $(x^{-1})^{-1} = x$.
Si x et $y \in A$ sont inversibles alors xy est inversible et $(xy)^{-1} = y^{-1}x^{-1}$.

Théorème

L'ensemble $U(A)$ des éléments inversibles de l'anneau $(A, +, \times)$ est un groupe multiplicatif.

Exemple $U(\mathbb{Z}) = \{1, -1\}, U(\mathbb{K}) = \mathbb{K}^\star$,
$U(\mathcal{M}_n(\mathbb{K})) = \mathrm{GL}_n(\mathbb{K})$ et $U(\mathcal{L}(E)) = \mathrm{GL}(E)$.

2.1.4 Produit fini d'anneaux

Soit $(A_1,+,\times)$,..., $(A_n,+,\times)$ des anneaux et $A = A_1 \times \ldots \times A_n$.
On définit des lois $+$ et $\times$ sur A en posant

$$(x_1,\ldots,x_n)+(y_1,\ldots,y_n) \underset{\text{déf}}{=} (x_1+y_1,\ldots,x_n+y_n)$$

et

$$(x_1,\ldots,x_n)\times(y_1,\ldots,y_n) \underset{\text{déf}}{=} (x_1\times y_1,\ldots,x_n\times y_n)$$

Théorème

L'ensemble A muni des lois $+$ et $\times$ définies ci-dessus est un anneau de neutres

$$0_A = (0_{A_1},\ldots,0_{A_n}) \text{ et } 1_A = (1_{A_1},\ldots,1_{A_n})$$

De plus, un élément $(a_1,\ldots,a_n) \in A$ est inversible si, et seulement si, les $a_1,\ldots,a_n$ le sont et son inverse est alors $(a_1^{-1},\ldots,a_n^{-1})$.

Corollaire

$U(A) = U(A_1) \times \ldots \times U(A_n)$.

Exemple $(A^n,+,\times)$ est un anneau de neutre $0_{A^n} = (0_A,\ldots,0_A)$ et $1_{A^n} = (1_A,\ldots,1_A)$.

Exemple $(\mathbb{Z}^2,+,\times)$ est un anneau commutatif où

$$(a,b)+(c,d) = (a+c,b+d) \text{ et } (a,b)\times(c,d) = (ac,bd)$$

On a

$$U\left(\mathbb{Z}^2\right) = \{(1,1),(1,-1),(-1,1),(-1,-1)\}$$

2.1.5 Sous-anneau

$(A,+,\times)$ désigne un anneau

Définition

On appelle sous-anneau de $(A,+,\times)$ toute partie B de A vérifiant :
1) $1_A \in B$;
2) $\forall x,y \in B, x-y \in B$;
3) $\forall x,y \in B, xy \in B$.

Attention : Vérifier $1_A \in B$ et non $0_A \in B$ ou seulement $B \neq \emptyset$.

Exemple $\mathbb{Z}$ est un sous-anneau de $(\mathbb{R},+,\times)$ mais pas $2\mathbb{Z}$ bien que stable par différence et produit

Exemple A est un sous-anneau de $(A,+,\times)$, mais généralement pas $\{0_A\}$.

Exemple On note $\mathcal{C}$ l'ensemble des suites réelles convergentes.
Montrons que $\mathcal{C}$ est un sous-anneau de $(\mathbb{R}^{\mathbb{N}}, +, \times)$.
On a évidemment $\mathcal{C} \subset \mathbb{R}^{\mathbb{N}}$, la suite constante égale à 1 est convergente et la différence et le produit de deux suites convergentes sont des suites convergentes.
En revanche, l'ensemble des suites réelles convergeant vers 0 n'est pas un sous-anneau.

Exemple Soit I un intervalle de $\mathbb{R}$ d'intérieur non vide et $k \in \mathbb{N} \cup \{\infty\}$.
Vérifions que $\mathcal{C}^k(I, \mathbb{R})$ est un sous-anneau de $(\mathcal{F}(I, \mathbb{R}), +, \times)$.
On a évidemment $\mathcal{C}^k(I, \mathbb{R}) \subset \mathcal{F}(I, \mathbb{R})$, la fonction constante égale à 1 est de classe $\mathcal{C}^k$ et la différence et le produit de deux fonctions de classe $\mathcal{C}^k$ sont des fonctions de classe $\mathcal{C}^k$.

Théorème

Si B est un sous-anneau de $(A, +, \times)$ alors B peut être muni des lois $+$ et $\times$ définies par restriction des lois sur A et $(B, +, \times)$ est alors un anneau de mêmes neutres que A.

dém. :
B est un sous-groupe du groupe abélien $(A, +)$ donc $(B, +)$ est un groupe abélien.
B est stable par $\times$ donc on peut définir la restriction de la loi $\times$ sur B.
Celle-ci est associative sur A et possède un neutre $1_A \in B$ donc $\times$ est associative sur B et y possède un neutre.
Enfin, $\times$ est distributive sur $+$ sur A donc a fortiori aussi sur B.
$\square$

Exemple Considérons
$$\mathbb{Z}[i] = \{a + ib / a, b \in \mathbb{Z}\}$$
et montrons que $(\mathbb{Z}[i], +, \times)$ est un anneau commutatif.
Montrons que $\mathbb{Z}[i]$ un sous-anneau de l'anneau commutatif $(\mathbb{C}, +, \times)$.
On a évidemment $\mathbb{Z}[i] \subset \mathbb{C}$.
$1 = 1 + i.0 \in \mathbb{Z}[i]$.
Pour $x, y \in \mathbb{Z}[i]$, on peut écrire $x = a + ib$ et $y = c + id$ avec $a, b, c, d \in \mathbb{Z}$.
On a
$$x - y = (a - c) + i(b - d) \in \mathbb{Z}[i]$$
car $a - c, b - d \in \mathbb{Z}$
et
$$xy = (ac - bd) + i(ad + bc) \in \mathbb{Z}[i]$$
Ainsi, $\mathbb{Z}[i]$ est un sous-anneau de $(\mathbb{C}, +, \times)$ et donc $(\mathbb{Z}[i], +, \times)$ est un anneau commutatif.

2.1.6 L'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$

Théorème

$(\mathbb{Z}/n\mathbb{Z}, +, \times)$ est un anneau commutatif de neutres $\bar{0}$ et $\bar{1}$.
De plus, dans $(\mathbb{Z}/n\mathbb{Z}, +, \times)$, $\bar{m}$ est inversible si, et seulement si, $m \wedge n = 1$.

dém. :
$(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe abélien de neutre $\bar{0}$.

On vérifie aisément que la loi $\times$ est commutative, associative sur $\mathbb{Z}/n\mathbb{Z}$ et possède un neutre $\bar{1}$. On vérifie aussi que la loi $\times$ est distributive sur $+$.
Soit $\bar{m} \in \mathbb{Z}/n\mathbb{Z}$.
$\bar{m}$ inversible si, et seulement si, il existe $\bar{k} \in \mathbb{Z}/n\mathbb{Z}$ vérifiant $\bar{k}\bar{m} = \bar{1}$ i.e. si, et seulement si, il existe $k \in \mathbb{Z}$ tel que $km \equiv 1 \quad [n]$. Ainsi $\bar{m}$ est inversible si, et seulement si, il existe $k, \ell \in \mathbb{Z}$ tels que

$$km + \ell n = 1$$

Par le théorème de Bézout, cela revient à affirmer $m \wedge n = 1$.
□

Remarque Si $m \wedge n = 1$ alors une égalité de Bézout $um + vn = 1$ fournit $\bar{m}^{-1} = \bar{u}$.

Exemple Résolvons l'équation $4x + 2 \equiv 0 \quad [11]$
Dans $\mathbb{Z}/11\mathbb{Z}$ l'équation dévient

$$\bar{4}\bar{x} + \bar{2} = \bar{0}$$

Par opérations

$$\bar{4}\bar{x} + \bar{2} = \bar{0} \Leftrightarrow \bar{4}\bar{x} = \bar{9}$$

Puisque $4 \wedge 11 = 1$, $\bar{4}$ est inversible dans $\mathbb{Z}/11\mathbb{Z}$ et on observe

$$\bar{4}^{-1} = \bar{3}$$

On a alors

$$\bar{4}\bar{x} = \bar{9} \Leftrightarrow \bar{x} = \bar{3} \times \bar{9}$$

Ainsi

$$\bar{4}\bar{x} + \bar{2} = \bar{0} \Leftrightarrow \bar{x} = \bar{5}$$

Les solutions de l'équation étudiées sont donc les $5 + 11k$ avec $k \in \mathbb{Z}$.

Exemple Résolvons l'équation $4x \equiv 6 \quad [10]$
Ici 4 et 10 ne sont pas premiers entre eux, mais l'équation est simplifiable par leur PGCD

$$4x \equiv 6 \quad [10] \Leftrightarrow \exists k \in \mathbb{Z}, 4x = 6 + 10k \Leftrightarrow \exists k \in \mathbb{Z}, 2x = 3 + 5k$$

ce qui nous ramène à l'équation $2x \equiv 3 \quad [5]$ avec $2 \wedge 5 = 1$ qu'on peut résoudre.

$$2x \equiv 3 \quad [5] \Leftrightarrow x \equiv 3 \times 3 = 4 \quad [5]$$

Les solutions sont les $4 + 5k$ avec $k \in \mathbb{Z}$.

Exemple Résolvons l'équation $4x \equiv 7 \quad [10]$
Ici 4 et 10 ne sont pas premiers entre eux et l'équation n'est pas simplifiable : il n'y a pas de solutions.

2.1.7 Anneaux intègres

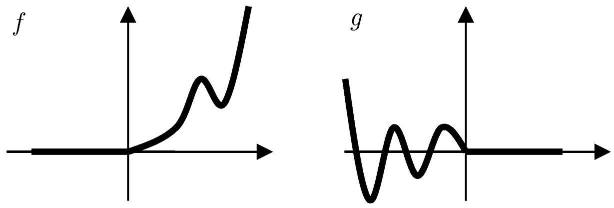
Soit $(A, +, \times)$ un anneau.

2.1.7.1 Diviseurs de zéro

Attention : On sait
$$\forall a, b \in A, a = 0_A \text{ ou } b = 0_A \Rightarrow ab = 0_A$$
La réciproque n'est pas toujours vraie !

Exemple Dans l'anneau $(\mathbb{Z}^2, +, \times)$, on a $(1, 0) \times (0, 1) = (0, 0)$ alors que $(1, 0), (0, 1) \neq (0, 0)$

Exemple Dans l'anneau $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +, \times)$, considérons les fonctions données par



On a $fg = \tilde{0}$ alors que $f, g \neq \tilde{0}$.

Exemple Dans $(\mathcal{M}_2(\mathbb{R}), +, \times)$, pour
$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ et } B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$
on a $AB = O_2$ alors que $A, B \neq O_2$.

Exemple Dans $(\mathbb{Z}/6\mathbb{Z}, +, \times)$, $\bar{2} \times \bar{3} = \bar{0}$ alors que $\bar{2}, \bar{3} \neq \bar{0}$.

Définition

Lorsque $a, b \in A$ vérifient $ab = 0_A$ avec $a, b \neq 0_A$, on dit que a et b sont des diviseurs de zéro.

Attention : On ne considère pas que 0_A est un diviseur de zéro.

Exemple En général, les anneaux $\mathcal{F}(X, \mathbb{K})$, $\mathcal{L}(E)$ et $\mathcal{M}_n(\mathbb{K})$ possèdent des diviseurs de zéros.

Exemple Les éléments inversibles d'un anneau ne sont pas diviseurs de zéros.
En effet, si $ab = 0_A$ avec a inversible alors
$$b = a^{-1} \times (ab) = a^{-1} \times 0_A = 0_A$$

Exemple Dans $(\mathbb{R}^2, +, \times)$ les diviseurs de zéros sont les $(x, 0)$ et $(0, x)$ avec $x \neq 0$.

2.1.7.2 Intégrité

Définition

Un anneau $(A, +, \times)$ est dit intègre si
1) A non réduit à $\{0_A\}$;
2) A ne possède pas de diviseurs de zéros.

Exemple $(\mathbb{Z}, +, \times)$ est un anneau intègre.

Proposition

Dans un anneau intègre $(A, +, \times)$

$$\forall a, b \in A, ab = 0_A \Rightarrow a = 0_A \text{ ou } b = 0_A$$

dém. :
C'est l'absence de diviseurs de zéro !
□

Proposition

Dans un anneau intègre $(A, +, \times)$:

$$\forall a, b, c \in A, (ab = ac \text{ et } a \neq 0_A) \Rightarrow b = c$$

et

$$\forall a, b, c \in A, (ba = ca \text{ et } a \neq 0_A) \Rightarrow b = c$$

dém. :
Si $ab = ac$ alors $ab - ac = 0_A$ et donc $a(b - c) = 0_A$.
Si de plus $a \neq 0_A$ alors, par intégrité, $b - c = 0_A$ et donc $b = c$.
□

Remarque Dans un anneau intègre l'équation $x^2 = 1$ a pour seules solutions 1 et -1 car

$$x^2 = 1_A \Leftrightarrow (x - 1_A)(x + 1_A) = 0_A$$

Dans $(\mathbb{R}^2, +, \times)$, l'équation $x^2 = 1_{\mathbb{R}^2}$ a pour solutions

$$(1, 1), (-1, -1), (1, -1), (-1, 1)$$

Dans $(\mathcal{M}_2(\mathbb{R}), +, \times)$, l'équation $A^2 = I_2$ a pour solutions

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & -3 \\ 1 & -2 \end{pmatrix}, \dots$$

2.1.7.3 Idempotence et nilpotence

Définition

Un élément $a \in A$ est dit idempotent si $a^2 = a$.

Exemple Dans un anneau intègre seuls 0_A et 1_A sont idempotents.

Exemple Dans $(\mathbb{R}^2, +, \times)$, $(1,0)$ et $(0,1)$ sont aussi idempotents.

Exemple Dans $(\mathbb{Z}/6\mathbb{Z}, +, \times)$, l'élément $\bar{3}$ est idempotent.

Exemple Dans $(\mathcal{L}(E), +, \circ)$ les éléments idempotents sont les projecteurs.

Définition

Un élément $a \in A$ est dit nilpotent s'il existe $n \in \mathbb{N}^\star$ tel que $a^n = 0_A$.

Exemple Dans un anneau intègre seul 0_A est nilpotent.

Exemple Dans $(\mathbb{Z}/8\mathbb{Z}, +, \times)$, l'élément $\bar{2}$ est nilpotent.

Exemple Montrons que si a est nilpotent alors $1_A - a \in U(A)$.
Puisque a est nilpotent, il existe $n \in \mathbb{N}$ vérifiant $a^n = 0_A$.
Puisque 1_A et a commutent,

$$1_A = 1_A^n - a^n = (1-a)\left(\sum_{k=0}^{n-1} a^k\right) = \left(\sum_{k=0}^{n-1} a^k\right)(1-a)$$

Ainsi, $1_A - a$ est inversible et

$$(1_A - a)^{-1} = \sum_{k=0}^{n-1} a^k$$

2.2 Corps

2.2.1 Définition

Définition

On appelle corps tout anneau $(K, +, \times)$ vérifiant
1) $(K, +, \times)$ est commutatif ;
2) K est non réduit à $\{0_K\}$ et
3) tous les éléments de K, sauf le nul, sont inversibles.

Exemple $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$ et $(\mathbb{K}(X), +, \times)$ sont des corps usuels.

Proposition

> Tout corps est intègre.

dém. :
Soit K un corps. K est commutatif et non réduit à $\{0_K\}$.
Pour $a, b \in K$, si $ab = 0_K$ et $a \neq 0_K$ alors on peut introduire a^{-1} et on a $b = a^{-1}(ab) = 0_K$.
Ainsi, K ne possède pas de diviseurs de zéro. Il est donc intègre.
□

2.2.2 Sous-corps

Soit $(K, +, \times)$ un corps.

Définition

> On appelle sous-corps d'un corps $(K, +, \times)$ toute partie L de K vérifiant :
> 1) L est un sous-anneau de $(K, +, \times)$;
> 2) $\forall x \in L, x \neq 0_K \Rightarrow x^{-1} \in L$.

Exemple $\mathbb{Q}$ est un sous-corps de $(\mathbb{R}, +, \times)$.

Théorème

> Si L est un sous-corps de $(K, +, \times)$ alors $(L, +, \times)$ est un corps.

dém. :
Puisque L est un sous-anneau de l'anneau commutatif $(K, +, \times)$, on peut affirmer que $(L, +, \times)$ est un anneau commutatif. Puisque $1_K \in L$, on peut affirmer que l'anneau $(L, +, \times)$ n'est pas réduit à 0. Enfin, puisque l'inverse d'un élément non nul de L est élément de L, on peut affirmer que tout élément non nul de l'anneau L est inversible dans celui-ci.
□

Exemple Considérons $\mathbb{Q}\left[\sqrt{2}\right] = \left\{a + b\sqrt{2}/a, b \in \mathbb{Q}\right\}$.
Montrons que $(\mathbb{Q}\left[\sqrt{2}\right], +, \times)$ est un corps.
Pour cela montrons que $\mathbb{Q}\left[\sqrt{2}\right]$ est un sous-corps du corps $(\mathbb{R}, +, \times)$.
On a évidemment $\mathbb{Q}\left[\sqrt{2}\right] \subset \mathbb{R}$.
$1 = 1 + 0 \times \sqrt{2} \in \mathbb{Q}\left[\sqrt{2}\right]$.
Pour $x, y \in \mathbb{Q}[i]$, on peut écrire $x = a + b\sqrt{2}$ et $y = c + d\sqrt{2}$ avec $a, b, c, d \in \mathbb{Q}$.
On a alors

$$x - y = (a - c) + (b - d)\sqrt{2} \in \mathbb{Q}\left[\sqrt{2}\right]$$

et

$$xy = (ab + 2dc) + \sqrt{2}(ad + bc) \in \mathbb{Q}\left[\sqrt{2}\right]$$

Enfin, si $x \neq 0$,

$$x^{-1} = \frac{1}{a + b\sqrt{2}} = \frac{a - b\sqrt{2}}{(a + b\sqrt{2})(a - b\sqrt{2})} = \frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2} \in \mathbb{Q}\left[\sqrt{2}\right]$$

car $\dfrac{a}{a^2 + b^2}, -\dfrac{b}{a^2 + b^2} \in \mathbb{Q}$.

2.2.3 Le corps $(\mathbb{Z}/p\mathbb{Z}, +, \times)$

Théorème

$(\mathbb{Z}/p\mathbb{Z}, +, \times)$ est un corps si, et seulement si, p est un nombre premier.

dém. :
Supposons que $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ soit un corps.
Pour tout $a \in \{2, \dots, p-1\}$, $\bar{a}$ est inversible dans $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ donc $a \wedge p = 1$ et par conséquent a ne divise pas p. On en déduit que p est un nombre premier.
Inversement, supposons p nombre premier.
$(\mathbb{Z}/p\mathbb{Z}, +, \times)$ est un anneau commutatif et $\mathbb{Z}/p\mathbb{Z} \neq \{\bar{0}\}$ car $p = \text{Card}(\mathbb{Z}/p\mathbb{Z}) \geqslant 2$.
Pour tout $\bar{m} \in \mathbb{Z}/p\mathbb{Z}$, si $\bar{m} \neq \bar{0}$ alors p ne divise pas m et donc, puisque p est un nombre premier,

$$m \wedge p = 1$$

On en déduit que $\bar{m}$ est inversible.
□

Remarque On note usuellement $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Exemple Soit $\mathbb{F}_2 = \{\bar{0}, \bar{1}\}$. $(\mathbb{F}_2, +, \times)$ est un corps pour les opérations suivantes

$+$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

et

$\times$	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$

Exemple Soit $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$. $(\mathbb{F}_3, +, \times)$ est un corps pour les opérations suivantes

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{0}$	$\bar{1}$

et

$\times$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{1}$

2.3 Morphismes d'anneaux

Soit $(A, +, \times)$ et $(A', +, \times)$ des anneaux.

2.3.1 Morphisme d'anneaux

Définition

On dit qu'une application $\varphi : A \to A'$ est un morphisme d'anneaux si
1) $\varphi(1_A) = 1_{A'}$;
2) $\forall x, y \in A, \varphi(x + y) = \varphi(x) + \varphi(y)$;
3) $\forall x, y \in A, \varphi(xy) = \varphi(x)\varphi(y)$.

Exemple L'application identité $\text{Id}_A : A \to A$ est un morphisme de l'anneau $(A, +, \times)$ vers lui-même.

Exemple Considérons $\mathcal{C}$ l'anneau des suites réelles convergentes.
L'application $\varphi : u \mapsto \lim\limits_{n\to+\infty} u_n$ est un morphisme d'anneaux de $\mathcal{C}$ vers $\mathbb{R}$.

Exemple L'application $\varphi : \mathbb{Z} \to A$ définie par $\varphi(k) = k.1_A$ est un morphisme d'anneaux de $(\mathbb{Z}, +, \times)$ vers $(A, +, \times)$.
En effet, $\varphi(1) = 1_A$, $\varphi(k+\ell) = (k+\ell).1_A = k.1_A + \ell.1_A = \varphi(k) + \varphi(\ell)$ et
$\varphi(k\ell) = (k\ell).1_A = (k.1_A) \times (\ell.1_A) = \varphi(k)\varphi(\ell)$.

Exemple Soit $a \in U(A)$ et $\tau : A \to A$ définie par $\tau(x) = axa^{-1}$.
Vérifions que τ est un morphisme d'anneaux bijectif.
$\tau(1_A) = a.1_A.a^{-1} = 1_A$, $\tau(x+y) = a(x+y)a^{-1} = axa^{-1} + aya^{-1} = \tau(x) + \tau(y)$ et
$\tau(xy) = axya^{-1} = ax(a^{-1}a)ya^{-1} = \tau(x)\tau(y)$.
Enfin,

$$y = \tau(x) \Leftrightarrow x = a^{-1}ya$$

donc

$\forall y \in A, \exists! x \in A, y = \tau(x)$
L'application τ est donc bijective.

Attention : Ne pas oublier d'étudier $\varphi(1_A)$!
L'application $x \in \mathbb{R} \mapsto (x, 0) \in \mathbb{R}^2$ n'est pas un morphisme d'anneaux !

2.3.2 Propriétés

Proposition

La composée de deux morphismes d'anneaux est un morphisme d'anneaux.

Proposition

Si $\varphi : A \to A'$ est un morphisme d'anneaux alors
a) $\varphi(0_A) = 0_{A'}$;
b)$\forall x \in A, \varphi(-x) = -\varphi(x)$;
c) $\forall x \in A, \forall n \in \mathbb{Z}, \varphi(n.x) = n.\varphi(x)$;
d) $\forall x \in A, \forall n \in \mathbb{N}, \varphi(x^n) = \varphi(x)^n$;
e) $\forall x \in A, x \in U(A) \Rightarrow \varphi(x) \in U(A')$ avec $\varphi(x)^{-1} = \varphi(x^{-1})$

dém. :
φ est un morphisme du groupe $(A, +)$ vers $(A', +)$ donc

$$\forall x \in A, \forall n \in \mathbb{Z}, \varphi(n.x) = n.\varphi(x)$$

Par récurrence, on obtient aisément

$$\forall x \in A, \forall n \in \mathbb{N}, \varphi(x^n) = \varphi(x)^n$$

Enfin, si $x \in U(A)$ alors $\varphi(xx^{-1}) = \varphi(1_A)$ donne $\varphi(x)\varphi(x^{-1}) = 1_{A'}$. Aussi $\varphi(x^{-1})\varphi(x) = 1_{A'}$ donc $\varphi(x) \in U(A')$ et $\varphi(x)^{-1} = \varphi(x^{-1})$.
□

2.3.3 Image et noyaux

Définition

Soit $\varphi : A \to A'$ un morphisme d'anneaux.
On appelle image et noyau du morphisme φ les ensembles

$$\mathrm{Im}\varphi = \varphi(A) \text{ et } \ker\varphi = \varphi^{-1}(\{0_{A'}\})$$

Remarque Ce sont en fait les images et noyaux de φ en tant que morphisme de groupes additifs.

Remarque On vérifie aisément que $\mathrm{Im}\varphi$ est un sous-anneaux de A'.
En revanche, $\ker\varphi$ n'est généralement pas un sous-anneau de $(A, +, \times)$.

Proposition

φ est injective si, et seulement si, $\ker\varphi = \{0_A\}$.
φ est surjective si, et seulement si, $\mathrm{Im}\varphi = A'$.

dém. :
Car φ est en particulier un morphisme de groupes additifs.
□

2.3.4 Isomorphisme d'anneaux

Définition

On dit qu'une application $\varphi : A \to A'$ est un isomorphisme d'anneaux si
a) φ est un morphisme d'anneaux ;
b) φ est bijective.

Proposition

La composée de deux isomorphismes d'anneaux est un isomorphisme d'anneaux.
L'application réciproque d'un isomorphisme d'anneaux et un isomorphisme d'anneaux.

Définition

On dit que deux anneaux A et A' sont isomorphes s'il existe un isomorphisme d'anneaux de l'un vers l'autre : ces deux anneaux possèdent alors les mêmes propriétés calculatoires.

Exemple Considérons $\varphi : \mathbb{C} \to \mathcal{M}_2(\mathbb{R})$ définie par

$$\varphi(a + i.b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

On vérifie que φ est un morphisme d'anneaux injectifs.

En conséquence

$$\mathrm{Im}\varphi = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} / a, b \in \mathbb{R}^2 \right\}$$

est un sous-anneau de $\mathcal{M}_2(\mathbb{R})$ isomorphe à $(\mathbb{C}, +, \times)$.

2.3.5 Théorème des restes chinois

Soit m et n deux entiers naturels non nuls. Pour $k \in \mathbb{Z}$, on note on note $\bar{k}$, $\hat{k}$ et $\dot{k}$ les classes d'équivalence de k dans $\mathbb{Z}/mn\mathbb{Z}$, $\mathbb{Z}/m\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z}$.

Théorème

> Si m et n sont premiers entre eux alors l'application
>
> $$\pi : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$
>
> définie par
>
> $$\pi(\bar{k}) = (\dot{k}, \hat{k})$$
>
> est un isomorphisme d'anneaux.

dém. :
L'application est bien définie car

$$k = \ell \quad [mn] \Rightarrow k = \ell \quad [m] \text{ et } k = \ell \quad [n]$$

et ainsi

$$\bar{k} = \bar{\ell} \Rightarrow \hat{k} = \hat{\ell} \text{ et } \dot{k} = \dot{\ell}$$

On vérifie aisément que cette application est un morphisme d'anneaux.
Etudions le noyau de π.
Si $\bar{x} \in \ker \pi$ alors $\pi(\bar{x}) = (\dot{0}, \hat{0})$ i.e. $\bar{x} = \bar{0}$ et $\dot{x} = \dot{0}$. On alors $m \mid x$ et $n \mid x$ donc $mn \mid x$ puisque $m \wedge n = 1$. Ainsi $\bar{x} = \bar{0}$ ce qui permet d'affirmer $\ker \pi = \{\bar{0}\}$.
Le morphisme π est donc injectif.
Puisque

$$\mathrm{Card}(\mathbb{Z}/nm\mathbb{Z}) = nm = \mathrm{Card}(\mathbb{Z}/n\mathbb{Z})\mathrm{Card}(\mathbb{Z}/m\mathbb{Z}) < +\infty$$

on peut affirmer par cardinalité que π est bijective et finalement π est un isomorphisme.
□

Remarque Soit à résoudre un système du type

$$\begin{cases} x \equiv a & [m] \\ x \equiv b & [n] \end{cases}$$

avec $m \wedge n = 1$. Par ce qui précède, ce système possède une unique solution modulo mn.
Pour la déterminer, il suffit de trouver x_1 et x_2 solutions respectives des systèmes

$$\begin{cases} x \equiv 1 & [m] \\ x \equiv 0 & [n] \end{cases} \text{ et } \begin{cases} x \equiv 0 & [m] \\ x \equiv 1 & [n] \end{cases}$$

Par morphisme, $x = ax_1 + bx_2$ est alors solution du système initial.
Pour déterminer x_1 et x_2, on part de la relation de Bézout

$$mu + nv = 1$$

et l'on prend $x_1 = nv$ et $x_2 = mu$.

Exemple Résolvons le système
$$\begin{cases} x \equiv 1 & [5] \\ x \equiv 7 & [9] \end{cases}$$
$5 \wedge 9 = 1$ avec la relation de Bézout $2 \times 5 - 9 = 1$.
-9 et 10 sont solutions des systèmes
$$\begin{cases} x \equiv 1 & [5] \\ x \equiv 0 & [9] \end{cases} \text{ et } \begin{cases} x \equiv 0 & [5] \\ x \equiv 1 & [9] \end{cases}$$
donc $x = 1 \times (-9) + 7 \times 10 = 61$ est solution du système posé.
La solution générale est alors
$$16 + 45k \text{ avec } k \in \mathbb{Z}$$

Exemple Résolvons le système
$$\begin{cases} 9x \equiv 3 & [21] \\ 5x \equiv 2 & [8] \end{cases}$$
$$9x \equiv 3 \quad [21] \Leftrightarrow 3x \equiv 1 \quad [7]$$
Puisque $3 \wedge 7 = 1$, $\bar{3}$ est inversible et $\bar{3}^{-1} = \bar{5}$ dans $\mathbb{Z}/7\mathbb{Z}$.
Ainsi
$$3x \equiv 1 \quad [7] \Leftrightarrow x \equiv 5 \quad [7]$$
De même
$$5x \equiv 2 \quad [8] \Leftrightarrow x \equiv 2 \quad [8]$$
car $\bar{5}^{-1} = \bar{5}$ dans $\mathbb{Z}/8\mathbb{Z}$
Ainsi
$$\begin{cases} 9x \equiv 3 & [21] \\ 5x \equiv 2 & [8] \end{cases} \Leftrightarrow \begin{cases} x \equiv 5 & [7] \\ x \equiv 2 & [8] \end{cases}$$
$7 \wedge 8 = 1$ avec la relation de Bézout $(-1) \times 7 + 8 = 1$.
$x = 5 \times 8 + 2 \times (-7) = 26$ est solution de ce système dont la solution générale est
$$x = 26 + 56k \text{ avec } k \in \mathbb{Z}$$

2.4 Idéal d'un anneau commutatif

Soit $(A, +, \times)$ un anneau commutatif.

2.4.1 Définition

Définition

On appelle idéal de l'anneau $(A, +, \times)$ toute partie I de A vérifiant :
1) $0_A \in I$;
2) $\forall x, y \in I, x + y \in I$;
3) $\forall a \in A, \forall x \in I, ax \in I$ [absorption].

Remarque Un idéal est en particulier un sous-groupe additif (il suffit d'exploiter l'absorption avec $a = -1$)

Exemple $\{0_A\}$ et A sont des idéaux de $(A, +, \times)$.

Exemple $n\mathbb{Z}$ est un idéal de $(\mathbb{Z}, +, \times)$.

Exemple Le noyau d'un morphisme d'anneaux $\varphi : A \to A'$ est un idéal de $(A, +, \times)$.
En effet, $\ker \varphi \subset A$, $0_A \in \ker \varphi$ car $\varphi(0_A) = 0_{A'}$.
Soit $x, y \in \ker \varphi$.
$\varphi(x + y) = \varphi(x) + \varphi(y) = 0_{A'} + 0_{A'} = 0_{A'}$ donc $x + y \in \ker \varphi$.
Soit de plus $a \in A$.
$\varphi(ax) = \varphi(a)\varphi(x) = \varphi(a) \times 0_{A'} = 0_{A'}$ donc $ax \in \ker \varphi$.

Proposition

Soit I un idéal de l'anneau $(A, +, \times)$
Si $1_A \in I$ alors $I = A$.
Si $I \cap U(A) \neq \emptyset$ alors $I = A$.

dém. :
Par absorption $1_A \in I$ entraîne $A \subset I$ puis $=$.
De même, par absorption, $I \cap U(A) \neq \emptyset$ entraîne $1_A \in I$ puis $I = A$.
□

Remarque Les seuls idéaux d'un corps sont $\{0_K\}$ et lui-même.

2.4.2 Opérations

Proposition

Si I et J sont deux idéaux de $(A, +, \times)$ alors $I \cap J$ est un idéal.
De plus, $I \cap J$ est inclus dans I et J et contient tout idéal inclus dans I et J.

dém. :
$I \cap J \subset A$, $0_A \in I$ et $0_A \in J$ donc $0_A \in I \cap J$.
Si $x, y \in I \cap J$ alors $x, y \in I$ donc $x + y \in I$. De même $x + y \in J$ donc $x + y \in I \cap J$.
Si $a \in A$ et $x \in I \cap J$ alors $x \in I$ donc $ax \in I$. De même $ax \in J$ donc $ax \in I \cap J$.
□

Proposition

Si I et J sont deux idéaux de $(A, +, \times)$ alors

$$I + J \underset{\text{déf}}{=} \{x + y / x \in I, y \in J\}$$

est un idéal.
De plus, $I + J$ contient I et J et est inclus dans tout idéal contenant I et J.

dém. :
Pour $x \in I$, $x = x + 0_A \in I + J$ car $0_A \in J$. Ainsi $I \subset I + J$ et de même $J \subset I + J$.
$0_A \in I + J$ car $0_A = 0_A + 0_A$ avec $0_A \in I, J$.

Pour $x, y \in I + J$, on peut écrire $x = x' + x''$ et $y = y' + y''$ avec $x', y' \in I$ et $x'', y'' \in J$.
On a alors $x + y = (x' + y') + (x'' + y'') \in I + J$ car $x' + y' \in I$ et $x'' + y'' \in J$.
Enfin, pour $a \in A$, $ax = (ax') + (ax'') \in I + J$ car $ax' \in I$ et $ax'' \in J$.
De plus, si K est un idéal contenant I et J alors K contient $I + J$ car stable pour l'addition.
□

2.4.3 Idéal engendré par un élément

Définition

On appelle idéal engendré par $x \in A$ l'ensemble

$$xA \underset{\text{déf}}{=} \{xu/u \in A\}$$

Théorème

xA est un idéal contenant l'élément x et inclus dans tout idéal contenant x.

dém. :
$x = x \times 1 \in xA$ et si I est un idéal contenant x alors par absorption, il contient xA.
Il reste à montrer que xA est un idéal.
On a $xA \subset A$ et $0_A = x \times 0_A \in xA$.
Pour $y, z \in xA$, on peut écrire $y = xu$ et $z = xv$ avec $u, v \in A$ et alors $y + z = x(u + v) \in xA$.
Enfin, pour $a \in A$, $ay = x(au) \in xA$.
□

2.4.4 Idéaux de $(\mathbb{Z}, +, \times)$

Théorème

Les idéaux de $(\mathbb{Z}, +, \times)$ sont de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.

dém. :
Les idéaux de $(\mathbb{Z}, +, \times)$ sont des sous-groupes de $(\mathbb{Z}, +)$ donc de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.
□

2.5 Application à l'arithmétique

Soit $(A, +, \times)$ un anneau intègre commutatif

2.5.1 Divisibilité dans un anneau intègre

Définition

On dit que $a \in A$ divise $b \in A$ s'il existe $u \in A$ tel que $b = au$. On note alors $a \mid b$.

Exemple 1_A divise a et a divise a.

Exemple a divise 0_A et $0_A \mid a \Rightarrow a = 0_A$.
La notion de diviseurs de zéro dans le cadre arithmétique ne doit pas être confondue avec celle du cadre de l'intégrité !

Théorème

On a équivalence entre :
(i) $a \mid b$;
(ii) $b \in aA$;
(iii) $bA \subset aA$.

dém. :
Par définition (i) $\Leftrightarrow$ (ii)
(ii) $\Rightarrow$ (iii) Si $b \in aA$ alors $bA \subset aA$ car aA est un idéal.
(iii) $\Rightarrow$ (ii) Supposons $bA \subset aA$. Puisque $b \in bA$, on a $b \in aA$.
□

Proposition

Soit $a, b, c \in A$.

$$a \mid b \text{ et } b \mid c \Rightarrow a \mid c$$

dém. :
$bA \subset aA$ et $cA \subset bA \Rightarrow cA \subset aA$.
□

Proposition

Soit $a, b, c \in A$.

$$a \mid b \text{ et } a \mid c \Rightarrow a \mid (b + c)$$

dém. :
$bA \subset aA$ et $cA \subset aA \Rightarrow (b + c)A \subset bA + cA \subset aA$ car aA est un idéal.
□

2.5.2 Association

Définition

On dit que $a \in A$ est associé à $b \in A$ si a et b se divise mutuellement.

Proposition

Ceci définit une relation d'équivalence sur A.

Théorème

Soit $a, b \in A$. On a équivalence entre :
(i) a et b sont associés ;
(ii) $aA = bA$;
(iii) $\exists u \in U(A), b = au$.

dém. :
(i) $\Leftrightarrow bA \subset aA$ et $aA \subset bA \Leftrightarrow$ (ii)
(i) $\Rightarrow$ (iii) Supposons a et b associés.
Il existe $u, v \in A$ tels que $b = au$ et $a = bv$.
On a alors $a = a(uv)$.
Cas $a = 0_A$: $b = au = 0_A$ et donc $b = a \times 1_A$.
Cas $a \neq 0_A$: Par intégrité, $uv = 1_A$ et donc $u \in U(A)$ puis $b = au$ avec $u \in U(A)$.

(iii) $\Rightarrow$ (i) Supposons qu'il existe $u \in U(A)$ tel que $b = au$.
On a donc $b \in aA$ puis $bA \subset aA$.
Aussi $a = bu^{-1}$ donc $aA \subset bA$ puis $=$.
□

Exemple Dans $\mathbb{Z}$, a et b sont associés si, et seulement si, $|a| = |b|$.
Ainsi, tout entier est associé à un unique entier naturel.

Exemple Dans $\mathbb{K}[X]$, A et B sont associés si, et seulement si,

$$\exists \lambda \in \mathbb{K}^\star, A = \lambda B$$

Ainsi, tout polynôme non nul est associé à un unique polynôme unitaire.

2.5.3 Arithmétique dans $\mathbb{Z}$

Par ce qui précède

$$a \mid b \Leftrightarrow b\mathbb{Z} \subset a\mathbb{Z}$$

Dans la suite nous exploitons cette interprétation pour revoir l'arithmétique des entiers.

2.5.3.1 PGCD et PPCM

Théorème

Soit $a, b \in \mathbb{Z}$. Il existe unique $d \in \mathbb{N}$ tel que

$$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$$

On a alors

$$d \mid a,\ d \mid b \text{ et } \forall c \in \mathbb{Z},\ (c \mid a \text{ et } c \mid b) \Rightarrow c \mid d$$

dém. :
$a\mathbb{Z}$ et $b\mathbb{Z}$ sont des idéaux de $\mathbb{Z}$ donc $a\mathbb{Z} + b\mathbb{Z}$ aussi.
Par suite, il existe $d \in \mathbb{N}$ unique vérifiant $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$.
Puisque $a\mathbb{Z} \subset a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$, on a $d \mid a$. De même $d \mid b$.
Si $c \mid a$ et $c \mid b$ alors $a\mathbb{Z} \subset c\mathbb{Z}$ et $b\mathbb{Z} \subset c\mathbb{Z}$ donc $d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z} \subset c\mathbb{Z}$ puis $c \mid d$.
□

Définition

Ce naturel d est appelé PGCD de a et b

$$d \underset{\text{déf}}{=} a \wedge b$$

Corollaire

Si $d = a \wedge b$ alors il existe $u, v \in \mathbb{Z}$ vérifiant $d = au + bv$.

Théorème

Soit $a, b \in \mathbb{Z}$. Il existe unique $m \in \mathbb{N}$ tel que

$$a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$$

On a alors

$$a \mid m, b \mid m \text{ et } \forall c \in \mathbb{Z}, (a \mid c \text{ et } b \mid c) \Rightarrow m \mid c$$

dém. :
$a\mathbb{Z}$ et $b\mathbb{Z}$ sont des idéaux de $\mathbb{Z}$ donc $a\mathbb{Z} \cap b\mathbb{Z}$ aussi. Par suite, il existe $m \in \mathbb{N}$ unique vérifiant $a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$.
Puisque $m\mathbb{Z} \subset a\mathbb{Z}$, on a $a \mid m$ et de même $b \mid m$.
Si $a \mid c$ et $b \mid c$ alors $c\mathbb{Z} \subset a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$ donc $m \mid c$.
□

Définition

Ce naturel m est appelé PPCM de a et b :

$$m \underset{\text{déf}}{=} a \vee b$$

Remarque On définit aussi le pgcd d et le ppcm m de plusieurs entiers $a_1, \ldots, a_n$ par

$$d\mathbb{Z} = a_1\mathbb{Z} + \cdots + a_n\mathbb{Z} \text{ et } m\mathbb{Z} = a_1\mathbb{Z} \cap \ldots \cap a_n\mathbb{Z}$$

2.5.3.2 Entiers premiers entre eux

Définition

Deux entiers a et b sont dits premiers entre eux si $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$ (autrement dit si leur PGCD vaut 1).
On note $a \wedge b = 1$.

Théorème

Soit $a, b \in \mathbb{Z}$. On a équivalence entre :
(i) a et b sont premiers entre eux ;
(ii) $\exists u, v \in \mathbb{Z}, au + bv = 1$.

dém. :
(i) $\Rightarrow$ (ii) via l'égalité de Bézout.
(ii) $\Rightarrow$ (i) via $1 \in a\mathbb{Z} + b\mathbb{Z}$ donc $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$.
□

Corollaire

On a

$$\forall a, b, c \in \mathbb{Z}, (a \wedge b = 1 \text{ et } a \wedge c = 1) \Rightarrow a \wedge (bc) = 1$$

$$\forall a, b \in \mathbb{Z}, a \wedge b = 1 \Rightarrow \forall \alpha, \beta \in \mathbb{N}, a^\alpha \wedge b^\beta = 1$$

Théorème

$$\forall a,b,c \in \mathbb{Z},\ (a \mid bc \text{ et } a \wedge b = 1) \Rightarrow a \mid c$$

dém. :
$c\mathbb{Z} = c(a\mathbb{Z} + b\mathbb{Z}) = ac\mathbb{Z} + bc\mathbb{Z} \subset a\mathbb{Z}$ donc $a \mid c$.
□

Théorème

$$\forall a,b,c \in \mathbb{Z},\ (a \wedge b = 1, a \mid c \text{ et } b \mid c) \Rightarrow ab \mid c$$

2.5.3.3 Nombre premiers

Définition

Un naturel $p \geqslant 2$ est dit premier si ses seuls diviseurs positifs sont 1 et lui-même.

Exemple Deux entiers a et b sont premiers entre eux si, et seulement si, ils ne possède pas de facteurs premiers en commun.

Théorème

Pour tout $a \in \mathbb{N}$ tel que $a \geqslant 2$ on peut écrire

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}$$

avec $N \in \mathbb{N}^\star$, $p_1, \dots, p_N$ nombres premiers deux à deux distincts et $\alpha_1, \dots, \alpha_n \in \mathbb{N}^\star$.
De plus, cette décomposition est unique à l'ordre près des facteurs.

Exemple Si $a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}$ et $b = p_1^{\beta_1} p_2^{\beta_2} \dots p_N^{\beta_N}$ (écriture qu'il est possible d'obtenir en autorisant les exposants à être nuls) alors

$$a \wedge b = \prod_{i=1}^{N} p_i^{\min(\alpha_i, \beta_i)} \text{ et } a \vee b = \prod_{i=1}^{N} p_i^{\max(\alpha_i, \beta_i)}$$

En particulier, on constate

$$(a \wedge b) \times (a \vee b) = ab$$

2.5.4 Fonction indicatrice d'Euler

Définition

On appelle fonction indicatrice d'Euler l'application $\varphi : \mathbb{N}^\star \to \mathbb{N}^\star$ définie par

$$\varphi(n) = \text{Card}\,\{k \in [\![1, n]\!] / k \wedge n = 1\}$$

Exemple $\varphi(12) = \text{Card}\,\{1, 5, 7, 11\} = 4$.

Remarque $\varphi(n)$ est aussi :
- le nombre de générateurs du groupe $(\mathbb{Z}/n\mathbb{Z}, +)$;
(c'est aussi le nombre de racines primitives n-ième de l'unité)
- le nombre d'éléments inversibles de l'anneau $(\mathbb{Z}/n\mathbb{Z}, +, \times)$.
(c'est donc le cardinal de $U(\mathbb{Z}/n\mathbb{Z})$)

Lemme

Si p est un nombre premier et $\alpha \in \mathbb{N}^\star$ alors

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$$

dém. :
Pour $k \in [\![1, p^\alpha]\!]$, le pgcd de k et p^α est un diviseur de p^α.
Puisque p est premier les naturels diviseurs de p^α sont $1, p, p^2, \ldots, p^\alpha$.
Par suite $\operatorname{pgcd}(k, p^\alpha) = 1, p, \ldots$ ou p^α.
On en déduit

$$k \wedge p^\alpha \neq 1 \Leftrightarrow p \mid k$$

Par suite, les entiers $k \in [\![1, p^\alpha]\!]$ qui ne sont pas premiers avec p^α sont ceux qui sont les multiples de p suivants

$$p, 2p, \ldots, p^\alpha$$

Il y en a $p^{\alpha-1}$ et donc

$$\varphi(p^\alpha) = \operatorname{Card}[\![1, p^\alpha]\!] - p^{\alpha-1} = p^\alpha - p^{\alpha-1}$$

□

Lemme

Si n et m sont deux entiers naturels non nuls premiers entre eux alors

$$\varphi(nm) = \varphi(n)\varphi(m)$$

dém. :
Par le théorème Chinois, l'anneau $\mathbb{Z}/mn\mathbb{Z}$ est isomorphe à $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Il y a donc autant d'éléments inversibles dans $\mathbb{Z}/mn\mathbb{Z}$ que dans $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.
Il y a exactement $\varphi(mn)$ éléments inversibles dans $\mathbb{Z}/mn\mathbb{Z}$.
Les éléments inversibles de $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ sont les couples formés par un élément inversible de $\mathbb{Z}/m\mathbb{Z}$ et un élément inversible de $\mathbb{Z}/n\mathbb{Z}$. Il y en a exactement $\varphi(m)\varphi(n)$.
Au final, on peut conclure

$$\varphi(mn) = \varphi(m)\varphi(n)$$

□

Théorème

Si $n \geqslant 2$ s'écrit

$$n = p_1^{\alpha_1} \dots p_N^{\alpha_N}$$

avec $p_1, \dots, p_N$ nombres premiers deux à deux distincts et $\alpha_1, \dots, \alpha_N \in \mathbb{N}^\star$ alors

$$\varphi(n) = n \prod_{i=}^{N} \left(1 - \frac{1}{p_i}\right)$$

dém. :
On a

$$\varphi(n) = \varphi(p_1^{\alpha_1} p_2^{\alpha_2} \dots p_N^{\alpha_N}) = \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2} \dots p_N^{\alpha_N})$$

car $p_1^{\alpha_1} \wedge (p_2^{\alpha_2} \dots p_N^{\alpha_N}) = 1$ puisque les nombres premiers p_i sont deux à deux distincts.
De même

$$\varphi(n) = \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2}) \dots \varphi(p_N^{\alpha_N}) = \prod_{i=1}^{N} \varphi(p_i^{\alpha_i})$$

Or

$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha(1 - 1/p)$$

donc

$$\varphi(n) = \prod_{i=1}^{N} p_i^{\alpha_i} \prod_{i=1}^{N} \left(1 - \frac{1}{p_i}\right) = n \prod_{i=1}^{N} \left(1 - \frac{1}{p_i}\right)$$

□

Exemple Les facteurs premiers de 12 sont 2 et 3.

$$\varphi(12) = 12 \times \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) = 4$$

2.5.5 Théorème d'Euler

Théorème

Si a est un entier premier avec n alors

$$a^{\varphi(n)} \equiv 1 \quad [n]$$

dém. :
$\bar{a}$ est un élément du groupe $(U(\mathbb{Z}/n\mathbb{Z}), \times)$. Ce groupe possède $\varphi(n)$ éléments donc

$$\bar{a}^{\varphi(n)} = \bar{1}$$

i.e.

$$a^{\varphi(n)} \equiv 1 \quad [n]$$

□

Remarque Si p est un nombre premier, $\varphi(p) = p - 1$ et l'on retrouve le petit théorème de Fermat

$$a \not\equiv 0 \quad [p] \Rightarrow a^{p-1} \equiv 1 \quad [p]$$

2.5.6 Musculations

2.5.6.1 Une relation

Proposition

$$\forall n \in \mathbb{N}^\star, n = \sum_{d|n} \varphi(d)$$

dém. :
Considérons les n nombres rationnels

$$\frac{1}{n}, \frac{2}{n}, \dots, \frac{k}{n}, \dots, \frac{n}{n}$$

L'écriture irréductible des ces nombres est de la forme

$$\frac{k}{n} = \frac{p}{d} \text{ avec } d \mid n \text{ et } p \wedge d = 1$$

Il y a exactement $\varphi(d)$ fractions qui se réduisent avec le dénominateur d et donc

$$\varphi(n) = \sum_{d|n} \varphi(d)$$

□

2.5.6.2 Nombre de diviseurs

Exemple Pour $n \in \mathbb{N}^\star$, notons

$$\text{Div}(n) = \{d \in \mathbb{N}^\star / d \mid n\} \text{ et } \delta(n) = \text{CardDiv}(n)$$

Pour $n = 6$, Div$(6) = \{1, 2, 3, 6\}$ et $\delta(6) = 4$.
De façon générale, exprimons $\delta(n)$.
Pour $n = p^\alpha$ avec p nombre premier on a

$$\text{Div}(p^\alpha) = \{1, p, \dots, p^\alpha\} \text{ et } \delta(p^\alpha) = \alpha + 1$$

Pour $m \wedge n = 1$, montrons $\delta(mn) = \delta(m)\delta(n)$.
Considérons l'application f : Div$(m) \times$ Div$(n) \to$ Div(mn) définie par $f(a, b) = ab$.
L'application considérée est bien définie par

$$(a \mid m \text{ et } b \mid n) \Rightarrow ab \mid mn$$

Montrons que f est bijective.
Supposons $f(a, b) = f(c, d)$. On a $ab = cd$.
a divise cd or $a \wedge d = 1$ (car a et d sont diviseurs de m et n premiers entre eux) donc a divise c.

De même c divise a et donc $a = c$ puis $b = d$.
Ainsi f est injective.
Soit $d \in \text{Div}(mn)$.
Posons $a = \text{pgcd}(d, m)$ et $b = \text{pgcd}(d, n)$.
On a $(a, b) \in \text{Div}(m) \times \text{Div}(n)$. Montrons que $f(a, b) = ab = d$.
On a $a \mid d$, $b \mid d$ et $a \wedge b = 1$ (car a et b sont diviseurs de m et n premiers entre eux) donc $ab \mid d$.
Inversement, par égalité de Bézout on peut écrire $a = du + mv$ et $b = du' + nv'$ donc
$ab = dw + mnvv'$. Puisque d divise mn alors d divise ab puis finalement $d = ab$.
Ainsi f est surjective et donc bijective.
De la bijectivité de f, on déduit

$$\delta(mn) = \delta(m)\delta(n)$$

Par suite, si

$$n = p_1^{\alpha_1} \dots p_N^{\alpha_N}$$

avec $p_1, \dots, p_N$ nombres premiers deux à deux distincts, on obtient

$$\delta(n) = (\alpha_1 + 1) \dots (\alpha_N + 1)$$

2.6 Polynômes en une indéterminée

K désigne un sous-corps de $(\mathbb{C}, +, \times)$ qui sera par exemple $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \dots$
Le cours de première année relatif aux polynômes à coefficients réels ou complexe s'étend au cadre des polynômes à coefficients dans K.

2.6.1 L'anneau $K[X]$

Définition

On appelle polynôme à coefficients dans K en une indéterminée toute expression de la forme

$$P = \sum_{n=0}^{+\infty} a_n X^n$$

où $(a_n)_{n\in\mathbb{N}}$ est une suite d'éléments K nulle à partir d'un certain rang.
On note $K[X]$ l'ensemble des polynômes à coefficients dans K en l'indéterminée X.

Définition

Lorsque $P = \sum_{n=0}^{+\infty} a_n X^n$ n'est pas le polynôme nul, on introduit son degré

$$\deg P = \max \{n \in \mathbb{N} / a_n \neq 0\}$$

Par convention, on pose $\deg 0 = -\infty$.

Définition

Pour $P = \sum_{n=0}^{+\infty} a_n X^n$ et $Q = \sum_{n=0}^{+\infty} b_n X^n$ éléments de $\mathbb{K}[X]$, on pose

$$P + Q = \sum_{n=0}^{+\infty} (a_n + b_n) X^n \text{ et } PQ = \sum_{n=0}^{+\infty} c_n X^n \text{ avec } c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

Théorème

$(K[X], +, \times)$ est un anneau intègre de neutres 0 et 1 dont les éléments inversibles sont les polynômes constants non nuls.

dém. :
L'intégrité et la description des inversibles découlent de la relation

$$\deg(PQ) = \deg P + \deg Q$$

□

Définition

On appelle valeur d'un polynôme $P = \sum_{n=0}^{N} a_n X^n$ en $x \in K$ le nombre

$$P(x) = \sum_{n=0}^{N} a_n x^n \in K$$

Exemple On dit que x est racine de P si $P(x) = 0$.

2.6.2 Divisibilité dans $K[X]$

Puisque que $K[X]$ est un anneau commutatif intègre, le vocabulaire de divisibilité se transpose aux polynômes.
Pour $A, B \in K[X]$, on obtient

$$A \mid B \Leftrightarrow \exists U \in K[X], B = AU \Leftrightarrow B.\mathbb{K}[X] \subset A.\mathbb{K}[X]$$

et

$$A \text{ et } B \text{ sont associés } \Leftrightarrow \exists \lambda \in K^\star, B = \lambda A$$

En particulier, tout polynôme non nul est associé à un unique polynôme unitaire.
De plus, on bénéficie dans $K[X]$ d'une division euclidienne

$$\forall (A, B) \in K[X] \times (K[X] \setminus \{0\}), \exists!(Q, R) \in K[X], A = BQ + R \text{ et } \deg R < \deg B$$

Exemple a est racine de $P \in K[X]$ si, et seulement si, $X - a$ divise P.

2.6.3 Idéaux de $(K[X], +, \times)$

Théorème

Les idéaux de $(K[X], +, \times)$ sont de la forme $P.K[X]$ avec $P \in K[X]$.

dém. :
Soit I un idéal de $K[X]$.
Si $I = \{0\}$ alors $I = P.K[X]$ avec $P = 0$.
Sinon, soit P un polynôme non nul de I de degré minimal.
Par absorption $P.K[X] \subset I$.
Pour $A \in I$, par division euclidienne $A = PQ + R$ avec $\deg R < \deg P$. $R = A - P \in I$ car $A \in I$ et $P \in P.K[X] \subset I$.
Or $\deg R < \deg P$ donc par minimalité du degré de P parmi les polynômes non nuls de I, on peut affirmer $R = 0$ et donc $A \in P.K[X]$. Ainsi $I \subset P.K[X]$ puis $I = P.K[X]$.
□

2.6.4 PGCD et PPCM

Théorème

Soit $A, B \in K[X]$. Il existe un unique polynôme unitaire ou nul $D \in K[X]$ vérifiant tel que

$$A.K[X] + B.K[X] = D.K[X]$$

On a alors

$$D \mid A,\ D \mid B \text{ et } \forall P \in K[X], (P \mid A \text{ et } P \mid B) \Rightarrow P \mid D$$

dém. :
Existence :
$A.K[X]$ et $B.K[X]$ sont des idéaux de $K[X]$ donc $A.K[X]+B.K[X]$ aussi. Il existe donc $D \in K[X]$ vérifiant

$$A.K[X] + B.K[X] = D.K[X]$$

Si le polynôme D n'est pas nul, on peut le remplacer par un polynôme associé et dès lors le choisir unitaire.
Unicité :
Si D et $\tilde{D}$ sont solutions alors ils sont associés et donc égaux car tous deux unitaires ou nuls.
□

Définition

Ce polynôme D est appelé PGCD des polynômes A et B.

$$D \underset{\text{déf}}{=} A \wedge B$$

Corollaire

Si $D = A \wedge B$ alors il existe $U, V \in K[X]$ vérifiant

$$D = AU + BV$$

Définition

De même, on définit le PPCM de deux polynômes $A, B \in K[X]$ comme l'unique polynôme $M \in K[X]$ unitaire ou nul vérifiant

$$AK[X] \cap BK[X] = MK[X]$$

On note

$$M = A \vee B$$

Remarque On peut aussi parler du PGCD D et du PPCM M d'une famille de plusieurs polynômes A_1, K, A_n définis par

$$D.K[X] = A_1.K[X] + \cdots + A_n.K[X] \text{ et } M.K[X] = A_1.K[X] \cap \cdots \cap A_n.K[X]$$

2.6.5 Polynômes premiers entre eux

Définition

On dit que deux polynômes $A, B \in K[X]$ sont premiers entre eux si

$$A.K[X] + B.K[X] = K[X]$$

autrement dit si $A \wedge B = 1$.

Exemple Si $a \neq b$ alors $X - a$ et $X - b$ sont premiers entre eux.

Théorème

Soit $A, B \in K[X]$. On a équivalence entre :
(i) A et B sont premiers entre eux ;
(ii) $\exists (U, V) \in K[X]^2, AU + BV = 1$.

Théorème

Soit $A, B, C \in K[X]$.

$$A \mid BC \text{ et } A \wedge B = 1 \Rightarrow A \mid C$$

Théorème

Soit $A, B, C \in K[X]$.

$$A \wedge B = 1, A \mid C \text{ et } B \mid C \Rightarrow AB \mid C$$

Exemple Si $a_1, \ldots, a_n \in K$ sont des racines deux à deux distinctes de P alors

$$(X - a_1) \ldots (X - a_n) \text{ divise } P$$

En particulier, si P n'est pas le polynôme nul, P possède au plus $\deg P$ racines.
Ce résultat peut être approfondi en introduisant la notion de multiplicité d'une racine.

Théorème

$A, B \in K[X]$ sont premiers entre eux si, et seulement si, A et B n'ont aucunes racines complexes en commun.

dém. :
($\Rightarrow$) Par contraposée
Si A et B ont une racine complexe z en commun alors celle-ci est racine de $D = A \wedge B$ en vertu de la relation de Bézout. Le polynôme D n'est alors pas constant égal à 1.
($\Leftarrow$) Par contraposée
Si A et B ne sont pas premiers entre eux alors $D = 0$ ou D n'est pas constant. Dans les deux cas D admet une racine complexe qui est alors racine commune aux polynômes A et B.
□

Corollaire

Le polynôme $P \in \mathbb{C}[X]$ est à racines simples si, et seulement si, $P \wedge P' = 1$.

2.6.6 Polynômes irréductibles

Définition

Un polynôme non constant $P \in K[X]$ est dit irréductible sur $K[X]$ s'il n'est divisible que par les polynômes constants et ses polynômes associés.

Exemple Le polynôme $X - a$ est irréductible dans $K[X]$.

Exemple Le polynôme $X^2 + 1$ est irréductible dans $\mathbb{R}[X]$ mais ne l'est pas dans $\mathbb{C}[X]$.

Théorème

Si P est un polynôme non constant de $\mathbb{K}[X]$, on peut écrire

$$P = \lambda \prod_{1 \leqslant i \leqslant N} P_i^{\alpha_i}$$

avec $\lambda \in \mathbb{K}^\star$, $N \in \mathbb{N}^\star$, $P_1, \ldots, P_N$ polynômes irréductibles unitaires deux à deux distincts et $\alpha_1, \ldots, \alpha_N \in \mathbb{N}^\star$.
De plus, cette décomposition est unique à l'ordre près des facteurs.

dém. :
Il suffit d'adapter la démonstration vue en première année.
□

Rappel :
Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.
Les polynômes irréductibles unitaires de $\mathbb{C}[X]$ sont les $X - a$ avec $a \in \mathbb{C}$.
Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et ceux de degré 2 sans racines réelles.
Les polynômes irréductibles unitaires sont les polynômes

$$X - a \text{ avec } a \in \mathbb{R} \text{ et } X^2 + pX + q \text{ avec } p, q \in \mathbb{R} \text{ vérifiant } p^2 - 4q < 0$$

Corollaire

Tout polynôme réel de degré impair possède au moins une racine réelle.

dém. :
Sa décomposition en facteurs irréductibles doit au moins faire apparaître un terme de degré ce qui détermine une racine du polynôme. Un argument de continuité en lien avec les limites en l'infini d'un polynôme de degré impair est aussi possible.
□

Remarque Les polynômes irréductibles de $\mathbb{Q}[X]$ sont plus variés...

Exemple Le polynôme $X^3 + X + 1$ est irréductible dans $\mathbb{Q}[X]$.
En effet, s'il était composé, il possèderait au moins une racine rationnelle $x = p/q$ avec $p \wedge q = 1$.
Or $x^3 + x + 1 = 0$ donne $p^3 + pq^2 + q^3 = 0$ et donc $q \mid p$ et $p \mid q$. Cela entraîne $x = \pm 1$ or ce nombre n'est pas racine du polynôme.

Chapitre 3

Espaces vectoriels

La théorie sur les espaces vectoriels présentées en MPSI dans le cas où le corps de base est $\mathbb{R}$ ou $\mathbb{C}$ s'étend pour l'essentiel au cas où le corps de base est un corps quelconque.
On se limite cependant dans ce cours au cas où $\mathbb{K}$ est un sous-corps de $\mathbb{C}$: $\mathbb{K} = \mathbb{C}, \mathbb{R}, \mathbb{Q}, \ldots$

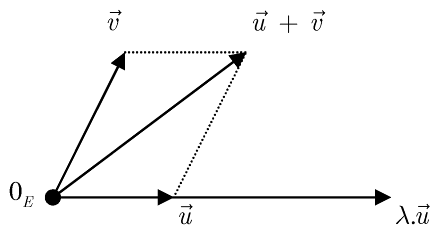
3.1 Structure d'espace vectoriel

3.1.1 Définition

Définition

On appelle $\mathbb{K}$-espace vectoriel tout triplet $(E, +, .)$ formé d'un ensemble E, d'une loi de composition interne $+$ sur E et d'un produit extérieur . opérant de $\mathbb{K}$ sur E vérifiant :
(1) $(E, +)$ est un groupe abélien ;
(2) $\forall x, y \in E, \forall \lambda, \mu \in \mathbb{K}$, $\lambda(x + y) = \lambda x + \lambda y$, $(\lambda + \mu)x = \lambda x + \mu x$, $\lambda(\mu x) = (\lambda\mu)x$ et $1.x = x$.
Les éléments de $\mathbb{K}$ sont appelés scalaires, ceux de E sont appelés vecteurs, en particulier le neutre additif de E est appelé vecteur nul et note 0_E.

Exemple On peut visualiser géométriquement les opérations à l'intérieur d'un espace vectoriel en commençant par visualiser le vecteur nul 0_E et en convenant que tout vecteur sera représenté en partant de celui-ci.



Exemple Espaces vectoriels usuels : $\mathbb{K}^n$, $\mathbb{K}[X]$, $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{F}(X, \mathbb{K})$.

Exemple $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel. Dans ce cas, vecteurs et scalaires se confondent et le produit extérieur correspond à la multiplication sur $\mathbb{K}$.

Proposition

Si $\mathbb{L}$ est un sous-corps de $\mathbb{K}$ alors, par restriction du produit extérieur, tout $\mathbb{K}$-espace vectoriel est encore un $\mathbb{L}$-espace vectoriel.

dém. :
La propriété (1) est conservée alors que la propriété (2) valant pour tout $\lambda, \mu \in \mathbb{K}$ vaut a fortiori pour tout $\lambda, \mu \in \mathbb{L}$.
□

Exemple Tout $\mathbb{C}$-espace vectoriel est aussi un $\mathbb{R}$-espace vectoriel.
En particulier $\mathbb{C}$ est un $\mathbb{R}$-espace vectoriel.

Exemple $\mathbb{R}$ est un $\mathbb{Q}$-espace vectoriel.

3.1.2 Produit d'un nombre fini d'espaces vectoriels

Proposition

Si $E_1, \ldots, E_n$ sont des $\mathbb{K}$-espaces vectoriels alors $E = E_1 \times \cdots \times E_n$ est un $\mathbb{K}$-espace vectoriel pour les lois + et . définies par :

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) \underset{\text{déf}}{=} (x_1 + y_1, \ldots, x_n + y_n) \text{ et } \lambda.(x_1, \ldots, x_n) \underset{\text{déf}}{=} (\lambda x_1, \ldots, \lambda x_n)$$

De plus le vecteur nul de E est alors $0_E = (0_{E_1}, \ldots, 0_{E_n})$.

Exemple On retrouve que $\mathbb{K}^n$ est un $\mathbb{K}$-espace vectoriel de nul $0_{\mathbb{K}^n} = (0, \ldots, 0)$

Exemple Si E et F sont deux $\mathbb{K}$-espaces vectoriels alors $E \times F$ est un $\mathbb{K}$-espace vectoriel.

3.1.3 Espace de fonctions

Soit X un ensemble quelconque

Proposition

Si E un $\mathbb{K}$-espace vectoriel alors $\mathcal{F}(X, E)$ est un $\mathbb{K}$-espace vectoriel pour les lois + et . définies par :

$$f + g : x \mapsto f(x) + g(x) \text{ et } \lambda.f : x \mapsto \lambda.f(x)$$

De plus, le vecteur nul de $\mathcal{F}(X, E)$ est la fonction nulle : $\tilde{0} : x \mapsto 0_E$.

Exemple On retrouve que $\mathcal{F}(X, \mathbb{K})$ est un $\mathbb{K}$-espace vectoriel.

3.2 Sous-espaces vectoriels

E désigne un $\mathbb{K}$-espace vectoriel.

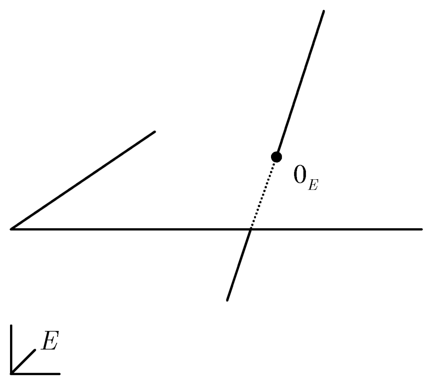
3.2.1 Définition

Définition

On appelle sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E toute partie F de E vérifiant :
1) $0_E \in F$;
2) $\forall \lambda, \mu \in \mathbb{K}, \forall x, y \in F$, $\lambda x + \mu y \in F$.

Exemple $\{0_E\}$ et E sont des sous-espaces vectoriels de E.

Exemple Géométriquement, les sous-espaces vectoriels non triviaux se visualisent comme des droites et des plans contenant le vecteur nul.



Théorème

Si F est un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E alors F est aussi un $\mathbb{K}$-espace vectoriel pour les lois restreintes.

Exemple $\mathbb{K}_n[X]$ est un $\mathbb{K}$-espace vectoriel.
C'est en effet un sous-espace vectoriel de $\mathbb{K}[X]$.

3.2.2 Opérations

Proposition

Si F et G sont deux sous-espaces vectoriels de E alors

$$F \cap G = \{x \in E / x \in F \text{ et } x \in G\}$$

est un sous-espace vectoriel de E.

dém. :
$F \cap G \subset E$.
$0_E \in F \cap G$ car $0_E \in F$ et $0_E \in G$.
Soit $\lambda, \mu \in \mathbb{K}$ et $x, y \in F \cap G$.
On a $\lambda x + \mu y \in F \cap G$ car $\lambda x + \mu y \in F$ puisque $x, y \in F$ et F est un sous-espace vectoriel et de même $\lambda x + \mu y \in G$.
□

Proposition

Si F et G sont deux sous-espaces vectoriels de E alors

$$F + G = \{a + b/a \in F, b \in G\}$$

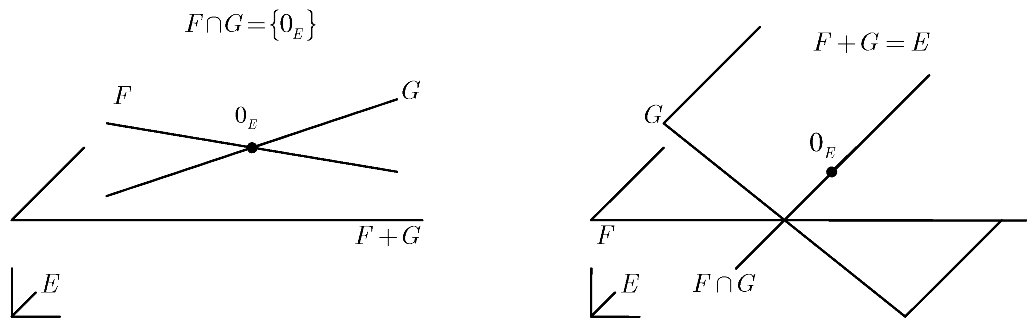
est un sous-espace vectoriel de E.

dém. :
$F + G \subset E$.
$0_E = 0_E + 0_E \in F + G$ car $0_E \in F$ et $0_E \in G$.
Soit $\lambda, \mu \in \mathbb{K}$ et $x, y \in F + G$.
On peut écrire $x = a + b$ et $y = a' + b'$ avec $a, a' \in F$ et $b + b' \in G$ donc

$$\lambda x + \mu y = (\lambda a + \mu a') + (\lambda b + \mu b') \in F + G$$

□

Exemple



Remarque Les opérations d'intersection et de somme de sous-espaces vectoriels :
? sont commutatives ;
? sont associatives ;
? possèdent des neutres E et $\{0_E\}$ respectivement.
En particulier, pour $F_1, \ldots, F_n$ des sous-espaces vectoriels de E, on peut introduire les sous-espaces vectoriels

$$\bigcap_{i=1}^{n} F_i = F_1 \cap \ldots \cap F_n \text{ et } \sum_{i=1}^{n} F_i = F_1 + \cdots + F_n = \left\{ \sum_{i=1}^{n} x_i / x_i \in F_i \right\}$$

3.2.3 Espace vectoriel engendré

Définition

On appelle espace vectoriel engendré par une partie A de E l'intersection Vect$A$ de tous les sous-espaces vectoriels de E contenant A.

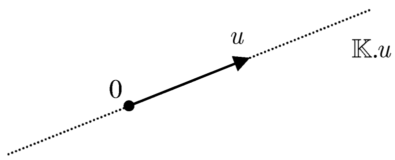
Théorème

Vect$A$ est un sous-espace vectoriel de E contenant A.
De plus, pour tout sous-espace vectoriel F de E :

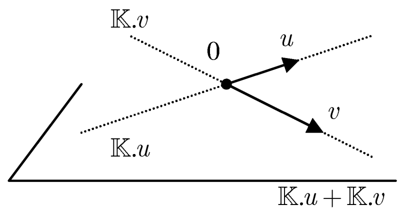
$$A \subset F \Rightarrow \text{Vect}A \subset F$$

Vect$A$ apparaît comme étant le plus petit sous-espace vectoriel de E contenant A.

Exemple Pour $A = \{u\}$, $\text{Vect}(u) = \mathbb{K}.u = \{\lambda.u/\lambda \in \mathbb{K}\}$.



Exemple $\text{Vect}(u, v) = \{\lambda u + \mu v/\lambda, \mu \in \mathbb{K}\} = \mathbb{K}.u + \mathbb{K}.v$.



Remarque Par récurrence

$$\text{Vect}(u_1, \ldots, u_n) = \{\lambda_1 u_1 + \cdots + \lambda_n u_n/\lambda_i \in \mathbb{K}\} = \mathbb{K}.u_1 + \cdots + \mathbb{K}.u_n$$

Exemple Si F et G sont des sous-espaces vectoriels

$$\text{Vect}(F \cup G) = F + G$$

3.2.4 Somme directe

Soit $F_1, \ldots, F_n$ des sous-espaces vectoriels de E.

Définition

Soit $F_1, \ldots, F_n$ des sous-espaces vectoriels. On dit que la somme $\sum_{i=1}^{n} F_i$ est directe si

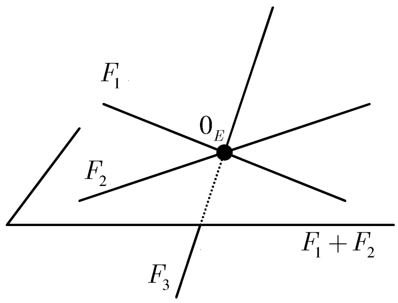
$$\forall x \in \sum_{i=1}^{n} F_i, \exists!(x_1, \ldots, x_n) \in F_1 \times \ldots \times F_n, x = x_1 + \cdots + x_n$$

Autrement dit, il y a unicité dans l'écriture de la décomposition d'un vecteur de la somme. La somme $\sum_{i=1}^{n} F_i$ est alors notée

$$\bigoplus_{i=1}^{n} F_i \text{ ou } F_1 \oplus \cdots \oplus F_n$$

Remarque Si F et G sont en somme directe et si $F + G$ est en somme directe avec H alors F, G, H sont en somme directe. On dispose ainsi de la relation d'associativité

$$(F \oplus G) \oplus H = F \oplus G \oplus H$$



Théorème

Les espaces $F_1, \ldots, F_n$ sont en somme directe si, et seulement si,

$$\forall (x_1, \ldots, x_n) \in F_1 \times \ldots \times F_n, x_1 + \cdots + x_n = 0_E \Rightarrow \forall 1 \leqslant i \leqslant n, x_i = 0_E$$

Ce qui revient à signifier l'unicité de la décomposition du vecteur nul.

Remarque Si l'on se limite à deux sous-espaces vectoriels F et G, on a aussi

$$F \text{ et } G \text{ sont en somme directe } \Leftrightarrow F \cap G = \{0_E\}$$

3.2.5 Sous-espaces vectoriels supplémentaires

Soit F et G deux sous-espaces vectoriels de E.

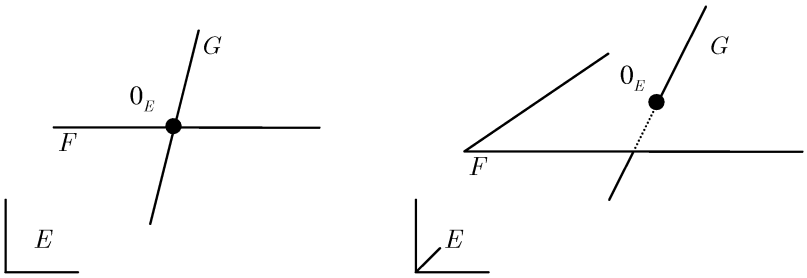
Définition

On dit que les espaces F et G sont supplémentaires si

$$\forall x \in E, \exists!(a, b) \in F \times G, x = a + b$$

Exemple E et $\{0_E\}$ sont supplémentaires dans E.

Exemple



Théorème

> Les espaces F et G sont supplémentaires si, et seulement si, $F \cap G = \{0_E\}$ et $F + G = E$.
> Autrement dit, si, et seulement si, $E = F \oplus G$.

Exemple On note $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ les sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{R})$ formés des matrices symétriques et antisymétriques. Montrer que $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont des sous-espaces vectoriels supplémentaires.
On a $\mathcal{S}_n(\mathbb{R}) \cap \mathcal{A}_n(\mathbb{R}) = \{O_n\}$ car

$$^tM = M \text{ et } {^tM} = -M \Rightarrow M = O_n$$

Aussi $\mathcal{S}_n(\mathbb{R}) + \mathcal{A}_n(\mathbb{R}) = \mathcal{M}_n(\mathbb{R})$ car

$$M = \frac{1}{2}\left(M + {^tM}\right) + \frac{1}{2}\left(M - {^tM}\right)$$

avec

$$\frac{1}{2}\left(M + {^tM}\right) \in \mathcal{S}_n(\mathbb{R}) \text{ et } \frac{1}{2}\left(M - {^tM}\right) \in \mathcal{A}_n(\mathbb{R})$$

Exemple Soit $E = \mathcal{C}([-1,1],\mathbb{R})$,

$$F_1 = \{x \in [-1,1] \mapsto ax + b/a, b \in \mathbb{R}\} \text{ et } F_2 = \{f \in F/f(-1) = f(1) = 0\}$$

Montrons que F_1 et F_2 sont des sous-espaces vectoriels supplémentaires.
F_1 et F_2 sont évidemment des sous-espaces vectoriels de E.
Etudions $F_1 \cap F_2$.
Soit $f \in F_1 \cap F_2$. Il existe $a, b \in \mathbb{R}$ tels que $f(x) = ax + b$ pour tout $x \in [-1,1]$.
Or $f(1) = f(-1) = 0$ donc $a + b = a - b = 0$ puis $a = b = 0$ et enfin $f = \tilde{0}$.
Ainsi $A \cap B \subset \{0\}$ puis $=$.
Etudions $F_1 + F_2$.
Analyse :
On suppose $f = g + h$ avec $g \in F_1$ et $h \in F_2$.
Il existe $a, b \in \mathbb{R}$ tel que $g(x) = ax + b$ et on a $h(1) = h(-1) = 0$.

On en déduit $a+b=f(1)$ et $a-b=f(-1)$ puis

$$a=\frac{1}{2}(f(1)+f(-1)) \text{ et } b=\frac{1}{2}(f(1)-f(-1))$$

Ceci détermine g puis $h=f-g$.
Synthèse :
Soit $f\in E$. Posons

$$a=\frac{1}{2}(f(1)+f(-1)) \text{ et } b=\frac{1}{2}(f(1)-f(-1))$$

Considérons ensuite $g:x\in[-1,1]\mapsto ax+b$ et $h=f-g$.
On a $f=g+h$ avec $g\in F_1$.
De plus $f(1)=a+b+h(1)$ donne $h(1)=0$ et, de même, on obtient $h(-1)=0$. Ainsi $h\in F_2$.
Finalement $E\subset F_1+F_2$ puis $=$. On peut conclure

$$E=F_1\oplus F_2$$

3.2.6 Sous-espace affine

Définition

On appelle sous-espace affine passant $a\in E$ et dirigé par un sous-espace vectoriel F de E l'ensemble

$$V=a+F=\{a+x/x\in F\}$$

Exemple Géométriquement les sous-espaces affines se visualisent comme étant des points, des droites ou des plans ne passant pas nécessairement par 0_E.

Proposition

Si V est un sous-espace affine de direction F et si $b\in V$ alors

$$V=b+F$$

dém. :
Ecrivons $V=a+F$.
Puisque $b\in V$, on a $b-a\in F$ et donc

$$b+F=\{b+x/x\in F\}=\{a+x'/x'\in F\}=a+F$$

□

Proposition

L'intersection de deux sous-espaces affines V et W de directions F et G est soit vide, soit égal à un sous-espace affine de direction $F\cap G$.

dém. :
Supposons $V\cap W\neq\emptyset$. Considérons $a\in V\cap W$. On a $V=a+F$ et $W=a+G$.
Par suite, pour $x\in E$, $x\in V\cap W\Leftrightarrow x-a\in F\cap G$ et ainsi $V\cap W=a+F\cap G$.
□

3.3 Dimension

I désigne un ensemble, éventuellement infini.
E désigne un $\mathbb{K}$-espace vectoriel.

3.3.1 Combinaisons linéaires

Définition

Une famille de scalaires $(\lambda_i)_{i\in I}$ est dite presque nulle si

$$\{i \in I/\lambda_i = 0\} \text{ est fini}$$

On note $\mathbb{K}^{(I)}$ l'ensemble de ces familles.

Exemple Si I est un ensemble fini alors $\mathbb{K}^I = \mathbb{K}^{(I)}$.

Exemple Une suite nulle à partir d'un certain rang est une famille presque nulle de $\mathbb{K}^{\mathbb{N}}$.
Ainsi

$$\mathbb{K}^{(\mathbb{N})} = \{u = (u_n)_{n\in\mathbb{N}}/\exists N \in \mathbb{N}, \forall n \geqslant N, u_n = 0\}$$

Définition

On appelle combinaison linéaire d'une famille $(x_i)_{i\in I}$ de vecteurs de E tout vecteur de E pouvant s'écrire

$$\sum_{i\in I} \lambda_i x_i$$

avec $(\lambda_i)_{i\in I}$ une famille de scalaire presque nulle.

Remarque Bien que la somme porte sur l'ensemble I pouvant être infini, la somme a du sens car elle ne comporte qu'un nombre fini de termes non nuls.

Exemple Cas $I = \emptyset$:
Seul le vecteur nul est combinaison linéaire de la famille vide.
Cas Card$I = 1$:
Les combinaisons linéaires de (x) sont les λx avec $\lambda \in \mathbb{K}$.
Cas Card$I = n$:
Quitte à réindexer, on peut supposer $I = \{1, \ldots, n\}$.
Les combinaisons linéaires de $(x_i)_{1\leqslant i\leqslant n}$ sont les $\lambda_1 x_1 + \cdots + \lambda_n x_n$ avec $\lambda_i \in \mathbb{K}$.
Cas Card$I = +\infty$:
Les combinaisons linéaires de la famille $(x_i)_{i\in I}$ correspondent aux combinaisons linéaires de ses sous-familles finies.

Exemple Dans $\mathbb{K}[X]$, les combinaisons linéaires des monômes X^k avec $k \in \mathbb{N}$ sont exactement les polynômes.

Remarque Si A est une partie de E alors Vect(A) est l'ensemble des combinaisons linéaires (finies) d'éléments de A.

Proposition

Si F est un sous-espace vectoriel de E alors toute combinaison linéaire d'une famille de vecteurs de F est élément de F.

3.3.2 Famille génératrice

Définition

On note Vect$(x_i)_{i\in I}$ l'espace vectoriel engendré par la partie $\{x_i/i \in I\}$.

Théorème

Vect$(x_i)_{i\in I}$ est l'ensemble des combinaisons linéaires de la famille $(x_i)_{i\in I}$.

Définition

Une famille $(x_i)_{i\in I}$ de vecteurs de E est dite génératrice si Vect$(x_i)_{i\in I} = E$ ce qui signifie que tout vecteur de E est combinaison linéaire de cette famille

$$\forall x \in E, \exists(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}, x = \sum_{i\in I} \lambda_i x_i$$

Exemple La famille vide est génératrice de $\{0_E\}$.

Exemple Dans $\mathbb{K}^n$, considérons $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$.
La famille $(e_i)_{1\leqslant i\leqslant n}$ est génératrice.

Exemple Dans $\mathbb{K}[X]$, la famille $(X^k)_{k\in\mathbb{N}}$ est génératrice.

3.3.3 Famille libre

Définition

Une famille $(x_i)_{i\in I}$ de vecteurs de E est dite libre si

$$\forall(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}, \sum_{i\in I} \lambda_i x_i = 0_E \Rightarrow \forall i \in I, \lambda_i = 0$$

Sinon, la famille est dite liée et toute égalité $\sum_{i\in I} \lambda_i x_i = 0_E$ avec $(\lambda_i)_{i\in I} \neq 0$ est appelée relation linéaire sur la famille $(x_i)_{i\in I}$.

Exemple La famille vide est libre.

Exemple (x) est libre si, et seulement si, $x \neq 0_E$.

Exemple (x, y) est liée si, et seulement si, il existe $(\alpha, \beta) \neq (0, 0)$ tel que $\alpha x + \beta y = 0_E$.
Cela équivaut encore à dire

$$\exists \lambda \in \mathbb{K}, x = \lambda y \text{ ou } \exists \mu \in \mathbb{K}, x = \mu y$$

Attention : (x, y) liée n'implique pas qu'il existe $\lambda \in \mathbb{K}$ tel que $y = \lambda x$ (prendre $x = 0_E$ et $y \neq 0_E$ quelconque)
Cependant

$$(x, y) \text{ liée et } x \neq 0_E \Rightarrow \exists \lambda \in \mathbb{K}, y = \lambda x$$

Exemple Dans $\mathbb{K}^n$, la famille $(e_i)_{1 \leqslant i \leqslant n}$ est libre.

Exemple Une famille infinie est libre si, et seulement si, toutes ses sous-familles finies le sont.

Exemple La famille $(X^n)_{n \in \mathbb{N}}$ est libre car

$$\forall n \in \mathbb{N}, (X^k)_{0 \leqslant k \leqslant n} \text{ est libre}$$

Exemple $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$. Pour $a \in \mathbb{R}$, on note e_a l'application de $\mathbb{R}$ vers $\mathbb{R}$ définie par $e_a(t) = \mathrm{e}^{at}$.
Montrons que $(e_a)_{a \in \mathbb{R}}$ est une famille libre d'éléments de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
Soit $a_1, \ldots, a_n$ des réels deux à deux distincts.
Supposons

$$\lambda_1 e_{a_1} + \cdots + \lambda_n e_{a_n} = 0$$

Pour tout $t \in \mathbb{R}$,

$$\lambda_1 \mathrm{e}^{a_1 t} + \lambda_2 \mathrm{e}^{a_2 t} + \cdots + \lambda_n \mathrm{e}^{a_n t} = 0$$

Quitte à réindexer, on peut supposer $a_1 < a_2 < \ldots < a_n$
En multipliant la relation par $\mathrm{e}^{-a_1 t}$, on obtient

$$\lambda_1 + \lambda_2 \mathrm{e}^{(a_2 - a_1)t} + \cdots + \lambda_n \mathrm{e}^{(a_n - a_1)t} = 0$$

Quand $t \to -\infty$, la relation précédente donne $\lambda_1 = 0$.
On obtient alors

$$\lambda_2 \mathrm{e}^{a_2 t} + \cdots + \lambda_n \mathrm{e}^{a_n t} = 0$$

pour tout $t \in \mathbb{R}$ et on peut reprendre la démarche pour obtenir successivement $\lambda_2 = \ldots = \lambda_n = 0$.
Ainsi, la famille $(e_{a_1}, \ldots, e_{a_n})$ est libre et puisque toutes ses sous-familles finies sont libres, la famille $(e_a)_{a \in \mathbb{R}}$ est libre.

3.3.4 Base

Définition

On appelle base de E toute famille $(e_i)_{i\in I}$ de vecteurs de E à la fois libre et génératrice.

Exemple La famille vide est base de $E = \{0_E\}$.

Exemple $(e_i)_{1\leqslant i\leqslant n}$ est une base de $\mathbb{K}^n$ (dite canonique).

Exemple $(X^k)_{k\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$ (dite canonique).

Exemple (1) est base de $\mathbb{K}$ (dite canonique).

Exemple $(1, i)$ est base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$ (dite canonique).

Théorème

Si $(e_i)_{i\in I}$ est une base de E alors

$$\forall x \in E, \exists!(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}, x = \sum_{i\in I} \lambda_i e_i$$

Définition

La famille $(\lambda_i)_{i\in I}$ est alors appelée famille des coordonnées (ou composantes) de x dans la base $(e_i)_{i\in I}$.

Exemple Les coordonnées de $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ dans la base canonique sont ses éléments x_i.

Exemple Les coordonnées de $P \in \mathbb{K}[X]$ dans la base canonique de $\mathbb{K}[X]$ sont ses coefficients.

Exemple Soit $j \in I$. On peut écrire

$$e_j = \sum_{i\in I} \delta_{i,j} e_i \text{ avec } \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

La famille $(\delta_{i,j})_{i\in I}$ est donc la famille des coordonnées de e_j dans la base $(e_i)_{i\in I}$.

3.3.5 Dimension

Définition

> On dit qu'un $\mathbb{K}$-espace vectoriel est de dimension finie s'il possède une famille génératrice finie. On sait qu'un tel espace possède alors une base finie et que toute base de cet espace est formée du même nombre de vecteurs qu'on appelle la dimension de celui-ci.

Exemple $\dim\{0_E\} = 0$, $\dim \mathbb{K}^n = n$, $\dim \mathcal{M}_{n,p}(\mathbb{K}) = np$, $\dim \mathbb{K}_n[X] = n+1$, $\dim \mathbb{K} = 1$, $\dim_{\mathbb{C}} \mathbb{C} = 1$ et $\dim_{\mathbb{R}} \mathbb{C} = 2$.

Définition

> Si un $\mathbb{K}$-espace vectoriel E n'est pas de dimension finie, on pose $\dim E = +\infty$.

Exemple $\dim \mathbb{K}[X] = +\infty$.

3.3.6 Construction de bases

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie.

Théorème

> De toute famille génératrice de E on peut extraire une base

Théorème

> Toute famille libre de vecteurs de E peut être complétée en une base.

Théorème

> Soit E est un $\mathbb{K}$-espace vectoriel de dimension finie et $(e_i)_{1\leqslant i\leqslant n}$ une famille de vecteurs de E.
> On suppose
> $$n = \dim E$$
> On a équivalence entre :
> (i) $(e_i)_{1\leqslant i\leqslant n}$ est une base de E ;
> (ii) $(e_i)_{1\leqslant i\leqslant n}$ est une famille libre ;
> (iii) $(e_i)_{1\leqslant i\leqslant n}$ est une famille génératrice de E.

Exemple Soit $(P_n)_{n\in\mathbb{N}} \in \mathbb{K}[X]^{\mathbb{N}}$ une famille de polynômes de degrés étagés (i.e. $\forall n \in \mathbb{N}, \deg P_n = n$)
Montrons que $(P_n)_{n\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$.
Commençons par étudier la sous-famille $(P_k)_{0\leqslant k\leqslant n}$.
Supposons
$$\lambda_0 P_0 + \cdots + \lambda_n P_n = 0$$
On a
$$\lambda_n P_n = -(\lambda_0 P_0 + \cdots + \lambda_{n-1} P_{n-1})$$
donc $\deg(\lambda_n P_n) < n$ puis $\lambda_n = 0$.
En reprenant le procédé, on obtient successivement $\lambda_{n-1} = 0$,..., $\lambda_0 = 0$.
Ainsi, la famille $(P_k)_{0\leqslant k\leqslant n}$ est libre, or cette famille est formée de $n+1 = \dim \mathbb{K}_n[X]$ vecteurs de $\mathbb{K}_n[X]$ c'est donc une base de $\mathbb{K}_n[X]$.

La famille $(P_n)_{n\in\mathbb{N}}$ est alors libre car chacune de ses sous-familles finies est libre. Elle est de plus génératrice car pour tout $P \in \mathbb{K}[X]$, il existe $n \in \mathbb{N}$ tel que $P \in \mathbb{K}_n[X]$ ce qui permet d'écrire

$$P = \sum_{k=0}^{n} \lambda_k P_k = \sum_{k=0}^{+\infty} \lambda_k P_k \text{ en posant } \lambda_k = 0 \text{ pour } k > n$$

Finalement, la famille $(P_n)_{n\in\mathbb{N}}$ est une base de $\mathbb{K}[X]$.

3.3.7 Dimension d'un sous-espace vectoriel

3.3.7.1 Sous-espace vectoriel en dimension finie

Théorème

Si F est un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E de dimension finie alors F est de dimension finie et

$$\dim F \leqslant \dim E$$

De plus

$$\dim F = \dim E \Leftrightarrow F = E$$

3.3.7.2 Formule de Grassmann

Théorème

Si F et G sont des sous-espaces vectoriels de dimensions finies d'un $\mathbb{K}$-espace vectoriel E alors $F+G$ et $F\cap G$ sont de dimensions finies et

$$\dim(F+G) = \dim F + \dim G - \dim(F\cap G)$$

dém. :
On complète une base de $F\cap G$, d'une part, en une base de F et, d'autre part, en une base de G puis on forme une base de $F+G$ en considérant la famille de tous ses vecteurs.

□

Corollaire

Si F et G sont en somme directe alors

$$\dim(F\oplus G) = \dim F + \dim G$$

3.3.7.3 Supplémentarité en dimension finie

Théorème

Tout sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel de dimension finie admet au moins un supplémentaire et tous ses supplémentaires sont d'égales dimensions.

Théorème

Si F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E de dimension finie vérifiant

$$\dim E = \dim F + \dim G$$

alors on a équivalence entre :
(i) F et G sont supplémentaires ;
(ii) $F \cap G = \{0_E\}$;
(iii) $F + G = E$.

Exemple Soit H un sous-espace vectoriel de dimension $n-1$ d'un $\mathbb{K}$-espace vectoriel E de dimension $n \in \mathbb{N}^\star$ (autrement dit H est hyperplan). Pour tout vecteur $a \in E \backslash H$, on a

$$H \oplus \text{Vect}(a) = E$$

Exemple On peut obtenir rapidement la supplémentarité se $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ en exploitant un argument de dimension.

3.3.7.4 Somme de plusieurs sous-espaces vectoriels

Théorème

Si $F_1, \ldots, F_m$ sont des sous-espaces vectoriels de dimensions finies alors $\sum_{k=1}^{m} F_k$ est de dimension finie et

$$\dim \sum_{k=1}^{m} F_k \leqslant \sum_{k=1}^{m} \dim F_k$$

De plus, il y a égalité si, et seulement si, les sous-espaces vectoriels $F_1, \ldots, F_m$ sont en somme directe.
Ainsi

$$\dim \bigoplus_{k=1}^{m} F_k = \sum_{k=1}^{m} \dim F_k$$

Théorème

On suppose

$$E = \bigoplus_{k=1}^{m} F_k$$

En accolant des bases des sous-espaces vectoriels $F_1, \ldots, F_m$, on forme une base de E.

Définition

Une telle base est dite adaptée à la décomposition $E = \bigoplus_{k=1}^{m} F_k$.

Exemple Supposons F et G supplémentaires dans E.
Si $(e_1, \ldots, e_p)$ est une base de F et $(e_{p+1}, \ldots, e_n)$ une base de G alors $(e_1, \ldots, e_n)$ détermine une base de E adaptée à la supplémentarité $E = F \oplus G$.

3.4 Applications linéaires

Soit E et E' des $\mathbb{K}$-espaces vectoriels.

3.4.1 Définition

Définition

On appelle application linéaire de E vers E' toute application $u : E \to E'$ vérifiant :

$$\forall \lambda, \mu \in \mathbb{K}, \forall x, y \in E, u(\lambda x + \mu y) = \lambda u(x) + \mu u(y)$$

Théorème

L'ensemble $\mathcal{L}(E, E')$ des applications linéaires de E vers E' est un espace vectoriel pour les lois usuelles de neutre l'application linéaire nulle $\tilde{o}$.

Définition

Lorsque $E' = \mathbb{K}$, on parle de forme linéaire et on note $E^\star$ au lieu de $\mathcal{L}(E, \mathbb{K})$.
L'espace $E^\star$ est appelé espace dual de E.

Définition

Lorsque $E' = E$, on parle d'endomorphisme et on note $\mathcal{L}(E)$ au lieu de $\mathcal{L}(E, E)$.
$\mathcal{L}(E)$ est un anneau pour les lois + et $\circ$ de neutres $\tilde{0}$ et Id_E.

Définition

Lorsque u est bijective, on parle d'isomorphisme et on dit que les espaces E et E' sont isomorphes.
On note $\mathrm{GL}(E, E')$ l'ensemble des isomorphismes de E vers E'.

Définition

Lorsque u est bijective et $E' = E$, on parle d'automorphisme et on note $\mathrm{GL}(E) = \mathrm{GL}(E, E)$ l'ensemble des automorphismes de E. $(\mathrm{GL}(E), \circ)$ est le groupe des inversibles de l'anneau $(\mathcal{L}(E), +, \circ)$, on l'appelle groupe linéaire de E.

3.4.2 Propriétés

Proposition

Si $u \in \mathcal{L}(E, E')$ alors

$$u(0_E) = 0_{E'}$$

Théorème

L'image directe (resp. réciproque) d'un sous-espace vectoriel par une application linéaire est un sous-espace vectoriel.

Exemple Si $u \in \mathcal{L}(E, E')$ et $A \subset E$ alors $u(\mathrm{Vect}(A)) = \mathrm{Vect}(u(A))$.
En effet, $A \subset \mathrm{Vect}A$ donc $u(A) \subset u(\mathrm{Vect}A)$.
Or $u(\mathrm{Vect}A)$ est un sous-espace vectoriel donc $\mathrm{Vect}u(A) \subset u(\mathrm{Vect}A)$.
Inversement, $u(A) \subset \mathrm{Vect}u(A)$ donc $u^{-1}(u(A)) \subset u^{-1}(\mathrm{Vect}u(A))$.

Or $A \subset u^{-1}(u(A))$ donc $A \subset u^{-1}(\text{Vect}u(A))$.
Mais $u^{-1}(\text{Vect}u(A))$ est un sous-espace vectoriel donc Vect$A \subset u^{-1}(\text{Vect}u(A))$ puis $u(\text{Vect}A) \subset u(u^{-1}(\text{Vect}A))$.
Enfin $u(u^{-1}(\text{Vect}u(A))) \subset \text{Vect}u(A)$ donc $u(\text{Vect}A) \subset \text{Vect}u(A)$.

3.4.3 Noyau et image

Définition

On appelle noyau et image d'une application linéaire u de E vers E' les ensembles

$$\ker u = u^{-1}(\{0_{E'}\}) \text{ et } \text{Im}u = u(E)$$

Ce sont respectivement des sous-espaces vectoriels de E et E'.

Théorème

Soit $u \in \mathcal{L}(E, E')$.
a) u est injective si, et seulement si, $\ker u = \{0\}$,
b) u est surjective si, et seulement si, Im$u = E'$.

Exemple Soit $u, v \in \mathcal{L}(E)$. Montrons

$$v \circ u = \tilde{0} \Leftrightarrow \text{Im}u \subset \ker v$$

($\Leftarrow$) Supposons Im$u \subset \ker v$.
Pour tout $x \in E$, $u(x) \in \text{Im}u$ donc $u(x) \in \ker v$ puis $v(u(x)) = 0$. Ainsi $v \circ u = \tilde{0}$
($\Rightarrow$) Supposons $v \circ u = \tilde{0}$.
Pour tout $y \in \text{Im}u$, on peut écrire $y = u(x)$ avec $x \in E$. Mezalor $v(y) = v(u(x)) = 0$ donc $y \in \ker v$.

Exemple Soit $u \in \mathcal{L}(E)$. Comparons $\ker u$ et $\ker u^2$.
Soit $x \in \ker u$. On a $u(x) = 0$ donc $u^2(x) = u(u(x)) = u(0) = 0$. Ainsi $\ker u \subset \ker u^2$.
Comparons Imu et Imu^2.
Soit $y \in \text{Im}u^2$. On peut écrire $y = u^2(x)$ donc $y = u(u(x)) \in \text{Im}u$. Ainsi Im$u^2 \subset$ Imu.
Plus généralement, on montre $\ker u^n \subset \ker u^{n+1}$ et Im$u^{n+1} \subset$ Imu^n.

3.4.4 Equations linéaires

On considère l'équation $u(x) = y$ avec $u \in \mathcal{L}(E, E')$, $y \in E'$ et d'inconnue $x \in E$:
- si $y \notin \text{Im}u$: l'équation n'est pas compatible ;
- si $y \in \text{Im}u$, l'ensemble des solutions est un sous-espace affine de direction $\ker u$.

Protocole de résolution d'une équation linéaire compatible :
- on résout l'équation homogène (ce qui détermine $\ker u$) ;
- on détermine une solution particulière ;
- on exprime la solution générale comme somme de la solution particulière et de la solution générale de l'équation homogène.

3.4.5 Image linéaire d'une famille de vecteurs

Proposition

Si $u \in \mathcal{L}(E, E')$ alors

$$\forall (x_i)_{i\in I} \in E^I, \forall (\lambda_i) \in \mathbb{K}^{(I)}, u\left(\sum_{i\in I} \lambda_i x_i\right) = \sum_{i\in I} \lambda_i u(x_i)$$

Proposition

Si $(x_i)_{i\in I}$ une famille génératrice de vecteurs de E et si $u \in \mathcal{L}(E, E')$ est surjective alors $(u(x_i))_{i\in I}$ est une famille de vecteurs de E' génératrice.

dém. :
Pour tout $y \in F$, il existe $x \in E$ tel que $y = u(x)$.
Or, il existe aussi $(\lambda_i) \in \mathbb{K}^{(I)}$ telle que $x = \sum_{i\in I} \lambda_i x_i$ et alors $y = \sum_{i\in I} \lambda_i u(x_i)$.
Ainsi, $(u(x_i))_{i\in I}$ est génératrice.
□

Proposition

Si $(x_i)_{i\in I}$ une famille libre de vecteurs de E et si $u \in \mathcal{L}(E, E')$ est injective alors $(u(x_i))_{i\in I}$ est une famille libre de E'.

dém. :
Supposons $\sum_{i\in I} \lambda_i u(x_i) = 0_{E'}$.
On a $u(\sum_{i\in I} \lambda_i x_i) = 0$ donc $\sum_{i\in I} \lambda_i x_i \in \ker u = \{0_E\}$ puis $\sum_{i\in I} \lambda_i x_i = 0_E$.
Or la famille $(x_i)_{i\in I}$ est libre donc

$$\forall i \in I, \lambda_i = 0$$

Ainsi $(u(x_i))_{i\in I}$ est libre.
□

Théorème

Soit $u \in \mathcal{L}(E, E')$ et $(e_i)_{i\in I}$ une base de E.
1) u est injective si, et seulement si, $(u(e_i))_{i\in I}$ est libre.
2) u est surjective si, et seulement si, $(u(e_i))_{i\in I}$ est génératrice de E'.
3) u est un isomorphisme si, et seulement si, $(u(e_i))_{i\in I}$ est une base de E'.

dém. :
1) ($\Rightarrow$) ci-dessus.
($\Leftarrow$) Supposons $(u(e_i))_{i\in I}$ libre.
Soit $x = \sum_{i\in I} \lambda_i e_i$ tel que $u(x) = 0_{E'}$. On a $\sum_{i\in I} \lambda_i u(e_i) = 0_{E'}$ donc $\lambda_i = 0$ pour tout i puis $x = 0_E$.
2) ($\Rightarrow$) ci-dessus.
($\Leftarrow$) Supposons $(u(e_i))_{i\in I}$ génératrice.
Pour tout $y \in F$, on peut écrire $y = \sum_{i\in I} \lambda_i u(e_i)$ et donc $y = u(e)$ avec $e = \sum_{i\in I} \lambda_i e_i$.
3) via 1) et 2)
□

Corollaire

Si deux espaces vectoriels sont isomorphes, ils sont d'égales dimension.

3.4.6 Construction d'une application linéaire

3.4.6.1 Par l'image d'une base

Théorème

Si $(e_i)_{i\in I}$ est une base de E et $(e'_i)_{i\in I}$ une famille de vecteurs de E' alors il existe une unique application linéaire $u : E \to E'$ vérifiant

$$\forall i \in I, u(e_i) = e'_i$$

dém. :
Analyse / Unicité : Supposons u solution.
Pour $e \in E$, on peut écrire $e = \sum_{i\in I} \lambda_i e_i$ avec $(\lambda_i)_{i\in I} \in \mathbb{K}^{(I)}$ et alors

$$u(e) = \sum_{i\in I} \lambda_i u(e_i) = \sum_{i\in I} \lambda_i e'_i$$

ce qui détermine entièrement u.
Synthèse / Existence : Considérons l'application u qui à $e = \sum_{i\in I} \lambda_i e_i$ associe

$$u(e) = \sum_{i\in I} \lambda_i e'_i$$

On vérifie aisément que u est linéaire et transforme e_i en e'_i.
□

Corollaire

Si deux applications linéaires $u, v \in \mathcal{L}(E, E')$ sont égales sur chacun des vecteurs d'une base de E alors elles sont égales sur E.

Corollaire

Deux espaces de dimensions finies égales sont isomorphes.

3.4.6.2 Par ses restrictions linéaires

On suppose

$$E = \overset{m}{\underset{k=1}{\oplus}} F_k$$

Théorème

Si, pour tout $k \in \{1, \dots, m\}$, u_k désigne une application linéaire de F_k vers E' alors il existe une unique application linéaire u de E vers E' prolongeant les u_k i.e. vérifiant

$$\forall 1 \leqslant k \leqslant m, \forall x \in F_k, u(x) = u_k(x)$$

dém. :
Analyse / Unicité :
Supposons u solution.
Pour $x \in E$, on peut écrire $x = \sum_{k=1}^{m} x_k$ avec $x_k \in F_k$ et alors par linéarité,

$$u(x) = \sum_{k=1}^{m} u(x_k) = \sum_{k=1}^{m} u_k(x_k)$$

ce qui détermine entièrement u.
Synthèse / Existence :
Considérons l'application qui à $x = \sum_{k=1}^{m} x_k$ (avec $x_k \in F_k$) associe

$$u(x) = \sum_{k=1}^{m} u_k(x_k)$$

On vérifie aisément que u est linéaire et que sa restriction à E_k vaut u_k.
□

Corollaire

Si deux applications linéaires sont égales sur chacun des espaces E_i alors elles sont égales sur E.

Exemple On suppose la supplémentarité

$$E = F \oplus G$$

On appelle projection vectorielle sur F parallèlement à G l'endomorphisme $p \in \mathcal{L}(E)$ déterminé par

$$\forall x \in F, p(x) = x \text{ et } \forall x \in G, p(x) = 0_E$$

L'endomorphisme p vérifie

$$p^2 = p \text{ , Im}p = F \text{ et } \ker p = G$$

Remarque Inversement, si p est un endomorphisme p vérifiant $p^2 = p$ alors
a) $F = \text{Im}p$ et $G = \ker p$ sont des sous-espaces vectoriels supplémentaires de E ;
b) p est la projection sur F parallèlement à G.

3.4.7 Rang d'une application linéaire

Définition

On appelle rang d'une application linéaire u la dimension de son image

$$\text{rg}u \underset{\text{déf}}{=} \dim \text{Im}u$$

Proposition

> Soit $u \in \mathcal{L}(E, E')$ avec $\dim E < +\infty$
> On a rg$u \leqslant \dim E$ avec égalité si, et seulement si, u injective.

dém. :
Introduisons $(e_1, \ldots, e_n)$ une base de E avec $n = \dim E$
rg$u = \dim \text{Im} u = \dim u(E)$, or

$$u(E) = u(\text{Vect}(e_1, \ldots, e_n)) = \text{Vect}(u(e_1), \ldots, u(e_n))$$

Par suite rg$u \leqslant n$ avec égalité si, et seulement si, $(u(e_1), \ldots, u(e_n))$ est libre i.e. u injective.
□

Proposition

> Soit $u \in \mathcal{L}(E, E')$ avec $\dim E' < +\infty$
> On a rg$u \leqslant \dim E'$ avec égalité si, et seulement si, u surjective.

dém. :
rg$u = \dim \text{Im} u$ avec Im$u \subset F$.
Par suite rg$u \leqslant \dim F$ avec égalité si, et seulement si, Im$u = F$ i.e. u surjective.
□

Théorème

> Soit $u \in \mathcal{L}(E, E')$ et $v \in \mathcal{L}(E', E'')$. On a
>
> $$\text{rg}(v \circ u) \leqslant \min(\text{rg}u, \text{rg}v)$$

dém. :
rg$(v \circ u) = \dim \text{Im}(v \circ u) = \dim v(u(E))$.
D'une part, $v(u(E)) = \text{Im} v_{\restriction u(E)}$ donc rg$(v \circ u) = $ rg $v|_{u(E)} \leqslant \dim u(E) = $ rg$u$.
D'autre part, $v(u(E)) \subset v(F) = $ Imv donc rg$(v \circ u) \leqslant$ rg$v$.
□

Corollaire

> On ne modifie pas le rang d'une application linéaire en composant celle-ci avec un isomorphisme.

dém. :
Si φ est un isomorphisme alors

$$\text{rg}(\varphi \circ u) \leqslant \text{rg}u \text{ et } \text{rg}u = \text{rg}(\varphi^{-1} \circ \varphi \circ u) \leqslant \text{rg}(\varphi \circ u)$$

Ainsi rg$u = $ rg$(\varphi \circ u)$ et de même rg$u = $ rg$(u \circ \varphi)$
□

3.4.8 Théorème du rang

Théorème

> Si $u \in \mathcal{L}(E, E')$ et si S est un sous-espace vectoriel supplémentaire de $\ker u$ dans E alors E induit un isomorphisme de S sur Imu.

dém. :
Considérons la restriction $v : S \to$ Imu définie par $v(x) = u(x)$.
L'application v est bien définie et linéaire.

Pour $x \in \ker v$, on a $x \in \ker u \cap S = \{0_E\}$ donc $x = 0_E$. L'application linéaire v est injective.
Pour $y \in \mathrm{Im}u$, on peut écrire $y = u(x)$ avec $x \in E$. On peut aussi écrire $x = a + b$ avec $a \in \ker u$ et $b \in S$. On a alors

$$y = u(x) = u(a) + u(b) = 0_{E'} + v(b) = v(b)$$

Ainsi v est surjective et c'est donc un isomorphisme.
□

Corollaire

Si $\dim E < +\infty$ alors

$$\dim E = \mathrm{rg}u + \dim \ker u$$

Exemple Les hyperplans sont par définition les noyaux des formes linéaires non nulles : ils correspondent aussi aux sous-espaces vectoriels de dimension $n-1$.
Supposons $\dim E = n \in \mathbb{N}^\star$ et considérons $\varphi \in \mathcal{L}(E, \mathbb{K})$ une forme linéaire non nulle.
On a $\mathrm{Im}\varphi = \mathbb{K}$ et donc $\dim \ker \varphi = n - 1$
Un hyperplan de E est donc un espace dimension $n-1$.
La réciproque est aussi vraie.

Exemple On peut retrouver la formule de Grassman en appliquant la formule du rang à l'application $F \times G \to F + G$ définie par $(x, y) \mapsto x + y$.

3.4.9 Théorème d'isomorphisme

Théorème

On suppose

$$n = \dim E = \dim E' < +\infty$$

Pour $f \in \mathcal{L}(E, E')$, on a équivalence entre :
(i) f est un isomorphisme ;
(ii) f est injective ;
(iii) f est surjective ;
(iv) $\mathrm{rg}f = n$;
(v) $\exists g \in \mathcal{L}(E', E), g \circ f = \mathrm{Id}_E$;
(vi) $\exists h \in \mathcal{L}(E', E), f \circ h = \mathrm{Id}_{E'}$.
De plus, si tel est le cas

$$f^{-1} = g = h$$

dém. :
(i) $\Leftrightarrow$ (ii) et (iii)
(ii) $\Rightarrow$ (iv) car $\mathrm{rg}f = \dim E - \dim \ker f = n$.
(iv) $\Rightarrow$ (iii) car $\mathrm{rg}f = n = \dim F$ donc f surjective.
(iii) $\Rightarrow$ (ii) car $\dim \ker f = \dim E - \mathrm{rg}f = n - n = 0$
(i) $\Rightarrow$ (v) et (vi) ok
(v) $\Rightarrow$ (ii) car $g \circ f$ injective entraîne f injective.
(vi) $\Rightarrow$ (iii) car $f \circ h$ surjective entraîne f surjective.
□

Corollaire

Si $\dim E < +\infty$, ce qui précède permet de caractériser les automorphismes de E.

Exemple Soit $a_0, \ldots, a_n$ des éléments de $\mathbb{K}$ deux à deux distincts.
L'application $\varphi : \mathbb{K}_n[X] \to \mathbb{K}^{n+1}$ définie par

$$\varphi(P) = (P(a_0), \ldots, P(a_n))$$

est un isomorphisme de $\mathbb{K}$-espaces vectoriels.
En effet, φ est évidemment linéaire et

$$\dim \mathbb{K}_n[X] = n + 1 = \dim \mathbb{K}^{n+1} < +\infty$$

Soit $P \in \ker \varphi$. On a $P(a_0) = \ldots = P(a_n) = 0$.
Ainsi, le polynôme P admet au moins $n + 1$ racines, or $\deg P \leqslant n$ donc $P = 0$. Ainsi $\ker \varphi = \{0\}$ puis, par le théorème d'isomorphisme, φ est un isomorphisme.
En conséquence

$$\forall (b_0, \ldots, b_n) \in \mathbb{K}^{n+1}, \exists! P \in \mathbb{K}_n[X], \forall i \in [\![0, n]\!], P(a_i) = b_i$$

Pour décrire, un polynôme P solutions, on introduit

$$L_k = \prod_{i \neq k} \frac{X - a_i}{a_k - a_i}$$

On a $\varphi(L_k) = e_k$ avec $(e_0, \ldots, e_n)$ la base canonique de $\mathbb{K}^{n+1}$.
Par linéarité, le polynôme $P \in \mathbb{K}_n[X]$ vérifiant

$$\forall 0 \leqslant i \leqslant n, P(a_i) = b_i$$

est

$$P = \sum_{i=0}^{n} b_i L_i$$

3.5 Structure d'algèbre

3.5.1 Définition

Définition

On appelle $\mathbb{K}$-algèbre tout quadruplet $(A, +, \times, .)$ formé d'un ensemble A, de deux lois de composition internes $+, \times$ sur A et d'un produit extérieur opérant de $\mathbb{K}$ sur A vérifiant :
(1) $(A, +, .)$ est un $\mathbb{K}$-espace vectoriel ;
(2) $(A, +, \times)$ est un anneau ;
(3) $\forall \lambda \in \mathbb{K}$, $\forall x, y \in A$, $(\lambda.x)y = \lambda.(xy) = x(\lambda.y)$.

Exemple $\mathbb{K}$, $\mathbb{K}[X]$, $\mathcal{F}(X, \mathbb{K})$ sont des $\mathbb{K}$-algèbres commutatives.

Exemple $\mathcal{M}_n(\mathbb{K})$ et $\mathcal{L}(E)$ sont des $\mathbb{K}$-algèbres.

Remarque Si $\mathbb{L}$ est un sous-corps de $\mathbb{K}$ alors toute $\mathbb{K}$-algèbre est aussi par restriction une $\mathbb{L}$-algèbre.

Exemple $\mathbb{C}$ est une $\mathbb{C}$-algèbre, mais aussi une $\mathbb{R}$-algèbre.

3.5.2 Sous-algèbre

Définition

On appelle sous-algèbre d'une $\mathbb{K}$-algèbre A toute partie B de A vérifiant :
1) $1_A \in B$;
2) $\forall \lambda, \mu \in \mathbb{K}, \forall x, y \in B, \lambda x + \mu y \in B$;
3) $\forall x, y \in B$, $xy \in B$.

Remarque sous-algèbre = sous-espace vectoriel + sous-anneau.

Exemple Soit I un intervalle de $\mathbb{R}$ et $k \in \mathbb{N} \cup \{\infty\}$.
L'ensemble $\mathcal{C}^k(I, \mathbb{K})$ est une sous-algèbre de $\mathcal{F}(I, \mathbb{K})$.

Exemple $\mathbb{R}^{\mathbb{N}} = \mathcal{F}(\mathbb{N}, \mathbb{R})$ est une $\mathbb{R}$-algèbre.
$\mathcal{C} = \left\{(u_n) \in \mathbb{R}^{\mathbb{N}} / (u_n) \text{ converge}\right\}$ est une sous-algèbre de $\mathbb{R}^{\mathbb{N}}$.
$\mathcal{C}_0 = \left\{(u_n) \in \mathbb{R}^{\mathbb{N}} / u_n \to 0\right\}$ n'est pas une sous-algèbre de $\mathbb{R}^{\mathbb{N}}$ car ne contient par la suite $(1)_{n \in \mathbb{N}}$.

Exemple Soit $u \in \mathcal{L}(E)$.
L'ensemble $\mathcal{C} = \{v \in \mathcal{L}(E) / u \circ v = v \circ u\}$ est une sous-algèbre de $\mathcal{L}(E)$.

Théorème

Une sous-algèbre est une $\mathbb{K}$-algèbre pour les lois restreintes possédant les mêmes neutres.

dém. :
C'est un sous-espace vectoriel et un sous-anneau et la propriété calculatoire 3) est évidemment conservée.
□

3.5.3 Morphisme d'algèbres

Définition

Soit A et A' deux $\mathbb{K}$-algèbres. On appelle morphisme d'algèbres de A vers A' toute application $\varphi : A \to A'$ vérifiant :
1) $\varphi(1_A) = 1_{A'}$;
2) $\forall \lambda, \mu \in \mathbb{K}, \forall x, y \in A, \varphi(\lambda x + \mu y) = \varphi(x) + \varphi(y)$;
3) $\forall x, y \in A$, $\varphi(xy) = \varphi(x)\varphi(y)$.

Remarque morphisme d'algèbre = application linéaire + morphisme d'anneaux.
Le noyau d'un morphisme d'algèbre est en particulier un sous-espace vectoriel et un idéal.

Exemple L'application $z \in \mathbb{C} \mapsto \bar{z}$ est un morphisme de la $\mathbb{R}$-algèbre $\mathbb{C}$ dans elle-même.

Exemple Pour $P \in \mathrm{GL}_n(\mathbb{K})$, l'application $M \mapsto PMP^{-1}$ est un morphisme bijectif de la $\mathbb{K}$-algèbre $\mathcal{M}_n(\mathbb{K})$ dans elle-même.

Chapitre 4

Calculs matriciels

La théorie sur les matrices présentées en MPSI dans le cas où le corps de base est $\mathbb{R}$ ou $\mathbb{C}$ s'étend pour l'essentiel au cas où le corps de base est un corps quelconque.
On se limite cependant dans ce cours au cas où $\mathbb{K}$ est un sous-corps de $\mathbb{C}$: $\mathbb{K} = \mathbb{C}, \mathbb{R}, \mathbb{Q}, \ldots$

4.1 Calcul matriciel

4.1.1 Matrices rectangles

Définition

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n,p) à coefficients dans $\mathbb{K}$ i.e. l'ensemble des familles $A = (a_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p}$ d'éléments de $\mathbb{K}$. Une telle matrice est généralement figurée par un tableau

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$$

Exemple On note

$$E_{i,j} = \begin{pmatrix} 0 & & 0 \\ & 1 & \\ 0 & & 0 \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$$

appelée matrice élémentaire d'indice (i,j) de $\mathcal{M}_{n,p}(\mathbb{K})$.

Théorème

$\mathcal{M}_{n,p}(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel de dimension np et d'élément nul $O_{n,p}$.
La famille des matrices élémentaires $(E_{i,j})_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$

Définition

Pour $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$, on pose $AB = (c_{i,k}) \in \mathcal{M}_{n,q}(\mathbb{K})$ avec

$$c_{i,k} \underset{\text{déf}}{=} \sum_{j=1}^{p} a_{i,j} b_{j,k}$$

Exemple Pour

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

on a

$$AX = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,p}x_p \\ \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p \end{pmatrix}$$

Exemple Pour $E_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $E_{k,\ell} \in \mathcal{M}_{p,q}(\mathbb{K})$, on a $E_{i,j} \times E_{k,\ell} = \delta_{j,k} E_{i,\ell}$.
En effet,
- si $j \neq k$ alors $E_{i,j} \times E_{k,\ell} = O_{n,q}$ car les 1 ne se croisent pas.
- si $j = k$ alors $E_{i,j} E_{k,\ell} = E_{i,\ell} \in \mathcal{M}_{n,q}(\mathbb{K})$ car les 1 se croisent lors du calcul du coefficient d'indice (i,ℓ).
On retient $E_{i,j} \times E_{k,\ell} = \delta_{j,k} E_{i,\ell}$.

Remarque Les opérations matricielles peuvent aussi être conduites en raisonnant « par blocs » .

Exemple Calcul de A^2 pour $A = \begin{pmatrix} O_n & -I_n \\ I_n & O_n \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{R})$.
Le produit par blocs se pose comme un produit de matrice à coefficients (en prenant garde à l'ordre des facteurs).

$$A^2 = \begin{pmatrix} -I_n & O_n \\ O_n & -I_n \end{pmatrix} = -I_{2n}$$

Exemple Calcul de MX avec

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ avec } A, B, C, D \in \mathcal{M}_n(\mathbb{K}) \text{ et } X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \text{ avec } X_1, X_2 \in \mathcal{M}_{n,1}(\mathbb{K})$$

On obtient

$$MX = \begin{pmatrix} AX_1 + BX_2 \\ CX_1 + DX_2 \end{pmatrix}$$

Exemple Calcul des puissances de

$$M = \begin{pmatrix} A & B \\ O_n & A \end{pmatrix} \text{ avec } A, B \in \mathcal{M}_n(\mathbb{K}) \text{ commutant}$$

On a

$$M^2 = \begin{pmatrix} A^2 & AB + BA \\ O_n & A^2 \end{pmatrix}$$

Puisque $AB = BA$, on simplifie

$$M^2 = \begin{pmatrix} A^2 & 2AB \\ O_n & A^2 \end{pmatrix}$$

Par récurrence, on montre

$$\forall k \in \mathbb{N}^\star, M^k = \begin{pmatrix} A^k & kA^{k-1}B \\ O_n & A^k \end{pmatrix}$$

4.1.2 Matrices carrées

Définition

On note $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.

Théorème

$\mathcal{M}_n(\mathbb{K})$ est une $\mathbb{K}$-algèbre de dimension n^2 de neutres O_n et I_n.
Celle-ci est non commutative dès que $n \geqslant 2$.

Exemple L'ensemble $D_n(\mathbb{K})$ formé des matrices diagonales est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$.
On observe

$$\begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \begin{pmatrix} \mu_1 & & (0) \\ & \ddots & \\ (0) & & \mu_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n\mu_n \end{pmatrix}$$

Exemple L'ensemble $T_n^+(\mathbb{K})$ formé des matrices triangulaires supérieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$.
On observe

$$\begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \begin{pmatrix} \mu_1 & & \star' \\ & \ddots & \\ (0) & & \mu_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mu_1 & & \star'' \\ & \ddots & \\ (0) & & \lambda_n\mu_n \end{pmatrix}$$

4.1.3 Problèmes de commutation

Proposition

Les matrices commutant avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$ sont les matrices scalaires i.e. les matrices λI_n avec $\lambda \in \mathbb{K}$.

dém. :
Les matrices scalaires commutent avec toute matrice de $\mathcal{M}_n(\mathbb{K})$.
Inversement, soit $A = (a_{i,j})$ une matrice commutant avec tout élément de $\mathcal{M}_n(\mathbb{K})$

$$\forall M \in \mathcal{M}_n(\mathbb{K}), AM = MA$$

Pour $M = E_{i,j}$ avec $i \neq j$, on a $E_{i,j}A = AE_{i,j}$.
Or $[E_{i,j}A]_{i,j} = a_{i,i}$ et $[AE_{i,j}]_{i,j} = a_{j,j}$ donc $a_{i,i} = a_{j,j}$.

Aussi $[E_{i,j}A]_{i,i} = a_{j,i}$ et $[AE_{i,j}]_{i,i} = 0$ donc $a_{j,i} = 0$.
Ainsi, la matrice A est diagonale de diagonale constante.
□

Proposition

Soit D une matrice diagonale à coefficients diagonaux deux à deux distincts.
Les matrices commutant avec D sont les matrices diagonales.

dém. :
On peut écrire $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ avec $\lambda_1, \ldots, \lambda_n$ deux à deux distincts.
Pour $M = (m_{i,j})_{1\leqslant i,j\leqslant n} \in \mathcal{M}_n(\mathbb{K})$, on a

$$DM = (\lambda_i m_{i,j})_{1\leqslant i,j\leqslant n} \text{ et } MD = (\lambda_j m_{i,j})_{1\leqslant i,j\leqslant n}$$

et donc

$$MD = DM \Leftrightarrow \forall 1 \leqslant i,j \leqslant n, (\lambda_i - \lambda_j)m_{i,j} = 0$$

Cette dernière condition est vérifiée si, et seulement si, M est diagonale.
□

Remarque Ce résultat peut être étendu en raisonnant par blocs : les matrices commutant avec

$$D = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix} \text{ avec } \lambda \neq \mu$$

sont les matrices de la forme

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix}$$

4.1.4 Noyau, image et rang d'une matrice

On identifie les tuples éléments de $\mathbb{K}^n$ avec les colonnes éléments de $\mathcal{M}_{n,1}(\mathbb{K})$ via l'isomorphisme

$$\begin{array}{ccc} \mathbb{K}^n & \to & \mathcal{M}_{n,1}(\mathbb{K}) \\ x = (x_1, \ldots, x_n) & \mapsto & X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array}$$

Définition

Pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$, on appelle application linaire canoniquement associée à la matrice A l'application $u : \mathbb{K}^p \mapsto \mathbb{K}^n$ qui à $x \in \mathbb{K}^p$ associe $y \in \mathbb{K}^n$ définie par

$$y = Ax$$

Exemple Précisons l'application linéaire canoniquement associée à la matrice
$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathcal{M}_{3,2}(\mathbb{R})$.

Par produit matriciel avec la colonne X de coefficients x_1, x_2, x_3, on obtient l'application linéaire

$$\begin{aligned}&\mathbb{R}^3 \to \mathbb{R}^2\\&(x_1, x_2, x_3) \mapsto (x_1 + 2x_2 - x_3, x_2 + x_3)\end{aligned}$$

Définition

On définit le noyau, l'image et le rang de la matrice A par
- $\ker A = \ker u = \{x \in \mathbb{K}^p / Ax = 0\}$;
- $\mathrm{Im}A = \mathrm{Im}u = \{y \in \mathbb{K}^n / \exists x \in \mathbb{K}^p, y = Ax\}$;
- $\mathrm{rg}A = \dim \mathrm{Im}A$.

Proposition

Si $C_1, \ldots, C_p$ désignent les colonnes de A alors

$$\mathrm{Im}A = \mathrm{Vect}(C_1, \ldots, C_p) \text{ et } \mathrm{rg}A = \mathrm{rg}(C_1, \ldots, C_p)$$

dém. :

$$\mathrm{Im}A = \{Ax / x \in \mathbb{K}^p\} = \{x_1C_1 + \cdots + x_pC_p / x_1, \ldots, x_p \in \mathbb{K}\}$$

donc

$$\mathrm{Im}A = \mathrm{Vect}(C_1, \ldots, C_p) \text{ puis } \mathrm{rg}A = \mathrm{rg}(C_1, \ldots, C_p)$$

□

Proposition

$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \mathrm{rg}(A) \leqslant \min(n, p)$,
$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \forall B \in \mathcal{M}_{p,q}(\mathbb{K}), \mathrm{rg}(AB) \leqslant \min(\mathrm{rg}A, \mathrm{rg}B)$.

dém. :

$$\mathrm{rg}A = \mathrm{rg}u \leqslant \min(\dim \mathcal{M}_{p,1}(\mathbb{K}), \dim \mathcal{M}_{n,1}(\mathbb{K})) = \min(p, n)$$

Notons aussi v et w les applications linéaires canoniquement associées aux matrices B et AB. On vérifie aisément $w = u \circ v$.

$$\mathrm{rg}(AB) = \mathrm{rg}(u \circ v) \leqslant \min(\mathrm{rg}u, \mathrm{rg}v) = \min(\mathrm{rg}A, \mathrm{rg}B)$$

□

Théorème

On a la formule du rang

$$\mathrm{rg}A + \dim \ker A = p$$

Exemple Déterminons image, noyau et rang de

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

On a
$$A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 + x_3 = 0 \\ x_2 + x_3 = 0 \\ x_1 - x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_2 = x_1 \\ x_3 = -x_1 \end{cases}$$
Donc
$$\ker A = \{(x_1, x_1, -x_1)/x_1 \in \mathbb{R}\} = \text{Vect}(1, 1, -1)$$
Par la formule du rang rg$A = 2$.
Puisque les vecteurs
$$y_1 = (1, 0, 1) = Ae_1, y_2 = (1, 1, -1) = Ae_2$$
appartiennent à l'image de A et puisqu'ils sont aussi indépendantes
$$\text{Im}A = \text{Vect}(y_1, y_2)$$

4.1.5 Matrices inversibles

Définition

On dit que $A \in \mathcal{M}_n(\mathbb{K})$ est inversible s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ vérifiant
$$AB = BA = I_n$$
Cette matrice B est unique, on l'appelle inverse de A et on la note A^{-1}.

Exemple Une matrice triangulaire supérieure est inversible si, et seulement si, ses coefficients diagonaux sont non nuls et alors
$$\begin{pmatrix} a_1 & & \star \\ & \ddots & \\ (0) & & a_n \end{pmatrix}^{-1} = \begin{pmatrix} 1/a_1 & & \star \\ & \ddots & \\ (0) & & 1/a_n \end{pmatrix}$$

Théorème

L'ensemble $\text{GL}_n(\mathbb{K})$ des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est un groupe multiplicatif de neutre I_n.

dém. :
C'est le groupe des inversibles de $\mathcal{M}_n(\mathbb{K})$.
□

Attention : $(AB)^{-1} = B^{-1}A^{-1}$.

Proposition

On ne modifie pas le rang d'une matrice en la multipliant par une matrice inversible.

dém. :
Soit $P \in GL_n(\mathbb{K})$ et $A \in \mathcal{M}_{n,p}(\mathbb{K})$.
On a rg$(PA) \leqslant A$ et rg$A = $ rg$(P^{-1}PA) \leqslant$ rg(PA) puis $=$.
□

Théorème

Pour $A \in \mathcal{M}_n(\mathbb{K})$, on a équivalence entre :
(i) A est inversible ;
(ii) $\ker A = \{0\}$;
(iii) $\mathrm{Im}A = \mathbb{K}^n$;
(iv) $\mathrm{rg}A = n$;
(v) $\exists B \in \mathcal{M}_n(\mathbb{K}), AB = I_n$;
(vi) $\exists C \in \mathcal{M}_n(\mathbb{K}), CA = I_n$.
De plus, si tel est le cas

$$B = C = A^{-1}$$

dém. :
(i) $\Leftrightarrow$ (iv) est connue et le reste est alors immédiat.
□

Exemple Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $A + B = AB$. Montrons $AB = BA$.
On a $(I_n - A)(I_n - B) = I_n - (A + B) + AB = I_n$ donc $I_n - A$ est inversible d'inverse $I_n - B$.
Par suite $(I_n - B)(I_n - A) = I_n$ donc $BA = A + B = AB$.

Exemple Inversons

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix}$$

Par la méthode du pivot, on opère sur les lignes d'une matrice de blocs A et I_n pour transformer A en I_n. On sait qu'alors le bloc I_n sera transformé en A^{-1}.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array}\right)$$

On conclut

$$A^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$$

4.1.6 Transposition

Définition

Pour $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$, on pose ${}^tA = (a'_{j,i}) \in \mathcal{M}_{p,n}(\mathbb{K})$ avec

$$a'_{j,i} \underset{\text{déf}}{=} a_{i,j}$$

Remarque Si $A = (a_{i,j})_{i,j}$ alors ${}^tA = (a_{i,j})_{j,i}$.

Proposition

$\forall \lambda, \mu \in \mathbb{K}, \forall A, B \in \mathcal{M}_{n,p}(\mathbb{K}), {}^t(\lambda A + \mu B) = \lambda^t A + \mu^t B$
$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \forall B \in \mathcal{M}_{p,q}(\mathbb{K}), {}^t(AB) = {}^tB^tA.$
$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), {}^t\left({}^tA\right) = A$
$\forall A \in \mathrm{GL}_n(\mathbb{K}), {}^tA \in \mathrm{GL}_n(\mathbb{K})$ et $\left({}^tA\right)^{-1} = {}^t\left(A^{-1}\right)$

Définition

Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est dite symétrique (resp. antisymétrique) si ${}^tM = M$ (resp. ${}^tM = -M$)

Théorème

Les ensembles $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ formés des matrices symétriques et antisymétriques de $\mathcal{M}_n(\mathbb{R})$ sont des sous-espaces vectoriels supplémentaires et

$$\dim \mathcal{S}_n(\mathbb{R}) = \frac{n(n+1)}{2} \text{ et } \dim \mathcal{A}_n(\mathbb{R}) = \frac{n(n-1)}{2}$$

4.2 Représentations matricielles

4.2.1 Matrices des coordonnées d'un vecteur

Soit E un $\mathbb{K}$-espace vectoriel de dimension n.
On considère une base $e = (e_1, \dots, e_n)$ de E. On a

$$\forall x \in E, \exists!(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, x = \lambda_1.e_1 + \cdots + \lambda_n.e_n$$

Définition

On note

$$\mathrm{Mat}_e(x) \underset{\text{déf}}{=} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$$

la matrice des coordonnées de x dans la base e.

Exemple $\text{Mat}_e(e_i) = \begin{pmatrix} (0) \\ 1 \\ (0) \end{pmatrix} = E_i$.

Théorème

L'application $x \mapsto \text{Mat}_e(x)$ est un isomorphisme du $\mathbb{K}$-espace vectoriel E vers $\mathcal{M}_{n,1}(\mathbb{K})$.

Définition

Soit $x_1, \ldots, x_p \in E$. On note

$$\text{Mat}_e(x_1, \ldots, x_p) \in \mathcal{M}_{n,p}(\mathbb{K})$$

la matrice dont les colonnes sont

$$\text{Mat}_e(x_1), \ldots, \text{Mat}_e(x_p)$$

Exemple $\text{Mat}_e e = (E_1 \mid \ldots \mid E_n) = I_n$.

Proposition

Si $A = \text{Mat}_e(x_1, \ldots, x_p)$ alors $\text{rg}A = \text{rg}(x_1, \ldots, x_p)$.

dém. :
Notons φ l'isomorphisme $x \in E \mapsto \text{Mat}_e(x)$.
Les colonnes $C_1, \ldots, C_p$ de A sont données pas $C_j = \varphi(x_j)$.

$$\text{rg}A = \text{rg}(C_1, \ldots, C_p) = \dim \text{Vect}(C_1, \ldots, C_p)$$

donc

$$\text{rg}A = \dim \text{Vect}(\varphi(x_1), \ldots, \varphi(x_p)) = \dim \varphi(\text{Vect}(x_1, \ldots, x_p))$$

Mais l'application φ est un isomorphisme donc

$$\text{rg}A = \dim \varphi(\text{Vect}(x_1, \ldots, x_p)) = \dim \text{Vect}(x_1, \ldots, x_p) = \text{rg}(x_1, \ldots, x_p)$$

□

4.2.2 Matrice d'une application linéaire

Soit E et F des $\mathbb{K}$-espaces vectoriels de dimensions p et n.
On considères deux bases $e = (e_1, \ldots, e_p)$ et $f = (f_1, \ldots, f_n)$ des espaces E et F.

Définition

Pour $u \in \mathcal{L}(E, F)$, on note

$$\text{Mat}_{e,f}(u) \underset{\text{déf}}{=} \text{Mat}_f \left(u(e_1), \ldots, u(e_p)\right) \in \mathcal{M}_{n,p}(\mathbb{K})$$

la matrice de l'application linéaire u relative aux base e et f.

Exemple Soit $a_0, \ldots, a_n \in \mathbb{K}$ deux à deux distincts.
Etudions quelques représentations matricielles de l'application linéaire $\varphi : \mathbb{K}_n[X] \to \mathbb{K}^{n+1}$ définie par

$$\varphi(P) = (P(a_0), \ldots, P(a_n))$$

Soit $(1, X, \ldots, X^n)$ et $c = (c_0, \ldots, c_n)$ les bases canoniques de $\mathbb{K}_n[X]$ et $\mathbb{K}^{n+1}$.
Formons

$$A = \mathrm{Mat}_{(1,X,\ldots,X^n),c}(\varphi)$$

On a $\varphi(X^k) = \left(a_0^k, \ldots, a_n^k\right)$ donc

$$\mathrm{Mat}_c(\varphi(X^k)) = \begin{pmatrix} a_0^k \\ \vdots \\ a_n^k \end{pmatrix}$$

et alors

$$A = \begin{pmatrix} 1 & a_0 & a_0^2 & \cdots & a_0^n \\ 1 & a_1 & a_1^2 & \cdots & a_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^n \end{pmatrix}$$

Soit $(L_0, \ldots, L_n)$ la base de $\mathbb{K}_n[X]$ formée des polynômes d'interpolation de Lagrange en $a_0, \ldots, a_n$.
Puisque $\varphi(L_k) = c_k$, la matrice de φ dans $(L_0, \ldots, L_n)$ et $\mathcal{C}$ est I_{n+1}.

Exemple Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. La matrice de l'application linéaire canoniquement associée à A dans les bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$ est A.
En effet, $\varphi_A(e_j) = Ae_j$ correspond à la j-ème colonne de A.

Théorème

Soit $u \in \mathcal{L}(E, F)$.
La matrice $\mathrm{Mat}_{e,f}(u)$ est l'unique matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ vérifiant

$$\forall x \in E, \forall y \in F, y = u(x) \Leftrightarrow Y = AX$$

avec $A = \mathrm{Mat}_{e,f}(u)$ $X = \mathrm{Mat}_e(x)$ et $Y = \mathrm{Mat}_f(y)$.

Théorème

L'application $u \in \mathcal{L}(E, F) \mapsto \mathrm{Mat}_{e,f}(u) \in \mathcal{M}_{n,p}(\mathbb{K})$ est un isomorphisme de $\mathbb{K}$-espaces vectoriels.

4.2.3 Matrice d'un endomorphisme

Soit E un $\mathbb{K}$-espace vectoriel de dimension n.
On considère $e = (e_1, \ldots, e_n)$ une base de E.

Définition

Pour $u \in \mathcal{L}(E)$, on note

$$\mathrm{Mat}_e(u) \underset{\text{déf}}{=} \mathrm{Mat}_{e,e}(u) \in \mathcal{M}_n(\mathbb{K})$$

la matrice de l'endomorphisme u dans la base e.

Exemple $\mathrm{Mat}_e(\mathrm{Id}_E) = I_n$.

Théorème

L'application $u \in \mathcal{L}(E) \mapsto \mathrm{Mat}_e(u) \in \mathcal{M}_p(\mathbb{K})$ est un isomorphisme de $\mathbb{K}$-algèbres.

4.2.4 Transport du vectoriel au matriciel

Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimension p et n munis de bases e et f.

Vecteur	Matrice colonne
$x \in E$	$X \in \mathcal{M}_{p,1}(\mathbb{K})$
0	$O_{p,1}$
$\lambda x + \mu x'$	$\lambda X + \mu X'$
Application linéaire	Matrice rectangle
$u \in \mathcal{L}(E, F)$	$A \in \mathcal{M}_{n,p}(\mathbb{K})$
$\tilde{o}$	$O_{n,p}$
$y = u(x)$	$Y = AX$
$\lambda u + \mu v$	$\lambda A + \mu B$
$u \circ v$	AB
u isomorphisme, u^{-1}	A inversible, A^{-1}
Imu, ker u et rg$u$	ImA, ker A et rg$A$
Endomorphisme	Matrice carrée
$u \in \mathcal{L}(E)$	$A \in \mathcal{M}_p(\mathbb{K})$
Id_E	I_p
u^n	A^n
$u \in \mathrm{GL}(E)$, u^{-1}	$A \in \mathrm{GL}_p(\mathbb{K})$, A^{-1}
$\det u$	$\det A$
Formes linéaires	Matrice ligne
$\varphi \in E^\star$	$L \in \mathcal{M}_{1,p}(\mathbb{K})$
$y = \varphi(x) \in \mathbb{K}$	$(y) = LX$

Exemple Déterminons les endomorphismes d'un $\mathbb{K}$-espace vectoriel E de dimension n commutant avec tout autre endomorphisme.
Soit $u \in \mathcal{L}(E)$.
Considérons e une base de E et $A = \mathrm{Mat}_e(u) \in \mathcal{M}_n(\mathbb{K})$.
u commute avec tout endomorphisme de E si, et seulement si,

$$\forall B \in \mathcal{M}_n(\mathbb{K}), AB = BA$$

i.e. A scalaire. Ainsi, les endomorphismes recherchés sont les homothéties.

Exemple Calcul des puissances de

$$J = \begin{pmatrix} 0 & & (0) & 1 \\ 1 & \ddots & & (0) \\ & \ddots & \ddots & \\ (0) & & 1 & 0 \end{pmatrix}$$

On introduit $E = \mathbb{K}^n$ et u l'endomorphisme canoniquement associé à J.
On a $u(e_1) = e_2$, $u(e_2) = e_3$,..., $u(e_{n-1}) = e_n$ et $u(e_n) = e_1$.
On en déduit $u^k(e_i) = e_{i+k}$ avec $e_i = e_j$ si $i \equiv j \quad [n]$.
On peut alors exprimer J^k.

4.2.5 Formules de changement de bases

4.2.5.1 Matrice de passage

Soit E un $\mathbb{K}$-espace vectoriel de dimension n.
On considère e et e' deux bases de E.

Définition

On appelle matrice de passage de e à e' la matrice

$$P_e^{e'} = \text{Mat}_e e' \in \mathcal{M}_n(\mathbb{K})$$

Proposition

$P_e^{e'} = \text{Mat}_{e',e}(\text{Id}_E) \in \text{GL}_n(\mathbb{K})$ et $\left(P_e^{e'}\right)^{-1} = P_{e'}^e$

4.2.5.2 Nouvelles coordonnées d'un vecteur

Théorème

Si P est la matrice de passage d'une base e à une base e' d'un $\mathbb{K}$-espace vectoriel E alors

$$\forall x \in E, X = PX'$$

avec $X = \text{Mat}_e(x)$ et $X' = \text{Mat}_{e'}(x)$.

dém. :

$$\text{Mat}_e(x) = \text{Mat}_e\left(\text{Id}_E(x)\right) = \text{Mat}_{e',e}(\text{Id}_E) \times \text{Mat}_{e'}(x) = PX'$$

□

4.2.5.3 Nouvelle matrice d'une application linéaire

Théorème

Si P est la matrice de passage d'une base e à une base e' d'un $\mathbb{K}$-espace vectoriel E et si Q est la matrice de passage d'une base f à une base f' d'un $\mathbb{K}$-espace vectoriel F alors

$$\forall u \in \mathcal{L}(E, F), A' = Q^{-1}AP$$

avec $A = \text{Mat}_{e,f}(u)$ et $A' = \text{Mat}_{e',f'}(u)$.

dém. :
Soit $x \in E$ et $y \in F$. On note

$$X = \text{Mat}_e(x),\ X' = \text{Mat}_{e'}(x),\ Y = \text{Mat}_f(y) \text{ et } Y' = \text{Mat}_{f'}(y)$$

On a $X = PX'$ et $Y = QY'$. Si $y = u(x)$ alors

$$Y = AX \text{ et } Y' = A'X'$$

donc $AX = QA'X'$ puis

$$AX = QA'P^{-1}X$$

Or ceci doit être valable pour toute colonne X donc

$$A = QA'P^{-1}$$

□

Corollaire

On a

$$\forall u \in \mathcal{L}(E), A' = P^{-1}AP$$

avec $A = \text{Mat}_e(u)$, $A' = \text{Mat}_{e'}(u)$.

4.2.6 Matrices équivalentes

Définition

On dit qu'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est équivalente à une matrice $B \in \mathcal{M}_{n,p}(\mathbb{K})$ s'il existe $P \in \text{GL}_p(\mathbb{K})$ et $Q \in \text{GL}_n(\mathbb{K})$ telles que

$$B = Q^{-1}AP$$

Exemple Les matrices d'une même application linéaire sont équivalentes.

Proposition

L'équivalence de matrice est une relation d'équivalence sur $\mathcal{M}_{n,p}(\mathbb{K})$.

Théorème

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $r \in \mathbb{N}$ avec $0 \leqslant r \leqslant \min(n,p)$.

$$\text{rg}A = r \Leftrightarrow A \text{ est équivalente à } J_r$$

avec

$$J_r = \left(\begin{array}{c|c} I_r & O_{r,p-r} \\ \hline O_{n-r,r} & O_{n-r,p-r} \end{array}\right) \in \mathcal{M}_{n,p}(\mathbb{K})$$

dém. :

($\Leftarrow$) Car $\text{rg}(J_r) = r$ et l'on ne modifie pas le rang en multipliant par des matrices inversibles.

($\Rightarrow$) Soit E et F deux $\mathbb{K}$-espaces vectoriels de dimensions p et n munis de bases e et f.

On considère $u \in \mathcal{L}(E,F)$ déterminée par

$$\text{Mat}_{e,f}(u) = A$$

Si $r = \text{rg}A$ alors $r = \text{rg}u$ et donc $\dim\ker u = p - r$.

Soit G un supplémentaire de $\ker u$ dans E :

$$E = G \oplus \ker u$$

avec $\dim G = r$.
Soit une base $e' = (e'_1, \ldots, e'_r, e'_{r+1}, \ldots, e'_p)$ adaptée à la décomposition $E = G \oplus \ker u$.
L'application $u_{|G} : G \to \text{Im}u$ est un isomorphisme de $\mathbb{K}$-espaces vectoriels.
Posons
$$f'_1 = u(e'_1), \ldots, f'_r = u(e'_r)$$
La famille $(f'_1, \ldots, f'_r)$ est base de Imu, on peut la compléter en une base $f' = (f'_1, \ldots, f'_p)$ de F.
On obtient $\text{Mat}_{e',f'}(u) = J_r$ donc A et J_r sont équivalentes car représentent la même application linéaire.
□

Corollaire

Deux matrices sont équivalentes si, et seulement si, elles ont le même rang.

Exemple Soit $A \in \mathcal{M}_n(\mathbb{K})$ de rang 1.
Montrons qu'il existe $X, Y \in \mathcal{M}_{n,1}(\mathbb{K})$ tels que $A = Y^tX$.
(1) Analyse : Si $A = Y^tX$ alors
$$A = \begin{pmatrix} x_1y_1 & \cdots & x_ny_1 \\ \vdots & & \vdots \\ x_1y_n & \cdots & x_ny_n \end{pmatrix} = (x_1Y \ldots x_nY)$$
et donc les colonnes de A sont colinéaires à une même colonne Y, les coefficients de colinéarité formant la matrice X.
Synthèse :
rg$A = 1$ donc ImA est une droite vectorielle.
Soit $Y \neq 0$ élément de ImA :
$$\text{Im}A = \text{Vect}Y$$
Notons $C_1, \ldots, C_n$ les colonnes de A.
Puisque $C_1, \ldots, C_n \in \text{Im}A$, il existe $x_1, \ldots, x_n \in \mathbb{K}$ tels que $C_j = x_jY$.
Pour ${}^tX = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$, on a
$$Y^tX = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix} = A$$
(2) A est équivalente à J_1 donc on peut écrire
$A = QJ_1P$ avec $P, Q \in \text{GL}_n(\mathbb{K})$.
On observe que $J_1 = E_1{}^tE_1$ donc $A = Y^tX$ avec $Y = QE_1$ et ${}^tX = {}^tE_1P$ i.e. $X = {}^tPE_1$.

4.2.7 Matrices semblables

Définition

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est semblable à une matrice $B \in \mathcal{M}_n(\mathbb{K})$ s'il existe $P \in \text{GL}_n(\mathbb{K})$ telle que
$$B = P^{-1}AP$$

Exemple Les matrices d'un même endomorphisme sont semblables.

Exemple Si A est semblable à une matrice scalaire λI_n alors il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A = P^{-1}(\lambda I_n)P$ et donc $A = \lambda P^{-1}P = \lambda I_n$.

Proposition

La similitude définit une relation d'équivalence sur $\mathcal{M}_n(\mathbb{K})$.

Proposition

Deux matrices semblables sont équivalentes et ont donc même rang.
La réciproque est fausse.

Protocole :
Pour montrer qu'une matrice A de $\mathcal{M}_n(\mathbb{K})$ est semblable à une matrice B simple, il est fréquent de transposer le problème en termes vectoriels.
- on introduit u l'endomorphisme canoniquement associé à la matrice A ;
- on détermine (souvent par analyse-synthèse) une nouvelle base de $\mathbb{K}^n$ dans laquelle u est représenté par B.

Exemple Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $A^{n-1} \neq O$ et $A^n = O$.
Montrons que A est semblable à

$$B = \begin{pmatrix} 0 & & & (0) \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ (0) & & 1 & 0 \end{pmatrix}$$

Soit u l'application linéaire canoniquement associée à la matrice A.
On a $u^n = \tilde{0}$ et $u^{n-1} \neq \tilde{0}$.
Déterminons une base $e = (e_1, \ldots, e_n)$ de $\mathbb{K}^n$ dans laquelle u est représenté par B.
Analyse :
Supposons $e = (e_1, \ldots, e_n)$ convenable.
On a $u(e_1) = e_2, \ldots, u(e_{n-1}) = e_n$ et $u(e_n) = 0_E$.
On en déduit $e_2 = u(e_1)$, $e_3 = u^2(e_1)$,…, $e_n = u^{n-1}(e_1)$.
Notons que la propriété $u(e_n) = 0$ sera obtenue et que nécessairement $e_1 \notin \ker u^{n-1}$ pour que $e_n \neq 0_E$.
Synthèse :
Soit $e_1 \notin \ker u^{n-1}$ et $e = (e_1, \ldots, e_n)$ avec $e_2 = u(e_1)$, $e_3 = u^2(e_1)$,…, $e_n = u^{n-1}(e_1)$.
On a $u(e_1) = e_2, \ldots, u(e_{n-1}) = e_n$ et $u(e_n) = 0_E$.
Il reste à montrer que e est une base de E.
Supposons $\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n = 0_E$.
On a $\lambda_1 e_1 + \lambda_2 u(e_1) + \cdots + \lambda_n u^{n-1}(e_1) = 0_E$.
En appliquant f plusieurs fois, on obtient successivement
$\lambda_1 u(e_1) + \cdots + \lambda_{n-1} u^{n-1}(e_1) = 0_E$,…, $\lambda_1 u^{n-2}(e_1) + \lambda_2 u^{n-1}(e_1) = 0_E$ et $\lambda_1 u^{n-1}(e_1) = 0_E$.
Or $u^{n-1}(e_1) \neq 0_E$ donc on résout le système triangulaire formé pour obtenir $\lambda_1 = \ldots = \lambda_n = 0$.
Finalement, e est une famille libre formée de $n = \dim E$ vecteurs de E, c'est donc une base de E.

4.2.8 Traces

4.2.8.1 Trace d'une matrice carrée

Définition

On appelle trace d'une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ le scalaire

$$\text{tr}A = a_{1,1} + \cdots + a_{n,n}$$

Proposition

La trace définit une forme linéaire non nulle sur $\mathcal{M}_n(\mathbb{K})$.

dém. :
On vérifier aisément que l'application trace est linéaire et non nulle.
□

Exemple L'ensemble des matrices de trace nulle de $\mathcal{M}_n(\mathbb{K})$ est un hyperplan car noyau d'une forme linéaire non nulle.

Théorème

$$\forall A \in \mathcal{M}_{n,p}(\mathbb{K}), \forall B \in \mathcal{M}_{p,n}(\mathbb{K}), \text{tr}(AB) = \text{tr}(BA)$$

dém. :
Introduisons les coefficients des matrices A et B : $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,i}) \in \mathcal{M}_{p,n}(\mathbb{K})$.
Les matrices AB et BA sont carrées donc on peut calculer leur trace et on a

$$\text{tr}(AB) = \sum_{i=1}^{n} [AB]_{i,i} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}b_{j,i}$$

et

$$\text{tr}(BA) = \sum_{j=1}^{p} [BA]_{j,j} = \sum_{j=1}^{p}\sum_{i=1}^{n} b_{j,i}a_{i,j}$$

En permutant les deux sommes, on obtient $\text{tr}(BA) = \text{tr}(AB)$.
□

Corollaire

Deux matrices semblables ont même trace.

dém. :
Si $B = P^{-1}AP$ alors $\text{tr}B = \text{tr}\left(P^{-1}(AP)\right) = \text{tr}\left((AP)P^{-1}\right) = \text{tr}A$
□

4.2.8.2 Trace d'un endomorphisme

Définition

On appelle trace d'un endomorphisme d'un $\mathbb{K}$-espace vectoriel de dimension finie la trace commune aux matrices représentant cet endomorphisme.

Exemple $\text{tr}(\text{Id}_E) = n = \dim E$

Théorème

La trace définit une forme linéaire sur $\mathcal{L}(E)$ vérifiant

$$\forall u \in \mathcal{L}(E, F), \forall v \in \mathcal{L}(F, E), \text{tr}(u \circ v) = \text{tr}(v \circ u)$$

Théorème

Si p est une projection vectorielle d'un $\mathbb{K}$-espace vectoriel E de dimension finie alors

$$\text{tr}p = \text{rg}p$$

dém. :
On sait

$$E = \text{Im}p \oplus \ker p$$

Dans une base adaptée à cette décomposition, la matrice de p est de la forme

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

avec $r = \dim \text{Im}p = \text{rg}p$. Par suite $\text{tr}p = \text{rg}p$.
□

4.3 Déterminants

4.3.1 Définitions

4.3.1.1 Déterminant d'une matrice carrée

Définition

On appelle déterminant d'une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ le scalaire

$$\det A \underset{\text{déf}}{=} \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

encore noté

$$\begin{vmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{vmatrix}_{[n]}$$

Exemple Un déterminant d'ordre 0 vaut 1.

Exemple Un déterminant d'ordre 1 est égal à son coefficient.

Exemple Un déterminant d'ordre 2 se calcule par un produit en croix

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Exemple Un déterminant d'ordre 3 peut se calculer par la règle de Sarrus.

Exemple Si $A = (a_{i,j}) \in T_n^+(\mathbb{K})$ alors $\det A = \prod_{i=1}^{n} a_{i,i}$.

En effet, pour $i > j$, $a_{i,j} = 0$ donc $\prod_{i=1}^{n} a_{\sigma(i),i} = 0$ dès qu'il existe i vérifiant $\sigma(i) > i$.
En simplifiant les termes correspondants de la somme définissant le déterminant, il ne reste que les permutations σ vérifiant

$$\forall i \in \{1, \dots, n\}, \sigma(i) \leqslant i$$

Or pour une telle permutation $\sigma(1) \leqslant 1$ donc $\sigma(1) = 1$ puis $\sigma(2) \leqslant 2$ donc $\sigma(2) = 2$ car σ est injective, etc. Au final $\sigma = \mathrm{Id}$ et il ne reste qu'un terme dans la somme donnant le déterminant de A d'où la formule.

Proposition

$$\forall A \in \mathcal{M}_n(\mathbb{K}),\ \det\left({}^tA\right) = \det A$$

et donc

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

Théorème

Pour tout $A, B \in \mathcal{M}_n(\mathbb{K})$

$$\det(AB) = \det(A).\det(B)$$

De plus A est inversible si, et seulement si, $\det A \neq 0$ et alors $\det A^{-1} = 1/\det A$.

Attention : $\det(A + B) = ??$ et $\det(\lambda A) = \lambda^n \det A$.

Corollaire

$\mathrm{SL}_n(\mathbb{K}) = \{A \in \mathcal{M}_n(\mathbb{K}) / \det A = 1\}$ est un sous-groupe de $(\mathrm{GL}_n(\mathbb{K}), \times)$ appelé groupe spécial linéaire d'ordre n.

dém. :
$\mathrm{SL}_n(\mathbb{K})$ est le noyau du morphisme de groupes $\mathrm{GL}_n(\mathbb{K}) \to \mathbb{K}^\star$ qui envoie A sur $\det A$.
□

Corollaire

Deux matrices semblables ont même déterminant.

dém. :
Si $B = P^{-1}AP$ avec $P \in \mathrm{GL}_n(\mathbb{K})$ alors $\det B = \det P^{-1} \det A \det P = \det A$.
□

4.3.1.2 Déterminant d'un endomorphisme

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^\star$.

Définition

On appelle déterminant de $u \in \mathcal{L}(E)$ la valeur commune des déterminants des matrices représentant l'endomorphisme u .

Exemple $\det(\mathrm{Id}_E) = \det(I_n) = 1$.

Théorème

Pour tout $u, v \in \mathcal{L}(E)$,

$$\det(u \circ v) = \det u \det v$$

De plus, u est inversible si, et seulement si, $\det u \neq 0$ et alors $\det u^{-1} = 1/\det u$.

Corollaire

$\mathrm{SL}(E) = \{u \in \mathcal{L}(E) / \det u = 1\}$ est un sous groupe de $(\mathrm{GL}(E), \circ)$ appelé groupe spécial linéaire de E.

4.3.1.3 Déterminant d'une famille de vecteurs

Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^\star$ muni d'une base $e = (e_1, \dots, e_n)$.

Définition

On appelle déterminant dans la base e de la famille $(x_1, \dots, x_n)$ de vecteurs de E le scalaire

$$\det_e(x_1, \dots, x_n) \underset{\text{déf}}{=} \det \mathrm{Mat}_e(x_1, \dots, x_n)$$

Exemple $\det_e e = \det \mathrm{Mat}_e e = \det I_n = 1$.

Proposition

Si $e' = (e'_1, \dots, e'_n)$ est une autre base de E alors

$$\det_e(x_1, \dots, x_n) = \det_e e' \det_{e'}(x_1, \dots, x_n)$$

dém. :
Soit P la matrice de passage de e à e' et $A = \mathrm{Mat}_e(x_1, \dots, x_n)$, $A' = \mathrm{Mat}_{e'}(x_1, \dots, x_n)$.
Notons $X_1, \dots, X_n$ les colonnes de A et $X'_1, \dots, X'_n$ celles de A'.
Par formule de changement de bases : $X_j = PX'_j$ donc $A = PA'$.
En effet

$$PA' = P \begin{pmatrix} X'_1 & \cdots & X'_n \end{pmatrix} = \begin{pmatrix} PX'_1 & \cdots & PX'_n \end{pmatrix} = \begin{pmatrix} X_1 & \cdots & X_n \end{pmatrix} = A$$

Par suite $\det A = \det P \det A'$ puis la relation proposée.
□

Théorème

> L'application
> $$\begin{aligned} E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\mapsto \det_e(x_1, \ldots, x_n) \end{aligned}$$
> est une forme n-linéaire alternée (donc antisymétrique)
> De plus, la famille $(x_1, \ldots, x_n)$ est une base de E si, et seulement si, $\det_e(x_1, \ldots, x_n) \neq 0$.

Rappel :
Pour $\varphi : E^n \to F$ multilinéaire :
alternée signifie :
$$\exists i \neq j, x_i = x_j \Rightarrow \varphi(x_1, \ldots, x_n) = 0_F$$
antisymétrique signifie :
$$\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma)\varphi(x_1, \ldots, x_n)$$
pour tout $\sigma \in \mathcal{S}_n$.

Remarque Soit $A \in \mathcal{M}_n(\mathbb{K})$ de colonnes $C_1, \ldots, C_n \in \mathcal{M}_{n,1}(\mathbb{K})$.
On introduit $\mathcal{B} = (E_1, \ldots, E_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$.
La matrice des coordonnées dans $\mathcal{B}$ d'une colonne C_j est exactement C_j.
Il en découle
$$A = \text{Mat}_{\mathcal{B}}(C_1, \ldots, C_n)$$
puis
$$\det A = \det_{\mathcal{B}}(C_1, \ldots, C_n)$$
Ainsi, le déterminant d'une matrice est une forme n-linéaire alternée de ses colonnes.
Par transposition, on peut aussi dire que le déterminant d'une matrice est une forme n-linéaire alternée de ses lignes.

Exemple Pour $n \geqslant 3$, calcul de
$$D_n = \begin{vmatrix} 1 & & 0 & 1 \\ & \ddots & & 0 \\ & & \ddots & \\ (1) & & & 1 \end{vmatrix}_{[n]}$$
En décomposant la dernière colonne en somme de deux colonnes :
$$D_n = \begin{vmatrix} 1 & & 0 & 1 \\ & \ddots & & 0 \\ & & \ddots & \\ (1) & & & 1 \end{vmatrix}_{[n]} = \begin{vmatrix} 1 & & & (0) \\ & \ddots & & \\ & & \ddots & \\ (1) & & & 1 \end{vmatrix}_{[n]} + \begin{vmatrix} 1 & & 0 & 1 \\ & \ddots & & 0 \\ & & 1 & \\ (1) & & & 0 \end{vmatrix} = 1 + 0$$
car le dernier déterminant présente deux lignes identiques.

4.3.2 Opérations élémentaires sur les déterminants

Théorème

Les transvections $C_i \leftarrow C_i + \lambda C_j$ et $L_i \leftarrow L_i + \lambda L_j$ ne modifient pas le déterminant.
Les dilatations $C_i \leftarrow \alpha C_i$ et $L_i \leftarrow \alpha L_i$ multiplient par α le déterminant.
La permutation des lignes ou des colonnes d'une matrice selon une permutation σ multiplie son déterminant par $\varepsilon(\sigma)$.

dém. :
L'application $\det_{(E_1,\dots,E_n)}$ étant une forme linéaire alternée et antisymétrique

$$\det_{(E_1,\dots,E_n)}(C_1,\dots,C_i+\lambda C_j,\dots,C_n)=\det_{(E_1,\dots,E_n)}(C_1,\dots,C_i,\dots,C_n)+\lambda\det_{(E_1,\dots,E_n)}(C_1,\dots,C_j,\dots,C_n)$$

puis

$$\det_{(E_1,\dots,E_n)}(C_1,\dots,C_i+\lambda C_j,\dots,C_n)=\det_{(E_1,\dots,E_n)}(C_1,\dots,C_i,\dots,C_n)$$

car le déterminant multipliant λ possède la colonne C_j positionnée aux indices i et j.

$$\det_{(E_1,\dots,E_n)}(C_1,\dots,\alpha C_i,\dots,C_n)=\alpha\det_{(E_1,\dots,E_n)}(C_1,\dots,C_i,\dots,C_n)$$

et

$$\det_{(E_1,\dots,E_n)}(C_{\sigma(1)},\dots,C_{\sigma(n)})=\varepsilon(\sigma)\det_{(E_1,\dots,E_n)}(C_1,\dots,C_n)$$

On obtient les relations analogues sur les lignes.
□

Attention : L'opération $C_i \leftarrow C_j + \lambda C_i$ modifie le déterminant : c'est la combinaison de deux opérations élémentaires.

Attention : Les opérations élémentaires sont à réaliser successivement et non simultanément. Les opérations $C_1 \leftarrow C_1 + C_2$ et $C_2 \leftarrow C_1 + C_2$ transforment $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ en $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ et non en $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Exemple Calcul de

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix}$$

En retranchant à chaque ligne la précédente (en commençant par la dernière)

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix}=\begin{vmatrix} 1 & & 1 \\ & \ddots & \\ 0 & & 1 \end{vmatrix}=1$$

Exemple Soit $a, b \in \mathbb{K}$, $n \geqslant 2$. Calculons

$$D_n = \begin{vmatrix} a & & b \\ & \ddots & \\ b & & a \end{vmatrix}_{[n]}$$

En ajoutant toutes les colonnes à la première

$$D_n = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ a+(n-1)b & a & & (b) \\ \vdots & & \ddots & \\ a+(n-1)b & (b) & & a \end{vmatrix}$$

En retranchant la première ligne à chaque autre

$$D_n = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ 0 & a-b & & (0) \\ \vdots & & \ddots & \\ 0 & (0) & & a-b \end{vmatrix}$$

Finalement

$$D_n = (a+(n-1)b)(a-b)^{n-1}$$

Remarque On peut aussi raisonner par blocs comme dans l'exemple ci-dessous.

Exemple Pour $A, B \in \mathcal{M}_n(\mathbb{K})$, expression du déterminant de $\begin{pmatrix} A & B \\ B & A \end{pmatrix} \in \mathcal{M}_{2n}(\mathbb{K})$.
Via les opérations $C_1 \leftarrow C_1 + C_{n+1}, \ldots, C_n \leftarrow C_n + C_{2n}$,

$$\det \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det \begin{pmatrix} A+B & B \\ B+A & A \end{pmatrix}$$

Via les opérations $L_{n+1} \leftarrow L_{n+1} - L_1, \ldots, L_{2n} \leftarrow L_{2n} - L_{n+1}$,

$$\det \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det \begin{pmatrix} A+B & B \\ O & A-B \end{pmatrix} = \det(A+B)\det(A-B)$$

Si A et B commutent, on obtient

$$\det \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det\left(A^2 - B^2\right)$$

4.3.3 Développement d'un déterminant selon une rangée

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$.
Pour $i, j \in \{1, \ldots, n\}$, on appelle mineur d'indice (i, j) de A le scalaire

$$\Delta_{i,j} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \hat{a}_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{[n-1]}$$

et cofacteur d'indice (i, j) de A le scalaire

$$A_{i,j} = (-1)^{i+j} \times \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \hat{a}_{i,j} & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{[n-1]}$$

Théorème

Développement de $\det A$ selon sa i-ème ligne :

$$\det A = \sum_{j=1}^{n} a_{i,j} A_{i,j} = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \Delta_{i,j}$$

Développement de $\det A$ selon sa j-ème colonne :

$$\det A = \sum_{i=1}^{n} a_{i,j} A_{i,j} = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \Delta_{i,j}$$

Remarque Le signe de $(-1)^{i+j}$ est donné par le tableau

$$\begin{pmatrix} + & - & + & & (-1)^{n+1} \\ - & + & - & & \\ + & - & + & & \\ & & & \ddots & \\ (-1)^{n+1} & & & & + \end{pmatrix}$$

Exemple Pour $n \geqslant 2$, calcul de

$$D_n = \begin{vmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & (0) \\ 1 & (0) & 1 \end{vmatrix}_{[n]}$$

En développant selon la dernière ligne

$$D_n = (-1)^{n+1} \begin{vmatrix} 1 & \cdots & \cdots & 1 \\ 1 & & (0) & 0 \\ & \ddots & & \vdots \\ (0) & & 1 & 0 \end{vmatrix}_{[n-1]} + D_{n-1}$$

En permutant les colonnes selon le cycle $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 \end{pmatrix}$

$$D_n = (-1)^{n+1} \times (-1)^{n-2} \times \begin{vmatrix} 1 & \cdots & 1 \\ & \ddots & (0) \\ (0) & & 1 \end{vmatrix}_{[n-1]} + D_{n-1} = -1 + D_{n-1}$$

Puisque $D_2 = 2$, on obtient $D_n = 2 - n$.

4.3.4 Déterminant tridiagonal

Exemple Soit $a, b, c \in \mathbb{K}$. Calcul de

$$D_n = \begin{vmatrix} a & b & & (0) \\ c & \ddots & \ddots & \\ & \ddots & \ddots & b \\ (0) & & c & a \end{vmatrix}_{[n]}$$

En développant selon la première colonne,

$$D_n = aD_{n-1} - c \begin{vmatrix} b & 0 & \cdots & \cdots & 0 \\ c & a & b & & (0) \\ 0 & c & a & \ddots & \\ \vdots & & \ddots & \ddots & b \\ 0 & (0) & & c & a \end{vmatrix}_{[n-1]}$$

puis en développant le second déterminant selon la première ligne,

$$D_n = aD_{n-1} - bcD_{n-2}$$

Ainsi, (D_n) est une suite récurrente linéaire d'ordre 2.

Rappel :
On appelle suite récurrente linéaire d'ordre 2 toute suite $(u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ vérifiant

$$\forall n \in \mathbb{N}, u_{n+2} + pu_{n+1} + qu_n = 0$$

avec $(p, q) \in \mathbb{K} \times \mathbb{K}^\star$.
Pour exprimer son terme général, on introduit l'équation caractéristique associée

$$r^2 + pr + q = 0$$

de discriminant Δ.
Cas $\mathbb{K} = \mathbb{C}$.
Si $\Delta \neq 0$: 2 racines r_1, r_2 et $u_n = \lambda r_1^n + \mu r_2^n$ avec $\lambda, \mu \in \mathbb{C}$.
Si $\Delta = 0$: 1 racine double r et $u_n = (\lambda n + \mu) r^n$ avec $\lambda, \mu \in \mathbb{C}$.
Cas $\mathbb{K} = \mathbb{R}$.
Si $\Delta > 0$ ou $\Delta = 0$: semblable avec $\lambda, \mu \in \mathbb{R}$.
Si $\Delta < 0$: 2 racines conjuguées $r\mathrm{e}^{\pm i\theta}$ et $u_n = (\lambda\cos(n\theta) + \mu\sin(n\theta))\, r^n$ avec $\lambda, \mu \in \mathbb{R}$.
Dans chaque cas, λ, μ se déterminent à partir des deux rangs initiaux de la suite (u_n).

4.3.5 Déterminant de Vandermonde

Pour $a_1, \ldots, a_n \in \mathbb{K}$, on pose

$$V_n(a_1, ..., a_n) = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

Théorème

$$V_n(a_1, ..., a_n) = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i)$$

dém. :
Par récurrence sur $n \geqslant 1$.
Cas $n = 1$: ok
Supposons la propriété vraie au rang $n \geqslant 1$.
Soit $a_1, \ldots, a_n, a_{n+1} \in \mathbb{K}$
Cas : les $a_1, \ldots, a_n$ ne sont pas deux à deux distincts

$$V_{n+1}(a_1, \ldots, a_n, a_{n+1}) = 0 = \prod_{1 \leqslant i < j \leqslant n+1} (a_j - a_i)$$

Cas : les $a_1, \ldots, a_n$ sont deux à deux distincts.
Considérons la fonction

$$f : x \mapsto V_{n+1}(a_1, \ldots, a_n, x)$$

En développant selon la dernière ligne

$$f(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n \text{ avec } \alpha_n = V_n(a_1, \ldots, a_n)$$

Or $f(x) = 0$ pour $x \in \{a_1, \ldots, a_n\}$ car le déterminant comporte deux lignes égales.
On peut donc factoriser le polynôme

$$f(x) = \alpha_n \prod_{i=1}^{n} (x - a_i)$$

et ainsi on affirme

$$V_{n+1}(a_1, \ldots, a_n, a_{n+1}) = V_n(a_1, \ldots, a_n) \prod_{i=1}^{n} (a_{n+1} - a_i)$$

Récurrence établie.
□

4.3.6 Comatrice

Définition

On appelle comatrice de $A \in \mathcal{M}_n(\mathbb{K})$ la matrice des cofacteurs de A, on la note

$$\text{com}A \underset{\text{déf}}{=} (A_{i,j})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{K})$$

Théorème

$$\forall A \in \mathcal{M}_n(\mathbb{K}), {}^t(\text{com}A)\, A = A^t\, (\text{com}A) = \det(A)I_n$$

dém. :

$$\left[{}^t(\text{com}A)A\right]_{i,j} = \sum_{k=1}^{n} A'_{i,k}a_{k,j} = \sum_{k=1}^{n} a_{k,j}A_{k,i} = \det A.\delta_{i,j}$$

car se comprend comme le développement selon la i-ème colonne de la matrice obtenue en remplaçant dans A sa i-ème colonne par sa j-ème colonne.

□

Corollaire

Si $A \in \text{GL}_n(\mathbb{K})$ alors

$$A^{-1} = \frac{1}{\det A}{}^t\,(\text{com}A)$$

4.3.7 Musculation

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Etudions $\text{rg}(\text{com}A)$.
Si $\text{rg}A = n$ alors A est inversible donc ${}^t\text{com}A$ aussi puis

$$\text{rg}(\text{com}A) = n$$

Rappel : Le rang d'une matrice est l'ordre maximal des matrices carrées inversibles extraites de celle-ci
Si $\text{rg}A \leqslant n-2$ alors aucune matrice carrée d'ordre $n-1$ extraite de A n'est inversible. On en déduit que tous les mineurs de A sont nuls et donc $\text{com}A = O_n$ puis

$$\text{rg}(\text{com}A) = 0$$

Si $\text{rg}A = n-1$ alors $A^t\text{com}A = O_n$ donne

$$\text{Im}^t\text{com}A \subset \ker A$$

Or $\dim\ker A = 1$ donc $\text{rgcom}A \leqslant 1$.
Or $\text{com}A \neq O_n$ car A possède un mineur non nul puisque la matrice A possède une matrice extraite carrée d'ordre $n-1$ inversible. On conclut

$$\text{rg}(\text{com}A) = 1$$

Chapitre 5

Réduction géométrique

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$ et E un $\mathbb{K}$-espace vectoriel.

5.1 Sous-espaces stables

5.1.1 Définition

Définition

Un sous-espace vectoriel F de E est dit stable par $u \in \mathcal{L}(E)$ si $u(F) \subset F$ i.e.

$$\forall x \in F, u(x) \in F$$

Exemple $\{0_E\}$ et E sont stables par u.
F est stable par $\tilde{0}$, par Id_E et, plus généralement, par $\lambda \mathrm{Id}_E$ pour tout $\lambda \in \mathbb{K}$.

Exemple $E = \mathbb{K}[X]$, $D : P \mapsto P'$, $D \in \mathcal{L}(\mathbb{K}[X])$
$\mathbb{K}_n[X]$ est stable par D.
En effet,

$$\forall P \in \mathbb{K}[X], \deg P' \leqslant \deg P$$

Exemple $E = \mathbb{R}^{\mathbb{N}}$, $T : (u_n) \mapsto (u_{n+1})$, $T \in \mathcal{L}(\mathbb{R}^{\mathbb{N}})$.
Le sous-espace vectoriel $\mathcal{B}(\mathbb{N}, \mathbb{R})$ des suites réelles bornées est stable par T.

Proposition

Si F et G sont stables par u alors $F + G$ et $F \cap G$ aussi.

dém. :
$u(F + G) = u(F) + u(G) \subset F + G$.
$u(F \cap G) \subset u(F) \cap u(G) \subset F \cap G$.
□

Théorème

Si u et v commutent alors $\mathrm{Im} u$ et $\ker u$ sont stables par v.

dém. :
Pour tout $x \in \ker u$, $u(v(x)) = v(u(x)) = v(0_E) = 0_E$ donc $v(x) \in \ker u$.
Pour tout $y \in \text{Im}u$, on peut écrire $y = u(x)$ et alors $v(y) = v(u(x)) = u(v(x)) \in \text{Im}u$.
□

Exemple $\text{Im}u$ et $\ker u$ sont stables par u.
Pour $\lambda \in \mathbb{K}$, $\text{Im}(u - \lambda \text{Id}_E)$ et $\ker(u - \lambda \text{Id}_E)$ sont stables par u.

5.1.2 Endomorphisme induit

Définition

Si F est un sous-espace vectoriel stable par $u \in \mathcal{L}(E)$, on peut considérer l'application restreinte $u_F : F \to F$ qui définit un endomorphisme de F. On l'appelle endomorphisme induit par u sur F.

Exemple $\ker u$ est stable par u, on peut introduire $u_{\ker u}$ et l'on a $u_{\ker u} = \tilde{0}$.

Exemple $\text{Im}u$ est stable par u et on peut introduire $u_{\text{Im}u}$.
Cependant $u_{\text{Im}u}$ peut ne pas être surjectif.
En fait, $u_{\text{Im}u}$ est surjectif si, et seulement si, $\text{Im}u^2 = \text{Im}u$ car $\text{Im}u_{\text{Im}u} = \text{Im}u^2$

Exemple Soit $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ et $D : f \mapsto f'$.
$F = \text{Vect}(\cos, \sin)$ est stable par D car $D(\cos), D(\sin) \in F$ et

$$\text{Mat}_{(\cos,\sin)}(D_F) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = R_{-\pi/2}$$

Théorème

Si F est stable par u et $v \in \mathcal{L}(E)$ alors pour tout $\lambda \in \mathbb{K}$, F est stable par λu, $u + v$ et $u \circ v$.
De plus

$$(\lambda u)_F = \lambda u_F,\ (u + v)_F = u_F + v_F \text{ et } (u \circ v)_F = u_F \circ v_F$$

dém. :
$(\lambda u)(F) = \lambda u(F) \subset \lambda F \subset F$.
$(u + v)(F) \subset u(F) + v(F) \subset F + F \subset F$.
$(u \circ v)(F) = u(v(F)) \subset u(F) \subset F$.
Pour tout $x \in F$
$(\lambda u)_F(x) = (\lambda u)(x) = \lambda u(x) = \lambda u_F(x) = (\lambda u_F)(x)$.
$(u + v)_F(x) = (u + v)(x) = u(x) + v(x) = u_F(x) + v_F(x) = (u_F + v_F)(x)$.
$(u \circ v)_F(x) = (u \circ v)(x) = u(v(x)) = u(v_F(x)) = u_F(v_F(x)) = (u_F \circ v_F)(x)$.
□

Corollaire

L'ensemble des endomorphismes stabilisant F est une sous-algèbre de $\mathcal{L}(E)$ et l'application $u \mapsto u_F$ y définit un morphisme d'algèbres.

Proposition

Si F est stable par u alors

$$\ker u_F = \ker u \cap F \text{ et } \mathrm{Im} u_F \subset \mathrm{Im} u \cap F$$

dém. :
Soit $x \in \ker u_F$. On a $x \in F$ et $u(x) = u_F(x) = 0$ donc $x \in \ker u \cap F$.
Soit $x \in \ker u \cap F$. On a $u_F(x) = u(x) = 0$ donc $x \in \ker u_F$.
$\mathrm{Im} u_F \subset \mathrm{Im} u$ car u_F est restriction de u et $\mathrm{Im} u_F \subset F$ car F est stable par u.
□

Remarque Si u est injectif alors u_F est injectif.

Remarque Si u est surjectif, on ne peut rien dire a priori sur u_F.
Par exemple, la dérivation sur $\mathbb{K}[X]$ est surjective, mais l'endomorphisme induit sur $\mathbb{K}_n[X]$ ne l'est pas.

5.1.3 Visualisation en dimension finie

Ici E désigne un $\mathbb{K}$-espace vectoriel de dimension finie.

Théorème

Soit F un sous-espace vectoriel de dimension p muni d'une base $f = (e_1, \dots, e_p)$ complétée en une base $e = (e_1, \dots, e_n)$ de E. Pour $u \in \mathcal{L}(E)$, on a équivalence entre :
(i) F est stable par u ;
(ii) la matrice de u dans e est de la forme

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix} \text{ avec } A \in \mathcal{M}_p(\mathbb{K})$$

De plus, si tel est le cas, A est alors de la matrice de u_F dans la base f.

dém. :
(i) $\Rightarrow$ (ii) Supposons F stable par u. On peut introduire $A = \mathrm{Mat}_f(u_F) = (a_{i,j})$ et on a

$$\forall 1 \leqslant j \leqslant p, u(e_j) = \sum_{i=1}^{p} a_{i,j} e_j$$

et alors la matrice de u dans e est de la forme

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,p} & \\ \vdots & & \vdots & (\star) \\ a_{p,1} & \cdots & a_{p,p} & \\ & (0) & & (\star) \end{pmatrix} = \begin{pmatrix} A & B \\ O & C \end{pmatrix}$$

(ii) $\Rightarrow$ (i) Supposons la matrice de u dans e de la forme

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix}$$

avec $A \in \mathcal{M}_p(\mathbb{K})$. Pour tout $1 \leqslant j \leqslant p$, $u(e_j) \in \text{Vect}(e_1, \dots, e_p)$ donc $u(e_j) \in F$ puis, par linéarité, pour tout $x \in F$, $u(x) \in F$.
□

Théorème

On suppose $E = F_1 \oplus \dots \oplus F_m$ et on note e une base de E adaptée à cette décomposition.
Pour $u \in \mathcal{L}(E)$, on a équivalence entre :
(i) chaque F_k est stable par u ;
(ii) la matrice de u dans la base e est de la forme

$$\begin{pmatrix} A_1 & & O \\ & \ddots & \\ O & & A_m \end{pmatrix}$$

avec $A_k \in \mathcal{M}_{\alpha_k}(\mathbb{K})$ où $\alpha_k = \dim F_k$.

Remarque La réduction d'un endomorphisme u de E consiste à écrire

$$E = \mathop{\oplus}_{k=1}^{m} F_k$$

avec F_k stable par u et u_{F_k} « simple » .
En dimension finie, la réduction d'un endomorphisme correspond à l'obtention d'une représentation matricielle simple (la plus diagonale possible).

5.2 Eléments propres

E désigne un $\mathbb{K}$-espace vectoriel non réduit à $\{0_E\}$ de dimension quelconque et u un endomorphisme de E.

5.2.1 Valeur propre et vecteur propre

Proposition

Soit $x \in E \setminus \{0_E\}$ et $D = \text{Vect}(x)$ la droite vectorielle engendrée par x.
On a équivalence entre :
(i) D est stable pour $u \in \mathcal{L}(E)$;
(ii) il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$.

dém. :
(i) ⇒ (ii) Si D est stable par u alors $u(x) \in D$ et donc il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$.
(ii) ⇒ (i) Si $u(x) = \lambda x$ alors $u(D) = u(\text{Vect}x) = \text{Vect}u(x) \subset \text{Vect}x$.
□

Définition

On dit que $x \in E$ est vecteur propre de u si

$$x \neq 0_E \text{ et } \exists \lambda \in \mathbb{K}, u(x) = \lambda x$$

Attention : Par définition un vecteur propre est un vecteur non nul.

Remarque Il y a alors unicité de la valeur λ car

$$\lambda x = \mu x \text{ avec } x \neq 0_E \Rightarrow \lambda = \mu$$

On dit alors λ est la valeur propre associée au vecteur propre x.

Définition

On appelle valeur propre de u tout $\lambda \in \mathbb{K}$ vérifiant

$$\exists x \neq 0_E, u(x) = \lambda x$$

On appelle spectre de u l'ensemble des valeurs propres de u, on le note $\mathrm{Sp}u$.

Exemple On a

$$0 \in \mathrm{Sp}u \Leftrightarrow \exists x \neq 0_E, u(x) = 0_E$$

Ainsi

$$0 \in \mathrm{Sp}u \Leftrightarrow u \text{ non injectif}$$

5.2.2 Sous-espace propre

Définition

Pour $\lambda \in \mathbb{K}$ et $u \in \mathcal{L}(E)$, on note

$$E_\lambda(u) = \ker(u - \lambda \mathrm{Id}_E)$$

le sous-espace vectoriel formé des vecteurs $x \in E$ solutions de l'équation

$$u(x) = \lambda x$$

Exemple $E_0(u) = \ker u$.
$E_1(u) = \{x \in E / u(x) = x\}$. C'est l'espace des vecteurs invariants par u.

Théorème

On a équivalence entre :
(i) λ est valeur propre de u ;
(ii) $E_\lambda(u) \neq \{0_E\}$;
(iii) l'endomorphisme $u - \lambda\mathrm{Id}$ n'est pas injectif.

Définition

Si λ est valeur propre de u alors $E_\lambda(u)$ est appelé sous-espace propre associé à la valeur propre λ.

Remarque Si $\lambda \notin \mathrm{Sp}(u)$ alors $E_\lambda(u) = \{0_E\}$.
Si $\lambda \in \mathrm{Sp}(u)$ alors $E_\lambda(u) = \{0_E\} \cup \{\text{vecteur propre associé à la valeur propre } \lambda\}$.

5.2.3 Stabilité des sous-espaces propres

Théorème

Les sous-espaces propres de $u \in \mathcal{L}(E)$ sont stables par u et

$$\forall \lambda \in \text{Sp}u, u_{E_\lambda(u)} = \lambda \text{Id}$$

dém. :
u et $u - \lambda \text{Id}_E$ commutent donc $E_\lambda(u) = \ker(u - \lambda \text{Id}_E)$ est stable par u.
De plus, pour tout $x \in E_\lambda(u)$, $u(x) = \lambda x$ donc

$$u_{E_\lambda(u)} = \lambda \text{Id}$$

□

Corollaire

Si u et v commutent alors les sous-espaces propres de u sont stables pas v.

dém. :
En effet, $E_\lambda(u) = \ker(u - \lambda \text{Id})$ et $u - \lambda \text{Id}$ commute avec v.
□

5.2.4 Les sous-espaces propres sont en somme directe

Théorème

Des sous-espaces propres de $u \in \mathcal{L}(E)$ associés à des valeurs propres deux à deux distinctes sont en somme directe.

dém. :
Par récurrence sur $m \in \mathbb{N}^\star$, montrons que la somme de m sous-espace propres de u est directe.
Cas $m = 1$: il n'y a rien à démontrer.
Supposons la propriété établie au rang $m \geqslant 1$.
Soit $E_{\lambda_1}(u), \dots, E_{\lambda_m}(u), E_{\lambda_{m+1}}(u)$ des sous-espaces propres de u associés à des valeurs propres deux à deux distinctes.
Supposons $x_1 + \cdots + x_m + x_{m+1} = 0_E$ avec $x_k \in E_{\lambda_k}(u)$.
En appliquant u, on obtient $\lambda_1 x_1 + \cdots + \lambda_m x_m + \lambda_{m+1} x_{m+1} = 0_E$.
Par combinaison de ces deux équations, on obtient $(\lambda_1 - \lambda_{m+1})x_1 + \cdots + (\lambda_m - \lambda_{m+1})x_m = 0_E$.
Cette équation est de la forme $y_1 + \cdots + y_m = 0_E$ avec $y_k = (\lambda_k - \lambda_{m+1})x_k \in E_{\lambda_k}(u)$.
Par hypothèse de récurrence, les espaces $E_{\lambda_1}(u), \dots, E_{\lambda_m}(u)$ sont en somme directe donc

$$\forall 1 \leqslant k \leqslant m, y_k = 0_E$$

ce qui fournit

$$\forall 1 \leqslant k \leqslant m, x_k = 0_E$$

car

$$\lambda_k - \lambda_{m+1} \neq 0$$

Enfin, en reprenant l'équation initiale, on a aussi $x_{m+1} = 0_E$.
Récurrence établie.
□

Corollaire

> Une famille de vecteurs propres associés à des valeurs propres deux à deux distinctes est libre.

dém. :
Cas d'une famille finie :
Soit $x_1, \dots, x_m$ des vecteurs propres associés à des valeurs propres $\lambda_1, \dots, \lambda_m$ deux à deux distinctes.
Supposons $\alpha_1 x_1 + \cdots + \alpha_m x_m = 0_E$.
Puisque $\alpha_k x_k \in E_{\lambda_k}(u)$ et puisque les sous-espaces vectoriels $E_{\lambda_1}(u), \dots, E_{\lambda_m}(u)$ sont en somme directe, on a

$$\forall k \in \{1, \dots, m\}, \alpha_k x_k = 0_E$$

Or $x_k \neq 0_E$ (car c'est un vecteur propre) donc $\alpha_k = 0$.
Cas d'une famille infinie :
Celle-ci est libre car ses sous-familles finies le sont par l'argumentaire précédent.
□

Corollaire

> En dimension finie égale à n, un endomorphisme ne peut admettre plus de n valeurs propres.

dém. :
Si $\lambda_1, \dots, \lambda_m$ sont des valeurs propres de $u \in \mathcal{L}(E)$ avec $\dim E = n$ alors

$$\underset{k=1}{\overset{m}{\oplus}} E_{\lambda_k}(u) \subset E \text{ avec } \dim E_{\lambda_k}(u) \geqslant 1$$

donne $m \leqslant \dim E$.
□

Remarque En dimension infinie, il peut y avoir une infinité de valeurs propres.

5.2.5 Détermination pratique

Protocole :
Pour déterminer les valeurs propres de u, on étudie pour quels scalaires $\lambda \in \mathbb{K}$, l'équation

$$u(x) = \lambda x$$

possède d'autres solutions que la solution nulle.
Cette équation est appelée l'équation aux éléments propres associée à u.

Exemple Soit $E = \mathbb{K}[X]$ et $\varphi \in \mathcal{L}(E)$ défini par $\varphi(P) = XP'(X)$. Déterminons Spφ.
Soit $\lambda \in \mathbb{K}$ et $P \in \mathbb{K}[X]$.

$$\varphi(P) = \lambda P \Leftrightarrow XP'(X) = \lambda P(X)$$

Analyse :
Si cette équation possède une solution $P \neq 0$ alors en posant $n = \deg P$, on peut écrire
$P = a_n X^n + \cdots + a_1 X + a_0$ avec $a_n \neq 0$. L'équation $XP'(X) = \lambda P(X)$ donne

$$\forall 0 \leqslant k \leqslant n, \lambda a_k = n a_k$$

Sachant $a_n \neq 0$, on obtient $\lambda = n$ et $a_{n-1} = \ldots = a_1 = a_0 = 0$.
Ainsi

$$\lambda \in \mathbb{N} \text{ et } P = a_\lambda X^\lambda$$

Synthèse :

Pour $\lambda \in \mathbb{N}$ et $P = a_\lambda X^\lambda$ avec $a_\lambda \neq 0$, on vérifie $XP'(X) = \lambda P(X)$ avec $P \neq 0$ donc $\lambda \in \text{Sp}\varphi$.
Finalement $\text{Sp}\varphi = \mathbb{N}$ et

$$\forall \lambda \in \mathbb{N}, E_\lambda(\varphi) = \text{Vect}(X^\lambda)$$

Exemple Soit $E = \mathbb{K}[X]$ et $\psi \in \mathcal{L}(E)$ défini par $\psi(P) = XP(X)$. Déterminons $\text{Sp}\psi$.
Soit $\lambda \in \mathbb{K}$ et $P \in \mathbb{K}[X]$.

$$\psi(P) = \lambda P(X) \Leftrightarrow XP(X) = \lambda P(X) \Leftrightarrow (X - \lambda)P(X) = 0 \Leftrightarrow P(X) = 0$$

donc $\text{Sp}\psi = \emptyset$.

Exemple Soit $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ et $D : f \mapsto f'$. Déterminons $\text{Sp}D$.
Soit $\lambda \in \mathbb{C}$ et $f \in E$.

$$D(f) = \lambda f \Leftrightarrow f' = \lambda f \Leftrightarrow f \in \text{Vect}(e_\lambda)$$

avec $e_\lambda : t \mapsto \mathrm{e}^{\lambda t}$ fonction non nulle.
On en déduit $\text{Sp}D = \mathbb{C}$ et

$$\forall \lambda \in \mathbb{C}, E_\lambda(D) = \text{Vect}(e_\lambda)$$

Exemple Soit $E = \mathcal{B}(\mathbb{N}, \mathbb{R})$ et $T : (u_n)_{n\in\mathbb{N}} \mapsto (u_{n+1})_{n\in\mathbb{N}}$. Déterminons $\text{Sp}T$.
Soit $\lambda \in \mathbb{R}$ et $u = (u_n) \in E$.

$$T(u) = \lambda u \Leftrightarrow \forall n \in \mathbb{N}, u_{n+1} = \lambda u_n \Leftrightarrow \forall n \in \mathbb{N}, u_n = \lambda^n u_0$$

Si $|\lambda| > 1$ alors la suite $(\lambda^n u_0)$ est bornée si, et seulement si, $u_0 = 0$ et c'est alors la suite nulle.
Si $|\lambda| \leqslant 1$ alors la suite $(\lambda^n u_0)$ est bornée et non nulle pour tout $u_0 \neq 0$.
Finalement $\text{Sp}T = [-1, 1]$ et

$$\forall \lambda \in [-1, 1], E_\lambda(T) = \text{Vect}\left((\lambda^n)_{n\in\mathbb{N}}\right)$$

5.3 Eléments propres en dimension finie

E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^\star$ et u un endomorphisme de E.

5.3.1 Eléments propres d'une matrice carrée

Définition

On dit que $\lambda \in \mathbb{K}$ est valeur propre de $A \in \mathcal{M}_n(\mathbb{K})$ s'il existe $X \in \mathcal{M}_{n,1}(\mathbb{K})$ vérifiant

$$AX = \lambda X \text{ et } X \neq 0$$

On dit alors que la colonne X est vecteur propre associé à la valeur propre λ.
On appelle spectre de la matrice A l'ensemble $\text{Sp}A$ formé des valeurs propres de A.

Définition

Pour $\lambda \in \mathbb{K}$, on note $E_\lambda(A) = \ker(A - \lambda I_n)$ l'espace des solutions de l'équation $AX = \lambda X$.
Lorsque λ est valeur propre de A, $E_\lambda(A)$ est appelé sous-espace propre de A associé à la valeur propre λ.

Remarque En identifiant tuple et colonne, les éléments propres de A correspondent aux éléments propres de l'endomorphisme canoniquement associé à A défini par

$$x \in \mathbb{K}^n \mapsto y = Ax \in \mathbb{K}^n$$

Remarque Pour déterminer, les valeurs propres de A, on étudie l'équation aux éléments propres $AX = \lambda X$.

Théorème

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie non nulle et e une base de E.
Pour $u \in \mathcal{L}(E)$ et $x \in E$, en notant $A = \mathrm{Mat}_e(u)$ et $X = \mathrm{Mat}_e(x)$, on a

$$\mathrm{Sp}A = \mathrm{Sp}u \text{ et } \forall \lambda \in \mathrm{Sp}u, x \in E_\lambda(u) \Leftrightarrow X \in E_\lambda(A)$$

dém. :
On a

$$u(x) = \lambda x \Leftrightarrow AX = \lambda X \text{ et } x \neq 0_E \Leftrightarrow X \neq 0$$

□

Corollaire

Deux matrices semblables ont le même spectre.

dém. :
Car elles représentent le même endomorphisme.
□

5.3.2 Polynôme caractéristique d'une matrice carrée

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Pour tout $\lambda \in \mathbb{K}$, l'expression

$$\det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n,1} & \cdots & -a_{n,n-1} & \lambda - a_{n,n} \end{vmatrix}$$

est un polynôme en λ.

Définition

On appelle polynôme caractéristique de A, le polynôme $\chi_A \in \mathbb{K}[X]$ déterminé par la propriété

$$\forall \lambda \in \mathbb{K}, \chi_A(\lambda) = \det(\lambda I_n - A)$$

Exemple Polynôme caractéristique de $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$$\det(\lambda I_2 - A) = \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix} = (\lambda - a)(\lambda - d) - bc$$

et donc
$$\chi_A(X) = X^2 - (a+d)X + (ad - bc)$$

Exemple Polynôme caractéristique de
$$A = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
Comme déterminant diagonal, on obtient
$$\chi_A = \prod_{i=1}^{n} (X - \lambda_i)$$

Théorème

Le polynôme caractéristique de $A \in \mathcal{M}_n(\mathbb{K})$ est unitaire, de degré n et possède les coefficients remarquables suivants
$$\chi_A(X) = X^n - \text{tr}(A)X^{n-1} + \cdots + (-1)^n \det(A)$$

dém. :
Par la formule des déterminants
$$\chi_A(\lambda) = \det(\lambda I_n - A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} \left(\lambda \delta_{\sigma(i),i} - a_{\sigma(i),i}\right)$$
Pour tout $\sigma \in \mathcal{S}_n$, posons
$$P_\sigma(\lambda) = \prod_{i=1}^{n} \left(\lambda \delta_{\sigma(i),i} - a_{\sigma(i),i}\right)$$
P_σ est une fonction polynôme de degré $\leqslant n$.
Si $\sigma \neq \text{Id}_{\mathbb{N}_n}$, il existe au moins deux indices i, j tels que $\sigma(i) \neq i$ et $\sigma(j) \neq j$, la fonction polynôme P_σ est alors de degré $\leqslant n - 2$.
Si $\sigma = \text{Id}_{\mathbb{N}_n}$
$$P_{\text{Id}}(\lambda) = \prod_{i=1}^{n} (\lambda - a_{i,i}) = \lambda^n - (a_{1,1} + \cdots + a_{n,n})\lambda^{n-1} + \cdots$$
Ainsi
$$\det(\lambda I_n - A) = \lambda^n - \text{tr}(A)\lambda^{n-1} + \cdots$$
Enfin, le coefficient constant de χ_A est $\chi_A(0) = (-1)^n \det(A)$.
□

Exemple Soit $P = X^n - a_{n-1}X^{n-1} - \cdots - a_1 X - a_0$ et
$$A = \begin{pmatrix} 0 & & (0) & a_0 \\ 1 & 0 & & \vdots \\ & \ddots & \ddots & a_{n-2} \\ (0) & & 1 & a_{n-1} \end{pmatrix}$$

Calculons le polynôme caractéristique de A.

$$\det(\lambda I_n - A) = \begin{vmatrix} \lambda & & (0) & -a_0 \\ -1 & \lambda & & \vdots \\ & \ddots & \ddots & -a_{n-2} \\ (0) & & -1 & \lambda - a_{n-1} \end{vmatrix}$$

En développant selon la dernière colonne

$$\chi_A(\lambda) = P(\lambda)$$

5.3.3 Polynôme caractéristique et valeurs propres

Théorème

Les valeurs propres de A sont exactement les racines de χ_A.

dém. :

$$\lambda \in \text{Sp}A \Leftrightarrow \ker(A - \lambda I_n) \neq \{0\} \Leftrightarrow A - \lambda I_n \text{ non inversible} \Leftrightarrow \det(A - \lambda I_n) = 0$$

Or

$$\det(A - \lambda I_n) = (-1)^n \det(\lambda I_n - A) = (-1)^n \chi_A(\lambda)$$

donc

$$\lambda \in \text{Sp}A \Leftrightarrow \chi_A(\lambda) = 0$$

□

Exemple Si

$$A = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

alors

$$\text{Sp}A = \{\lambda_1, \ldots, \lambda_n\}$$

Corollaire

$A \in \mathcal{M}_n(\mathbb{K})$ possède au plus n valeurs propres.

dém. :
Car un polynôme de degré n admet au plus n racines.
□

Corollaire

$A \in \mathcal{M}_n(\mathbb{C})$ possède au moins une valeur propre complexe.

dém. :
$\chi_A \in \mathbb{C}[X]$ est un polynôme non constant, il possède donc au moins une racine dans $\mathbb{C}$.
□

Remarque Aussi $A \in \mathcal{M}_{2n+1}(\mathbb{R})$ possède au moins une valeur propre réelle.

Exemple Etude des éléments propres de

$$A = \begin{pmatrix} -2 & 1 & -1 \\ 1 & -2 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

On a

$$\chi_A(\lambda) = (-1)^3 \det(A - \lambda I_3) = -\begin{vmatrix} -2-\lambda & 1 & -1 \\ 1 & -2-\lambda & 1 \\ 1 & -1 & -\lambda \end{vmatrix} = -\begin{vmatrix} -2-\lambda & 1 & -1 \\ -1-\lambda & -1-\lambda & 0 \\ -1-\lambda & 0 & -1-\lambda \end{vmatrix}$$

$$\det(\lambda I_3 - A) = -(\lambda+1)^2 \begin{vmatrix} -(2+\lambda) & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = (\lambda+1)^2(\lambda+2)$$

Ainsi

$$\chi_A(X) = (X+1)^2(X+2)$$

Ainsi $\text{Sp}(A) = \{-1, -2\}$
Etudions $E_{-2}(A)$

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-2}(A) \Leftrightarrow AX = -2X \Leftrightarrow (A + 2I_3)X = 0 \Leftrightarrow \begin{cases} y - z = 0 \\ x + z = 0 \\ x - y + 2z = 0 \end{cases}$$

donc $E_{-2}(A) = \text{Vect}(1, -1, -1)$
Etudions $E_{-1}(A)$

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_{-1}(A) \Leftrightarrow (A + I_3)X = 0 \Leftrightarrow \begin{cases} -x + y - z = 0 \\ x - y + z = 0 \\ x - y + z = 0 \end{cases}$$

donc $E_{-1} = \text{Vect}\{(1,1,0), (0,1,1)\}$

Exemple Etude des éléments propres de

$$A = \begin{pmatrix} 0 & & (1) \\ & \ddots & \\ (1) & & 0 \end{pmatrix}$$

Via $C_1 \leftarrow C_1 + \cdots + C_n$

$$\chi_A(\lambda) = \det(\lambda I_n - A) = \begin{vmatrix} \lambda & & (-1) \\ & \ddots & \\ (-1) & & \lambda \end{vmatrix} = (\begin{vmatrix} \lambda - (n-1) & -1 & \cdots & -1 \\ \vdots & \lambda & & (-1) \\ \vdots & & \ddots & \\ \lambda - (n-1) & (-1) & & \lambda \end{vmatrix}$$

puis via $L_2 \leftarrow L_2 - L_1, \dots, L_n \leftarrow L_n - L_1$

$$\chi_A(\lambda) = \begin{vmatrix} \lambda - (n-1) & -1 & \cdots & -1 \\ 0 & \lambda + 1 & & (0) \\ \vdots & & \ddots & \\ 0 & (0) & & \lambda + 1 \end{vmatrix}$$

Ainsi $\chi_A(\lambda) = (\lambda - (n-1))(\lambda + 1)^{n-1}$ et donc

$$\chi_A(X) = (X - (n-1))\,(X+1)^{n-1}$$

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E_{-1}(A) \Leftrightarrow (A + I_n)X = 0$$

$$\Leftrightarrow \begin{cases} x_1 + \cdots + x_n = 0 \\ \qquad\vdots \\ x_1 + \cdots + x_n = 0 \end{cases} \Leftrightarrow x_1 + \cdots + x_n = 0$$

Ainsi $E_{-1}(A)$ est l'hyperplan d'équation $x_1 + \cdots + x_n = 0$.

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E_{n-1}(A) \Leftrightarrow (A + I_n)X = nX$$

$$\Leftrightarrow \begin{cases} x_1 + \cdots + x_n = nx_1 \\ \qquad\vdots \\ x_1 + \cdots + x_n = nx_n \end{cases} \Leftrightarrow x_1 = \ldots = x_n$$

Ainsi $E_{n-1}(A) = \text{Vect}\,\{(1, \dots, 1)\}$.

5.3.4 Polynôme caractéristique d'un endomorphisme

Proposition

Si $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblables alors $\chi_A = \chi_B$.

dém. :
Si $B = P^{-1}AP$ avec $P \in \text{GL}_n(\mathbb{K})$ alors $\chi_B(\lambda) = \det(\lambda I_n - P^{-1}AP) = \det\left(P^{-1}(\lambda I_n - A)P\right) = \det(\lambda I_n - A) = \chi_A(\lambda)$.
□

Définition

On appelle polynôme caractéristique de $u \in \mathcal{L}(E)$, le polynôme caractéristique commun aux matrices représentant l'endomorphisme u ; on le note χ_u.

Exemple $\chi_{\lambda \text{Id}_E} = \chi_{\lambda I_n} = (X - \lambda)^n$ avec $n = \dim E$.

Exemple Supposons $E = F \oplus G$. Déterminons le polynôme caractéristique de la projection sur F parallèlement à G.
Dans une base adaptée à la décomposition $E = F \oplus G$, la matrice de p est de la forme

$$\begin{pmatrix} I_r & O \\ O & O_{n-r} \end{pmatrix}$$

avec $r = \dim F$. On a alors

$$\chi_p(X) = (X-1)^r X^{n-r}$$

Théorème

Pour $u \in \mathcal{L}(E)$, χ_u est un polynôme unitaire de degré exactement $n = \dim E$ de la forme

$$\chi_u(\lambda) = X^n - \text{tr}(u)X^{n-1} + \cdots + (-1)^n \det(u)$$

De plus, les valeurs propres de u sont exactement les racines de χ_u.

dém. :
Si $A \in \mathcal{M}_n(\mathbb{K})$ est la matrice de u dans une base de E, $\chi_u = \chi_A$ avec tr$A$ = tr$u$ et $\det A = \det u$.
De plus, Sp(u) = Sp(A) et donc les racines de χ_u correspondent aux valeurs propres de u.
□

Corollaire

Un endomorphisme $u \in \mathcal{L}(E)$ possède au plus $\dim E$ valeurs propres.

Corollaire

Si E est un $\mathbb{C}$-espace vectoriel de dimension finie alors tout $u \in \mathcal{L}(E)$ possède au moins une valeur propre.

Remarque Si E est un $\mathbb{R}$-espace vectoriel de dimension impaire alors tout $u \in \mathcal{L}(E)$ possède au moins une valeur propre.

5.3.5 Multiplicité d'une valeur propre

Rappel :
Si $P \in \mathbb{K}[X]$ est un polynôme non nul, on appelle ordre de multiplicité de λ en tant que racine de P le plus grand $\alpha \in \mathbb{N}$ tel que

$$(X-\lambda)^\alpha \mid P$$

Ceci équivaut encore à

$$P(\lambda) = P'(\lambda) = \ldots = P^{(\alpha-1)}(\lambda) = 0 \text{ et } P^{(\alpha)}(\lambda) \neq 0$$

Rappel :
Un polynôme P non constant est dit scindé dans $\mathbb{K}[X]$ si, et seulement si, on peut le factoriser sous la forme

$$P = \mu \prod_{i=1}^{n} (X - \lambda_i)$$

Les scalaires $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ correspondent alors à ses racines comptées avec multiplicité.
En regroupant les racines égales, on obtient l'écriture

$$P = \mu \prod_{k=1}^{m} (X - \lambda_k)^{\alpha_k}$$

avec $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ deux à deux distincts et $\alpha_1, \ldots, \alpha_m$ leurs multiplicités respectives.

Définition

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. On appelle multiplicité de λ en tant que valeur propre de $u \in \mathcal{L}(E)$, l'ordre de multiplicité de λ en tant que racine de χ_u ; on la note $m_\lambda(u)$ (idem en $A \in \mathcal{M}_n(\mathbb{K})$ pour $m_\lambda(A)$)

Remarque Abusivement, λ valeur propre de multiplicité 0 signifie que λ n'est pas valeur propre.

Exemple Valeurs de propres de

$$A = \begin{pmatrix} \lambda & & \star \\ & \lambda & \\ 0 & & \mu \end{pmatrix} \text{ avec } \lambda \neq \mu$$

On a $\chi_A = (X - \lambda)^2(X - \mu)$
λ est valeur propre double et μ est valeur propre simple de A.

Exemple Valeurs propres de

$$A = \begin{pmatrix} \lambda_1 & & (\star) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

On a $\chi_A = \prod_{i=1}^{n} (X - \lambda_i)$.
Les valeurs propres de A sont les $\lambda_1, \ldots, \lambda_n$ comptées avec multiplicité.

Théorème

$$\forall u \in \mathcal{L}(E), \sum_{\lambda \in \mathrm{Sp} u} m_\lambda(u) \leqslant \dim E$$

avec égalité si, et seulement si, le polynôme χ_u est scindé dans $\mathbb{K}[X]$ (idem pour $A \in \mathcal{M}_n(\mathbb{K})$).

dém. :
La somme des multiplicités des racines d'un polynôme non nul est inférieure à son degré avec égalité si, et seulement si, ce polynôme est scindé.
□

Corollaire

Si $\mathbb{K} = \mathbb{C}$ alors $u \in \mathcal{L}(E)$ possède exactement n valeurs propres comptées avec multiplicité (idem en $A \in \mathcal{M}_n(\mathbb{C})$).

dém. :
Dans $\mathbb{C}[X]$, tout polynôme non constant est scindé.
□

5.3.6 Multiplicité et dimension des sous-espaces propres

Théorème

Si F est un sous-espace vectoriel stable par u alors le polynôme caractéristique de l'endomorphisme induit par u sur F divise le polynôme caractéristique de u.

dém. :
Dans une base adaptée à F, la matrice de u est de la forme

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix}$$

avec A matrice de u_F. On a alors $\chi_u = \chi_A \chi_C$ avec $\chi_A = \chi_{u_F}$.
□

Théorème

$$\forall \lambda \in \text{Sp}(u), 1 \leqslant \dim E_\lambda(u) \leqslant m_\lambda(u)$$

(idem avec $A \in \mathcal{M}_n(\mathbb{K})$).

dém. :
Soit $\lambda \in \text{Sp}u$.
D'une part, $F = E_\lambda(u) = \ker(u - \lambda\text{Id}) \neq \{0_E\}$ donc $\dim F \geqslant 1$.
D'autre part, F est stable par u donc $\chi_{u_F} \mid \chi_u$.
Or $\chi_{u_F} = (X - \lambda)^{\dim F}$ car $u_F = \text{Id}_F$ donc λ est racine de multiplicité au moins $\dim F$ de χ_u.
□

Corollaire

Si λ est une valeur propre simple alors le sous-espace propre associé est de dimension 1.

5.3.7 Changement de corps

Supposons $\mathbb{L}$ un sous-corps de $\mathbb{K}$.
Pour $A \in \mathcal{M}_n(\mathbb{L})$, on peut aussi comprendre $A \in \mathcal{M}_n(\mathbb{K})$.
On peut donc parler de valeurs propres de A dans $\mathbb{L}$, mais aussi dans $\mathbb{K}$. Bien évidemment

$$\text{Sp}_{\mathbb{L}}(A) \subset \text{Sp}_{\mathbb{K}}(A)$$

En particulier, on peut parler des valeurs propres complexes d'une matrice réelle.

Exemple Considérons

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$$

On a $\chi_A = X^2 + 1$ donc $\text{Sp}_{\mathbb{R}} A = \emptyset$ et $\text{Sp}_{\mathbb{C}} = \{i, -i\}$.

Théorème

Les valeurs propres complexes d'une matrice réelle sont deux à deux conjuguées.
De plus, deux racines complexes conjuguées ont même multiplicité et les sous-espaces propres associés se correspondent par conjugaison.

dém. :
Soit $A \in \mathcal{M}_n(\mathbb{R})$. Le polynôme caractéristique de A est réel. Ses racines complexes sont donc deux à deux conjuguées et deux racines conjuguées ont même multiplicité. Aussi

$$AX = \lambda X \Leftrightarrow A\bar{X} = \bar{\lambda}\bar{X}$$

L'application $X \mapsto \bar{X}$ définit alors une bijection de $E_\lambda(A)$ vers $E_{\bar{\lambda}}(A)$.
□

Remarque Par conjugaison, une base de $E_\lambda(A)$ est transformée en une base de $E_{\bar{\lambda}}(A)$: ces deux sous-espaces propres sont d'égales dimensions.

5.4 Diagonalisabilité

E désigne un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^\star$

5.4.1 Endomorphisme diagonalisable

Définition

Un endomorphisme $u \in \mathcal{L}(E)$ est dit diagonalisable s'il existe une base de E dans laquelle sa matrice est diagonale. Une telle base est appelée base de diagonalisation de u.

Exemple Id_E est diagonalisable et n'importe quelle base de E est base de diagonalisation.

Exemple Les projections vectorielles sont diagonalisables.
En effet, si $E = F \oplus G$ alors la projection p sur F parallèlement à G a pour matrice

$$\begin{pmatrix} I_r & O \\ O & O_{n-r} \end{pmatrix} \text{ avec } r = \dim F$$

dans une base adaptée à la décomposition $E = F \oplus G$.
Aussi, les symétries vectorielles sont diagonalisables.

Théorème

Pour $u \in \mathcal{L}(E)$, on a équivalence entre :
(i) u est diagonalisable ;
(ii) il existe une base de E formée de vecteurs propres de u.
Une base de diagonalisation est aussi appelée une base propre.

dém. :
(i) $\Rightarrow$ (ii) Supposons u diagonalisable et considérons $e = (e_1, \ldots, e_n)$ une base de diagonalisation de u.
La matrice de u dans e est de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Pour tout $i \in \{1, \ldots, n\}$, on a $u(e_i) = \lambda_i e_i$ avec $e_i \neq 0_E$ donc e_i vecteur propre de u.
La famille e est donc une base de vecteurs propres de u.

(ii) $\Rightarrow$ (i) Supposons l'existence d'une base $e = (e_1, \dots, e_n)$ de vecteurs propres de u.
Pour tout $i \in \{1, \dots, n\}$, on a $u(e_i) = \lambda_i e_i$ avec λ_i la valeur propre associée au vecteur propre e_i.
La matrice de u dans la base e est alors de la forme

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

□

Exemple Un endomorphisme diagonalisable possède au moins une valeur propre.

Exemple Si u est diagonalisable et si u ne possède qu'une valeur propre λ alors $u = \lambda \mathrm{Id}_E$.
En effet, la matrice de u dans une base propre est λI_n et donc $u = \lambda \mathrm{Id}_E$.

5.4.2 Une condition suffisante de diagonalisabilité

Théorème

Si $u \in \mathcal{L}(E)$ possède $n = \dim E$ valeurs propres distinctes alors u est diagonalisable et ses sous-espaces propres sont tous des droites vectorielles.

dém. :
Soit $\lambda_1, \dots, \lambda_n$ les valeurs propres deux à deux distinctes de u.
Soit $e_1, \dots, e_n$ des vecteurs propres associés.
La famille $e = (e_1, \dots, e_n)$ est libre car formée de vecteurs propres associés à des valeurs propres deux à deux distinctes. Etant formée de $n = \dim E$ vecteurs de E, c'est une base de E diagonalisant u.
On a alors

$$\mathrm{Mat}_e u = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = D$$

et donc

$$\chi_u = \chi_D = (-1)^n \prod_{i=1}^{n} (X - \lambda_i)$$

Puisque les $\lambda_1, \dots, \lambda_n$ sont deux à deux distincts, les valeurs propres de u sont toutes simples et les sous-espaces propres sont donc de dimension 1.
□

Exemple Considérons l'application $\varphi : \mathbb{K}_n[X] \to \mathbb{K}_n[X]$ définie par $\varphi(P) = nXP - (X^2 - 1)P'$.
Etudions la diagonalisabilité de φ.
L'application φ est bien définie car si $P = aX^n + \cdots$, $nXP = aX^{n+1} + \cdots$,
$n(X^2 - 1)P' = naX^{n+1} + \cdots$ et donc $\varphi(P) = 0.X^{n+1} + \cdots \in \mathbb{K}_n[X]$.
Puisque $\varphi(X^k) = (n-k)X^{k+1} + kX^{k-1}$, la matrice de φ dans $(1, X, \dots, X^n)$ est

$$\begin{pmatrix} 0 & 1 & & \\ n & \ddots & \ddots & \\ & \ddots & \ddots & n \\ & & 1 & 0 \end{pmatrix}$$

Le calcul du polynôme caractéristique n'est alors pas simple.
Considérons alors la base de Taylor $\mathcal{B} = (1, (X-1), \dots, (X-1)^n)$.
Puisque $\varphi\left((X-1)^k\right) = (n-k)(X-1)^{k+1} + (n-2k)(X-1)^k$, la matrice de φ dans $\mathcal{B}$ est

$$\begin{pmatrix} n & & & (0) \\ n & n-2 & & \\ & \ddots & \ddots & \\ (0) & & 1 & -n \end{pmatrix}$$

On en déduit $\chi_\varphi = \prod_{k=0}^{n} (n-2k-X) = (-1)^n \prod_{k=0}^{n} (X-(n-2k))$ et $\mathrm{Sp}\varphi = \{n-2k / k \in [\![0,n]\!]\}$.
Puisque $\mathrm{CardSp}\varphi = n+1 = \dim \mathbb{K}_n[X]$, l'endomorphisme φ est diagonalisable et sous-espaces propres sont des droites vectorielles.

5.4.3 Diagonalisabilité et sous-espaces propres

Théorème

Soit $u \in \mathcal{L}(E)$. On a équivalence entre :
(i) u est diagonalisable ;
(ii) E est la somme directe des sous-espaces propres de u i.e. :

$$E = \underset{\lambda \in \mathrm{Sp}(u)}{\oplus} E_\lambda(u)$$

(iii) $\sum_{\lambda \in \mathrm{Sp}(u)} \dim E_\lambda(u) = \dim E$.

dém. :
Rappelons que l'on sait déjà que les sous-espaces propres d'un endomorphisme sont en somme directe.
(i) $\Rightarrow$ (ii) Supposons u diagonalisable.
Soit $e = (e_1, \dots, e_n)$ une base propre de u. Pour tout $i \in \{1, \dots, n\}$, e_i est vecteur propre de u donc

$$e_i \in \underset{\lambda \in \mathrm{Sp}(u)}{\oplus} E_\lambda(u)$$

puis $E \subset \underset{\lambda \in \mathrm{Sp}(u)}{\oplus} E_\lambda(u)$ et enfin $E = \underset{\lambda \in \mathrm{Sp}(u)}{\oplus} E_\lambda(u)$.
(ii) $\Rightarrow$ (iii) Car l'on sait $\dim \overset{m}{\underset{i=1}{\oplus}} F_i = \sum_{i=1}^{m} \dim F_i$.
(iii) $\Rightarrow$ (i) Une famille formée par concaténation de bases des espaces propres $E_\lambda(u)$ est une famille libre formée de $\dim E$ vecteurs, c'est donc une base de vecteurs propres.
□

Corollaire

Soit $u \in \mathcal{L}(E)$. On a équivalence entre :
(i) u est diagonalisable ;
(ii) χ_u est scindé dans $\mathbb{K}[X]$ et, pour tout $\lambda \in \mathrm{Sp}(u)$, $\dim E_\lambda(u) = m_\lambda(u)$.

dém. :
(i) $\Rightarrow$ (ii) Supposons u diagonalisable.

Notons $\lambda_1, \ldots, \lambda_m$ les valeurs propres de u.
Dans une base adaptée à l'écriture $E = \overset{m}{\underset{j=1}{\oplus}} E_{\lambda_j}(u)$ la matrice de u est

$$\begin{pmatrix} \lambda_1 I_{\alpha_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_m I_{\alpha_m} \end{pmatrix}$$

avec $\alpha_k = \dim E_{\lambda_k}(u)$. On a alors

$$\chi_u = \prod_{k=1}^{m} (X - \lambda_k)^{\alpha_k}$$

χ_u est scindé et pour tout $k \in \{1, \ldots, m\}$, λ_k est racine de χ_u de multiplicité $n_k = \dim E_{\lambda_k}(u)$.
(ii) $\Rightarrow$ (i) Supposons (ii)
Puisque χ_u est scindé, la somme des multiplicités de ses racines égale son degré.
Ainsi $\deg \chi_u = \sum_{\lambda \in \mathrm{Sp}(u)} m_\lambda(u)$ et donc $\sum_{\lambda \in \mathrm{Sp}(u)} \dim E_\lambda(u) = \dim E$ ce qui entraîne la diagonalisabilité de u.
□

5.4.4 Matrice diagonalisable

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite diagonalisable si elle est semblable à une matrice diagonale i.e. il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $D \in D_n(\mathbb{K})$ vérifiant

$$P^{-1}AP = D \text{ ou, et c'est équivalent, } A = PDP^{-1}$$

Théorème

Soit A la matrice d'un endomorphisme u dans une base e de E. On a équivalence entre :
(i) A est diagonalisable ;
(ii) u est diagonalisable.

dém. :
Les matrices semblables à A correspondent à celles pouvant représenter l'endomorphisme u.
□

Exemple En particulier, A est diagonalisable si l'endomorphisme canoniquement associé à la matrice A l'est.

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{K})$. On a équivalence entre :
(i) A est diagonalisable ;
(ii) $\overset{m}{\underset{i=1}{\oplus}} E_{\lambda_i}(A) = \mathcal{M}_{n,1}(\mathbb{K})$ (ou $\mathbb{K}^n$) ;
(iii) $n = \sum_{\lambda \in \mathrm{Sp}(A)} \dim E_\lambda(A)$;
(iv) χ_A est scindé dans $\mathbb{K}[X]$ et pour tout $\lambda \in \mathrm{Sp}(A)$, $\dim E_\lambda(A) = m_\lambda(A)$.
De plus, les matrices diagonales semblables à A sont celles dont les coefficients diagonaux sont les valeurs propres de A comptées avec multiplicité.

dém. :
On transite par l'endomorphisme canoniquement associé.
□

Théorème

Si $A \in \mathcal{M}_n(\mathbb{K})$ admet n valeurs propres distinctes alors A est diagonalisable et, de plus, ses sous-espaces propres sont des droites vectorielles.

Exemple Une matrice triangulaire à coefficients diagonaux distincts est assurément diagonalisable.

Exemple Soit $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$.
a) Diagonalisabilité si $\mathbb{K} = \mathbb{R}$.
b) Diagonalisabilité si $\mathbb{K} = \mathbb{C}$.
$\chi_A = X^2 - 2X + 2$.
Dans $\mathcal{M}_2(\mathbb{R})$, A n'est pas diagonalisable car χ_A n'est pas scindé.
Dans $\mathcal{M}_n(\mathbb{C})$, A est diagonalisable car admet deux valeurs propres $1 + i$ et $1 - i$.
La matrice A est alors semblable à

$$\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}$$

Exemple Diagonalisabilité de $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.
$\chi_A(X) = (1 - X)^2$, Sp$A = \{1\}$.
Si A est diagonalisable alors A est semblable à I_2 donc égale à I_2.
Ainsi A est diagonalisable si, et seulement si, $a = 0$.

Exemple Diagonalisabilité de $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$.
$\chi_A = (X - 1)^2(X - 2)$, Sp$A = \{1, 2\}$.
$\dim E_1(A) = 3 - \text{rg}(A - I_3)$, or $\text{rg}(A - I_3) = \text{rg} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = 2$ donc
$\dim E_1(A) = 1 < 2 = m_1(A)$.
La matrice A n'est donc pas diagonalisable.

Exemple Diagonalisabilité de

$$A = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}) \text{ (avec } n \geqslant 2)$$

$\chi_A = (X - n)X^{n-1}$, Sp$(A) = \{0, n\}$.
$\dim E_0(A) = n - \text{rg}A = n - 1$ et $\dim E_1(A) = 1$ (valeur propre simple).
Puisque $\dim E_0(A) + \dim E_n(A) = n$, A est diagonalisable semblable à $D = \text{diag}(n, 0, \ldots, 0)$.

Bilan :
- n valeurs propres distinctes $\Rightarrow A$ diagonalisable ;
- $\sum \dim E_\lambda(A) = n \Rightarrow A$ diagonalisable ;
- χ_A non scindé $\Rightarrow A$ non diagonalisable ;
- $\exists \lambda \in \mathrm{Sp}A, \dim E_\lambda(A) < m_\lambda(A) \Rightarrow A$ non diagonalisable.

5.4.5 Diagonalisation

5.4.5.1 D'un endomorphisme

Soit $u \in \mathcal{L}(E)$ diagonalisable.
Pour diagonaliser l'endomorphisme u, il suffit d'exhiber une base propre en considérant, par exemple, une base adaptée à la décomposition

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(u)} E_\lambda(u)$$

Exemple Soit E un $\mathbb{R}$-espace vectoriel de dimension 3 muni d'une base $e = (e_1, e_2, e_3)$.
Diagonalisation de $u \in \mathcal{L}(E)$ dont la matrice dans e est

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$\chi_u = X(X-1)(X-2)$, $\mathrm{Sp}u = \{0, 1, 2\}$.
$\mathrm{CardSp}u = 3 = \dim E$ donc u est diagonalisable.
$E_0(u) = ?$

Soit $x = x_1e_1 + x_2e_2 + x_3e_3 \in E$ et $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

$$u(x) = 0 \Leftrightarrow AX = 0 \Leftrightarrow \begin{cases} x_1 + x_2 - x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} x_2 = -x_1 \\ x_3 = 0 \end{cases}$$

Ainsi $E_0(u) = \mathrm{Vect}(e_1 - e_2)$ et de même on obtient $E_1(u) = \mathrm{Vect}(-e_1 + e_2 + e_3)$,
$E_2(u) = \mathrm{Vect}(e_2 + e_3)$.
Soit $\varepsilon_1 = e_1 - e_2$, $\varepsilon_2 = -e_1 + e_2 + e_3$ et $\varepsilon_3 = e_2 + e_3$.
La famille $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est une base de E (famille de vecteurs propres associés à des valeurs propres distinctes ou base adaptée à la décomposition de E en somme directe de sous-espaces propres).
La matrice de u dans ε est

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

En notant P la matrice de passage de e à ε, on a $A = PDP^{-1}$.
Ici

$$P = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \text{ et } P^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

5.4.5.2 D'une matrice

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice diagonalisable.
Notons $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{K}^n$. L'endomorphisme u canoniquement associé à la matrice A est diagonalisable. On peut introduire $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ base de vecteurs propres de u.

$$u(\varepsilon_j) = \lambda_j \varepsilon_j$$

La matrice de u dans la base ε est

$$D = \text{diag}(\lambda_1, \ldots, \lambda_n)$$

Par formule de changement de base

$$A = PDP^{-1} \text{ avec } P = \text{Mat}_e \varepsilon$$

Bilan : On forme une matrice de passage P diagonalisant A en prenant pour colonnes les vecteurs propres de A. La matrice diagonale D obtenue a pour coefficients diagonaux les valeurs propres respectives des colonnes formant P.

Exemple Diagonalisation de

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$\chi_A = (X-1)^2(X+1)^2$ via $C_1 \leftarrow C_1 + C_4$ et $C_2 \leftarrow C_2 + C_3$.
$\text{Sp}(A) = \{1, -1\}$.

$$E_1(A) = \text{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}\right), E_{-1}(A) = \text{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}\right)$$

$\dim E_1(A) + \dim E_{-1}(A) = 4$ donc A est diagonalisable.
Pour

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix} \text{ et } D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

on a $A = PDP^{-1}$.

Exemple Soit $\theta \neq 0 \quad [\pi]$.
Diagonalisation de $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$.
$\chi_{R(\theta)} = X^2 - 2\cos\theta X + 1$.
$\Delta = -4\sin^2\theta < 0$
Cas $\mathbb{K} = \mathbb{R}$
La matrice $R(\theta)$ n'est pas diagonalisable car $\chi_{R(\theta)}$ non scindé.
Cas $\mathbb{K} = \mathbb{C}$

On a

$$\mathrm{Sp}_{\mathbb{C}}(R_\theta) = \left\{ \mathrm{e}^{i\theta}, \mathrm{e}^{-i\theta} \right\}$$

et

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \in E_{\mathrm{e}^{i\theta}}(R(\theta)) \Leftrightarrow \begin{cases} \cos\theta x - \sin\theta y = \mathrm{e}^{i\theta} x \\ \sin\theta x + \cos\theta y = \mathrm{e}^{i\theta} y \end{cases} \Leftrightarrow ix + y = 0$$

On en déduit

$$E_{\mathrm{e}^{i\theta}}(R(\theta)) = \mathrm{Vect} \begin{pmatrix} i \\ 1 \end{pmatrix}$$

Par conjugaison

$$E_{\mathrm{e}^{-i\theta}}(R(\theta)) = \mathrm{Vect} \begin{pmatrix} -i \\ 1 \end{pmatrix}$$

Pour $P = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$, on a $R(\theta) = PD(\theta)P^{-1}$ avec

$$D(\theta) = \begin{pmatrix} \mathrm{e}^{i\theta} & 0 \\ 0 & \mathrm{e}^{-i\theta} \end{pmatrix}$$

5.4.6 Applications

5.4.6.1 Calcul des puissances d'une matrice

Si A est diagonalisable, on peut écrire $A = PDP^{-1}$ avec $P \in \mathrm{GL}_n(\mathbb{K})$ et D diagonale. On a alors

$$\forall k \in \mathbb{N}, A^k = PD^kP^{-1}$$

Exemple Calcul des puissances de

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$$

$\chi_A = X^2 - 5X + 6$. $\mathrm{Sp}A = \{2, 3\}$.
Après résolution

$$E_2(A) = \mathrm{Vect} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ et } E_3(A) = \mathrm{Vect} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$A = PDP^{-1}$ avec $P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ et $P^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

$$A^n = PD^nP^{-1} = P \begin{pmatrix} 2^n & 0 \\ 0 & 0 \end{pmatrix} P^{-1} + P \begin{pmatrix} 3^n & 0 \\ 0 & 0 \end{pmatrix} P^{-1} = 2^n \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix} + 3^n \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix}$$

Remarque Si l'on étudie un couple (u_n, v_n) de suites réelles vérifiant

$$\forall n \in \mathbb{N}, \begin{cases} u_{n+1} = u_n + 2v_n \\ v_{n+1} = -u_n + 4v_n \end{cases}$$

l'étude qui précède permet d'exprimer (u_n, v_n) en fonction de (u_0, v_0).
En effet, en introduisant $X_n = {}^t\begin{pmatrix} u_n & v_n \end{pmatrix}$, on a $X_{n+1} = AX_n$ et donc $X_n = A^nX_0$.

5.4.6.2 Commutant d'un endomorphisme diagonalisable

Théorème

Soit $u \in \mathcal{L}(E)$ un endomorphisme diagonalisable et $v \in \mathcal{L}(E)$.
On a équivalence entre :
(i) v commute avec u ;
(ii) les sous-espaces propres de u sont stables par v.

dém. :
(i) $\Rightarrow$ (ii) déjà vue.
(ii) $\Leftarrow$ (i) Supposons (ii).
Puisque u est diagonalisable

$$E = \underset{\lambda \in \operatorname{Sp} u}{\oplus} E_\lambda(u)$$

Pour $\lambda \in \operatorname{Sp}u$ et $x \in E_\lambda(u)$:

$$(v \circ u)(x) = v(u(x)) = v(\lambda x) = \lambda v(x)$$

et

$$(u \circ v)(x) = u(v(x)) = \lambda v(x)$$

car $v(x) \in E_\lambda(u)$.
Ainsi, les endomorphismes $u \circ v$ et $v \circ u$ coïncident sur tous les sous-espaces propres de u.
Puisque $E = \underset{\lambda \in \operatorname{Sp} u}{\oplus} E_\lambda(u)$, ces endomorphismes sont égaux sur E.
□

5.4.6.3 Résolution d'équation matricielle

Exemple Résolvons l'équation matricielle

$$M^2 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$$

Posons $D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.
Si M est solution alors $MD = M^3 = DM$.
Les solutions sont à rechercher parmi les matrices commutant avec D.
Pour $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, la relation $MD = DM$ donne $\begin{pmatrix} a & 4b \\ c & 4d \end{pmatrix} = \begin{pmatrix} a & b \\ 4c & 4d \end{pmatrix}$ et donc $b = c = 0$.
Ainsi, la matrice M est diagonale.
Pour $M = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$, l'équation $M^2 = D$ équivaut à $\begin{cases} a^2 = 1 \\ d^2 = 4 \end{cases}$.
Ainsi, les solutions de l'équation sont

$$D_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, D_2 = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}, D_3 = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \text{ et } D_4 = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$$

Remarque L'équation de degré 2 ici résolue possède plus de deux solutions car l'anneau $\mathcal{M}_n(\mathbb{K})$ n'est pas intègre.

Exemple Résolvons l'équation matricielle

$$M^2 = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$$

Posons $A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$. $\chi_A = X^2 - 5X + 4 = (X-1)(X-4)$.
Sp$A = \{1, 4\}$ et A est diagonalisable.
$E_1(A) = \text{Vect}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $E_4(A) = \text{Vect}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.
Pour $P = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$, $A = PDP^{-1}$ avec $D = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.
$M^2 = A \Leftrightarrow M^2 = PDP^{-1} \Leftrightarrow P^{-1}M^2P = D \Leftrightarrow (P^{-1}MP)^2 = D$.
Ainsi, les solutions de l'équation étudiée sont $PD_1P^{-1}, PD_2P^{-1}, PD_3P^{-1}$ et PD_4P^{-1}.

5.5 Trigonalisabilité

E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^\star$.

5.5.1 Endomorphisme trigonalisable

Définition

Un endomorphisme u de E est dit trigonalisable s'il existe une base de E dans laquelle la matrice de u est triangulaire supérieure. Une telle base est dite base de trigonalisation de l'endomorphisme u.

Exemple Un endomorphisme diagonalisable est a fortiori trigonalisable.

Théorème

Soit $e = (e_1, \ldots, e_n)$ une base de l'espace E.
On a équivalence entre :
(i) la base e trigonalise un endomorphisme u ;
(ii) $\forall 1 \leqslant k \leqslant n, \text{Vect}(e_1, \ldots, e_k)$ est stable par u

dém. :
(i) $\Rightarrow$ (ii) Si la matrice $A = (a_{i,j})$ de u dans la base e est triangulaire supérieure alors

$$\forall 1 \leqslant k \leqslant n, u(e_k) \in \text{Vect}(e_1, \ldots, e_k)$$

On en déduit

$$\forall 1 \leqslant k \leqslant n, u(e_1), \ldots, u(e_k) \in \text{Vect}(e_1, \ldots, e_k) \text{ stable par } u$$

puis (ii) par combinaison linéaire.
(ii) $\Rightarrow$ (i) Supposons (ii). On a en particulier

$$\forall 1 \leqslant k \leqslant n, u(e_k) \in \text{Vect}(e_1, \ldots, e_k)$$

et donc la matrice de u dans e est triangulaire supérieure.
□

Corollaire

Le premier vecteur d'une base de trigonalisation est un vecteur propre de l'endomorphisme.

5.5.2 Matrice trigonalisable

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite trigonalisable si elle est semblable à une matrice triangulaire supérieure.

Théorème

Soit A la matrice d'un endomorphisme u dans une base e de E. On a équivalence entre :
(i) A est trigonalisable ;
(ii) u est trigonalisable.

dém. :
Les matrices semblables à A correspondent à celles pouvant représenter l'endomorphisme u.
□

Exemple En particulier, A est trigonalisable si l'endomorphisme canoniquement associé à la matrice A l'est.

5.5.3 Caractérisation

Théorème

Pour $u \in \mathcal{L}(E)$, on a équivalence entre :
(i) u est trigonalisable ;
(ii) χ_u est scindé dans $\mathbb{K}[X]$;
On a un critère analogue pour $A \in \mathcal{M}_n(\mathbb{K})$.

dém. :
(i) $\Rightarrow$ (ii) Supposons u trigonalisable. Il existe une base de E dans laquelle la matrice de u est de la forme

$$T = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

On a alors

$$\chi_u(X) = \chi_T(X) = \prod_{i=1}^{n} (X - \lambda_i)$$

Ainsi χ_u est scindé dans $\mathbb{K}[X]$ (et les coefficients diagonaux de T sont les valeurs propres de u comptées avec multiplicité).
(ii) $\Rightarrow$ (i) Raisonnons matriciellement. Par récurrence sur $n \in \mathbb{N}^\star$, montrons que si le polynôme caractéristique de $A \in \mathcal{M}_n(\mathbb{K})$ est scindé alors A est semblable à une matrice triangulaire supérieure.
Cas $n = 1$: C'est immédiat, une matrice $A \in \mathcal{M}_1(\mathbb{K})$ étant déjà triangulaire supérieure.
Supposons la propriété établie au rang $n - 1 \geqslant 1$.
Soit $A \in \mathcal{M}_n(\mathbb{K})$ de polynôme caractéristique χ_A scindé.
Le polynôme χ_A possède au moins une racine λ_1 est celle-ci est valeur propre de A. Soit $e_1 \in \mathbb{K}^n$ un vecteur propre associé. On complète ce vecteur en une base de $\mathbb{K}^n$ de la forme $e = (e_1, e_2, \dots, e_n)$. La matrice de l'endomorphisme u canoniquement associé à la matrice A dans la base e est de la forme

$$B = \begin{pmatrix} \lambda_1 & \star \\ 0 & A' \end{pmatrix}$$

On a alors

$$\chi_A(X) = (X - \lambda)\chi_{A'}(X)$$

et donc le polynôme caractéristique de A' est scindé. Par hypothèse de récurrence, il existe $P' \in \mathrm{GL}_{n-1}(\mathbb{K})$ telle que la matrice $P'^{-1}A'P'$ soit triangulaire supérieure. Considérons alors la matrice

$$P = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

La matrice P est inversible avec

$$P^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & P'^{-1} \end{pmatrix}$$

Par produit par blocs

$$P^{-1}BP = \begin{pmatrix} \lambda_1 & \star' \\ 0 & P'^{-1}A'P' \end{pmatrix}$$

est triangulaire supérieure.
Finalement, A est semblable à une matrice triangulaire supérieure.
Récurrence établie.
□

Corollaire

Tout endomorphisme d'un $\mathbb{C}$-espace vectoriel E de dimension finie est trigonalisable.
Toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable.

dém. :
Car de polynôme caractéristique scindé.
□

Corollaire

Si χ_u est scindé dans $\mathbb{K}[X]$ alors tr(u) et $\det(u)$ sont la somme et le produit des valeurs propres comptées avec multiplicité.
Idem pour $A \in \mathcal{M}_n(\mathbb{K})$

dém. :
u est trigonalisable et peut donc être représenté par une matrice de la forme

$$\begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

Le polynôme caractéristique de u est alors

$$\prod_{k=1}^{n} (X - \lambda_k)$$

Les $\lambda_1, \ldots, \lambda_m$ sont alors les valeurs propres comptées avec multiplicité.
Parallèlement tr$(u) = \lambda_1 + \cdots + \lambda_n$ et $\det(u) = \lambda_1 \ldots \lambda_n$.
□

Remarque Ce résultat peut aussi se voir comme une conséquence de l'écriture

$$\chi_u(\lambda) = \lambda^n - \mathrm{tr}(u)\lambda^{n-1} + \cdots + (-1)^n \det(u)$$

Remarque Pour $A \in \mathcal{M}_n(\mathbb{C})$ le résultat qui précède s'applique automatiquement.
Pour $A \in \mathcal{M}_n(\mathbb{R})$, on peut interpréter $A \in \mathcal{M}_n(\mathbb{C})$ et affirmer que tr$A$ et $\det A$ sont la somme et le produit des valeurs propres complexes de A comptées avec multiplicité.

Exemple Déterminons les valeurs propres de

$$A = \begin{pmatrix} a_1 & \cdots & a_1 \\ \vdots & & \vdots \\ a_n & \cdots & a_n \end{pmatrix} \neq O_n$$

La matrice A est de rang 1 donc $\dim E_0(A) = \dim \ker A = n - 1$.
0 est alors valeur propre de A de multiplicité au moins $n-1$. Le polynôme χ_A s'écrit alors

$$\chi_A = (-1)^n X^{n-1}(X - \lambda)$$

Il est donc scindé dans $\mathbb{K}[X]$ et la trace de A est alors la somme des valeurs propres de A. On en déduit

$$\mathrm{Sp}A = \{0, a_1 + \cdots + a_n\}$$

5.5.4 Trigonalisation

Soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que χ_A soit scindé dans $\mathbb{K}[X]$.
Protocole :
Pour trigonaliser A, on détermine λ_1 valeur propre de A et e_1 vecteur propre associé.
Le vecteur e_1 définit la première colonne d'une matrice de passage Q que l'on construit inversible. On a alors

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & \star \\ 0 & A' \end{pmatrix}$$

avec A' trigonalisable. En déterminant P' inversible telle que

$$P'^{-1}A'P' = \begin{pmatrix} \lambda_2 & & \star \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

on forme alors

$$P = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}$$

et alors

$$P^{-1}Q^{-1}AQP = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

de sorte que $R = QP$ trigonalise la matrice A.

Exemple Trigonalisation de

$$A = \begin{pmatrix} -1 & 0 & -1 \\ 2 & -3 & -5 \\ -1 & 1 & 1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

$\chi_A = -(X+1)^3$. Sp$A = \{-1\}$
La matrice A est trigonalisable sans être diagonalisable car $A \neq -I_3$.
$E_{-1}(A) =$ Vecte_1 avec $e_1 = (1,1,0)$.
Considérons

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ avec } Q^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

On a

$$Q^{-1}AQ = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -3 & -4 \\ 0 & 1 & 1 \end{pmatrix}$$

et l'on considère

$$A' = \begin{pmatrix} -3 & -4 \\ 1 & 1 \end{pmatrix}$$

$E_{-1}(A') =$ Vect$(2,-1)$
Considérons

$$P' = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}, P'^{-1} = \begin{pmatrix} 1/2 & 0 \\ 1/2 & 1 \end{pmatrix}$$

puis

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix}, R^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}$$

On obtient

$$P^{-1}AP = \begin{pmatrix} -1 & 1 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix} \text{ avec } P = QR = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

Exemple Trigonalisation de

$$A = \begin{pmatrix} -1 & -3 & -1 \\ -1 & 1 & 1 \\ -2 & -3 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$$

$\chi_A = -(X+2)(X-1)^2$.
1 est valeur propre double et -2 est valeur propre simple
$E_{-2}(A) =$ Vect$(1,0,1)$, $E_1(A) =$ Vect$(1,-1,1)$
La matrice A n'est pas diagonalisable, cependant elle est trigonalisable.
Considérons

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ avec } P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

On obtient

$$P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

5.5.5 Nilpotence

Définition

Un endomorphisme $u \in \mathcal{L}(E)$ est dit nilpotent s'il existe $p \in \mathbb{N}$ vérifiant $u^p = \tilde{0}$.
Le plus petit p vérifiant cette identité est appelé indice de nilpotence de u.
Ce vocabulaire se transpose aux matrices

Exemple Si $A = \text{Mat}_e u$ alors la matrice A est nilpotente si, et seulement si, l'endomorphisme u l'est.

Exemple La matrice $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ est nilpotente car $A^2 = O_2$.

Exemple Soit A une matrice triangulaire supérieure stricte de $\mathcal{M}_n(\mathbb{K})$.

$$A = \begin{pmatrix} 0 & & & \star \\ & \ddots & & \\ & & \ddots & \\ (0) & & & 0 \end{pmatrix}, A^2 = \begin{pmatrix} 0 & 0 & & \star \\ & \ddots & \ddots & \\ & & \ddots & 0 \\ (0) & & & 0 \end{pmatrix}, \text{etc}$$

Montrons (proprement) que $A^n = O_n$.
Soit u l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A.
Notons $e = (e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$.
On a $u(e_1) = 0$ et pour tout $2 \leqslant i \leqslant n$, on a $u(e_i) \in \text{Vect}(e_1, \dots, e_{i-1})$.
Par suite

$$\text{Im}u = \text{Vect}(u(e_1), \dots, u(e_n)) \subset \text{Vect}(e_1, \dots, e_{n-1})$$

puis

$$\text{Im}u^2 \subset u(\text{Vect}(e_1, \dots, e_{n-1})) = \text{Vect}(u(e_1), \dots, u(e_{n-1})) \subset \text{Vect}(e_1, \dots, e_{n-2})$$

Par récurrence, on obtient

$$\forall 1 \leqslant k \leqslant n-1, \text{Im}u^k = \text{Vect}(e_1, \dots, e_{n-k-1})$$

En particulier $\text{Im}u^{n-1} \subset \text{Vect}(e_1)$ puis $\text{Im}u^n \subset \{0_E\}$ ce qui donne $u^n = \tilde{0}$.
On peut alors conclure $A^n = O_n$.

Théorème

Soit $u \in \mathcal{L}(E)$. On a équivalence entre :
(i) u est nilpotent ;
(ii) u est trigonalisable avec 0 pour seule valeur propre.
Ce résultat se transpose aux matrices de la façon suivante :
$A \in \mathcal{M}_n(\mathbb{K})$ est nilpotente si, et seulement si, A est semblable à une matrice triangulaire supérieure stricte

dém. :
(ii) $\Rightarrow$ (i) Car une matrice triangulaire supérieure figurant u a pour coefficients diagonaux les valeurs propres de u, elle est donc triangulaire supérieure stricte et par conséquent nilpotente.

(i) $\Rightarrow$ (ii)
Raisonnons matriciellement. Par récurrence sur $n \in \mathbb{N}^\star$, montrons que si $A \in \mathcal{M}_n(\mathbb{K})$ est nilpotente, alors A est semblable à une matrice triangulaire supérieure stricte.
Cas $n = 1$: Une matrice nilpotente de taille 1 est nécessairement nulle.
Supposons la propriété établie au rang $n - 1 \geqslant 1$.
Soit $A \in \mathcal{M}_n(\mathbb{K})$ nilpotente.
La matrice A ne peut être inversible et donc $\ker A \neq \{0\}$. Soit e_1 un vecteur non nul de $\ker A$. On complète ce vecteur e_1 en une base de $\mathbb{K}^n$ de la forme $e = (e_1, \ldots, e_n)$.
La matrice de l'endomorphisme canoniquement associé à A dans la base e est de la forme

$$B = \begin{pmatrix} 0 & \star \\ 0 & A' \end{pmatrix} \text{ avec } A' \in \mathcal{M}_{n-1}(\mathbb{K})$$

La matrice B est semblable à A et donc elle aussi nilpotente. On en déduit que le bloc A' est nilpotent.
Par hypothèse de récurrence, il existe $P' \in \mathrm{GL}_{n-1}(\mathbb{K})$ telle que $P'^{-1}A'P'$ soit triangulaire supérieure stricte. Formons alors

$$P = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix} \in \mathrm{GL}_n(\mathbb{K})$$

Par produit par blocs

$$P^{-1}BP = \begin{pmatrix} 0 & \star' \\ 0 & P'^{-1}A'P' \end{pmatrix}$$

est triangulaire supérieure stricte.
Finalement, A est semblable à une matrice triangulaire supérieure stricte.
Récurrence établie.
□

Remarque Le polynôme caractéristique de u (ou de A) est alors X^n.

Corollaire

Si u est un endomorphisme nilpotent d'un $\mathbb{K}$-espace vectoriel E de dimension n alors

$$u^n = \tilde{0}$$

Si $A \in \mathcal{M}_n(\mathbb{K})$ est nilpotente alors $A^n = O_n$.

Chapitre 6

Réduction algébrique

$\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$ et E un $\mathbb{K}$-espace vectoriel.

6.1 Polynômes en un endomorphisme

6.1.1 Valeur d'un polynôme en un endomorphisme

Définition

On appelle valeur d'un polynôme $P = \sum_{k=0}^{N} a_k X^k \in \mathbb{K}[X]$ en un endomorphisme $u \in \mathcal{L}(E)$ l'application

$$P(u) \underset{\text{déf}}{=} \sum_{k=0}^{N} a_k u^k \in \mathcal{L}(E)$$

Exemple La valeur de $P = X^3$ en u est $P(u) = u^3$.
La valeur de $P = X^3 + 2X - 1$ en u est $P(u) = u^3 + 2u - \text{Id}$.

Attention : La valeur de $P(u)$ en $x \in E$ est notée $P(u)(x)$ à comprendre $[P(u)](x)$.
Ecrire $P(u(x))$ n'a pas de sens.

Théorème

L'application $\varphi_u : \mathbb{K}[X] \to \mathcal{L}(E)$ définie par $\varphi_u(P) = P(u)$ est un morphisme de $\mathbb{K}$-algèbres.

dém. :
L'application φ_u est bien définie entre deux $\mathbb{K}$-algèbres.
$\varphi_u(1) = \text{Id}_E$.
Soit $\lambda, \mu \in \mathbb{K}$ et $P, Q \in \mathbb{K}[X]$.
On peut écrire $P = \sum_{k=0}^{N} a_k X^k$ et $Q = \sum_{k=0}^{M} b_k X^k$.

Quitte à adjoindre des coefficients nuls, on peut supposer $M = N$. On a

$$\varphi_u(\lambda P + MQ) = \sum_{k=0}^{N} (\lambda a_k + M b_k) u^k = \lambda \sum_{k=0}^{N} a_k u^k + M \sum_{k=0}^{N} b_k u^k = \lambda \varphi_u(P) + M \varphi_u(Q)$$

Aussi

$$\varphi_u(PQ) = (PQ)(u) = \left(\sum_{k=0}^{N} a_k X^k Q \right)(u) = \sum_{k=0}^{N} a_k (X^k Q)(u)$$

la dernière égalité étant justifiée par linéarité de φ_u. Or, pour $k \in \{0, 1, \ldots, N\}$, on a

$$(X^k Q)(u) = \sum_{\ell=0}^{N} b_\ell u^{k+\ell} = u^k \circ Q(u)$$

donc

$$(PQ)(u) = \left(\sum_{k=0}^{N} a_k X^k Q \right)(u) = \sum_{k=0}^{N} a_k u^k \circ Q(u)$$

puis

$$\varphi_u(PQ) = (PQ)(u) = P(u) \circ Q(u) = \varphi_u(P) \circ \varphi_u(Q)$$

□

Remarque Par ce morphisme, toute identité polynomiale se transpose aux endomorphismes.

Exemple Puisque

$$X^3 - 2X + 1 = (X - 1)(X^2 + X - 1)$$

on a

$$u^3 - 2u + \mathrm{Id}_E = (u - \mathrm{Id}_E)(u^2 + u - \mathrm{Id}_E)$$

Exemple Soit $P = X^n + a_{n-1} X^{n-1} + \cdots + a_0 \in \mathbb{C}[X]$. En notant $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ les racines de P comptées avec multiplicité

$$P = X^n + a_{n-1} X^{n-1} + \cdots + a_0 = \prod_{k=1}^{n} (X - \lambda_k)$$

alors

$$P(u) = \prod_{k=1}^{n} (u - \lambda_k \mathrm{Id}_E)$$

6.1.2 Polynôme d'endomorphisme

Définition

On dit que $v \in \mathcal{L}(E)$ est un polynôme en $u \in \mathcal{L}(E)$ s'il existe $P \in \mathbb{K}[X]$ tel que $v = P(u)$.
On note $\mathbb{K}[u]$ l'ensemble des polynômes en u :

$$\mathbb{K}[u] \underset{\text{déf}}{=} \{P(u)/P \in \mathbb{K}[X]\}$$

Exemple $u^3 + 3u + \mathrm{Id}_E$ et $(u - \lambda \mathrm{Id}_E)^\alpha$ sont des polynômes en u.

Théorème

$\mathbb{K}[u]$ est une sous-algèbre commutative de $\mathcal{L}(E)$.
De plus, si A est une sous-algèbre de $\mathcal{L}(E)$,

$$u \in A \Rightarrow \mathbb{K}[u] \subset A$$

Ainsi, $\mathbb{K}[u]$ est la plus petite sous-algèbre de $\mathcal{L}(E)$ contenant u, on l'appelle algèbre engendrée par u.

dém. :
$\mathbb{K}[u] \subset \mathcal{L}(E)$, $\mathrm{Id}_E \in \mathbb{K}[u]$ car pour $P(X) = 1$ on a $P(u) = \mathrm{Id}_E$.
Soit $\lambda, \mu \in \mathbb{K}$ et $v, w \in \mathbb{K}[u]$. Il existe $P, Q \in \mathbb{K}[X]$ tels que $v = P(u)$ et $w = Q(u)$.
On a alors $\lambda v + \mu w = (\lambda P + \mu Q)(u) \in \mathbb{K}[u]$ et $v \circ w = (PQ)(u) \in \mathbb{K}[u]$ donc $\mathbb{K}[u]$ est une sous-algèbre de $\mathcal{L}(E)$.
De plus, $w \circ v = (QP)(u) = (PQ)(u) = v \circ w$ donc $\mathbb{K}[u]$ est une sous-algèbre commutative de $\mathcal{L}(E)$.
Si A est une sous-algèbre de $\mathcal{L}(E)$ contenant u alors par récurrence

$$\forall n \in \mathbb{N}, u^n \in A$$

puis

$$\mathbb{K}[u] = \text{Vect}\left\{u^k/k \in \mathbb{N}\right\} \subset A$$

□

Exemple Si $P \in \mathbb{K}[X]$ alors $\mathrm{Im}P(u)$ et $\ker P(u)$ sont stables par u.
En effet, les $P(u)$ et u commutent. On retrouve en particulier que les sous-espaces propres de u sont stables par u.

6.1.3 Polynôme annulateur

Définition

On appelle polynôme annulateur de $u \in \mathcal{L}(E)$ tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(u) = \tilde{0}$.

Exemple Le polynôme nul annule tout endomorphisme.

Exemple Le polynôme $X-\lambda$ annule l'endomorphisme $\lambda\mathrm{Id}_E$

Exemple Le polynôme X^2-X est annulateur des projections vectorielles.

Théorème

> L'ensemble des polynômes annulateurs de $u\in\mathcal{L}(E)$ est un sous-espace vectoriel et un idéal de $\mathbb{K}[X]$.

dém. :
Notons $I=\left\{P\in\mathbb{K}[X]/P(u)=\tilde{0}\right\}$ l'ensemble des polynômes annulateurs de u.
I est le noyau du morphisme d'algèbres φ_u, c'est donc un sous-espace vectoriel et un idéal de $\mathbb{K}[X]$.
*Cor :*Si P annule u et si $P\mid Q$ alors Q annule u.
□

6.1.4 Polynôme annulateur et valeur propre

Lemme

> Si λ est valeur propre de $u\in\mathcal{L}(E)$ alors, pour tout $P\in\mathbb{K}[X]$, $P(\lambda)$ est valeur propre de $P(u)$.

dém. :
Soit λ une valeur propre de u. Il existe $x\neq 0_E$ tel que $u(x)=\lambda x$.
On a $u^2(x)=u(\lambda x)=\lambda^2x$,..., $u^n(x)=\lambda^n x$.
Soit $P=a_nX^n+\cdots+a_1X+a_0\in\mathbb{K}[X]$.
On a $P(u)(x)=(a_nu^n+\cdots+a_1u+a_0\mathrm{Id})(x)=(a_n\lambda^nx+\cdots+a_1\lambda x+a_0x)=P(\lambda)x$ avec $x\neq 0_E$ donc $P(\lambda)$ est valeur propre de $P(u)$.
□

Théorème

> Les valeurs propres de $u\in\mathcal{L}(E)$ figurent parmi les racines des polynômes annulateurs de u.

dém. :
Soit $P(X)$ un polynôme annulateur de u et λ une valeur propre de u.
On a $P(\lambda)$ valeur propre de $P(u)=\tilde{0}$ donc $P(\lambda)=0$.
□

Attention : Des racines d'un polynôme annulateur peuvent ne pas être valeur propre.

Exemple Si p est une projection vectorielle alors $X^2-X=X(X-1)$ est annulateur de p et donc $\mathrm{Sp}p\subset\{0,1\}$.

Exemple 0 est la seule valeur propre d'un endomorphisme nilpotent.
En effet,Soit $u\in\mathcal{L}(E)$ nilpotent.
Il existe $p\in\mathbb{N}^\star$ tel que $u^p=\tilde{0}$.
Le polynôme X^p est annulateur de u et donc $\mathrm{Sp}u\subset\{0\}$.
L'endomorphisme u ne peut être injectif (car u^p ne l'est pas) et donc $0\in\mathrm{Sp}u$.

6.2 Polynôme d'une matrice

6.2.1 Valeur d'un polynôme en une matrice carrée

Définition

On appelle valeur de $P = \sum_{k=0}^{N} a_k X^k \in \mathbb{K}[X]$ en $M \in \mathcal{M}_n(\mathbb{K})$ la matrice

$$P(M) \underset{\text{déf}}{=} \sum_{k=0}^{N} a_k M^k \in \mathcal{M}_n(\mathbb{K})$$

Exemple La valeur de $P = X^3 - 3X + 1$ en $M \in \mathcal{M}_n(\mathbb{K})$ est $P(M) = M^3 - 3M + I_n$.

Exemple Soit $u \in \mathcal{L}(E)$ et $e = (e_1, \ldots, e_n)$ une base de E.
Si $M = \text{Mat}_e u$ alors $\forall P \in \mathbb{K}[X]$, $P(M) = \text{Mat}_e P(u)$

Exemple Calcul de $P(M)$ pour

$$M = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

On vérifie par récurrence

$$\forall k \in \mathbb{N}, M^k = \begin{pmatrix} \lambda_1^k & & (0) \\ & \ddots & \\ (0) & & \lambda_n^k \end{pmatrix}$$

puis par linéarité

$$\forall P \in \mathbb{K}[X], P(M) = \begin{pmatrix} P(\lambda_1) & & (0) \\ & \ddots & \\ (0) & & P(\lambda_n) \end{pmatrix}$$

Exemple Expression de $P(M)$ pour

$$M = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

On vérifie par récurrence

$$\forall k \in \mathbb{N}, M^k = \begin{pmatrix} \lambda_1^k & & \star' \\ & \ddots & \\ (0) & & \lambda_n^k \end{pmatrix}$$

puis par linéarité

$$\forall P \in \mathbb{K}[X], P(M) = \begin{pmatrix} P(\lambda_1) & & \star'' \\ & \ddots & \\ (0) & & P(\lambda_n) \end{pmatrix}$$

Exemple Expression de $P(M)$ pour

$$M = \begin{pmatrix} A & \star \\ O & B \end{pmatrix} \text{ (avec } A, B \text{ matrices carrées)}$$

Comme au dessus, on obtient

$$\forall P \in \mathbb{K}[X], P(M) = \begin{pmatrix} P(A) & \star' \\ O & P(B) \end{pmatrix}$$

Exemple On a

$$\forall P \in \mathbb{K}[X], P({}^tM) = {}^tP(M)$$

En effet

$$\forall k \in \mathbb{N}, ({}^tM)^k = {}^t(M^k)$$

puis on conclut par linéarité

Théorème

L'application $\varphi_M : \mathbb{K}[X] \to \mathcal{M}_n(\mathbb{K})$ définie par $\varphi_M(P) = P(M)$ est un morphisme de $\mathbb{K}$-algèbres.

6.2.2 Polynôme en une matrice carrée

Définition

On dit que $A \in \mathcal{M}_n(\mathbb{K})$ est un polynôme en $M \in \mathcal{M}_n(\mathbb{K})$ s'il existe $P \in \mathbb{K}[X]$ tel que $A = P(M)$. On note

$$\mathbb{K}[M] \underset{\text{déf}}{=} \{P(M)/P \in \mathbb{K}[X]\}$$

l'ensemble des polynômes en M

Théorème

$\mathbb{K}[M]$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$ incluse dans toute sous-algèbre de $\mathcal{M}_n(\mathbb{K})$ contenant M ; on l'appelle algèbre engendrée par M.

6.2.3 Polynôme annulateur

Définition

On appelle polynôme annulateur de $M \in \mathcal{M}_n(\mathbb{K})$ tout polynôme $P \in \mathbb{K}[X]$ vérifiant $P(M) = O_n$.

Exemple Si $M = \mathrm{Mat}_e u$ alors les polynômes annulateurs de M et de u se correspondent.

Exemple Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$.
On vérifie par le calcul que $P = X^2 - (a+d)X + (ad - bc)$ est annulateur de M.

Exemple $P = (X - \lambda_1) \ldots (X - \lambda_n)$ est annulateur de

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$$

En effet

$$P(D) = \begin{pmatrix} P(\lambda_1) & & (0) \\ & \ddots & \\ (0) & & P(\lambda_n) \end{pmatrix} = O_n$$

Remarque Si A est diagonalisable semblable à D alors P est aussi annulateur de A. Plus généralement :

Proposition

Si $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblables alors A et B ont les mêmes polynômes annulateurs.

dém. :
Par le calcul à partir de la relation de similitude $B = Q^{-1}AQ$ ou simplement parce que les matrices A et B représentent le même endomorphisme.
□

Théorème

L'ensemble des polynômes annulateurs de $M \in \mathcal{M}_n(\mathbb{K})$ est un sous-espace vectoriel et un idéal de $\mathbb{K}[X]$.

Corollaire

Si P annule M et si $P \mid Q$ alors Q annule M.

6.2.4 Valeurs propres et polynômes annulateurs

Théorème

Les valeurs propres de $M \in \mathcal{M}_n(\mathbb{K})$ figurent parmi les racines des polynômes annulateurs de M.

Exemple Soit $A \in \mathcal{M}_3(\mathbb{R})$ vérifiant $A^3 = I_n$.
a) Valeurs propres réelles.
b) Valeurs propres complexes.
Le polynôme $X^3 - 1$ est annulateur de A.
Dans $\mathbb{R}$, $X^3 - 1 = (X - 1)(X^2 + X + 1)$ donc $\mathrm{Sp}_{\mathbb{R}} A \subset \{1\}$.
Or A est une matrice réelle de taille impaire donc $\mathrm{Sp}_{\mathbb{R}} A \neq \emptyset$ puis

$$\mathrm{Sp}_{\mathbb{R}} A = \{1\}$$

Dans $\mathbb{C}$, $X^3-1=(X-1)(X-j)(X-j^2)$ donc $\mathrm{Sp}_{\mathbb{C}}A\subset\left\{1,j,j^2\right\}$.
Puisque 1 est valeur propre et puisque les valeurs propres de A sont deux à deux conjuguées

$$\mathrm{Sp}_{\mathbb{C}}A=\{1\}\text{ ou }\mathrm{Sp}_{\mathbb{C}}A=\left\{1,j,j^2\right\}$$

On en déduit $\mathrm{tr}(A)=3$ ou $\mathrm{tr}(A)=0$ et $\det A=1$ (car χ_A est scindé et donc A trigonalisable)

6.3 Polynômes annulateurs en dimension finie

E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n\in\mathbb{N}^\star$.

6.3.1 Théorème de Cayley Hamilton

Théorème

Le polynôme caractéristique χ_u de $u\in\mathcal{L}(E)$ est annulateur de u.
Cet énoncé se transpose aux matrices $A\in\mathcal{M}_n(\mathbb{K})$.

Exemple Pour $A\in\mathcal{M}_2(\mathbb{K})$, le polynôme $\chi_A=X^2-\mathrm{tr}(A)X+\det(A)$ est annulateur de A.

6.3.2 Polynôme minimal

Théorème

Pour tout $u\in\mathcal{L}(E)$, il existe un unique polynôme Π_u vérifiant :
1) Π_u est annulateur de u ;
2) Π_u est unitaire ;
3) $\forall P\in\mathbb{K}\left[X\right], P(u)=\tilde{0}\Rightarrow\Pi_u\mid P$.
Ce polynôme Π_u est appelé polynôme minimal de l'endomorphisme u.
Cet énoncé se transpose aux matrices $A\in\mathcal{M}_n(\mathbb{K})$ ce qui définit le polynôme minimal Π_A

dém. :
Existence :
Considérons $I=\left\{P\in\mathbb{K}\left[X\right]/P(u)=\tilde{0}\right\}$.
Puisque I est un idéal de $\mathbb{K}\left[X\right]$, il existe un polynôme $Q\in\mathbb{K}\left[X\right]$ tel que $I=Q.\mathbb{K}\left[X\right]$.
Puisque $\chi_u\in I$, l'idéal I est non nul et donc $Q\neq 0$.
Notons λ le coefficient dominant de u et considérons $\Pi_u=Q/\lambda$. Le polynôme Π_u est unitaire et vérifie $I=\Pi_u.\mathbb{K}\left[X\right]$.
Unicité :
Supposons Π_u et $\tilde{\Pi}_u$ solutions.
Puisque $\tilde{\Pi}_u(u)=\tilde{0}$, $\Pi_u\mid\tilde{\Pi}_u$. De façon symétrique, $\tilde{\Pi}_u\mid\Pi_u$ et donc Π_u et $\tilde{\Pi}_u$ sont associés.
Or ils sont tous deux unitaires donc égaux.
□

Remarque Le polynôme Π_u est non constant.

Exemple Polynôme minimal de $u=\lambda\mathrm{Id}_E$.
$X-\lambda$ annule u et donc $\Pi_u\mid X-\lambda$.
Puisque Π_u est unitaire non constant, on obtient

$$\Pi_u=X-\lambda$$

Exemple Polynôme minimal de p projection autre que $\tilde{0}$ et Id_E.
On a $p^2 = p$ donne $\Pi_p \mid X(X-1)$.
Puisque Π_p est unitaire non constant

$$\Pi_p = X, X-1 \text{ ou } \Pi_p = X(X-1)$$

Puisque $p \neq \tilde{0}$ et $p \neq \mathrm{Id}_E$, les cas $\Pi_p = X$ et $\Pi_p = X-1$ sont à exclure.
Il reste

$$\Pi_p = X(X-1)$$

Exemple Polynôme minimal de $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.
$\chi_A = X^2 - 5X + 6 = (X-2)(X-3)$ est annulateur de A donc $\Pi_A \mid \chi_A$.
Par conséquent

$$\Pi_A = X-1, X-2 \text{ ou } (X-2)(X-3)$$

Les cas $\Pi_A = X-1$ ou $X-2$ sont à exclure et il reste

$$\Pi_A = (X-2)(X-3)$$

Exemple Polynôme minimal de $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$.
Cette fois-ci

$$\chi_D = (X-1)^2(X-2) \text{ et } \Pi_D = (X-1)(X-2)$$

6.3.3 Polynôme minimal et valeurs propres

Théorème

Les valeurs propres de $u \in \mathcal{L}(E)$ sont exactement les racines de son polynôme minimal.
Ce résultat se transpose aux matrices carrées.

dém. :
On sait déjà que les valeurs propres de u sont racines de Π_u car Π_u est annulateur.
Inversement, si λ est racine de Π_u alors λ est aussi racine de χ_u donc λ est valeur propre de u.
□

Exemple Le polynôme $\prod_{\lambda \in \mathrm{Sp} u} (X - \lambda)$ divise Π_u.

6.3.4 Application : calcul des puissances d'un endomorphisme

Soit $u \in \mathcal{L}(E)$. On introduit son polynôme minimal Π_u de degré d (avec $d \leqslant \dim E$ car Π_u divise χ_u).
On écrit

$$\Pi_u = X^d - (a_{d-1}X^{d-1} + \cdots + a_1 X + a_0)$$

et alors
$$u^d = a_0\mathrm{Id}_E + a_1 u + \cdots + a_{d-1}u^{d-1}$$
Puisque $u^{d+1} = u \circ u^d$
$$u^{d+1} = a_0 u + a_1 u^2 + \cdots + a_{d-1}u^d$$
et en exploitant la relation au dessus, on obtient une expression
$$u^{d+1} = a_0'\mathrm{Id}_E + a_1' u + \cdots + a_{d-1}' u^{d-1}$$
On peut répéter ce processus… Plus généralement :

Théorème

Si $d = \deg \Pi_u$ alors la famille $(u^k)_{0\leqslant k\leqslant d-1}$ est une base $\mathbb{K}[u]$.
Ce résultat se transpose aux matrices carrées.

dém. :
Commençons par montrer $\mathbb{K}[u] = \mathrm{Vect}(\mathrm{Id}, u, \ldots, u^{d-1})$
On a déjà l'inclusion $\mathrm{Vect}(\mathrm{Id}_E, u, \ldots, u^{d-1}) \subset \mathbb{K}[u]$.
Inversement, soit $P \in \mathbb{K}[X]$.
Par division euclidienne, on peut écrire $P = Q\Pi_u + R$ avec $\deg R < d$.
On a alors
$$P(u) = Q(u) \circ \Pi_u(u) + R(u) = R(u) \in \mathrm{Vect}(\mathrm{Id}_E, u, \ldots, u^{d-1})$$
Ainsi $\mathbb{K}[u] \subset \mathrm{Vect}(\mathrm{Id}_E, u, \ldots, u^{d-1})$ puis l'égalité.
Montrons maintenant que la famille $(\mathrm{Id}_E, u, \ldots, u^{d-1})$ est libre.
Supposons
$$a_0\mathrm{Id}_E + a_1 u + \cdots + a_{d-1}u^{d-1} = \tilde{0}$$
Pour $P = a_0 + a_1 X + \cdots + a_{d-1}X^{d-1}$, on a $P(u) = \tilde{0}$.
Or $\deg P < \deg \Pi_u$ donc $P = 0$ puis $a_0 = a_1 = \ldots = a_{d-1} = 0$.
Ainsi, la famille $(\mathrm{Id}, u, \ldots, u^{d-1})$ est libre et c'est donc une base de $\mathbb{K}[u]$.
□

Corollaire

$\dim \mathbb{K}[u] \leqslant \dim E$ et $\dim \mathbb{K}[A] \leqslant n$.

dém. :
Car le polynôme minimal est diviseur du polynôme caractéristique donc de degré inférieur à n.
□

Exemple Calculons les puissances de $A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} \in \mathcal{M}_2(\mathbb{R})$.
On sait $\Pi_A = (X-2)(X-3)$.
Par division euclidienne
$$X^n = \Pi_A(X)Q(X) + \alpha X + \beta \text{ (1)}$$
En évaluant la relation (1) en 2 et en 3, on obtient
$$\begin{cases} 2\alpha + \beta = 2^n \\ 3\alpha + \beta = 3^n \end{cases} \text{ donc } \begin{cases} \alpha = 3^n - 2^n \\ \beta = 3.2^n - 2.3^n \end{cases}$$
En évaluant la relation (1) en A, on obtient
$$A^n = (3^n - 2^n)A + (3.2^n - 2.3^n)I_2$$

6.4 Réduction et polynômes annulateurs

E désigne un $\mathbb{K}$-espace vectoriel non nul

6.4.1 Lemme de décomposition des noyaux

Théorème

Soit $P, Q \in \mathbb{K}[X]$ et $u \in \mathcal{L}(E)$.
Si P et Q sont premiers entre eux alors

$$\ker(PQ)(u) = \ker P(u) \oplus \ker Q(u)$$

dém. :
Puisque $P \wedge Q = 1$, il existe des polynômes V et W tel que $VP + WQ = 1$.
On a alors $\mathrm{Id} = V(u) \circ P(u) + W(u) \circ Q(u)$.
Soit $x \in \ker P(u) \cap \ker Q(u)$
On a

$$x = (V(u) \circ P(u))(x) + (W(u) \circ Q(u))(x) = 0$$

donc $\ker P(u)$ et $\ker Q(u)$ sont en somme directe.
Montrons $\ker P(u) \oplus \ker Q(u) \subset \ker(PQ)(u)$
Puisque $(PQ)(u) = Q(u) \circ P(u)$ on a $\ker P(u) \subset \ker PQ(u)$.
De même $\ker Q(u) \subset \ker(PQ)(u)$ et donc $\ker P(u) \oplus \ker Q(u) \subset \ker(PQ)(u)$.
Inversement
Soit $x \in \ker(PQ)(u)$. On a

$$x = (W(u) \circ Q(u))(x) + (V(u) \circ P(u))(x) = a + b$$

avec $a = (W(u) \circ Q(u))(x)$ et $b = (V(u) \circ P(u))(x)$.
Or

$$P(u)(a) = (P(u) \circ W(u) \circ Q(u))(x) = (W(u) \circ (PQ)(u))(x) = 0$$

et de même $Q(u)(b) = 0$. Ainsi $a \in \ker P(u)$ et $b \in \ker Q(u)$ puis

$$\ker(PQ)(u) \subset \ker P(u) \oplus \ker Q(u)$$

et enfin l'égalité.
□

Corollaire

Si $P_1, \ldots, P_m \in \mathbb{K}[X]$ sont deux à deux premiers entre eux alors :

$$\ker\left(\prod_{k=1}^{m} P_k\right)(u) = \mathop{\oplus}_{k=1}^{m} \ker P_k(u)$$

Ce résultat se transpose aux matrices carrées

$$\ker\left(\prod_{k=1}^{m} P_k\right)(A) = \mathop{\oplus}_{k=1}^{m} \ker P_k(A)$$

dém. :
On raisonne par récurrence en exploitant

$$(P_1 \dots P_m) \wedge P_{m+1} = 1 \Rightarrow \ker\left(\prod_{k=1}^{m+1} P_k\right)(u) = \ker\left(\prod_{k=1}^{m} P_k\right)(u) \oplus \ker P_{m+1}(u)$$

□
Rappel :
Si $a \neq b$ alors $(X-a) \wedge (X-b) = 1$.
Plus généralement, $(X-a)^\alpha \wedge (X-b)^\beta = 1$ pour tout $\alpha, \beta \in \mathbb{N}$.
Encore plus généralement, deux polynômes de $\mathbb{K}[X]$ sont premiers entre eux si, et seulement si, ils n'ont pas de racines complexes en commun.

Exemple On appelle projecteur de E tout $p \in \mathcal{L}(E)$ vérifiant $p^2 = p$.
Les espaces $F = \ker(p - \mathrm{Id})$ et $G = \ker p$ sont supplémentaires et

$$\forall x \in F, p(x) = x \text{ et } \forall x \in G, p(x) = 0_E$$

En effet $p^2 - p = \tilde{0}$ donc $E = \ker\left(p^2 - p\right)$.
Or $X^2 - X = (X-1)X$ avec $(X-1) \wedge X = 1$
donc $E = \ker(p^2 - p) = \ker(p - \mathrm{Id}) \oplus \ker p$.
on reconnaît que p est la projection sur F parallèlement à G.

Exemple On appelle symétrie de E tout $s \in \mathcal{L}(E)$ vérifiant $s^2 = \mathrm{Id}_E$.
Les espaces $F = \ker(s - \mathrm{Id})$ et $G = \ker(s + \mathrm{Id})$ sont supplémentaires et

$$\forall x \in F, s(x) = x \text{ et } \forall x \in G, s(x) = -x$$

En effet, $s^2 - \mathrm{Id}_E = \tilde{0}$ donc $E = \ker\left(s^2 - \mathrm{Id}_E\right)$.
Or $X^2 - 1 = (X-1)(X+1)$ avec $(X-1) \wedge (X+1) = 1$ donc $E = \ker(s - \mathrm{Id}) \oplus \ker(s + \mathrm{Id})$.
Posons
on reconnaît que s est la symétrie par rapport à F et parallèlement à G.

Exemple Soit $\lambda_1, \dots, \lambda_m$ les valeurs propres deux à deux distinctes de $u \in \mathcal{L}(E)$.
Les polynômes $X - \lambda_k$ étant deux à deux premiers entre eux, on retrouve que les sous-espaces propres d'un endomorphisme sont en somme directe.

6.4.2 Diagonalisabilité

Théorème

On a équivalence entre :
(i) u est diagonalisable ;
(ii) u annule un polynôme scindé à racines simples ;
(iii) le polynôme minimal de u est scindé à racines simples.
De plus, le polynôme minimal de u est alors

$$\Pi_u = \prod_{\lambda \in \mathrm{Sp} u} (X - \lambda)$$

Ce résultat se transpose aux matrices carrées.

dém. :
Notons $\lambda_1, \ldots, \lambda_m$ les valeurs propres de u.
(i) $\Rightarrow$ (ii) Supposons u diagonalisable. On a

$$E = \underset{k=1}{\overset{m}{\oplus}} E_{\lambda_k}(u)$$

Dans une base adaptée à cette décomposition la matrice de u est de la forme

$$\begin{pmatrix} \lambda_1 I_{\alpha_1} & & (0) \\ & \ddots & \\ (0) & & \lambda_m I_{\alpha_m} \end{pmatrix} \text{ avec } \alpha_k = \dim E_{\lambda_k}(u)$$

Considérons le polynôme

$$P = \prod_{k=1}^{m} (X - \lambda_k)$$

Dans la base précédente, la matrice de $P(u)$ est

$$\begin{pmatrix} P(\lambda_1) I_{\alpha_1} & & (0) \\ & \ddots & \\ (0) & & P(\lambda_m) I_{\alpha_m} \end{pmatrix} = O_n$$

u annule le polynôme P qui est scindé à racines simples.
(ii) $\Rightarrow$ (iii) Si u annule un polynôme scindé à racines simples alors Π_u le divise et est donc lui-même scindé à racines simples.
(iii) $\Rightarrow$ (i) Supposons Π_u scindé à racines simples. Puisque les racines de Π_u sont exactement les valeurs propres de u, on peut écrire

$$\Pi_u = \prod_{k=1}^{m} (X - \lambda_k)$$

Or les facteurs $(X - \lambda_k)$ étant premiers entre eux, le lemme de décomposition des noyaux donne

$$E = \ker \Pi_u(u) = \underset{k=1}{\overset{m}{\oplus}} \ker(u - \lambda_k \mathrm{Id}_E) = \underset{k=1}{\overset{m}{\oplus}} E_{\lambda_k}(u)$$

□

Définition

> On dit qu'un polynôme de $\mathbb{K}[X]$ est scindé simple lorsqu'il est scindé dans $\mathbb{K}[X]$ à racines simples

Exemple Diagonalisation de $T : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$ définie par $T(M) = {}^tM$.
On a $T^2 = \mathrm{Id}$ donc $X^2 - 1$ annule T.
Puisque le polynôme $X^2 - 1$ est scindé simple, l'endomorphisme T est diagonalisable.
De plus

$$\mathrm{Sp}T \subset \{1, -1\}, E_1(T) = \ker(T - \mathrm{Id}) = \mathcal{S}_n(\mathbb{R}) \text{ et } E_{-1}(T) = \ker(I + \mathrm{Id}) = \mathcal{A}_n(\mathbb{R})$$

On en déduit $\mathrm{tr}T = \dim \mathcal{S}_n(\mathbb{R}) - \dim \mathcal{A}_n(\mathbb{R}) = n$ et $\det T = (-1)^{\dim \mathcal{A}_n(\mathbb{R})} = (-1)^{n(n-1)/2}$.
En fait, l'endomorphisme T est la symétrie vectorielle par rapport à $\mathcal{S}_n(\mathbb{R})$ et parallèlement à $\mathcal{A}_n(\mathbb{R})$.

Exemple Soit $A \in \mathcal{M}_n(\mathbb{R})$ telle que $A^2 + I = 0$.
Montrons que n est pair et calculons $\det A$ et tr$A$.
A annule $X^2 + 1 = (X - i)(X + i)$ scindé simple donc A est diagonalisable dans $\mathcal{M}_n(\mathbb{C})$.
De plus,

$$\text{Sp}A \subset \{i, -i\}$$

Or Sp$A \neq \emptyset$ et les valeurs propres de A sont conjuguées car $A \in \mathcal{M}_n(\mathbb{R})$ donc

$$\text{Sp}A = \{i, -i\}$$

Enfin, les multiplicités des valeurs propres conjuguées sont égales car $\chi_A \in \mathbb{R}[X]$ donc $\dim E_i(A) = \dim E_{-i}(A)$.
En posant p cette valeur commune, on peut affirmer que A est semblable dans $\mathcal{M}_n(\mathbb{C})$ à
$\begin{pmatrix} iI_p & O \\ O & -iI_p \end{pmatrix}$
On en déduit $n = 2p$, $\det A = 1$ et tr$A = 0$.

6.4.3 Réduction d'un endomorphisme induit par un endomorphisme diagonalisable

Lemme

Si F est un sous-espace vectoriel stable par $u \in \mathcal{L}(E)$ alors F est stable par tout polynôme en u et

$$\forall P \in \mathbb{K}[X], P(u)_F = P(u_F)$$

dém. :
Puisque F est stable par u, il l'est aussi par $u^2, \ldots, u^n, \ldots$ et

$$\forall n \in \mathbb{N}, (u_F)^n = (u^n)_F$$

Par combinaison linéaire, F est encore stable par les polynômes en u et

$$\forall P \in \mathbb{K}[X], P(u)_F = P(u_F)$$

Si u est diagonalisable alors u annule un polynôme scindé simple P et alors $P(u_F) = (P(u))_F = \tilde{0}$ donc u_F annule un polynôme scindé simple et est donc diagonalisable.
□

Proposition

Si F est stable par $u \in \mathcal{L}(E)$ alors le polynôme minimal de u_F divise le polynôme minimal de u.

dém. :
Le polynôme minimal de u est annulateur de u_F.
□

Théorème

Si $u \in \mathcal{L}(E)$ est diagonalisable et si F est un sous-espace vectoriel stable par u alors u_F est diagonalisable.

dém. :
Π_u est scindé à racines simples dont Π_{u_F} l'est aussi.
□

Corollaire

> Soit $u \in \mathcal{L}(E)$ diagonalisable.
> Les sous-espaces vectoriels stables par u sont ceux admettant une base de vecteurs propres.

dém. :
Si F est stable par u alors u_F est diagonalisable donc F admet une base de vecteurs propres de u_F. Ceux-ci sont aussi vecteurs propres de u.
Inversement, si $(e_1, \dots, e_p)$ est une base de F formée de vecteurs propres alors pour tout $j \in \{1, \dots, p\}$, $u(e_j) \in \text{Vect}(e_j) \subset F$ et donc F est stable par u.
□

Exemple Soit u et $v \in \mathcal{L}(E)$ diagonalisables.
Montrons que si u et v commutent alors il existe une base de E formée de vecteurs propres communs à u et v.
Puisque u est diagonalisable, $E = \underset{\lambda \in \text{Sp}u}{\oplus} E_\lambda(u)$.
Pour $\lambda \in \text{Sp}u$, $E_\lambda(u)$ est stable par v, or v est diagonalisable donc $v_{E_\lambda(u)}$ l'est aussi. Ainsi, il existe une base $\mathcal{B}_\lambda$ de $E_\lambda(u)$ formée de vecteurs propres de v. Cette base est a fortiori formée de vecteurs propres de u. En accolant les bases $\mathcal{B}_\lambda$, on forme une base de E formée de vecteurs propres communs à u et v.
Matriciellement, on a obtenu que si $A, B \in \mathcal{M}_n(\mathbb{K})$ sont diagonalisables et commutent alors il existe $P \in \text{GL}_n(\mathbb{K})$ vérifiant $P^{-1}AP$ et $P^{-1}BP$ diagonales.

6.4.4 Trigonalisabilité

Théorème

> On a équivalence entre :
> (i) u est trigonalisable ;
> (ii) u annule un polynôme scindé dans $\mathbb{K}[X]$;
> (iii) le polynôme minimal de u est scindé dans $\mathbb{K}[X]$.
> De plus, l'espace E est alors la somme directe de sous-espaces stables par u sur chacun desquels u induit la somme d'une homothétie et d'un endomorphisme nilpotent.

dém. :
(i) $\Rightarrow$ (ii) Car si u est trigonalisable alors u annule son polynôme caractéristique qui est scindé dans $\mathbb{K}[X]$.
(ii) $\Rightarrow$ (iii) Car le polynôme minimal divise un polynôme scindé.
(iii) $\Rightarrow$ (i) Supposons le polynôme minimal Π_u de u scindé dans $\mathbb{K}[X]$. On peut écrire

$$\Pi_u = \prod_{k=1}^{m} (X - \lambda_k)^{\alpha_k}$$

avec $\lambda_1, \dots, \lambda_m$ les valeurs propres distinctes de u. Par le lemme de décomposition des noyaux

$$E = \overset{m}{\underset{k=1}{\oplus}} \ker (u - \lambda_k \text{Id}_E)^{\alpha_k}$$

Etudions $F = \ker(u - \lambda_k \text{Id}_E)^{\alpha_k}$. L'espace F est stable par u car u et $(u - \lambda_k \text{Id}_E)^{\alpha_k}$ commutent. On peut introduire $n_k = u_F - \lambda \text{Id}_F \in \mathcal{L}(F)$ et on a $n_k^{\alpha_k} = \tilde{0}$ car $F = \ker(u - \lambda_k \text{Id}_E)^{\alpha_k}$. Ainsi $u_F = \lambda_k \text{Id}_F + n_k$

avec n_k nilpotent. Enfin, puisque n_k est nilpotent, il existe une base F_k dans laquelle la matrice de n_k est triangulaire supérieure. En accolant ces bases, on obtient une base de E dans laquelle la matrice de u est triangulaire supérieure.
□

Remarque Les espaces $\ker\left(u - \lambda_k \mathrm{Id}_E\right)^{\alpha_k}$ s'appellent espaces caractéristiques de l'endomorphisme u.

Corollaire

Si $A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable alors A est semblable à une matrice diagonale par blocs où chaque bloc diagonal est de la forme

$$\lambda I_\alpha + N$$

avec N une matrice nilpotente.

Remarque Ce résultat s'applique automatiquement lorsque $\mathbb{K} = \mathbb{C}$ et l'on retrouve que toute matrice de $\mathcal{M}_n(\mathbb{C})$ est trigonalisable.

Corollaire

Si $u \in \mathcal{L}(E)$ est trigonalisable et si F est un sous-espace vectoriel stable par u alors u_F est trigonalisable.

dém. :
Π_u est scindé donc Π_{u_F} l'est aussi.
□

6.4.5 Musculation : décomposition de Dunford

Théorème

Soit $u \in \mathcal{L}(E)$ avec Π_u scindé dans $\mathbb{K}[X]$.
On peut écrire $u = d + n$ avec d diagonalisable, n nilpotent et $d \circ n = n \circ d$.

dém. :
On introduit $\lambda_1, \ldots, \lambda_m$ les valeurs propres deux à deux distincts de u.

$$\Pi_u = \prod_{k=1}^{m} (X - \lambda_k)^{\mu_k} \text{ et } E = \bigoplus_{k=1}^{m} \ker(u - \lambda_k \mathrm{Id}_E)^{n_k}$$

Posons d l'endomorphisme déterminé par

$$\forall 1 \leqslant k \leqslant m, \forall x \in \ker(u - \lambda_k \mathrm{Id}_E)^{n_k}, d(x) = \lambda_k x$$

L'endomorphisme d est évidemment diagonalisable, ses sous-espaces propres sont les espaces caractéristiques.
Posons n l'endomorphisme donné par $n = u - d$.

$$\forall 1 \leqslant k \leqslant m, \forall x \in \ker(u - \lambda_k \mathrm{Id}_E)^{n_k}, n^{n_k}(x) = 0_E$$

Pour $N = \max(n_1, \ldots, n_m)$, on obtient

$$\forall 1 \leqslant k \leqslant m, \forall x \in \ker(u - \lambda_k \mathrm{Id}_E)^{n_k}, n^{N}(x) = 0_E$$

L'endomorphisme n est donc nilpotent.
Enfin

$$\forall 1 \leqslant k \leqslant m, \forall x \in \ker(u - \lambda_k \mathrm{Id}_E)^{n_k}, (n \circ d)(x) = \lambda_k n(x) = (d \circ n)(x)$$

et donc les endomorphismes d et n commutent.
On peut aussi montrer qu'il y a unicité des endomorphismes d et n de cette décomposition.
Supposons d et n solutions.
d commute avec n donc aussi avec $u = d + n$.
L'espace caractéristique $F = \ker(u - \lambda \mathrm{Id}_E)^n$ est alors stable par d.
L'endomorphisme induit par d sur F est diagonalisable.
Soit μ une valeur propre de celui-ci et $G \subset F$ l'espace propre associé.
G est stable par u et donc aussi par $n = u - d$ et l'on a

$$u_G = \mu \mathrm{Id}_G + n_G$$

Puisque n_G est nilpotent, on peut calculer χ_{u_G} dans une base trigonalisant n_G et affirmer que μ est alors valeur propre de u_G donc de u_F. Or λ est la seule valeur propre de u_F et donc $\mu = \lambda$. On en déduit que λ est la seule valeur propre de l'endomorphisme diagonalisable d_F et ainsi

$$\forall x \in F, d(x) = \lambda x$$

L'endomorphisme d est alors déterminé de façon unique sur les espaces caractéristiques de u.
L'endomorphisme $n = u - d$ est alors aussi unique.
□

Remarque La décomposition de Dunford est utile pour calculer les puissances de u car la formule du binôme peut lui être appliquée.

Chapitre 7

Espaces préhilbertiens réels

E désigne un $\mathbb{R}$-espace vectoriel.

7.1 Produit scalaire

7.1.1 Définition

Définition

On appelle produit scalaire sur un $\mathbb{R}$-espace vectoriel E toute application $\varphi : E \times E \to \mathbb{R}$ vérifiant :
1) φ est bilinéaire ;
2) φ est symétrique i.e. $\forall x, y \in E, \varphi(y, x) = \varphi(x, y)$;
3) φ est positive i.e. $\forall x \in E, \varphi(x, x) \geqslant 0$;
4) φ est définie i.e. $\forall x \in E, \varphi(x, x) = 0 \Rightarrow x = 0_E$.
On dit qu'un produit scalaire est une forme bilinéaire symétrique définie positive.

Remarque Les points 3) et 4) peuvent être avantageusement remplacés par

$$\forall x \in E \backslash \{0_E\}, \varphi(x, x) > 0$$

Définition

On appelle espace préhilbertien réel tout couple (E, φ) formé d'un $\mathbb{R}$-espace vectoriel E et d'un produit scalaire φ sur E. Il est alors usuel de noter $(x \mid y)$, $\langle x, y\rangle$ ou $x.y$ au lieu de $\varphi(x, y)$ le produit scalaire de deux vecteurs de E.

Exemple Sur $E = \mathbb{R}^n$, $\langle x, y\rangle = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + \cdots + x_n y_n$ définit un produit scalaire.
$\langle ., .\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ est bien définie.
Soit $\lambda, \mu \in \mathbb{R}, x, y, z \in \mathbb{R}^n$.

$$\langle x, \lambda y + \mu z\rangle = \sum_{k=1}^{n} x_k(\lambda y_k + \mu z_k) = \lambda \langle x, y\rangle + \mu \langle x, z\rangle$$

$\langle ., .\rangle$ est linéaire en sa deuxième variable.

$$\langle y, x\rangle = \sum_{k=1}^{n} y_k x_k = \langle x, y\rangle$$

$\langle .,. \rangle$ est symétrique et donc bilinéaire.
Enfin
$$\langle x, x \rangle = \sum_{k=1}^{n} x_k^2 \geqslant 0$$
et
$$\langle x, x \rangle = 0 \Rightarrow x = 0_{\mathbb{R}^n}$$
Finalement $\langle .,. \rangle$ est un produit scalaire.

Exemple Sur $E = \mathcal{M}_{n,p}(\mathbb{R})$, $(A \mid B) = \text{tr}({}^tAB)$ définit un produit scalaire.
$(. \mid .) : E \times E \to \mathbb{R}$ est bien définie car ${}^tAB$ est une matrice carrée.
Soit $\lambda, \mu \in \mathbb{R}$ et $A, B, C \in \mathcal{M}_{n,p}(\mathbb{R})$.
$$(A \mid \lambda B + \mu C) = \text{tr}\left({}^tA(\lambda B + \mu C)\right) = \lambda(A \mid B) + \mu(A \mid C)$$
$$(B \mid A) = \text{tr}\left({}^tBA\right) = \text{tr}^t\left({}^tBA\right) = \text{tr}\left({}^tAB\right) = (A \mid B)$$
Ainsi $(. \mid .)$ est une forme bilinéaire symétrique.
$$(A \mid A) = \text{tr}\left({}^tAA\right) = \sum_{j=1}^{p} \left[{}^tAA\right]_{j,j}$$
Or
$$\left[{}^tAA\right]_{j,j} = \sum_{i=1}^{n} \left[{}^tA\right]_{j,i} [A]_{i,j} = \sum_{i=1}^{n} a_{i,j}^2$$
en notant $a_{i,j}$ les coefficients de A.
$$(A \mid A) = \sum_{j=1}^{p} \sum_{i=1}^{n} a_{i,j}^2$$
Ainsi $(A \mid A) \geqslant 0$ et $(A \mid A) = 0 \Rightarrow A = O_{n,p}$.
φ est donc définie positive et par suite c'est un produit scalaire.
En fait
$$(A \mid B) = \text{tr}({}^tAB) = \sum_{j=1}^{p} \left[{}^tAB\right]_{j,j} = \sum_{j=1}^{p} \sum_{i=1}^{n} a_{i,j} b_{i,j}$$
Le produit scalaire introduit est analogue à celui défini ci-dessus sur $\mathbb{R}^n$.

Remarque Sur $E = \mathcal{M}_{n,1}(\mathbb{R})$,
$$(X \mid Y) = \text{tr}({}^tXY) = {}^tXY$$
car ${}^tXY$ est une matrice uni-coefficient.
Ainsi, le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ est donné par
$$\varphi(X, Y) = {}^tXY = x_1 y_1 + \cdots + x_n y_n$$
avec $X = {}^t\begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$ et $Y = {}^t\begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix}$.
Par l'identification des colonnes et des tuples, les produits scalaires canoniques se correspondent.
L'action de ce produit scalaire est la même que celle du produit scalaire sur $\mathbb{R}^n$.

Exemple Soit $a < b$ deux réels et $E = \mathcal{C}([a,b],\mathbb{R})$. $(f \mid g) = \int_a^b f(t)g(t)\,\mathrm{d}t$ définit un produit scalaire sur E.
En effet, l'application $(.\mid .) : E \times E \to \mathbb{R}$ est bien définie et clairement bilinéaire symétrique et pour $f \in E$, on a

$$(f \mid f) = \int_a^b f(t)^2\,\mathrm{d}t \geqslant 0$$

et

$$(f \mid f) = 0 \Rightarrow f = \tilde{0}$$

car seule la fonction nulle est une fonction continue positive d'intégrale nulle.

Remarque Si l'on considère $\omega : [a,b] \to \mathbb{R}^{+\star}$ continue, on définit aussi un produit scalaire sur E en posant

$$\langle f, g\rangle = \int_a^b f(t)g(t)\omega(t)\,\mathrm{d}t$$

Le résultat est encore vrai pour ω s'annulant un nombre fini de fois.

Remarque On peut aussi définir des produits scalaires sur $\mathbb{R}[X]$ parmi lesquels les fameux suivants

$$\langle P, Q\rangle = \int_0^1 P(t)Q(t)\,\mathrm{d}t,\ \langle P, Q\rangle = \int_0^{+\infty} P(t)Q(t)\mathrm{e}^{-t}\,\mathrm{d}t \text{ ou } \langle P, Q\rangle = \int_{-1}^1 \frac{P(t)Q(t)}{\sqrt{1-t^2}}\,\mathrm{d}t$$

7.1.2 Norme euclidienne

E désigne un espace préhilbertien réel et $(.\mid .)$ désigne son produit scalaire.

Définition

On appelle norme euclidienne sur E l'application $\|.\| : E \to \mathbb{R}^+$ définie par

$$\|x\| = \sqrt{(x \mid x)}$$

Exemple Sur $E = \mathbb{R}^n$ muni du produit scalaire canonique

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2} = \|x\|_2$$

Dans le cas $n = 1$, $\|x\| = \sqrt{x^2} = |x|$.

Exemple Sur $E = \mathcal{M}_{n,p}(\mathbb{R})$ muni du produit scalaire canonique

$$\|A\| = \sqrt{\mathrm{tr}({}^tAA)} = \left(\sum_{i=1}^n \sum_{j=1}^p a_{i,j}^2\right)^{1/2} = \|A\|_2$$

Exemple Sur $E = \mathcal{C}([a,b], \mathbb{R})$,

$$\|f\| = \left(\int_a^b f(t)^2 \, \mathrm{d}t\right)^{1/2} = \|f\|_2$$

Proposition

$\forall x \in E, \|x\| = 0 \Rightarrow x = 0_E$.
$\forall \lambda \in \mathbb{R}, \forall x \in E, \|\lambda x\| = |\lambda| \, \|x\|$.

dém. :
$\|x\| = 0 \Rightarrow (x \mid x) = 0$ donc $\|x\| = 0 \Rightarrow x = 0_E$.
$\|\lambda x\|^2 = (\lambda x \mid \lambda x) = \lambda^2 (x \mid x) = \lambda^2 \|x\|^2$ donc $\|\lambda x\| = |\lambda| \, \|x\|$.
□

Proposition

$\forall a, b \in E, \|a+b\|^2 = \|a\|^2 + 2(a \mid b) + \|b\|^2$,
$\forall a, b \in E, \|a-b\|^2 = \|a\|^2 - 2(a \mid b) + \|b\|^2$,
$\forall a, b \in E, (a - b \mid a + b) = \|a\|^2 - \|b\|^2$.

dém. :
$\|a+b\|^2 = (a+b \mid a+b) = (a \mid a+b) + (b \mid a+b)$ par linéarité en la première variable.
$\|a+b\|^2 = (a \mid a) + (a \mid b) + (b \mid a) + (b \mid b)$ par linéarité en la deuxième variable.
$\|a+b\|^2 = \|a\|^2 + 2(a \mid b) + \|b\|^2$ par symétrie.
Les autres identités s'obtiennent de façon analogue.
□

Proposition

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

dém. :
Il suffit d'exploiter l'identité remarquable

$$\|x+y\|^2 = \|x\|^2 + 2(x \mid y) + \|y\|^2$$

□

Théorème

$$\forall x, y \in E, |(x \mid y)| \leqslant \|x\| \, . \, \|y\|$$

avec égalité si, et seulement si, la famille (x, y) est liée.

dém. :
Cas $x = 0_E$: immédiat.
Cas $x \neq 0_E$: Pour tout $\lambda \in \mathbb{R}$,

$$\|\lambda x + y\|^2 = \lambda^2 \|x\|^2 + 2\lambda(x \mid y) + \|y\|^2 = a\lambda^2 + b\lambda + c \geqslant 0$$

donc $\Delta = 4(x \mid y)^2 - 4\|x\|^2 \|y\|^2 \leqslant 0$. On en déduit $(x \mid y)^2 \leqslant \|x\|^2 \|y\|^2$.
De plus, il y a égalité si, et seulement si, $\Delta = 0$ c'est-à-dire si, et seulement si, il existe $\lambda \in \mathbb{R}$ vérifiant $\lambda x + y = 0$. Sachant $x \neq 0_E$, ceci équivaut à dire que la famille (x, y) est liée.
□

Exemple Sur $\mathbb{R}^n$,

$$\left|\sum_{k=1}^{n} x_k y_k\right| \leqslant \left(\sum_{k=1}^{n} x_k^2\right)^{1/2} \left(\sum_{k=1}^{n} y_k^2\right)^{1/2}$$

Exemple Sur $\mathcal{C}([a,b],\mathbb{R})$,

$$\left|\int_a^b f(t)g(t)\,\mathrm{d}t\right| \leqslant \left(\int_a^b f(t)^2\,\mathrm{d}t\right)^{1/2} \left(\int_a^b g(t)^2\,\mathrm{d}t\right)^{1/2}$$

Théorème

$$\forall x, y \in E,\ \|x+y\| \leqslant \|x\| + \|y\|$$

avec égalité si, et seulement si, x et y colinéaires et $(x \mid y) \geqslant 0$.
(on dit que x et y sont positivement liés)

dém. :
On a

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2(x \mid y) + \|y\|^2\\ &\leqslant \|x\|^2 + 2\,|(x \mid y)| + \|y\|^2\\ &\leqslant \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2\\ &= (\|x\| + \|y\|)^2\end{aligned}$$

avec égalité si, et seulement si, $(x \mid y) = |(x \mid y)| = \|x\|\,\|y\|$ i.e. x, y colinéaires et $(x \mid y) \geqslant 0$.
□

Corollaire

La norme euclidienne est une norme : tout espace préhilbertien réel est automatiquement un espace normé.

Théorème

Le produit scalaire est une application bilinéaire continue pour la norme euclidienne.

dém. :
$(.\mid .)$ est une application bilinéaire vérifiant $|(x \mid y)| \leqslant 1 \times \|x\|\,\|y\|$ elle est donc continue.
□

7.1.3 Vecteurs orthogonaux

E désigne un espace préhilbertien réel et $(.\mid .)$ désigne son produit scalaire.

Définition

Deux vecteurs x et y de E sont dits orthogonaux si $(x \mid y) = 0$.

Exemple Le vecteur nul est le seul vecteur orthogonal à lui-même :

$$(x \mid x) = 0 \Rightarrow x = 0_E$$

Exemple Le vecteur nul est le seul vecteur orthogonal à tout autre.

Définition

On dit qu'une famille $(e_i)_{i\in I}$ de vecteurs de E est orthogonale si elle est constituée de vecteurs deux à deux orthogonaux i.e.

$$\forall i, j \in I, i \neq j \Rightarrow (e_i \mid e_j) = 0$$

On dit que la famille est orthonormale si ses vecteurs sont de plus unitaires

$$\forall i, j \in I, (e_i \mid e_j) = \delta_{i,j}$$

Proposition

Toute famille orthogonale ne comportant pas le vecteur nul est libre.
En particulier, les familles orthonormales sont libres.

dém. :
Soit $(e_1, \ldots, e_n)$ une famille orthogonale finie ne comportant pas le vecteur nul.
Supposons $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0_E$.
Pour tout $1 \leqslant j \leqslant n$, $(e_j \mid \lambda_1 e_1 + \cdots + \lambda_n e_n) = (e_j \mid 0_E)$ donne $\lambda_j \left\|e_j\right\|^2 = 0$ et donc $\lambda_j = 0$.
On peut conclure que la famille est libre.
On étend le résultat aux familles infinies aisément car la liberté d'une famille infinie correspond à la liberté de ses sous-familles finies.
□

7.1.4 Algorithme d'orthonormalisation de Schmidt

Théorème

Si $(x_1, \ldots, x_n)$ est une famille libre de vecteurs de E alors il existe une unique famille orthonormale $(e_1, \ldots, e_n)$ vérifiant
1) $\forall 1 \leqslant k \leqslant n$, $\text{Vect}(x_1, \ldots, x_k) = \text{Vect}(e_1, \ldots, e_k)$;
2) $\forall 1 \leqslant k \leqslant n$, $(x_k \mid e_k) > 0$.
On dit que la famille $(e_1, \ldots, e_n)$ est la famille orthonormalisée de $(x_1, \ldots, x_n)$ par le procédé de Schmidt.

Dans la pratique pour orthonormaliser $(x_1, \ldots, x_n)$:
- Etape 1 : on pose $e_1 = x_1/\|x_1\|$;
- Etape 2 : on pose $u = x_2 + \lambda e_1$ et on détermine λ pour que $(e_1 \mid u) = 0$ puis on pose $e_2 = u/\|u\|$;
- Etape 3 : on pose $u = x_3 + \lambda e_1 + \mu e_2$ et on détermine λ et μ pour que $(e_1 \mid u) = (e_2 \mid u) = 0$ puis on pose $e_3 = u/\|u\|$;
- etc.

En fait

$$e_{k+1} = u/\|u\| \text{ avec } u = x_k - \sum_{i=1}^{k} (e_i \mid x_k) e_i$$

Exemple Dans $\mathbb{R}^3$ muni du produit scalaire canonique considérons la famille (x_1, x_2, x_3) avec

$$x_1 = (0, 1, 1), x_2 = (1, 0, 1), x_3 = (1, 1, 0)$$

La famille (x_1, x_2, x_3) est libre car

$$\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 2 \neq 0$$

$\|x_1\|^2 = 2$, $e_1 = \left(0, 1/\sqrt{2}, 1/\sqrt{2}\right)$
$u = x_2 + \lambda e_1$
$(e \mid e_1) = 0$ donne $\lambda = -1/\sqrt{2}$ puis $u = (1, -1/2, 1/2)$, $e_2 = \left(2/\sqrt{6}, -1/\sqrt{6}, 1/\sqrt{6}\right)$.
$u = x_3 + \lambda e_1 + \mu e_2$
$(e_3 \mid e_1) = 0$ donne $\lambda = -1/\sqrt{2}$,
$(e_3 \mid e_2) = 0$ donne $\mu = -1/\sqrt{6}$ puis $u = (2/3, 2/3, -2/3)$ et $e_3 = \left(1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}\right)$.

Exemple Dans $\mathcal{M}_2(\mathbb{R})$ muni du produit scalaire canonique $(A \mid B) = \text{tr}({}^tAB)$ considérons la famille (A_1, A_2, A_3) avec

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ et } A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

On vérifie aisément que cette famille est libre et le processus d'orthonormalisation de Schmidt donne

$$B_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ et } B_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

7.2 Espace euclidien

7.2.1 Définition

Définition

On appelle espace euclidien tout espace préhilbertien réel de dimension finie.

Exemple Pour leur produit scalaire canonique, $\mathbb{R}^n$ et $\mathcal{M}_{n,p}(\mathbb{R})$ sont des espaces euclidiens.

Définition

On appelle base orthonormale d'un espace euclidien E toute famille de vecteurs de E qui est à la fois une base et une famille orthonormale.

Exemple La base canonique de $\mathbb{R}^n$ est orthonormale pour le produit scalaire canonique.

Exemple La base canonique de $\mathcal{M}_{n,p}(\mathbb{R})$ est orthonormale pour le produit scalaire canonique.
En effet

$$(E_{i,j} \mid E_{k,\ell}) = \text{tr}({}^tE_{i,j}E_{k,\ell}) = \text{tr}(E_{j,i}E_{k,\ell}) = \text{tr}(\delta_{i,k}E_{j,\ell}) = \delta_{i,k}\delta_{j,\ell}$$

Théorème

Tout espace euclidien E possède une base orthonormale.

dém. :
Soit $(e_1, \dots, e_n)$ une base de E.
Par l'algorithme de Schmidt, on peut former une famille orthonormale $e' = (e'_1, \dots, e'_n)$.
Celle-ci est libre et constituée du bon nombre de vecteurs pour être une base.
□

Remarque Si $e' = (e'_1, \dots, e'_n)$ est une base orthonormale construite à partir d'une base $e = (e_1, \dots, e_n)$ par l'algorithme de Schmidt alors la matrice de passage de e à e' est triangulaire supérieure à coefficients diagonaux strictement positifs. En effet, on a

$$e'_k = u/\|u\| \text{ avec } u = e_k - \sum_{i=1}^{k-1} (e'_i \mid e_k)e'_i$$

et donc

$$e_k \in \text{Vect}(e'_1, \dots, e'_k)$$

Ainsi, la matrice de passage de e' à e est triangulaire supérieure, aussi l'est sa matrice inverse.

Théorème

Toute famille orthonormale d'un espace euclidien E peut être complétée en une base orthonormée.

dém. :
Soit $(x_1, \dots, x_p)$ une famille orthonormale de E.
Par le théorème de la base incomplète, on forme une base $(x_1, \dots, x_p, x_{p+1}, \dots, x_n)$.
En appliquant le procédé de Schmidt, on obtient une famille orthonormale $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$.
Or, par ce procédé, on a nécessairement $e_1 = x_1, \dots, e_p = x_p$ car la famille $(x_1, \dots, x_p)$ est déjà orthonormale.
On a ainsi obtenue une famille orthonormale de la forme $(x_1, \dots, x_p, e_{p+1}, \dots, e_n)$. Celle-ci est aussi une base de E car libre et constituée de $n = \dim E$ vecteurs de E.
□

7.2.2 Calcul des coordonnées dans une base orthonormale

Théorème

Les coordonnées $x_1, \dots, x_n$ d'un vecteur x de E dans la base orthonormée e sont données par

$$\forall k \in \{1, \dots, n\}, x_k = (e_k \mid x)$$

de sorte que

$$x = \sum_{k=1}^{n} (e_k \mid x)e_k$$

dém. :

On a $x = \sum_{k=1}^{n} x_k e_k$ donc

$$(e_k \mid x) = \left(e_k \mid \sum_{\ell=1}^{n} x_\ell e_\ell\right) = \sum_{\ell=1}^{n} x_\ell (e_k \mid e_\ell) = \sum_{\ell=1}^{n} x_\ell \delta_{k,\ell} = x_k$$

□

Corollaire

La matrice $A \in \mathcal{M}_n(\mathbb{K})$ d'un endomorphisme u de E dans une base orthonormale $e = (e_1, \ldots, e_n)$ a pour coefficient général

$$a_{i,j} = (e_i \mid u(e_j))$$

dém. :
Le coefficient d'indice (i, j) de A est la i-ème composante dans e du vecteur $u(e_j)$.
□

Exemple Si $e = (e_1, \ldots, e_n)$ est une base orthonormale, alors

$$\forall u \in \mathcal{L}(E), \mathrm{tr}u = \sum_{k=1}^{n} (e_k \mid u(e_k))$$

7.2.3 Expression du produit scalaire et de la norme

Théorème

Si $x, y \in E$ ont pour coordonnées $x_1, \ldots, x_n$ et $y_1, \ldots, y_n$ dans une base orthonormale $e = (e_1, \ldots, e_n)$ alors

$$(x \mid y) = x_1 y_1 + \cdots + x_n y_n = {}^t XY \text{ et } \|x\|^2 = x_1^2 + \cdots + x_n^2 = {}^t XX$$

dém. :
$x = \sum_{k=1}^{n} x_k e_k$ et $y = \sum_{k=1}^{n} y_k e_k$ donc

$$(x \mid y) = \left(\sum_{k=1}^{n} x_k e_k \mid \sum_{\ell=1}^{n} y_\ell e_\ell\right) = \sum_{k=1}^{n}\sum_{\ell=1}^{n} x_k y_\ell (e_k \mid e_\ell) = \sum_{k=1}^{n} x_k y_k$$

car $(e_k \mid e_\ell) = \delta_{k,\ell}$.
□

Remarque Considérons $\varphi : E \to \mathbb{K}^n$ définie par $\varphi(x) = (x_1, \ldots, x_n)$ avec $x_k = (e_k \mid x)$.
L'application φ est un isomorphisme de $\mathbb{K}$-espace vectoriel qui conserve le produit scalaire.
Ainsi, quand l'espace E est rapporté à une base orthonormée, il se comporte comme $\mathbb{K}^n$.

Exemple Soit $(x_1, \dots, x_p)$ une famille de p vecteurs d'un espace euclidien E muni d'une base orthonormale $e = (e_1, \dots, e_n)$. Notons $A = \text{Mat}_e(x_1, \dots, x_p)$. On a

$$^tAA = (\langle x_i, x_j \rangle)_{1 \leqslant i,j \leqslant p}$$

En effet,

$$\left[{}^tAA\right]_{i,j} = \sum_{k=1}^{p} a_{k,i}a_{k,j} = \langle x_i, x_j \rangle$$

car les $(a_{k,i})_{1 \leqslant k \leqslant n}$ sont les coordonnées de x_i dans la base orthonormale e.

7.2.4 Représentation d'une forme linéaire

Pour $a \in E$, l'application $\varphi_a : E \to \mathbb{R}$ définie par

$$\varphi_a(x) = (a \mid x)$$

est une forme linéaire.

Théorème

Si E est un espace euclidien alors

$$\forall \varphi \in E^\star, \exists! a \in E, \forall x \in E, \varphi(x) = (a \mid x)$$

dém. :

Considérons l'application $\Phi : E \to E^\star$ qui à $a \in E$ associe la forme linéaire $\varphi_a : x \mapsto (a \mid x)$.
L'application Φ est linéaire et injective car

$$(\forall x \in E, (a \mid x) = 0) \Rightarrow a = 0_E$$

Puisque $\dim E^\star = \dim E < +\infty$, l'application Φ est un isomorphisme.

□

Remarque Si $e = (e_1, \dots, e_n)$ est une base orthonormale de E et $\varphi \in E^\star$ alors le vecteur a pour lequel $\varphi = \varphi_a$ est

$$a = \sum_{k=1}^{n} \varphi(e_k)e_k$$

En effet, les coordonnées de a dans la base orthonormale E sont

$$a_k = (e_k \mid a) = \varphi(e_k)$$

Exemple Sur $E = \mathcal{M}_n(\mathbb{R})$, on considère le produit scalaire donné par $(A \mid B) = \text{tr}\left({}^tAB\right)$.
Si φ est une forme linéaire sur E alors il existe une matrice $A \in \mathcal{M}_n(\mathbb{R})$ vérifiant

$$\forall M \in \mathcal{M}_n(\mathbb{R}), \varphi(M) = \text{tr}(AM)$$

7.3 Sous-espaces vectoriels orthogonaux

7.3.1 Orthogonal d'une partie

Définition

On appelle orthogonal d'une partie A de E l'ensemble noté $A^\perp$ constitué des vecteurs de E orthogonaux à tous les vecteurs de A

$$A^\perp = \{x \in E/\forall a \in A, (a \mid x) = 0\}$$

Exemple $\{0_E\}^\perp = E$ et $E^\perp = \{0_E\}$.

Théorème

$A^\perp$ est un sous-espace vectoriel fermé de E.

dém. :
$A^\perp \subset E$ et $0_E \in A^\perp$ car 0_E est orthogonal à tous les vecteurs de E, notamment ceux de A.
Soit $\lambda, \mu \in \mathbb{K}$ et $x, y \in A^\perp$.
Pour tout $a \in A$, $(a \mid \lambda x + \mu y) = \lambda(a \mid x) + \mu(a \mid y) = 0$ donc $\lambda x + \mu y \in A^\perp$.
Soit $(x_n) \in (A^\perp)^{\mathbb{N}}$ convergeant vers un élément x_∞.
Soit $a \in A$. Pour tout $n \in \mathbb{N}$, $(a \mid x_n) = 0$ donc à la limite $(a \mid x_\infty) = 0$ car le produit scalaire est continue.
On en déduit $x_\infty \in A^\perp$.
□

Proposition

Pour $A, B \subset E$
a) $A \subset \left(A^\perp\right)^\perp$
b) $A \subset B \Rightarrow B^\perp \subset A^\perp$
c) $A^\perp = \text{Vect}(A)^\perp$

dém. :
a) Soit $x \in A$. Pour tout $y \in A^\perp$, $(x \mid y) = 0$ donc $x \in A^{\perp\perp}$.
b) Supposons $A \subset B$.
Soit $x \in B^\perp$. Pour tout $y \in A$ on a $(x \mid y) = 0$ car $x \in B^\perp$ et $y \in B$. Par suite $x \in A^\perp$.
Ainsi $A \subset B \Rightarrow B^\perp \subset A^\perp$.
c) $A \subset \text{Vect}(A)$ donc $\text{Vect}(A)^\perp \subset A^\perp$.
Aussi $A \subset A^{\perp\perp}$ donc $\text{Vect}(A) \subset A^{\perp\perp}$ puis $A^\perp \subset A^{\perp\perp\perp} \subset \text{Vect}(A)^\perp$
□

Proposition

Si $F = \text{Vect}(e_k)_{1 \leqslant k \leqslant m}$ alors

$$F^\perp = \{x \in E/\forall 1 \leqslant k \leqslant m, (e_k \mid x) = 0\}$$

dém. :
L'inclusion directe est immédiate, l'inclusion réciproque s'obtient par la propriété : si x est orthogonal à

une famille de vecteurs, il l'est aussi aux combinaisons linéaires de cette famille.
□

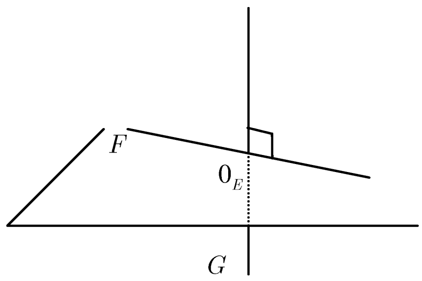
7.3.2 Sous-espaces vectoriels orthogonaux

Définition

Deux sous-espaces vectoriels F et G de E sont dits orthogonaux s'ils sont formés de vecteurs deux à deux orthogonaux i.e.

$$\forall (x,y) \in F \times G, (x \mid y) = 0$$

Exemple



Exemple F et $F^\perp$ sont des sous-espaces vectoriels orthogonaux.

Proposition

On a équivalence entre :
(i) F et G sont orthogonaux ;
(ii) $F \subset G^\perp$;
(iii) $G \subset F^\perp$.

dém. :
(i) $\Rightarrow$ (ii) Supposons F et G sont orthogonaux.
Soit $x \in F$. Pour tout $y \in G$, $(x \mid y) = 0$ donc $x \in G^\perp$. Ainsi $F \subset G^\perp$.
(ii) $\Rightarrow$ (i) Supposons $F \subset G^\perp$.
Pour tout $x \in F$ et $y \in G$, $(x \mid y) = 0$ car $x \in G^\perp$ et $y \in G$. Ainsi, les espaces F et G sont orthogonaux.
Par un argument de symétrie, on a aussi (i) $\Leftrightarrow$ (ii).
□

Remarque Une orthogonalité est une inclusion dans un orthogonal.

7.3.3 Somme directe orthogonale

Remarque Si F et G sont orthogonaux alors $F \cap G = \{0_E\}$ car

$$x \in F \cap G \Rightarrow (x \mid x) = 0$$

Ainsi deux sous-espaces vectoriels orthogonaux sont en somme directe. Plus généralement :

Théorème

> Si $F_1, \ldots, F_m$ sont des sous-espaces vectoriels de E deux à deux orthogonaux alors ceux-ci sont en somme directe.

dém. :
Supposons $x_1 + \cdots + x_m = 0_E$ avec chaque x_k dans F_k.
Pour tout $1 \leqslant k \leqslant m$,

$$(x_k \mid x_1 + \cdots + x_m) = (x_k \mid 0_E) = 0$$

donne $\|x_k\|^2 = 0$ car $(x_k \mid x_j) = 0$ pour $j \neq k$. Ainsi $x_k = 0_E$ pour tout $1 \leqslant k \leqslant m$.
□

Définition

> Lorsque les sous-espaces vectoriels $F_1, \ldots, F_m$ sont deux à deux orthogonaux, on dit qu'ils sont en somme directe orthogonale et leur somme est notée $\overset{n}{\underset{k=1}{\oplus}}{}^{\perp} F_k$.

Exemple Les espaces F et $F^\perp$ sont en somme directe orthogonale.

7.3.4 Supplémentaire orthogonal d'un sous-espace vectoriel de dimension finie

Théorème

> Si F est un sous-espace vectoriel de dimension finie alors l'espace $F^\perp$ est un supplémentaire de F dans E.
> On dit que $F^\perp$ est le supplémentaire orthogonal de F.

dém. :
On sait déjà que F et $F^\perp$ sont orthogonaux donc en somme directe.
Montrons $F + F^\perp = E$.
Soit $e = (e_1, \ldots, e_m)$ une base orthonormale de F.
Analyse : Soit $x \in E$. Supposons $x = a + b$ avec $a \in F$ et $b \in G$.
On a $a = \displaystyle\sum_{k=1}^{m} (e_k \mid a) e_k$ or $(e_k \mid a) = (e_k \mid x) - (e_k \mid b) = (e_k \mid x)$ car $e_k \in F$ et $b \in F^\perp$.
On en déduit $a = \displaystyle\sum_{k=1}^{m} (e_k \mid x) e_k$ et $b = x - a$.
Synthèse : Soit $x \in E$, $a = \displaystyle\sum_{k=1}^{m} (e_k \mid x) e_k$ et $b = x - a$.
On a évidemment $a \in F$ et $x = a + b$. Il reste à vérifier $b \in F^\perp$.
$F = \text{Vect}(e_1, \ldots, e_m)$ et $(e_k \mid b) = (e_k \mid x) - (e_k \mid a) = 0$ donc $b \in F^\perp$.
□

Corollaire

Si F est un sous-espace vectoriel d'un espace euclidien E alors

$$\dim F^{\perp} = \dim E - \dim F \text{ et } F = \left(F^{\perp}\right)^{\perp}$$

dém. :
$E = F \oplus F^{\perp}$ donne $\dim E = \dim F + \dim F^{\perp}$.
$F \subset \left(F^{\perp}\right)^{\perp}$ et l'égalité des dimensions donne $F = \left(F^{\perp}\right)^{\perp}$.
□

Exemple Dans $\mathcal{M}_n(\mathbb{R})$ muni du produit scalaire canonique, les sous-espaces vectoriels $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont supplémentaires orthogonaux.
En effet, Ceux-ci sont orthogonaux car pour $A \in \mathcal{S}_n(\mathbb{R})$ et $B \in \mathcal{A}_n(\mathbb{R})$

$$(A \mid B) = \text{tr}({}^tAB) = \text{tr}(AB)$$

et

$$(A \mid B) = (B \mid A) = \text{tr}({}^tBA) = -\text{tr}(BA) = \text{tr}(AB)$$

donc $(A \mid B) = 0$.
On en déduit

$$\mathcal{A}_n(\mathbb{R}) \subset \mathcal{S}_n(\mathbb{R})^{\perp}$$

puis, par égalité des dimensions,

$$\mathcal{A}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R})^{\perp} \text{ et aussi } \mathcal{A}_n(\mathbb{R})^{\perp} = \mathcal{S}_n(\mathbb{R})$$

7.3.5 Vecteur normal à un hyperplan en dimension finie

Soit H un hyperplan d'un espace euclidien E. Puisque $\dim H = \dim E - 1$, on obtient $\dim H^{\perp} = 1$.

Définition

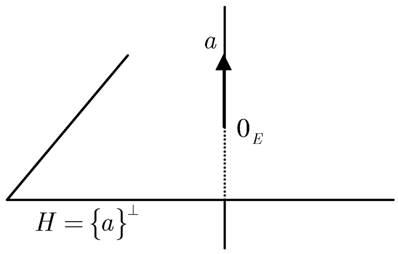
La droite $H^{\perp}$ est appelée droite normale à l'hyperplan H.

Pour tout $a \in H^{\perp}$ avec $a \neq 0_E$, on a

$$H = \left(H^{\perp}\right)^{\perp} = \text{Vect}(a)^{\perp} = \{a\}^{\perp}$$

et ainsi

$$\forall x \in E, x \in H \Leftrightarrow (a \mid x) = 0$$



Définition

Tout vecteur a non nul de $H^{\perp}$ est appelée vecteur normal de l'hyperplan H.

Exemple Considérons $E = \mathcal{M}_n(\mathbb{R})$ et $H = \{M \in \mathcal{M}_n(\mathbb{R})/\text{tr}(M) = 0\}$.
Déterminons un vecteur normal de H.
H est un hyperplan car noyau de la forme linéaire non nulle trace.
Puisque $\text{tr}(M) = \text{tr}({}^tI_nM) = (I_n \mid M)$, la matrice I_n est vecteur normal à H.

7.4 Projection orthogonale sur un sous-espace vectoriel de dimension finie

E désigne un espace préhilbertien réel de produit scalaire $(. \mid .)$.

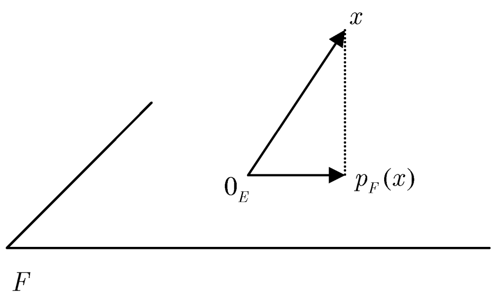
7.4.1 Projection orthogonale

Soit F un sous-espace vectoriel de dimension finie d'un espace préhilbertien E. On a

$$E = F \oplus F^{\perp}$$

Définition

On l'appelle projection orthogonale sur F la projection p_F sur F parallèlement à $F^{\perp}$.
On appelle symétrie orthogonale par rapport à F la symétrie s_F par rapport à F et parallèlement à $F^{\perp}$.



Exemple Si $F = \{0_E\}$ alors $p_F = \tilde{0}$.
Si $F = E$ alors $p_F = \text{Id}_E$.

Proposition

$p_F^2 = p_F$, $\text{Sp}(p_F) \subset \{0, 1\}$
$\ker(p_F - \text{Id}) = F = \text{Im}p_F$ et $\ker p_F = F^{\perp}$
De plus, $s_F = 2p_F - \text{Id}_E$ et $\text{Id} - p_F = p_{F^{\perp}}$.

dém. :
Ce sont les propriétés usuelles des projections qui sont ici particularisées.
□

Exemple Soit p un projecteur de E euclidien.
Montrer que p est une projection orthogonale si, et seulement si,

$$\forall x \in E, \|p(x)\| \leqslant \|x\|$$

($\Rightarrow$) Si p est la projection orthogonal sur F alors

$$x = p(x) + (x - p(x)) \text{ avec } p(x) \in F \text{ et } x - p(x) \in F^{\perp}$$

Par Pythagore

$$\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2 \geqslant \|p(x)\|^2$$

($\Leftarrow$) Si p est une projection sur un sous-espace vectoriel F parallèlement à un sous-espace vectoriel G, pour montrer que p est une projection orthogonale, il suffit de constater

$$\forall (a, b) \in F \times G, (a \mid b) = 0$$

Supposons

$$\forall x \in E, \|p(x)\| \leqslant \|x\|$$

Soit $\lambda \in \mathbb{R}$ et $x = a + \lambda b$. On a $p(x) = a$ et l'inégalité $\|p(x)\| \leqslant \|x\|$ fournit

$$\forall \lambda \in \mathbb{R}, 2\lambda\,(a \mid b) + \lambda^2 \|b\|^2 \geqslant 0$$

Si $(a \mid b) \neq 0$ alors

$$2\lambda\,(a \mid b) + \lambda^2 \|b\|^2 \underset{\lambda \to 0}{\sim} 2\lambda\,(a \mid b)$$

n'est pas de signe constant au voisinage de 0.
Nécessairement, $(a \mid b) = 0$.

7.4.2 Expression du projeté orthogonal

Théorème

Si $(e_1, \ldots, e_m)$ est une base orthonormale du sous-espace vectoriel F alors

$$\forall x \in E, p_F(x) = \sum_{k=1}^{m} (e_k \mid x)\, e_k$$

dém. :
Le vecteur $p_F(x)$ est élément de F. On peut donc écrire

$$p_F(x) = \sum_{k=1}^{m} (e_k \mid p_F(x))\, e_k$$

Pour tout $k \in \{1, \ldots, m\}$, $(e_k \mid x - p_F(x)) = 0$ car $x - p_F(x) \in F^{\perp}$ donc

$$(e_k \mid p_F(x)) = (e_k \mid x)$$

□

Exemple Soit $a \neq 0_E$ et $D = \text{Vect}(a)$.
$(a/\|a\|)$ forme une base orthonormale de D donc

$$\forall x \in E, p_D(x) = \frac{(a \mid x)}{\|a\|^2} a$$

Exemple Soit H hyperplan de vecteur normal a.
On a $H = \{a\}^{\perp} = D^{\perp}$ avec $D = \text{Vect}(a)$ et donc $p_H = \text{Id} - p_D$. Ainsi

$$\forall x \in E, p_H(x) = x - \frac{(a \mid x)}{\|a\|^2} a$$

Remarque Lors de la mise en place du procédé d'orthonormalisation de Schmidt d'une famille libre $(x_1, \ldots, x_n)$, le calcul

$$e_{k+1} = u/\|u\| \text{ avec } u = x_k - \sum_{i=1}^{k} (e_i \mid x_k) e_i$$

s'interprète comme l'obtention du vecteur complémentaire au projeté orthogonal.

7.4.3 Distance à un sous-espace vectoriel

Soit F un sous-espace vectoriel de E tel que F et $F^{\perp}$ sont supplémentaires.

Théorème

Soit $x \in E$.

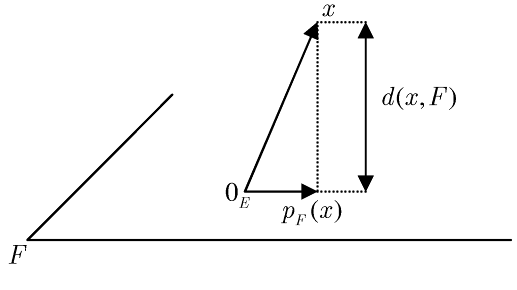
$$\forall y \in F, \|x - y\| \geqslant \|x - p_F(x)\|$$

avec égalité si, et seulement si, $y = p(x)$.

dém. :
$x - y = (x - p_F(x)) + (p_F(x) - y)$ avec $x - p_F(x) \in F^{\perp}$ et $p_F(x) - y \in F$.
Par Pythagore $\|x - y\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2 \geqslant \|x - p_F(x)\|^2$ avec égalité si, et seulement si, $y = p_F(x)$.
□

Corollaire

$d(x, F) = \|x - p_F(x)\|$.



dém. :
$d(x,F) = \inf_{y\in F} \|x-y\| = \min_{y\in F} \|x-y\| = \|x - p_F(x)\|$.
□

Corollaire

Soit $a \neq 0_E$ et $D = \text{Vect}(a)$.

$$\forall x \in E, d(x,D) = \left\| x - \frac{(a \mid x)}{\|a\|^2} a \right\|$$

$H = \text{Vect}(a)^{\perp}$.

$$\forall x \in E, d(x,H) = \frac{|(a \mid x)|}{\|a\|}$$

Exemple Soit $E = \mathcal{M}_2(\mathbb{R})$.
Calculons la distance de $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ à l'hyperplan H constitué des matrices de trace nulle.
Puisque I_2 est vecteur normal de H,

$$d(A,H) = \frac{|\text{tr}(A)|}{\|I_2\|} = \frac{5}{\sqrt{2}}$$

Exemple Calcul de

$$m = \inf_{(a,b)\in\mathbb{R}^2} \int_0^1 \left(t^2 - (at+b)\right)^2 \mathrm{d}t$$

Considérons $E = \mathbb{R}[X]$ muni du produit scalaire

$$(P,Q) \mapsto \int_0^1 P(t)Q(t)\,\mathrm{d}t$$

On a $m = d(X^2, \mathbb{R}_1[X])^2$.
Soit $p = p_{\mathbb{R}_1[X]}$. On a $m = \left\|X^2 - p(X^2)\right\|^2$.
Déterminons $p(X^2)$.
Pour cela formons une base orthonormée de $\mathbb{R}_1[X]$.
L'algorithme d'orthonormalisation de Schmidt donne la base orthonormée

$$P_1 = 1 \text{ et } P_2 = 2\sqrt{3}\left(X - \frac{1}{2}\right)$$

On en déduit

$$p(X^2) = \left(P_1 \mid X^2\right) P_1 + \left(P_2 \mid X^2\right) P_2 = X - 1/6$$

Après calculs

$$m = \left\|X^2 - X + 1/6\right\|^2 = 1/180$$

7.4.4 Inégalité de Bessel

Théorème

Si $(e_1, \ldots, e_n)$ est une famille orthonormale de vecteurs de E alors

$$\forall x \in E, \sum_{k=1}^{n} (e_k \mid x)^2 \leqslant \|x\|^2$$

dém. :

Soit $(e_1, \ldots, e_n)$ une famille orthonormale. Celle-ci est base orthonormale de l'espace $F = \text{Vect}(e_1, \ldots, e_n)$ et

$$p_F(x) = \sum_{k=1}^{n} (e_k \mid x) \, e_k$$

On a alors

$$\|p_F(x)\|^2 = \sum_{k=1}^{n} (e_k \mid x)^2$$

et la relation $\|p_F(x)\|^2 \leqslant \|x\|^2$ donne celle proposée.

□

Remarque Si $\dim E < +\infty$ et si $(e_1, \ldots, e_n)$ est une base orthonormale alors il y a égalité.
Si $\dim E = +\infty$ et si $(e_n)_{n \in \mathbb{N}}$ est une famille orthonormée de vecteurs de E alors pour tout $x \in E$, la série numérique $\sum (e_n \mid x)^2$ converge et

$$\sum_{n=0}^{+\infty} (e_n \mid x)^2 \leqslant \|x\|^2$$

En effet, les sommes partielles de la série à termes positifs $\sum |(e_n \mid x)|^2$ sont majorées par $\|x\|^2$.

7.4.5 Suite orthonormale de vecteurs d'un espace préhilbertien réel

Ici, E désigne un espace préhilbertien de dimension infinie.

Définition

On dit qu'une suite $(e_n)_{n \in \mathbb{N}}$ de vecteurs de E est totale si l'espace vectoriel qu'elle engendre est une partie dense de E i.e.

$$\overline{\text{Vect}\,\{e_n / n \in \mathbb{N}\}} = E$$

Exemple Soit $E = \mathcal{C}([-1, 1], \mathbb{R})$ muni du produit scalaire

$$\langle f, g \rangle = \int_{-1}^{1} f(t) g(t) \, \mathrm{d}t$$

La suite $(X^n)_{n \in \mathbb{N}}$ est totale (ou abusivement X^n désigne la fonction polynomiale $t \mapsto t^n$ définie sur $[-1, 1]$).

En effet, par le théorème de Weierstrass

$$\text{Vect}\,\{X^n/n\in\mathbb{N}\}=\mathbb{R}\,[X]$$

est une partie dense de E normé par $\|\,.\,\|_\infty$ donc, a fortiori, une partie dense de E normée par $\|\,.\,\|_2$ puisque

$$\|f\|_2\leqslant\sqrt{2}\,\|f\|_\infty$$

Théorème

Soit $(e_n)_{n\in\mathbb{N}}$ une suite orthonormale totale d'éléments de E.
En notant p_n la projection orthogonale sur l'espace $F_n=\text{Vect}(e_0,\dots,e_n)$ on a

$$\forall x\in E, p_n(x)\xrightarrow[n\to+\infty]{}x$$

dém. :
Commençons par remarquer

$$\text{Vect}\,\{e_n/n\in\mathbb{N}\}=\bigcup_{n\in\mathbb{N}}F_n$$

L'inclusion $(\supset)$ est immédiate. L'inclusion $(\subset)$ provient de ce que la réunion des F_n est un sous-espace vectoriel de E contenant tous les vecteurs e_n.
Soit $x\in E$.
Soit $\varepsilon>0$. Puisque $\text{Vect}\,\{e_n/n\in\mathbb{N}\}$ est une partie dense de E, il existe $y\in\text{Vect}\,\{e_n/n\in\mathbb{N}\}$ vérifiant $\|x-y\|\leqslant\varepsilon$. Par la remarque précédente, il existe $N\in\mathbb{N}$ tel que $y\in F_N$. Pour tout $n\geqslant N$, on a aussi $y\in F_n$ et donc

$$\|x-p_n(x)\|=d(x,F_n)\leqslant\|x-y\|\leqslant\varepsilon$$

□

Corollaire

Si $(e_n)_{n\in\mathbb{N}}$ est une suite orthonormale totale d'éléments de E alors

$$\forall x\in E, x=\sum_{n=0}^{+\infty}(e_n\mid x)\,e_n$$

dém. :
Il suffit d'exprimer $p_n(x)$ et d'observer

$$p_n(x)=\sum_{k=0}^{n}(e_k\mid x)\,e_k\xrightarrow[n\to+\infty]{}x$$

□

7.4.6 Musculations

7.4.6.1 Polynôme de Legendre

Soit $E = \mathcal{C}([-1,1], \mathbb{R})$ muni du produit scalaire

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$$

En orthonormalisant par l'algorithme de Schmidt, la famille $(X^n)_{n\in\mathbb{N}}$ on obtient une famille orthonormale totale, mais celle-ci est difficile à calculer...
Considérons

$$P_n = \left(\left(X^2-1\right)^n\right)^{(n)} = U_n^{(n)} \text{ avec } U_n = (X-1)^n (X+1)^n$$

Exemple $P_0 = 1$, $P_1 = 2X$, $P_2 = 4\left(3X^2-1\right)$

Proposition

$\deg P_n = n$ et $\forall Q \in \mathbb{R}_{n-1}[X], (P_n \mid Q) = 0$

dém. :
$\deg P_n = n$ car $\deg\left(X^2-1\right)^n = 2n$ et l'on dérive n fois
Par intégration par parties successives, on obtient

$$\forall Q \in \mathbb{R}_{n-1}[X], (P_n \mid Q) = (-1)\left(U_n^{(n-1)} \mid Q'\right) = \ldots = (-1)^n \left(U_n \mid Q^{(n)}\right) = 0$$

□

Théorème

La famille $(P_n/\|P_n\|)_{n\in\mathbb{N}}$ est une famille orthonormale totale de E et donc

$$f = \sum_{n=0}^{+\infty} \frac{(P_n \mid f)}{\|P_n\|^2} P_n$$

dém. :
La famille $(P_n)_{n\in\mathbb{N}}$ est orthogonale car $\forall m < n, (P_n \mid P_m) = 0$ en vertu de ce qui précède.
De plus, étant de degrés étagés, elle constitue une base de $\mathbb{R}[X]$ et c'est donc une famille totale comme cela a été vu au dessus.
□

Remarque La fonction polynôme

$$f_N = \sum_{n=0}^{N} \frac{(P_n \mid f)}{\|P_n\|^2} P_n$$

constitue alors la meilleure approximation euclidienne de f parmi les polynômes de degré inférieur à N.

7.4.6.2 Polynômes de Tchebychev

On a

$$\cos(2t) = 2\cos^2 t - 1,\ \cos(3t) = 4\cos^3 t - 3\cos t, \ldots$$

De façon générale, pour $n \in \mathbb{N}$, en développant

$$\cos(nt) = \text{Re}\left(e^{int}\right) = \text{Re}\left((\cos t + i\sin t)^n\right)$$

on obtient

$$\cos(nt) = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n}{2k} \cos^{n-2k}(t)\sin^{2k}(t)$$

et puisque $\sin^{2k}(t) = (1-\cos^2 t)^k$, cette expression est un polynôme en $\cos(t)$.

Définition

On appelle polynôme de Tchebychev, l'unique polynôme de $\mathbb{R}[X]$ vérifiant

$$\forall t \in \mathbb{R}, \cos(nt) = T_n(\cos t)$$

Exemple $T_0 = 1, T_1 = X, T_2 = 2X^2 - 1$ et $T_3 = 4X^3 - 3X$

En vertu des calculs qui précèdent

$$T_n(X) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} X^{n-2k}(X^2-1)^k$$

Proposition

$\forall n \in \mathbb{N}, T_{n+1} = 2XT_n - T_{n-1}$

dém. :

On a

$$\cos\left((n+1)t\right) + \cos\left((n-1)t\right) = 2\cos(t)\cos(nt)$$

donc

$$T_{n+1}(\cos t) = 2\cos(t)T_n(\cos t) - T_{n-1}(\cos t)$$

L'identité

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

étant vraie pour une infinité de valeurs (celles de $[-1,1]$) on peut affirmer l'identité polynomiale proposée.

□

Théorème

La famille $(T_n)_{n\in\mathbb{N}}$ est une famille orthogonale totale sur l'espace $E = \mathcal{C}\left([-1,1], \mathbb{R}\right)$ muni du produit scalaire

$$\langle f, g\rangle = \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}\,\mathrm{d}x$$

dém. :

On vérifie aisément que $\langle .,. \rangle$ définit un produit scalaire sur E.

La famille $(T_n)_{n\in\mathbb{N}}$ est une famille de polynômes de degrés étagés, c'est donc une base de $\mathbb{R}[X]$.
Par le théorème de Weierstrass et la comparaison

$$\|f\|_2^2 = \int_{-1}^{1} \frac{f(x)^2}{\sqrt{1-x^2}}\,\mathrm{d}x \leqslant \int_{-1}^{1} \frac{\|f\|_\infty^2}{\sqrt{1-x^2}}\,\mathrm{d}x = \pi\,\|f\|_\infty^2$$

on peut affirmer que cette famille est totale.
Enfin cette famille est orthogonale car pour $n \neq m$

$$\langle T_n \mid T_m\rangle = \int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} \underset{x=\cos t}{=} \int_0^\pi \cos(nt)\cos(mt)\,\mathrm{d}t = 0$$

On peut donc écrire dans l'espace préhilbertien E

$$f = \sum_{n=0}^{+\infty} \frac{\langle T_n, f\rangle}{\|T_n\|^2} T_n$$

□

7.4.6.3 Séries de Fourier

Soit E l'espace des fonctions réelles continues T-périodiques.
On définit un produit scalaire sur E en posant

$$(f \mid g) = \frac{1}{T}\int_0^T f(x)g(x)\,\mathrm{d}x$$

On définit les familles de fonctions $(c_n)_{n\in\mathbb{N}}$ et $(s_n)_{n\in\mathbb{N}^\star}$ par

$$c_0(x) = 1,\ c_n(x) = \cos(2\pi nx/T) \text{ et } s_n(x) = \sin(2\pi nx/T) \text{ pour } n \in \mathbb{N}^\star$$

Ces fonctions sont deux à deux orthogonales car

$$\forall n \neq m, (c_n \mid c_m) = \frac{1}{2T}\int_0^T \cos\left(2\pi(n+m)x/T\right) + \cos\left(2\pi(n-m)x/T\right)\,\mathrm{d}x = 0$$

et de façon analogue

$$\forall n \neq m, (s_n \mid s_m) = 0 \text{ et } \forall n, m, (c_n \mid s_m) = 0$$

On peut montrer (mais ce n'est pas immédiat) que la famille constituée de ces fonctions est une famille totale. On peut alors écrire dans l'espace préhilbertien E

$$f = \sum_{n=0}^{+\infty} \frac{(c_n \mid f)}{\|c_n\|^2} c_n + \sum_{n=1}^{+\infty} \frac{(s_n \mid f)}{\|s_n\|^2} s_n$$

On obtient ainsi l'écriture utilisée en sciences physiques

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos\left(\frac{2\pi nx}{T}\right) + b_n \sin\left(\frac{2\pi nx}{T}\right)$$

avec

$$a_n = \frac{2}{T}\int_0^T f(x)\cos\left(\frac{2\pi nx}{T}\right)\,\mathrm{d}x \text{ et } b_n = \frac{2}{T}\int_0^T f(x)\sin\left(\frac{2\pi nx}{T}\right)\,\mathrm{d}x$$

Chapitre 8

Endomorphismes des espaces euclidiens

E désigne un espace vectoriel euclidien de dimension $n \in \mathbb{N}^\star$.

8.1 Matrices orthogonales

8.1.1 Définition

Proposition

> Pour $A \in \mathcal{M}_n(\mathbb{R})$, on a équivalence entre :
> (i) A est inversible et $A^{-1} = {}^tA$;
> (ii) ${}^tAA = I_n$;
> (iii) $A^tA = I_n$.

dém. :
Il suffit d'appliquer le théorème d'inversibilité relatif aux matrices.
□

Définition

> On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si ${}^tAA = I_n$.

Exemple I_n et $-I_n$ sont des matrices orthogonales.

Théorème

> L'ensemble $\mathcal{O}_n(\mathbb{R})$ des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$ est un sous-groupe compact de $(\mathrm{GL}_n(\mathbb{R}), \times)$ appelé groupe orthogonal d'ordre n.

dém. :
$\mathcal{O}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$, $I_n \in \mathcal{O}_n(\mathbb{R})$.
Soit $A, B \in \mathcal{O}_n(\mathbb{R})$. $AB \in \mathcal{O}_n(\mathbb{R})$ car ${}^t(AB)\,AB = {}^tB^tAAB = {}^tBB = I_n$.
Soit $A \in \mathcal{O}_n(\mathbb{R})$. $A^{-1} \in \mathcal{O}_n(\mathbb{R})$ car ${}^t\left(A^{-1}\right)A^{-1} = {}^t\left({}^tA\right){}^tA = A^tA = I_n$.
Ainsi $\mathcal{O}_n(\mathbb{R})$ est un sous-groupe de $(\mathrm{GL}_n(\mathbb{R}), \times)$.
$\mathcal{O}_n(\mathbb{R}) = \left\{A \in \mathcal{M}_n(\mathbb{R}) / {}^tAA = I_n\right\} = f^{-1}(\{I_n\})$ avec $f : A \in \mathcal{M}_n(\mathbb{R}) \to {}^tAA$.
Puisque f est continue et $\{I_n\}$ fermé, $\mathcal{O}_n(\mathbb{R})$ est un fermé relatif à $\mathcal{M}_n(\mathbb{R})$ et c'est donc une partie fermée.
Enfin, considérons $\|\,.\,\|$ la norme euclidienne associée au produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$.

Pour $A \in \mathcal{O}_n(\mathbb{R})$, $\|A\| = \sqrt{\mathrm{tr}({}^tAA)} = \sqrt{\mathrm{tr}I_n} = \sqrt{n}$. Par suite $\mathcal{O}_n(\mathbb{R})$ est une partie bornée.
Puisque $\mathcal{M}_n(\mathbb{R})$ est de dimension finie, $\mathcal{O}_n(\mathbb{R})$ est une partie compacte car fermée et bornée.
□

Théorème

Soit $A \in \mathcal{M}_n(\mathbb{R})$ de colonnes $C_1, \dots, C_n$ et de lignes $L_1, \dots, L_n$.
On a équivalence entre :
(i) la matrice A est orthogonale ;
(ii) la famille $(C_1, \dots, C_n)$ est orthonormée ;
(iii) la famille $(L_1, \dots, L_n)$ est orthonormée.

dém. :
Etudions (i) $\Leftrightarrow$ (ii).
Sur $\mathcal{M}_{n,1}(\mathbb{R})$, le produit scalaire considéré est le produit scalaire canonique défini par

$$(X \mid Y) = {}^tXY = x_1y_1 + \cdots + x_ny_n$$

Pour tout $1 \leqslant i, j \leqslant n$,

$$\left[{}^tAA\right]_{i,j} = \sum_{k=1}^{n} \left[{}^tA\right]_{i,k} A_{k,j} = \sum_{k=1}^{n} a_{k,i}a_{k,j} = (C_i \mid C_j)$$

(i) $\Leftrightarrow {}^tAA = I_n \Leftrightarrow \forall 1 \leqslant i, j \leqslant n, \left[{}^tAA\right]_{i,j} = \delta_{i,j} = \forall 1 \leqslant i, j \leqslant n, (C_i \mid C_j) = \delta_{i,j} \Leftrightarrow$ (ii)
Etudions (i) $\Leftrightarrow$ (iii).
Sur $\mathcal{M}_{1,n}(\mathbb{R})$, le produit scalaire considéré est le produit scalaire canonique définie par

$$(L \mid L') = L^tL' = \ell_1\ell'_1 + \cdots + \ell_n\ell'_n$$

En remarquant que

$$\left[A^tA\right]_{i,j} = (L_j \mid L_i)$$

on démontre comme ci-dessus (i) $\Leftrightarrow$ (iii).
□

Exemple La matrice

$$A = \frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{pmatrix}$$

est orthogonale.
En effet, ses colonnes sont unitaires et deux à deux orthogonales.

8.1.2 Changement de bases orthonormales

Théorème

Soit $e = (e_1, \dots, e_n)$ une base orthonormale de E et $e' = (e'_1, \dots, e'_n)$ une famille de vecteurs de E.
On a équivalence entre :
(i) e' est orthonormale ;
(ii) $P = \mathrm{Mat}_e e'$ est orthogonale.
De plus, si tel est le cas,

$$\mathrm{Mat}_{e'} e = {}^tP$$

dém. :
Rappelons que si x et y sont des vecteurs de colonnes coordonnées X, Y dans une base orthonormale alors

$$(x \mid y) = {}^t XY$$

Notons $C_1, \dots, C_n$ les colonnes de P.
Les colonnes $C_1, \dots, C_n$ sont les colonnes des coordonnées des vecteurs $e'_1, \dots, e'_n$ dans la base orthonormale e et donc pour tout $1 \leqslant i, j \leqslant n$,

$$(e'_i \mid e'_j) = {}^t C_i C_j = (C_i \mid C_j)$$

Par suite, la famille e' est orthonormée si, et seulement si, la famille $(C_1, \dots, C_n)$ l'est. Cela équivaut à affirmer $P \in \mathcal{O}_n(\mathbb{R})$.
De plus, si tel est le cas, $\text{Mat}_{e'} e = P^{-1} = {}^t P$.
□

Corollaire

Si e et e' sont deux bases orthonormales de l'espace euclidien E alors la formule de changement de base relative aux endomorphismes s'écrit

$$A' = {}^t PAP$$

avec $A = \text{Mat}_e u$, $A' = \text{Mat}_{e'} u$ et $u \in \mathcal{L}(E)$.

Définition

On dit alors que les matrices A et A' sont orthogonalement semblables.

Remarque Deux matrices orthogonalement semblables sont a fortiori semblables.

8.1.3 Matrices orthogonales positives

Proposition

Si A est une matrice orthogonale alors $\det A = \pm 1$.

dém. :
${}^t AA = I_n$ donne $\det({}^t AA) = 1$ or $\det({}^t AA) = \det({}^t A) \det A = (\det A)^2$ donc $(\det A)^2 = 1$.
□

Définition

Les matrices orthogonales de déterminant 1 sont dite positives, les autres sont dites négatives.

Exemple I_n est une matrice orthogonale positive.
$-I_n$ est une matrice orthogonale positive si, et seulement si, n est pair.

Proposition

L'ensemble $\text{SO}_n(\mathbb{R})$ des matrices orthogonales positives de $\mathcal{M}_n(\mathbb{R})$ est un sous-groupe compact de $(\text{GL}_n(\mathbb{R}), \times)$.
On l'appelle groupe spécial orthogonal d'ordre n.

dém. :
$\text{SO}_n(\mathbb{R}) = \mathcal{O}_n(\mathbb{R}) \cap \text{SL}_n(\mathbb{R})$

avec $\mathcal{O}_n(\mathbb{R})$ sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ et
$\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) / \det A = 1\}$ sous-groupe fermé de $(\mathrm{GL}_n(\mathbb{R}), \times)$.
□

Proposition

Si e et e' sont deux bases orthonormées directes d'un espace euclidien orienté alors $\det_e e' = 1$.

dém. :
Puisque les bases e et e' ont même orientation $\det_e e' > 0$. Or $\det_e e' = \pm 1$ car $\mathrm{Mat}_e e' \in \mathcal{O}_n(\mathbb{R})$. On en déduit $\det_e e' = 1$
□

Remarque C'est cette relation qui permet de définir le produit mixte de $n = \dim E$ vecteurs d'un espace euclidien orienté comme égal au déterminant de cette famille dans n'importe quelle base orthonormale directe.

8.2 Isométries vectorielles

8.2.1 Définition

Définition

On appelle isométrie vectorielle de E tout endomorphisme $u \in \mathcal{L}(E)$ conservant la norme.

$$\forall x \in E, \|u(x)\| = \|x\|$$

Exemple $\mathrm{Id}_E, -\mathrm{Id}_E$ sont des isométries vectorielles.

Exemple Les symétries orthogonales sont des isométries vectorielles.
En effet, si s est une symétrie orthogonale par rapport à un sous-espace vectoriel F, pour $x = a + b$ avec $a \in F$ et $b \in F^\perp$ alors $s(x) = a - b$ et par le théorème de Pythagore

$$\|s(x)\|^2 = \|a\|^2 + \|b\|^2 = \|x\|^2$$

Proposition

Si u est une isométrie vectorielle alors $\mathrm{Sp}u \subset \{1, -1\}$.

dém. :
Soit $\lambda \in \mathrm{Sp}u$ et $x \neq 0_E$ vecteur propre associé.
D'une part $\|u(x)\| = \|\lambda x\| = |\lambda| \|x\|$, d'autre part $\|u(x)\| = \|x\|$. On en déduit $|\lambda| = 1$
□

Remarque En particulier $0 \notin \mathrm{Sp}u$ et donc u est un automorphisme.
On parle indifféremment d'automorphisme orthogonal ou d'isométrie vectorielle.

Théorème

Soit u un endomorphisme de E. On a équivalence entre :
(i) u est orthogonal ;
(ii) u conserve le produit scalaire i.e.

$$\forall x, y \in E, (u(x) \mid u(y)) = (x \mid y)$$

dém. :
(i) $\Rightarrow$ (ii) Supposons que pour tout $x \in E$, $\|u(x)\| = \|x\|$.
D'une part

$$\|u(x+y)\|^2 = \|u(x) + u(y)\|^2 = \|u(x)\|^2 + 2(u(x) \mid u(y)) + \|u(y)\|^2$$

et d'autre part

$$\|u(x+y)\|^2 = \|x+y\|^2 = \|x\|^2 + 2(x \mid y) + \|y\|^2$$

Or $\|u(x)\| = \|x\|$ et $\|u(y)\| = \|y\|$ donc

$$(u(x) \mid u(y)) = (x \mid y)$$

(ii) $\Rightarrow$ (i) Supposons que l'endomorphisme u conserve le produit scalaire.
Pour tout $x \in E$,

$$\|u(x)\|^2 = (u(x) \mid u(x)) = (x \mid x) = \|x\|^2$$

donc $\|u(x)\| = \|x\|$.
□

8.2.2 Matrice d'une isométrie en base orthonormale

Théorème

Soit $u \in \mathcal{L}(E)$ et $e = (e_1, \ldots, e_n)$ une base orthonormale de E.
On a équivalence entre :
(i) u est orthogonal ;
(ii) la famille $(u(e_1), \ldots, u(e_n))$ est une base orthonormale ;
(iii) $\text{Mat}_e u \in \mathcal{O}_n(\mathbb{R})$.

dém. :
(i) $\Rightarrow$ (ii) Supposons l'endomorphisme u orthogonal.
Pour tout $1 \leqslant i, j \leqslant n$,

$$(u(e_i) \mid u(e_j)) = (e_i \mid e_j) = \delta_{i,j}$$

donc la famille $(u(e_1), \ldots, u(e_n))$ est orthonormale et c'est donc une base orthonormée.
(ii) $\Rightarrow$ (iii) Supposons $(u(e_1), \ldots, u(e_n))$ orthonormale
Puisque $\text{Mat}_e u = \text{Mat}_e\,(u(e_1), \ldots, u(e_n))$, $\text{Mat}_e(u) \in \mathcal{O}_n(\mathbb{R})$ car matrice de passage entre deux bases orthonormales.
(iii) $\Rightarrow$ (i) Supposons $A = \text{Mat}_e u \in \mathcal{O}_n(\mathbb{R})$.
Soit x un vecteur de E de colonne coordonnées X dans la base e.
Puisque la base e est orthonormale $\|x\|^2 = {}^tXX$.
Puisque $u(x)$ a pour colonne coordonnées AX,

$$\|u(x)\|^2 = {}^t(AX)AX = {}^tX{}^tAAX = {}^tXX = \|x\|^2$$

Ainsi u conserve la norme et donc est une isométrie vectorielle.
□

Remarque Il est essentiel de vérifier que la base e est orthonormale pour exploiter ce résultat.

Corollaire

> L'ensemble $\mathcal{O}(E)$ des isométries vectorielles de E est un sous-groupe compact de $(\mathrm{GL}(E), \circ)$ appelé groupe orthogonal de E.

dém. :
Considérons e une base orthonormée de E et $\Phi : \mathcal{M}_n(\mathbb{R}) \to \mathcal{L}(E)$ l'application qui à $M \in \mathcal{M}_n(\mathbb{R})$ associe $u \in \mathcal{L}(E)$ déterminé par $\mathrm{Mat}_e u = M$. On a

$$\Phi(\mathcal{O}_n(\mathbb{R})) = \mathcal{O}(E)$$

Φ est continue (car linéaire au départ d'un espace de dimension finie) donc $\mathcal{O}(E)$ est compact.
Φ est un morphisme de groupe multiplicatif donc $\mathcal{O}(E)$ est un sous-groupe de $(\mathrm{GL}(E), \circ)$.
□

8.2.3 Isométries positives

Remarque Si $u \in \mathcal{O}(E)$ alors $\det u = \pm 1$.

Définition

> On appelle isométrie positive (ou isométrie directe) toute isométrie vectorielle de déterminant 1. On parle d'isométrie négative (ou indirecte) sinon.

Exemple Id_E est une isométrie positive
$-\mathrm{Id}_E$ est une isométrie positive si, et seulement si, $\dim E$ est pair.

Exemple On appelle réflexion toute symétrie orthogonale par rapport à un hyperplan.
Les réflexions sont des isométries négatives.

Proposition

> L'ensemble $\mathrm{SO}(E)$ des isométries positives de E est un sous-groupe compact de $(\mathrm{GL}(E), \circ)$ appelé groupe spécial orthogonal de E.

dém. :
$\mathrm{SO}(E) = \mathcal{O}(E) \cap \mathrm{SL}(E)$ avec $\mathrm{SL}(E) = \{u \in \mathcal{L}(E) / \det u = 1\}$ sous-groupe fermé.
□

8.2.4 Isométries du plan

Soit E un plan euclidien orienté.

8.2.4.1 Isométries positives

Théorème

Les matrices orthogonales positives de $\mathcal{M}_2(\mathbb{R})$ sont les matrices de la forme

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ avec } \theta \in \mathbb{R}.$$

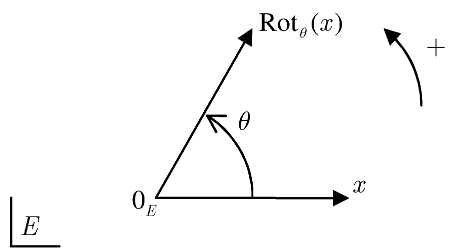
De plus, ces dernières commutent entre elles car

$$R(\theta)R(\theta') = R(\theta + \theta')$$

dém. :
Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R})$. On a $a^2 + c^2 = 1$ donc il existe $\theta \in \mathbb{R}$ vérifiant $a = \cos\theta$ et $b = \sin\theta$.
Puisque $(a-d)^2 + (b+c)^2 = 2 - 2(ad - bc) = 0$, on a aussi $c = -\sin\theta$ et $d = \cos\theta$.
Enfin, on vérifie par le calcul la relation $R(\theta)R(\theta') = R(\theta + \theta')$.
□

Corollaire

Une isométrie positive du plan a la même matrice dans toute base orthonormale directe.
Celle-ci est de la forme $R(\theta)$ avec $\theta \in \mathbb{R}$ unique à 2π près de sorte et on parle alors de rotation d'angle θ.



dém. :
Soit e et e' deux bases orthonormales du plan et $u \in \mathrm{SO}(E)$. On pose $A = \mathrm{Mat}_e u$ et $A' = \mathrm{Mat}_{e'} u$. Par formule de changement de base $A' = P^{-1}AP = AP^{-1}P = A$ car les matrices de $\mathrm{SO}_2(\mathbb{R})$ commutent entre elles.
□

8.2.4.2 Isométrie négatives

Théorème

Les matrices orthogonales négatives de $\mathcal{M}_2(\mathbb{R})$ sont les matrices de la forme

$$S(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \text{ avec } \theta \in \mathbb{R}.$$

Elles vérifient $(S(\theta))^2 = I_2$.

dém. :

Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R})$. On a $a^2 + c^2 = 1$ donc il existe $\theta \in \mathbb{R}$ vérifiant $a = \cos\theta$ et $b = \sin\theta$.
Puisque $(a + d)^2 + (b - c)^2 = 2 + 2(ad - bc) = 0$, on a aussi $c = \sin\theta$ et $d = -\cos\theta$.
Enfin, on vérifie par le calcul la relation $(S(\theta))^2 = I_2$.
□

Corollaire

Les isométries négatives du plan sont les symétries orthogonales par rapport à des droites.
Il existe une base orthonormale dans laquelle la symétrie est représentée par la matrice

$$S(0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

dém. :
On a

$$S(\theta) = S(0)R(\theta) = S(0)R(\theta/2)R(\theta/2) = R(-\theta/2)S(0)R(\theta/2)$$

donc $S(\theta)$ est semblable à $S(0)$ par le biais d'une matrice de passage orthogonale. Ainsi, une isométrie négative représentée initialement dans une base orthonormale par $S(\theta)$ peut aussi être représentée dans une base orthonormale par $S(0)$. On reconnaît alors une symétrie orthogonale.
□

8.2.5 Réduction d'une isométrie vectorielle

Lemme

Soit $u \in \mathcal{O}(E)$. Si F est un sous-espace vectoriel stable par u alors $F^\perp$ l'est aussi.

dém. :
On suppose F stable par u et donc $u(F) \subset F$. Or u est bijective donc conserve la dimension et par conséquent $u(F) = F$. Soit $x \in F^\perp$. Pour tout $y \in F$, on peut écrire $y = u(a)$ avec $a \in F$ et alors

$$(u(x) \mid y) = (u(x) \mid u(a)) = (x \mid a) = 0$$

Ainsi $u(x) \in F^\perp$.
□

Lemme

Si u est un endomorphisme d'un $\mathbb{R}$-espace vectoriel réel de dimension finie non nulle alors il existe au moins une droite vectorielle ou un plan stable par u.

dém. :
Soit $P \in \mathbb{R}[X]$ un polynôme unitaire annulateur de u (par exemple, son polynôme caractéristique ou minimal). On peut écrire $P = P_1 P_2 \ldots P_m$ avec P_k polynômes unitaires irréductibles de $\mathbb{R}[X]$.
Puisque $P(u) = \tilde{0}$, on a $P_1(u) \circ P_2(u) \circ \ldots \circ P_m(u) = \tilde{0}$ et par conséquent, au moins l'un des endomorphismes composés n'est pas injectif. Supposons que ce soit celui d'indice k. Le polynôme P_k est irréductible dans $\mathbb{R}[X]$, il est donc de l'une des deux formes suivantes :
Cas $P(X) = X - \lambda$
λ est alors valeur propre de u et tout vecteur propre associé engendre une droite vectorielle stable.
Cas $P(X) = X^2 + pX + q$ avec $\Delta = p^2 - 4q < 0$
Soit $x \in \ker P(u)$. On a $u^2(x) + pu(x) + qx = 0_E$ et donc $F = \mathrm{Vect}(x, u(x))$ est stable par u.
Dans les deux cas, u admet une droite ou un plan stable.
□

Théorème

Si $u \in \mathcal{O}(E)$ alors il existe une base orthonormale de E dans laquelle la matrice de u est diagonale par blocs de blocs diagonaux de la forme

$$(1), (-1) \text{ ou } \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ avec } \theta \in \mathbb{R}$$

Autrement dit, l'espace E est la somme directe orthogonale de $E_1(u)$, $E_{-1}(u)$ et de plans sur lesquels u opère comme une rotation.

dém. :
Par récurrence sur la dimension de E.
Cas $n = 1$:
u est une isométrie d'une droite et peut donc être représentée en base orthonormale par

$$(1) \text{ ou } (-1)$$

Cas $n = 2$:
u est une isométrie du plan et peut donc être représentée en base orthonormale par

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ ou } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Supposons la propriété établie jusqu'au rang n avec $n \geqslant 2$.
Soit E un espace euclidien de dimension $n + 1$ et $u \in \mathcal{O}(E)$.
Il existe une droite ou un plan F stable par u et $F^\perp$ est alors aussi stable par u.
Par hypothèse de récurrence, il existe une base orthonormale de $F^\perp$ telle que la matrice de u dans celle-ci soit de la forme voulue.
Par l'étude initiale, il existe une base orthonormale de F telle que la matrice de u dans celle-ci soit de la forme voulue.
En accolant ces deux, on forme une base orthonormale de E comme voulue.
Récurrence établie.
□

Corollaire

Toute matrice de $\mathcal{O}_n(\mathbb{R})$ est orthogonalement semblable à une matrice diagonale par blocs avec des blocs diagonaux de la forme

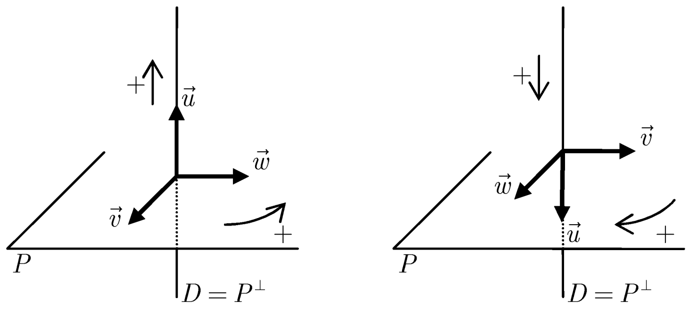
$$(1), (-1) \text{ ou } \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ avec } \theta \in \mathbb{R}$$

8.2.6 Réduction des isométries positives en dimension 3

Soit E un espace euclidien orienté de dimension 3.

8.2.6.1 Orientation induite

Soit P un plan de l'espace E et $D = P^\perp$ sa droite normale.
Il n'existe pas a priori d'orientation préférentielle ni sur P, ni sur D.
Choisissons une orientation sur D et soit $\vec{u}$ vecteur unitaire direct de D : on dit alors que D est un axe.
Complétons $\vec{u}$ en une base orthonormale directe $(\vec{u}, \vec{v}, \vec{w})$ de E.
La famille $(\vec{v}, \vec{w})$ est une base orthonormale de P. En choisissant celle-ci pour base orientée de référence, on dit qu'on a muni le plan P de l'orientation induite de celle de D. En effet, on peut montrer que cette orientation est indépendante de la manière dont on a complété u en une base orthonormée directe.



Remarque Si l'on inverse l'orientation sur D, l'orientation induite sur P est, elle aussi, inversée.

8.2.6.2 Rotation de l'espace

Une isométrie positive f de E autre que l'identité peut être représentée par la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

dans une base orthonormale $(\vec{u}, \vec{v}, \vec{w})$. Quitte à changer en son opposé le premier vecteur de base, on peut supposer la base orthonormale $(\vec{u}, \vec{v}, \vec{w})$ directe.
On introduit alors la droite $D = \text{Vect}(\vec{u})$ et le plan $P = \text{Vect}(\vec{v}, \vec{w})$ orienté par le vecteur normal $\vec{u}$. Pour $\vec{x} \in E$, on peut écrire

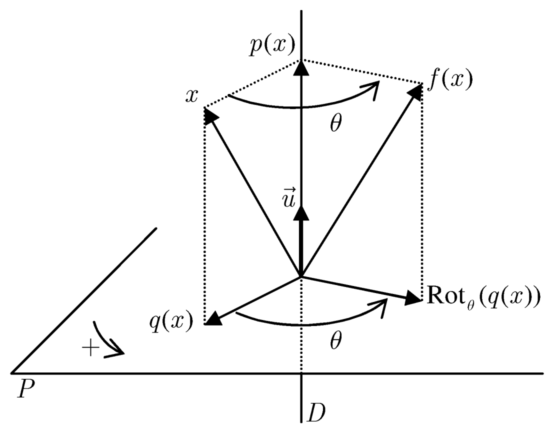
$$\vec{x} = p(\vec{x}) + q(\vec{x}) \text{ avec } p(\vec{x}) \in D \text{ et } q(\vec{x}) \in P$$

et alors

$$f(\vec{x}) = p(\vec{x}) + \text{Rot}_\theta(q(\vec{x}))$$

Définition

On dit alors que f est la rotation d'axe dirigé et orienté par $\vec{u}$ et d'angle θ. On la note $\text{Rot}_{\vec{u},\theta}$.



Proposition

$\forall \theta, \theta' \in \mathbb{R}$, $\text{Rot}_{u,\theta} = \text{Rot}_{u,\theta'} \Leftrightarrow \theta = \theta' \quad [2\pi]$
$\forall \theta, \theta' \in \mathbb{R}$, $\text{Rot}_{u,\theta} \circ \text{Rot}_{u,\theta'} = \text{Rot}_{u,\theta+\theta'} = \text{Rot}_{u,\theta'} \circ \text{Rot}_{u,\theta}$.
$\forall \theta, \theta' \in \mathbb{R}$, $\text{Rot}_{u,\theta}^{-1} = \text{Rot}_{u,-\theta}$.

dém. :
Immédiat par calcul matriciel.
□

Remarque Si l'on change le vecteur en son opposé, l'orientation induite sur P l'est aussi et les mesures angulaires dans P sont alors changées en leur opposée. Par suite

$$\text{Rot}_{u,\theta} = \text{Rot}_{-u,-\theta}$$

8.2.6.3 Réduction d'une rotation

Exemple Soit E un espace vectoriel euclidien muni d'une base orthonormée directe $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$. Déterminons l'endomorphisme f de E de matrice dans $\mathcal{B}$

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

La matrice A est orthogonale et $\det A = 1$ donc f est une rotation autre que l'identité.
Axe D :
L'axe D est formé des vecteurs invariants par f.
Pour $\vec{u} = x\vec{i} + y\vec{j} + z\vec{k}$, on a

$$f(\vec{u}) = \vec{u} \Leftrightarrow x = y = z$$

Par suite $D = \text{Vect}(\vec{i} + \vec{j} + \vec{k})$.
Orientons D par le vecteur $\vec{u} = \vec{i} + \vec{j} + \vec{k}$.
Angle θ de la rotation :
On a $\text{tr} f = 2\cos\theta + 1$ or $\text{tr} f = \text{tr} A = 0$ donc $\cos\theta = -1/2$.
Pour conclure, il reste à déterminer le signe de $\sin\theta$.
Soit $\vec{x} = \alpha\vec{u} + \beta\vec{v} + \gamma\vec{w} \notin D$. On a

$$[\vec{u}, \vec{x}, f(\vec{x})] = \begin{vmatrix} 1 & \alpha & \alpha \\ 0 & \beta & \beta\cos\theta - \gamma\sin\theta \\ 0 & \gamma & \beta\sin\theta + \gamma\cos\theta \end{vmatrix} = (\beta^2 + \gamma^2)\sin\theta$$

Ainsi, le signe de $\sin\theta$ est celui de

$$[\vec{u}, \vec{x}, f(\vec{x})]$$

En pratique, on détermine le signe de $\sin\theta$ en étudiant celui de

$$\left[\vec{u}, \vec{i}, f(\vec{i})\right]$$

Ici

$$\left[\vec{u}, \vec{i}, f(\vec{i})\right] = \begin{vmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 1 > 0$$

donc

$$\theta = 2\pi/3 \quad [2\pi]$$

Finalement, f est la rotation d'axe D dirigé et orienté par $\vec{u} = \vec{i} + \vec{j} + \vec{k}$ et d'angle $2\pi/3$.

8.3 Endomorphismes symétriques

8.3.1 Définition

Définition

Un endomorphisme $u \in \mathcal{L}(E)$ est dit symétrique si

$$\forall x, y \in E, (u(x) \mid y) = (x \mid u(y))$$

Exemple $\tilde{0}$ et Id sont symétriques.

Exemple Les projecteurs orthogonaux sont exactement les projecteurs symétriques.
En effet, soit p un projecteur orthogonal sur un sous-espace vectoriel F.
Pour tout $x, y \in E$,

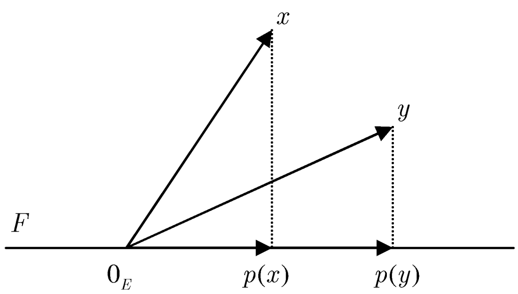
$$(p(x) \mid y) = (p(x) \mid p(y)) + (p(x) \mid y - p(y)) = (p(x) \mid p(y))$$

car $p(x) \in F$ et $y - p(y) \in F^{\perp}$. De même

$$(x \mid p(y)) = (p(x) \mid p(y)) + (x - p(x) \mid p(y)) = (p(x) \mid p(y))$$

Ainsi

$$(p(x) \mid y) = (x \mid p(y))$$



Inversement, si p est un projecteur sur un sous-espace vectoriel F parallèlement à un sous-espace vectoriel G et si celui-ci est symétrique alors pour tout $x \in F$ et $y \in G$ alors

$$(x \mid y) = (p(x) \mid y) = (x \mid p(y)) = (x \mid 0_E) = 0$$

Les espaces F et G sont donc orthogonaux et la projection p est orthogonale.
De même, les symétries orthogonales correspondent aux « symétries symétriques » .

Proposition

Si $u \in \mathcal{L}(E)$ est un endomorphisme symétrique alors

$$\mathrm{Im} u = (\ker u)^{\perp}$$

dém. :
Soit $x \in \ker u$ et $y \in \mathrm{Im} u$. On peut écrire $y = u(a)$ avec $a \in E$ et alors

$$(x \mid y) = (x \mid u(a)) = (u(x) \mid a) = (0_E \mid a) = 0$$

Ainsi, les espaces $\mathrm{Im} u$ et $\ker u$ sont orthogonaux et donc $\mathrm{Im} u \subset (\ker u)^{\perp}$ puis l'égalité par les dimensions.
□

8.3.2 Matrice d'un endomorphisme symétrique

Théorème

Soit $u \in \mathcal{L}(E)$ et $e = (e_1, \ldots, e_n)$ une base orthonormale de E.
On a équivalence entre :
(i) u est symétrique ;
(ii) la matrice $\mathrm{Mat}_e u$ est symétrique.

dém. :
(i) $\Rightarrow$ (ii) Supposons u symétrique et étudions $A = (a_{i,j}) = \mathrm{Mat}_e u$.
On a $a_{i,j} = (e_i \mid u(e_j))$ et donc par symétrie,

$$a_{i,j} = (u(e_i) \mid e_j) = (e_j \mid u(e_i)) = a_{j,i}$$

La matrice A est donc symétrique.
(ii) $\Rightarrow$ (i) Supposons $A = (a_{i,j}) = \mathrm{Mat}_e u$ symétrique.
Soit $x, y \in E$ de colonnes coordonnées X et Y dans la base e. Puisque la base e est orthonormale

$$(u(x) \mid y) = {}^t(AX)Y = {}^tX^tAY \text{ et } (x \mid u(y)) = {}^tXAY$$

Or ${}^tA = A$ donc $(u(x) \mid y) = (x \mid u(y))$.
□

Remarque Il est essentiel de vérifier que la base e est orthonormale pour exploiter ce résultat.

Corollaire

L'ensemble $\mathcal{S}(E)$ des endomorphismes symétriques de E est un sous-espace vectoriel de $\mathcal{L}(E)$ de dimension $\dfrac{n(n+1)}{2}$.

dém. :
$\mathcal{S}_n(\mathbb{R})$ et $\mathcal{S}(E)$ sont isomorphes via représentation matricielle dans la base orthonormée e.
□

8.3.3 Théorème spectral

Lemme

> Si F est un sous-espace vectoriel stable par $u \in \mathcal{L}(E)$ symétrique alors $F^{\perp}$ est aussi stable par u.
> De plus, les endomorphismes induits par u sur F et $F^{\perp}$ sont encore symétriques.

dém. :
Soit $x \in F^{\perp}$ et $y \in F$. On a

$$(u(x) \mid y) = (x \mid u(y)) = 0$$

car $x \in F^{\perp}$ et $u(y) \in F$.
De plus, pour tout $x, y \in F$,

$$(u_F(x) \mid y) = (u(x) \mid y) = (x \mid u(y)) = (x \mid u_F(y))$$

Ainsi, u_F est symétrique et il en est de même de $u_{F^{\perp}}$.
□

Lemme

> Les sous-espaces propres d'un endomorphisme symétrique sont deux à deux orthogonaux.

dém. :
Soit $\lambda, \mu \in \mathbb{R}$ distincts. Pour $x \in E_{\lambda}(u)$ et $y \in E_{\mu}(u)$:
D'une part, $(u(x) \mid y) = (\lambda x \mid y) = \lambda(x \mid y)$
D'autre part, $(u(x) \mid y) = (x \mid u(y)) = (x \mid \mu y) = \mu(x \mid y)$
On en déduit $\lambda(x \mid y) = \mu(x \mid y)$, or $\lambda \neq \mu$ donc $(x \mid y) = 0$.
□

Lemme

> Tout endomorphisme symétrique d'un espace euclidien non nul admet au moins une valeur propre réelle.

dém. :
Soit $u \in \mathcal{L}(E)$ un endomorphisme symétrique de E euclidien avec $\dim E > 0$.
Si $\dim E = 1$: les éléments non nuls de E sont vecteurs propres de u.
Si $\dim E = 2$: la matrice de u dans une base orthonormale de E est de la forme

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Son polynôme caractéristique est $\chi_u = X^2 - (a+c)X + (ac - b^2)$ de discriminant

$$\Delta = (a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 \geqslant 0$$

L'endomorphisme u admet donc au moins une valeur propre réelle.
Si $\dim E > 2$: l'endomorphisme u admet au moins une droite ou un plan stable. L'endomorphisme induit sur ce sous-espace vectoriel est encore symétrique et possède donc une valeur propre.
□

Théorème

> Tout endomorphisme symétrique est diagonalisable dans une base orthonormale.

dém. :
Soit $u \in \mathcal{S}(E)$ et

$$F = \underset{\lambda \in \mathrm{Sp}u}{\overset{\perp}{\oplus}} E_{\lambda}(u)$$

Le sous-espace vectoriel F est stable par u donc $F^\perp$ aussi.
Par l'absurde, supposons $F^\perp \neq \{0_E\}$. L'endomorphisme induit par u sur $F^\perp$ est symétrique, il possède donc au moins un vecteur propre. Or celui-ci est aussi vecteur propre de u et donc élément de F. C'est absurde car $F \cap F^\perp = \{0_E\}$.
Ainsi, E est la somme directe des sous-espaces propres de u et puisque ceux-ci sont deux à deux orthogonaux, on peut former une base orthonormale adaptée à cette décomposition, base qui diagonalise u.
□

Exemple Soit $u \in \mathcal{S}(E)$. Posons $\lambda_{\min} = \min \mathrm{Sp}u$ et $\lambda_{\max} = \max \mathrm{Sp}u$.
On a

$$\forall x \in E, \lambda_{\min} \|x\|^2 \leqslant (u(x) \mid x) \leqslant \lambda_{\max} \|x\|^2$$

En effet, soit $e = (e_1, \ldots, e_n)$ une base orthonormale diagonalisant u.
$\mathrm{Mat}_e(u) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ avec $\lambda_1, \ldots, \lambda_n$ les valeurs propres de u.
Pour $x \in E$, on peut écrire $x = \sum_{i=1}^{n} x_i e_i$ et on a alors $u(x) = \sum_{i=1}^{n} \lambda_i x_i e_i$.
Puisque la base e est orthonormale,

$$\|x\|^2 = \sum_{i=1}^{n} x_i^2 \text{ et } (u(x) \mid x) = \sum_{i=1}^{n} \lambda_i x_i^2$$

Or, pour tout $1 \leqslant i \leqslant n$, $\lambda_{\min} \leqslant \lambda_i \leqslant \lambda_{\max}$ donc

$$\lambda_{\min} \|x\|^2 \leqslant (u(x) \mid x) \leqslant \lambda_{\max} \|x\|^2$$

8.3.4 Diagonalisation des matrices symétriques réelles

Théorème

Toute matrice symétrique réelle est orthogonalement diagonalisable

$$\forall A \in \mathcal{S}_n(\mathbb{R}), \exists P \in \mathcal{O}_n(\mathbb{R}), \exists D \in D_n(\mathbb{R}), A = PDP^{-1} = PD{}^tP$$

dém. :
Soit $A \in \mathcal{S}_n(\mathbb{R})$. Munissons $E = \mathbb{R}^n$ du produit scalaire canonique et considérons u l'endomorphisme de $\mathbb{R}^n$ représenté par A dans la base canonique e.
Puisque A est symétrique et e orthonormale, l'endomorphisme u est autoadjoint. Il existe donc une base orthonormée e' diagonalisant u. Par changement de base, on a alors $A = PDP^{-1}$ avec D diagonale et P orthogonale car matrice de passage entre deux bases orthonormées.
□

Exemple Pour $A \in \mathcal{M}_n(\mathbb{R})$, ${}^tAA$ est diagonalisable car symétrique réelle.
Ses valeurs propres sont appelées valeurs singulières de A.

Attention : Une matrice symétrique complexe n'est pas nécessairement diagonalisable :

Exemple Pour $A = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$, $\chi_A = X^2$ donc $\mathrm{Sp}A = \{0\}$.
Puisque $A \neq O_2$, la matrice A n'est pas diagonalisable.

8.3.5 Musculation : positivité

8.3.5.1 Endomorphisme symétrique positif

Définition

Un endomorphisme symétrique u de E est dit positif si

$$\forall x \in E, (u(x) \mid x) \geqslant 0\,;$$

On le dit défini positif si de plus

$$\forall x \in E, (u(x) \mid x) = 0 \Rightarrow x = 0_E$$

On note $\mathcal{S}^+(E)$ (resp. $\mathcal{S}^{++}(E)$) l'ensemble des endomorphismes symétriques positifs (resp. définis et positifs).

Proposition

Soit u un endomorphisme symétrique de E.
On a équivalence entre ;
(i) u est positif (resp. défini positif) ;
(ii) $\mathrm{Sp}u \subset \mathbb{R}^+$ (resp. $\mathrm{Sp}u \subset \mathbb{R}^{+\star}$).

dém. :
(i) $\Rightarrow$ (ii) Supposons u positif.
Soit λ une valeur propre de u et x un vecteur propre associé.
$(u(x) \mid x) = (\lambda x \mid x) = \lambda \left\|x\right\|^2$ et $(u(x) \mid x) \geqslant 0$ donc $\lambda \left\|x\right\|^2 \geqslant 0$ puis $\lambda \geqslant 0$ car $\left\|x\right\|^2 > 0$.
(ii) $\Rightarrow$ (i) Supposons $\mathrm{Sp}(u) \subset \mathbb{R}^+$.
Par le théorème spectral, il existe une base orthonormale $e = (e_1, \ldots, e_n)$ diagonalisant u :

$$\mathrm{Mat}_e u = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix}$$

avec $\lambda_1, \ldots, \lambda_n$ les valeurs propres de u.
Pour tout $x \in E$, on peut écrire $x = \sum_{i=1}^{n} x_i e_i$ et alors

$$(u(x) \mid x) = \sum_{i=1}^{n} \lambda_i x_i^2 \geqslant 0$$

La démonstration s'adapté à l'étude des endomorphismes définis positifs.
□

Remarque On en déduit $\mathcal{S}^{++}(E) = \mathcal{S}^+(E) \cap \mathrm{GL}(E)$ car

$$0 \notin \mathrm{Sp}u \Leftrightarrow u \in \mathrm{GL}(E)$$

8.3.5.2 Matrice symétrique positive

Définition

Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ symétrique est dite positive si

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{K}), {}^tXAX \geqslant 0$$

On la dit définie positive si de plus

$$\forall X \in \mathcal{M}_{n,1}(\mathbb{K}), {}^tXAX = 0 \Rightarrow X = 0$$

On note $\mathcal{S}_n^+(\mathbb{R})$ (resp. $\mathcal{S}_n^{++}(\mathbb{R})$) l'ensemble des matrices symétriques positives (resp. définies positives).

Remarque Si l'on introduit le produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$ alors

$${}^tXAX = (AX \mid X)$$

De plus, il y a évidemment correspondance avec les endomorphismes symétriques positifs moyennant représentation en base orthonormale.

Exemple Si $M \in \mathcal{M}_n(\mathbb{R})$ alors $A = {}^tMM$ est symétrique positive.
${}^tA = {}^t\left({}^tMM\right) = {}^tMM = A$ donc A est symétrique et pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$,
${}^tXAX = {}^t(MX)MX = \|MX\|^2 \geqslant 0$
Si de plus $M \in \mathrm{GL}_n(\mathbb{R})$ alors $A = {}^tMM$ est définie positive.
En effet, ${}^tXAX = \|MX\|^2 = 0 \Rightarrow MX = 0$ donc ${}^tXAX = 0 \Rightarrow X = 0$ car M est inversible.

Proposition

Soit $A \in \mathcal{S}_n(\mathbb{R})$. On a équivalence entre :
(i) A est positive (resp. définie positive) ;
(ii) Sp$A \subset \mathbb{R}^+$ (resp. Sp$A \subset \mathbb{R}^{+\star}$).

dém. :
(i) $\Rightarrow$ (ii) Supposons A positive.
Soit $\lambda \in$ SpA et X vecteur propre associé.
${}^tXAX = \lambda {}^tXX = \lambda \|X\|^2 \geqslant 0$ avec $\|X\|^2 > 0$ donc $\lambda \geqslant 0$.
(ii) $\Rightarrow$ (i) Supposons Sp$A \subset \mathbb{R}^+$.
La matrice A est orthogonalement semblable à une matrice diagonale, donc il existe $P \in \mathcal{O}_n(\mathbb{R})$ telle que ${}^tPAP = D$ avec $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.
Pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$, ${}^tXAX = {}^t(PX)DPX = {}^tYDY$ avec $Y = PX$.
En notant $y_1, \ldots, y_n$ les coefficients de la colonne Y alors ${}^tXAX = \sum_{i=1}^n \lambda_i y_i^2 \geqslant 0$.

□

8.3.6 Musculation : matrice de Gram

Soit E un espace préhilbertien de produit scalaire $\langle ., . \rangle$.

Définition

On appelle matrice de Gram d'une famille $(a_1, \ldots, a_n)$ de vecteurs de E la matrice carrée

$$G(a_1, \ldots, a_n) = (\langle a_i, a_j \rangle)_{1 \leqslant i,j \leqslant n}$$

Exemple La famille $(a_1, \ldots, a_n)$ est orthogonale si, et seulement si, $G(a_1, \ldots, a_n)$ est diagonale.
La famille $(a_1, \ldots, a_n)$ est orthonormale si, et seulement si, $G(a_1, \ldots, a_n)$ est l'identité.

Théorème

La matrice de Gram $G(a_1, \ldots, a_n)$ est symétrique positive et inversible si, et seulement si, la famille $(a_1, \ldots, a_n)$ est libre.

dém. :
$A = G(a_1, \ldots, a_n)$ est symétrique car $\langle a_i, a_j \rangle = \langle a_j, a_i \rangle$.
Pour $X = {}^t\begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix}$, on observe

$${}^tXAX = \|\lambda_1 a_1 + \cdots + \lambda_n a_n\|^2 \geqslant 0$$

La matrice symétrique A est donc positive. Elle est définie positive si, et seulement si,

$${}^tXAX = 0 \Rightarrow X = 0$$

c'est-à-dire

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0_E \Rightarrow \lambda_1 = \ldots = \lambda_n = 0$$

ce qui correspond à la liberté de la famille $(a_1, \ldots a_n)$.
On en déduit que A est inversible si, et seulement si, $(a_1, \ldots a_n)$ est libre.
□

Théorème

Soit $x \in E$ et $(a_1, \ldots, a_n)$ une base d'un sous-espace vectoriel F de E.
On a

$$d(x, F) = \sqrt{\frac{\det G(a_1, \ldots, a_n, x)}{\det G(a_1, \ldots, a_n)}}$$

dém. :
On écrit $x = y + z$ avec $y \in F$ et $z \in F^\perp$. On sait $d(x, F) = \|z\|$. Puisque

$$\langle a_i, x \rangle = \langle a_i, y \rangle + \langle a_i, z \rangle = \langle a_i, y \rangle \text{ et } \langle x, x \rangle = \langle y, y \rangle + \langle z, z \rangle$$

on peut écrire

$$G(a_1, \ldots, a_n, x) = \begin{pmatrix} \langle a_1, a_1 \rangle & \cdots & \langle a_1, a_n \rangle & \langle a_1, y \rangle \\ \vdots & & \vdots & \vdots \\ \langle a_n, a_1 \rangle & \cdots & \langle a_n, a_n \rangle & \langle a_n, y \rangle \\ \langle y, a_1 \rangle & \cdots & \langle y, a_n \rangle & \langle y, y \rangle + \langle z, z \rangle \end{pmatrix}$$

En décomposant la dernière colonne en somme de deux colonnes

$$\det(G(a_1, \ldots, a_n, x)) = \det(G(a_1, \ldots, a_n, y)) + \det(G(a_1, \ldots, a_n)) \|z\|^2$$

La famille $(a_1, \ldots, a_n, y)$ étant liée, on obtient

$$\det\left(G(a_1, \ldots, a_n, x)\right) = \det\left(G(a_1, \ldots, a_n)\right) \|z\|^2$$

qui permet de conclure.
□

Deuxième partie

Analyse

Chapitre 9

Suites et séries numériques

$\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

9.1 Suites numériques

9.1.1 Limites

Définition

On dit qu'une suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{K}$ converge vers $\ell \in \mathbb{K}$ si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow |u_n - \ell| \leqslant \varepsilon$$

On note alors $u_n \to \ell$ ou $u_n \xrightarrow[n\to+\infty]{} \ell$.
Il y a alors unicité du nombre ℓ qui est appelée limite de la suite (u_n).

Définition

On dit qu'une suite réelle $(u_n)_{n\in\mathbb{N}}$ diverge vers $+\infty$ si

$$\forall A \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow u_n \geqslant A$$

On note alors $u_n \to +\infty$ ou $u_n \xrightarrow[n\to+\infty]{} +\infty$.
On définit de façon analogue la divergence vers $-\infty$.

Exemple Etudions $\lim\limits_{n\to+\infty} \left(1+\dfrac{1}{n}\right)^n$
On peut écrire

$$\left(1+\frac{1}{n}\right)^n = \exp\left(n\ln\left(1+\frac{1}{n}\right)\right)$$

Or

$$n\ln(1+1/n) \sim n \times \frac{1}{n} \to 1$$

donc

$$\left(1+\frac{1}{n}\right)^n \xrightarrow[n\to+\infty]{} \mathrm{e}$$

Exemple Soit $(u_n) \in (\mathbb{R}^+)^{\mathbb{N}}$. On suppose que $\sqrt[n]{u_n} \to \ell \in [0,1[$.
Montrons qu'alors $u_n \to 0$.
Introduisons $\varepsilon > 0$ (dont on précisera la valeur par la suite).
Puisque $\sqrt[n]{u_n} \to \ell$, pour n assez grand $\ell - \varepsilon \leqslant \sqrt[n]{u_n} \leqslant \ell + \varepsilon$ donc $0 \leqslant u_n \leqslant (\ell+\varepsilon)^n$.
Si l'on choisit initialement $\varepsilon > 0$ pour que $\ell + \varepsilon < 1$, on obtient $u_n \to 0$ par encadrement.
On montre de façon similaire, on montre

$$\sqrt[n]{u_n} \to \ell > 1 \Rightarrow u_n \to +\infty$$

9.1.2 Limites monotones

Théorème

a) Toute suite réelle croissante et majorée converge.
b) Toute suite réelle croissante, mais non majorée, diverge vers $+\infty$.

dém. :
Cas u croissante et majorée.
Posons $\ell = \sup\limits_{n\in\mathbb{N}} u_n \in \mathbb{R}$ et montrons $u_n \to \ell$.
On a déjà

$$\forall n \in \mathbb{N}, u_n \leqslant \ell$$

car $\ell = \sup\limits_{n\in\mathbb{N}} u_n$ majore la suite u.
Soit $\varepsilon > 0$. Comme $\ell - \varepsilon < \ell = \sup\limits_{n\in\mathbb{N}} u_n$, $\ell - \varepsilon$ n'est pas majorant de la suite u et donc il existe $N \in \mathbb{N}$ vérifiant $u_N > \ell - \varepsilon$. Par croissance de la suite u, on a alors

$$\forall n \geqslant N, u_n \geqslant u_N > \ell - \varepsilon$$

Alors, pour tout $n \geqslant N$, $\ell - \varepsilon \leqslant u_n \leqslant \ell$ donc $|u_n - \ell| \leqslant \varepsilon$. Finalement $u_n \to \ell$.
Cas u croissante non majorée.
Soit $A \in \mathbb{R}$. La suite u n'est pas majorée par A donc il existe $N \in \mathbb{N}$ vérifiant $u_N > A$.
Par croissance de la suite u on a alors

$$\forall n \geqslant N, u_n \geqslant u_N \geqslant A$$

Ainsi $u_n \to +\infty$
Les deux autres cas du théorème s'obtiennent par passage à l'opposé.
□

Exemple Etudions la convergence de $u_n = \sum\limits_{k=n+1}^{2n} \frac{1}{k}$
On a

$$u_{n+1} - u_n = \sum_{k=n+2}^{2n+2} \frac{1}{k} - \sum_{k=n+1}^{2n} \frac{1}{k} = \frac{1}{2n+1} + \frac{1}{2n+2} - \frac{1}{n+1} \geqslant 0$$

De plus

$$u_n \leqslant \sum_{k=n+1}^{2n} \frac{1}{n+1} = \frac{n}{n+1} \leqslant 1$$

La suite (u_n) est croissante et majorée, donc elle converge.

En fait, on peut montrer (par les sommes de Riemann, par exemple)

$$u_n \to \ln 2$$

9.1.3 Comparaisons asymptotiques

Définition

On dit que la suite (u_n) est dominée par la suite (v_n) et l'on écrit $u_n = O(v_n)$ s'il existe $M \in \mathbb{R}^+$ et $N \in \mathbb{N}$ vérifiant

$$\forall n \geqslant N, |u_n| \leqslant M\,|v_n|$$

Remarque Il revient au même de dire que l'on peut écrire à partir d'un certain rang

$$u_n = v_n b_n \text{ avec } (b_n) \text{ bornée}$$

Exemple On peut écrire

$$\frac{\cos(n)}{n^2+1} = O\left(\frac{1}{n^2}\right)$$

Définition

On dit que la suite (u_n) est négligeable devant (v_n) et l'on écrit $u_n = o(v_n)$ si, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ vérifiant

$$\forall n \geqslant N, |u_n| \leqslant \varepsilon\,|v_n|$$

Remarque Il revient au même de dire que l'on peut écrire à partir d'un certain rang

$$u_n = v_n \varepsilon_n \text{ avec } (\varepsilon_n) \text{ de limite nulle.}$$

Exemple En écrivant $u_n \llbracket v_n$ pour signifier $u_n = o(v_n)$, on peut proposer la hiérarchie suivante ;

$$\mathrm{e}^{-n} \llbracket \frac{1}{n^2} \llbracket \frac{1}{n} \llbracket \frac{1}{\ln n} \llbracket 1 \llbracket \ln n \llbracket \sqrt{n} \llbracket n \llbracket n^2 \llbracket \mathrm{e}^n$$

Définition

On dit que la suite (u_n) est équivalente à la suite (v_n) et l'on écrit $u_n \sim v_n$ si l'on peut écrire

$$u_n = v_n + o(v_n)$$

Remarque Il revient au même de dire que l'on peut écrire à partir d'un certain rang

$$u_n = v_n\varphi_n \text{ avec } (\varphi_n) \text{ de limite } 1$$

Exemple On peut écrire $\sin\left(\frac{1}{n}\right) \sim \frac{1}{n}$

9.1.4 Développements asymptotiques

Définition

Un développement asymptotique d'une suite est la décomposition de son terme général en somme de termes simples ordonnés en négligeabilité croissante.

Exemple Formons un DA à trois termes de $\left(1+\frac{1}{n}\right)^n$
Quand $n \to +\infty$.

$$\left(1+\frac{1}{n}\right)^n = \exp\left(n\ln(1+\frac{1}{n})\right) = \exp\left(1 - \frac{1}{2n} + \frac{1}{3n^2} + o\left(\frac{1}{n^2}\right)\right)$$

Par composition

$$\left(1+\frac{1}{n}\right)^n = \mathrm{e} - \frac{\mathrm{e}}{2n} + \frac{11\mathrm{e}}{24n^2} + o\left(\frac{1}{n^2}\right)$$

Exemple Soit $n \geqslant 2$. On considère l'équation $x^n = 1 + x$ d'inconnue $x \in [1, +\infty[$.
a) Montrons que celle-ci admet une unique solution x_n.
b) Déterminons la limite de $(x_n)_{n\geqslant 2}$.
c) Formons un développement asymptotique à deux termes de la suite $(x_n)_{n\geqslant 2}$.
Considérons $f_n : x \mapsto x^n - x - 1$ définie sur $[1, +\infty[$.
f_n est de classe $\mathcal{C}^\infty$ et $f_n'(x) = nx^{n-1} - 1 > 0$ sur $[1, +\infty[$. La fonction f est donc strictement croissante.
Puisque $f_n(1) = -1$ et $\lim\limits_{x\to+\infty} f_n(x) = +\infty$, la fonction f s'annule une unique fois sur $[1, +\infty[$.
Ceci définit $x_n \in [1, +\infty[$
On a

$$f_n(x_{n+1}) = x_{n+1}^n - x_{n+1} - 1 = x_{n+1}^n - x_{n+1}^{n+1} < 0$$

et donc $x_{n+1} < x_n$.
La suite (x_n) est décroissante et minorée (par 1), elle est donc convergente.
Posons ℓ sa limite. Puisque $x_n \in [1, +\infty[$, à la limite $\ell \in [1, +\infty[$.
Par l'absurde, si $\ell > 1$ alors $x_n^n \to +\infty$ car $x_n^n \geqslant \ell^n \to +\infty$.
Or $x_n^n = 1 + x_n \to 1 + \ell$. C'est absurde et on en déduit $\ell = 1$.
On peut alors écrire $x_n = 1 + \varepsilon_n$ avec $\varepsilon_n \to 0$.
Déterminons un équivalent de ε_n.
On a $(1+\varepsilon_n)^n = 2 + \varepsilon_n$ donc $n\ln(1+\varepsilon_n) = \ln(2+\varepsilon_n) \to \ln 2$ puis $n\varepsilon_n \sim \ln(2)$
On en déduit

$$x_n = 1 + \frac{\ln 2}{n} + o\left(\frac{1}{n}\right)$$

9.1.5 Suites récurrentes

Exemple Etudions la suite (u_n) déterminée par $u_0 > 0$ et $\forall n \in \mathbb{N}, u_{n+1} = \ln(1 + u_n)$
La fonction itératrice $f : x \mapsto \ln(1 + x)$ est définie sur $]-1, +\infty[$, il est facile d'en obtenir le tableau de variation.
Pour $\mathcal{D} =]0, +\infty[$, on a

$$u_0 \in \mathcal{D} \text{ et } \forall x \in \mathcal{D}, f(x) \in \mathcal{D}$$

On en déduit que la suite (u_n) est bien définie et

$$\forall n \in \mathbb{N}, u_n \in]0, +\infty[$$

Si (u_n) converge, sa limite ℓ appartient à $[0, +\infty[$.
De plus, en passant la relation de récurrence $u_{n+1} = \ln(1 + u_n)$ à la limite, on obtient $\ell = \ln(1 + \ell)$.
La seule solution de cette équation est $\ell = 0$.
En visualisant le comportement de (u_n) à partir d'une représentation de f, on est inspiré à étudier sa monotonie...
On a

$$u_{n+1} - u_n = \ln(1 + u_n) - u_n \leqslant 0$$

car on sait $\ln(1 + x) \leqslant x$ pour tout $x > -1$.
La suite (u_n) est donc décroissante et convergente car minorée par 0.
Puisque la seule limite finie possible est 0, on peut conclure que $u_n \to 0$.

Exemple Etudions la convergence de la suite (u_n) définie par $u_0 = 1$ et $u_{n+1} = \sqrt{3 - u_n}$
Considérons $f : x \mapsto \sqrt{3 - x}$ définie sur $]-\infty, 3]$

x	$-\infty$	0	3
$f(x)$	$+\infty$	$\sqrt{3}$	0

Pour $\mathcal{D} = [0, 3]$, on a $u_0 \in \mathcal{D}$ et pour tout $x \in \mathcal{D}$, $f(x) \in \mathcal{D}$.
La suite (u_n) est donc bien définie et pour tout $n \in \mathbb{N}$, $u_n \in [0, 3]$.
Supposons $u_n \to \ell \in \mathbb{R}$.
Puisque pour tout $n \in \mathbb{N}$, $0 \leqslant u_n \leqslant 3$, à la limite $\ell \in [0, 3]$.
En passant la relation de récurrence $u_{n+1} = \sqrt{3 - u_n}$ à la limite on obtient

$$\ell = \sqrt{3 - \ell}$$

ce qui donne

$$\ell = \frac{-1 + \sqrt{13}}{2}$$

car $\ell \geqslant 0$.
Notons

$$\alpha = \frac{-1 + \sqrt{13}}{2}$$

On a

$$|u_{n+1}-\alpha| = \left|\sqrt{3-u_n}-\sqrt{3-\alpha}\right| = \frac{|u_n-\alpha|}{\sqrt{3-u_n}+\sqrt{3-\alpha}} \leqslant \frac{|u_n-\alpha|}{\sqrt{3-\alpha}}$$

avec

$$q = \frac{1}{\sqrt{3-\alpha}} = \frac{1}{\alpha} \in [0,1[$$

Ainsi

$$|u_n-\alpha| \leqslant q^n\,|u_0-\alpha|$$

et donc $u_n \to \alpha$.

9.1.6 Théorème de Cesaro

Théorème

Si (u_n) est une suite numérique converge vers ℓ alors

$$v_n = \frac{u_1+\cdots+u_n}{n} \to \ell$$

dém. :
On a

$$v_n - \ell = \frac{1}{n}\left((u_1-\ell)+\cdots+(u_n-\ell)\right)$$

Pour $\varepsilon > 0$, il existe $N \in \mathbb{N}$ vérifiant

$$\forall n \geqslant N,\ |u_n-\ell| \leqslant \varepsilon$$

Pour $n \geqslant N$,

$$|v_n-\ell| \leqslant \frac{|u_1-\ell|+\cdots+|u_{N-1}-\ell|}{n} + \frac{n-N+1}{n}\varepsilon$$

donc

$$|v_n-\ell| \leqslant \frac{|u_1-\ell|+\cdots+|u_{N-1}-\ell|}{n} + \varepsilon$$

Or

$$\frac{|u_1-\ell|+\cdots+|u_{N-1}-\ell|}{n} = \frac{C^{te}}{n} \to 0$$

donc il existe $N' \in \mathbb{N}$ tel que pour $n \geqslant N'$,

$$\frac{|u_1-\ell|+\cdots+|u_{N-1}-\ell|}{n} \leqslant \varepsilon$$

Ainsi, pour $n \geqslant \max(N,N')$, $|v_n-\ell| \leqslant 2\varepsilon$ ce qui permet de conclure.
□

Exemple Déterminons un équivalent de (u_n) donnée par $u_0 > 0$ et $\forall n \in \mathbb{N}, u_{n+1} = \ln(1+u_n)$
On a déjà montré $u_n \to 0^+$. Déterminons maintenant un équivalent de (u_n).
On a

$$\frac{1}{u_{n+1}} - \frac{1}{u_n} = \frac{u_n-u_{n+1}}{u_n u_{n+1}} \sim \frac{\frac{1}{2}u_n^2}{u_n^2} \to \frac{1}{2}$$

Par le théorème de Cesaro

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{1}{u_{k+1}}-\frac{1}{u_k}\right)=\frac{1}{n}\left(\frac{1}{u_n}-\frac{1}{u_0}\right)\to\frac{1}{2}$$

On en déduit

$$u_n\sim\frac{2}{n}$$

9.2 Séries numériques

9.2.1 Définition

Définition

Soit $(u_n)_{n\geqslant n_0}$ une suite numérique. On appelle série de terme général u_n la suite $(S_n)_{n\geqslant n_0}$ avec

$$S_n=\sum_{k=n_0}^{n}u_k$$

Cette série est notée $\sum\limits_{n\geqslant n_0}u_n$ ou $\sum u_n$.

Le terme S_n est appelé somme partielle de rang n de cette série.

Remarque Une série est un cas particulier de suite, c'est une suite de sommes partielles.

Exemple La série $\sum\limits_{n\geqslant 0}n$ est la suite des sommes partielles

$$S_n=\sum_{k=0}^{n}k=\frac{n(n+1)}{2}$$

Exemple La série $\sum\limits_{n\geqslant 0}q^n$ est la suite des sommes partielles

$$S_n=\sum_{k=0}^{n}q^k=\frac{1-q^{n+1}}{1-q}\text{ (si }q\neq 1\text{)}$$

Exemple La série $\sum\limits_{n\geqslant 1}\frac{1}{n}$ est la suite des sommes partielles

$$S_n=\sum_{k=1}^{n}\frac{1}{k}\text{ (avec }n\geqslant 1\text{)}$$

Exemple Soit (v_n) une suite d'éléments de $\mathbb{K}$.
Posons $u_0 = v_0$ et $u_n = v_n - v_{n-1}$.
La série $\sum\limits_{n\geqslant 0} u_n$ est la suite des sommes partielles

$$\sum_{k=0}^{n} u_k = v_n$$

Ainsi, la suite (v_n) se confond avec la série $\sum u_n$.

On suppose désormais les séries étudiées définies à partir du rang $n_0 = 0$.
On peut s'y ramener quitte à poser les premiers termes de la série comme étant nuls si non définis.

9.2.2 Convergence d'une série numérique

9.2.2.1 Nature d'une série numérique

Définition

On dit que qu'une série $\sum u_n$ converge si la suite de ses sommes partielles converge.
On peut alors introduire la somme de la série

$$\sum_{k=0}^{+\infty} u_k \underset{\text{déf}}{=} \lim_{n\to+\infty} \sum_{k=0}^{n} u_k$$

Attention : Par essence, une somme de série numérique est une limite, pour la manipuler, il est indispensable de justifier a priori son existence, i.e. que la série soit convergente.

Exemple Etudions $\sum\limits_{n\geqslant 2} \dfrac{1}{n(n-1)}$
Pour $n \geqslant 2$,

$$\sum_{k=2}^{n} \frac{1}{k(k-1)} = \sum_{k=2}^{n} \frac{1}{k-1} - \frac{1}{k} = 1 - \frac{1}{n} \xrightarrow[n\to+\infty]{} 1$$

Ainsi la série $\sum\limits_{n\geqslant 2} \dfrac{1}{n(n-1)}$ converge et

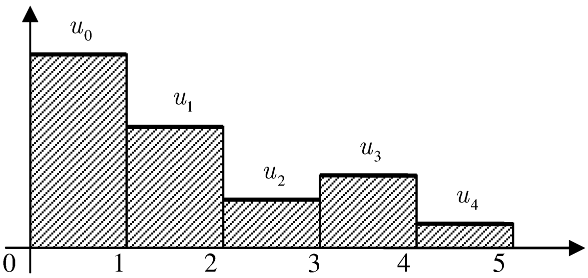
$$\sum_{n=2}^{+\infty} \frac{1}{n(n-1)} = 1$$

Exemple Etudions $\sum\limits_{n\geqslant 1} \dfrac{1}{n}$.
Pour $n \geqslant 1$, la fonction $t \mapsto 1/t$ étant décroissante, on a

$$\sum_{k=1}^{n} \frac{1}{k} \geqslant \sum_{k=1}^{n} \int_{k}^{k+1} \frac{\mathrm{d}t}{t} = \int_{1}^{n+1} \frac{\mathrm{d}t}{t} = \ln(n+1) \to +\infty$$

Ainsi la série $\sum\limits_{n\geqslant 1}\dfrac{1}{n}$ diverge.

Remarque $\sum u_n$ converge si, et seulement si, la somme des aires hachurées converge.



Exemple Etudions $\sum\limits_{n\geqslant 1}\dfrac{(-1)^{n-1}}{n}$.

Pour $n\geqslant 1$,

$$\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}=\sum_{k=1}^{n}(-1)^{k-1}\int_0^1 t^{k-1}\,\mathrm{d}t=\int_0^1\frac{1-(-t)^n}{1+t}\,\mathrm{d}t$$

Or

$$\int_0^1\frac{\mathrm{d}t}{1+t}=\ln 2 \text{ et } 0\leqslant\int_0^1\frac{t^n}{1+t}\,\mathrm{d}t\leqslant\int_0^1 t^n\,\mathrm{d}t=\frac{1}{n+1}$$

donc

$$\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k}\xrightarrow[n\to+\infty]{}\ln 2$$

Ainsi $\sum\limits_{n\geqslant 1}\dfrac{(-1)^{n-1}}{n}$ converge et

$$\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{n}=\ln 2$$

9.2.2.2 Reste d'une série convergente

Théorème

Soit $n_0\in\mathbb{N}$. On a équivalence entre :

(i) $\sum\limits_{n\geqslant 0}u_n$ converge ;

(ii) $\sum\limits_{n\geqslant n_0}u_n$ converge.

dém. :
Les sommes partielles de deux séries diffèrent d'une constante et donc l'une converge si, et seulement si, l'autre aussi.
□

Corollaire

On ne modifie pas la nature d'une série en en modifiant la valeur d'un nombre fini de termes.
En revanche, cela modifie évidemment la valeur de la somme…

Définition

Si la série $\sum u_n$ converge, on peut introduire la somme

$$R_n = \sum_{k=n+1}^{+\infty} u_k$$

Ce terme est appelé reste de rang n de cette série.

Attention : On ne peut introduire le reste d'une série qu'après avoir justifié sa convergence.

Théorème

Si $\sum u_n$ converge alors pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{+\infty} u_k = \sum_{k=0}^{n} u_k + \sum_{k=n+1}^{+\infty} u_k$$

De plus

$$R_n = \sum_{k=n+1}^{+\infty} u_k \xrightarrow[n\to+\infty]{} 0$$

dém. :
Soit $n \in \mathbb{N}$ fixé.
Pour $N \geqslant n$,

$$\sum_{k=0}^{N} u_k = \sum_{k=0}^{n} u_k + \sum_{k=n+1}^{N} u_k$$

Quand $N \to +\infty$, on obtient

$$\sum_{k=0}^{+\infty} u_k = \sum_{k=0}^{n} u_k + \sum_{k=n+1}^{+\infty} u_k$$

égalité qu'on écrit souvent $S = S_n + R_n$.
De plus, on a alors

$$R_n = S - S_n \xrightarrow[n\to+\infty]{} 0$$

□

9.2.3 Limite du terme d'une série convergente

Théorème

Si la série $\sum u_n$ converge alors $u_n \to 0$.

dém. :

Posons $S_n = \sum_{k=0}^{n} u_k$. Si (S_n) converge en posant S sa limite

$$u_n = S_n - S_{n-1} \to S - S = 0$$

□

Définition

Si (u_n) ne tend pas vers 0 alors on dit que la série de terme général u_n diverge grossièrement (DVG).

Exemple La série $\sum \cos(n)$ diverge grossièrement.
En effet, si $\cos(n) \to 0$ alors la relation $\cos(2n) = 2\cos^2(n) - 1$ donne à la limite l'absurdité $0 = -1$.

Exemple La série $\sum_{n\geqslant 1} \frac{1}{n}$ diverge, mais pas grossièrement.

Remarque Si $\sum u_n$ converge, alors

$$\sum_{k=n+1}^{2n} u_k = S_{2n} - S_n \xrightarrow[n\to+\infty]{} 0$$

On peut alors retrouver la divergence de $\sum 1/n$ en exploitant

$$\sum_{k=n+1}^{2n} \frac{1}{k} \geqslant n \times \frac{1}{2n} = \frac{1}{2}$$

9.2.4 Opérations sur les séries convergentes

9.2.4.1 Linéarité

Théorème

Si $\sum u_n$ et $\sum v_n$ sont convergentes alors pour tout $\lambda \in \mathbb{K}$, les séries $\sum \lambda u_n$ et $\sum u_n + v_n$ convergent et

$$\sum_{k=0}^{+\infty} \lambda u_k = \lambda \sum_{k=0}^{+\infty} u_k \text{ et } \sum_{k=0}^{+\infty} (u_k + v_k) = \sum_{k=0}^{+\infty} u_k + \sum_{k=0}^{+\infty} v_k$$

dém. :
Par opérations sur les limites.
□

Corollaire

L'ensemble constitué des suites $u = (u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ telles que la série $\sum u_n$ converge est un sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$. L'application $u \mapsto \sum_{n=0}^{+\infty} u_n$ y définit une forme linéaire.

Exemple Si $\sum u_n$ et $\sum (u_n + v_n)$ convergent alors $\sum v_n$ converge.
En effet, on peut écrire

$$v_n = (u_n + v_n) + (-1).u_n$$

Attention : Pour écrire

$$\sum_{k=0}^{+\infty} (u_k + v_k) = \sum_{k=0}^{+\infty} u_k + \sum_{k=0}^{+\infty} v_k$$

il faut vérifier la convergence d'au moins deux des séries engagées.
Ceci interdit d'écrire des aberrations du type

$$\sum_{n=0}^{+\infty} 0 = \sum_{n=0}^{+\infty} 1 + \sum_{n=0}^{+\infty} (-1)$$

Exemple Si $\sum u_n$ converge et $\sum v_n$ diverge alors $\sum (u_n + v_n)$ diverge.

Attention : Si $\sum u_n$ et $\sum v_n$ divergent, on ne peut rien conclure sur la nature de $\sum (u_n + v_n)$.

9.2.4.2 Positivité

Théorème

Soit (u_n) une suite réelle.
Si $\sum u_n$ converge et si tous les termes de la suite sont positifs alors

$$\sum_{n=0}^{+\infty} u_n \geqslant 0$$

dém. :
Pour tout $N \in \mathbb{N}$, on a

$$\sum_{n=0}^{N} u_n \geqslant 0$$

donc à la limite

$$\sum_{n=0}^{+\infty} u_n \geqslant 0$$

□

Corollaire

Soit (u_n) et (v_n) deux suites réelles vérifiant $u_n \leqslant v_n$ pour tout $n \in \mathbb{N}$.

Si $\sum u_n$ et $\sum v_n$ convergent alors $\sum_{n=0}^{+\infty} u_n \leqslant \sum_{n=0}^{+\infty} v_n$.

dém. :

On a, avec convergences,

$$\sum_{n=0}^{+\infty} v_n - \sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} (v_n - u_n) \geqslant 0$$

□

Théorème

Soit (u_n) une suite réelle.

Si $u_n \geqslant 0$ pour tout $n \in \mathbb{N}$, si $\sum u_n$ converge et si $\sum_{n=0}^{+\infty} u_n = 0$ alors

$$\forall n \in \mathbb{N}, u_n = 0$$

dém. :

La suite (S_n) des sommes partielles est croissante car

$$S_{n+1} - S_n = u_{n+1} \geqslant 0$$

Or celle-ci est aussi positive et tend vers 0 donc

$$\forall n \in \mathbb{N}, S_n = 0$$

puis

$$\forall n \in \mathbb{N}, u_n = 0$$

□

9.2.4.3 Conjugaison

Théorème

Soit (z_n) une suite complexe.

Si $\sum z_n$ converge alors $\sum \overline{z_n}$ aussi et

$$\overline{\sum_{k=0}^{+\infty} z_k} = \sum_{k=0}^{+\infty} \overline{z_k}$$

dém. :
Par conjugaison de limites.

□

Corollaire

On a équivalence entre :
(i) $\sum z_n$ converge ;
(ii) $\sum \text{Re}(z_n)$ et $\sum \text{Im}(z_n)$ convergent.
De plus, on a alors

$$\sum_{k=0}^{+\infty} z_k = \sum_{k=0}^{+\infty} \text{Re}(z_k) + i\sum_{k=0}^{+\infty} \text{Im}(z_k)$$

dém. :
(i) $\Rightarrow$ (ii) car $\text{Re}(z_n) = \frac{1}{2}(z_n + \overline{z_n})$ et $\text{Im}(z_n) = \frac{1}{2i}(z_n - \overline{z_n})$.
(ii) $\Rightarrow$ (i) car $z_n = \text{Re}(z_n) + i\text{Im}(z_n)$.

□

9.3 Convergence par comparaison à une série positive

9.3.1 Cas des séries à termes réels positifs

Définition

Une série à termes positifs est une série dont le terme général est élément de $\mathbb{R}^+$.

Théorème

Soit $\sum u_n$ une série à termes positifs. On a équivalence entre :
(i) $\sum u_n$ converge ;
(ii) $\exists M \in \mathbb{R}, \forall n \in \mathbb{N}, \sum_{k=0}^{n} u_k \leqslant M$.

dém. :
La suite (S_n) des sommes partielles est croissante car $S_n - S_{n-1} = u_n \geqslant 0$. Ainsi, cette suite converge si, et seulement si, elle est majorée.

□

Remarque Si $\sum u_n$ est une série à termes positifs divergente alors $\sum_{k=0}^{n} u_k \xrightarrow[n\to+\infty]{} +\infty$

9.3.2 Comparaison de séries à termes positifs

Théorème

Soit $\sum u_n$ et $\sum v_n$ deux séries à termes positifs vérifiant

$$\forall n \in \mathbb{N}, u_n \leqslant v_n$$

a) Si $\sum v_n$ converge alors $\sum u_n$ aussi.
b) Si $\sum u_n$ diverge alors $\sum v_n$ aussi.

dém. :
a) $\sum u_n$ converge car c'est une série à termes positifs aux sommes partielles majorées car

$$\sum_{k=0}^{n} u_k \leqslant \sum_{k=0}^{n} v_k \leqslant \sum_{k=0}^{+\infty} v_k = M$$

b) C'est la contraposée de a).
□

Remarque Le résultat demeure même si la comparaison ne vaut qu'à partir d'un certain rang.

Exemple Déterminons la nature de $\sum_{n\geqslant 1} \frac{1}{n^2}$

Pour $n \geqslant 2$,

$$\frac{1}{n^2} \leqslant \frac{1}{n(n-1)}$$

or $\sum \frac{1}{n(n-1)}$ converge donc, par comparaison de série à termes positifs, la série $\sum_{n\geqslant 2} \frac{1}{n^2}$ converge, puis la série $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge.

Exemple Déterminons la nature de $\sum_{n\geqslant 1} \frac{\ln n}{n+1}$

On a $n\frac{\ln n}{n+1} \sim \ln n \to +\infty$ donc pour n assez grand,

$$\frac{\ln n}{n+1} \geqslant \frac{1}{n}$$

Or $\sum \frac{1}{n}$ diverge donc, par comparaison de série à termes positifs, la série $\sum \frac{\ln n}{n+1}$ diverge.
Plus précisément, on peut même affirmer

$$\sum_{k=1}^{n} \frac{\ln k}{k+1} \xrightarrow[n\to+\infty]{} +\infty$$

car la suite des sommes partielles est croissante puisque ses termes sont positifs.

Théorème

> Soit $\sum u_n$ et $\sum v_n$ deux séries à termes positifs.
> Si $u_n \sim v_n$ alors les séries $\sum u_n$ et $\sum v_n$ ont même nature.

dém. :
A partir d'un certain rang n_0, on peut écrire

$$1/2v_n \leqslant u_n \leqslant 2v_n$$

Quitte à modifier les premiers termes des séries, on peut supposer l'encadrement vrai pour tout rang n.
Par cet encadrement, la convergence d'une série entraîne la convergence de l'autre.
□

Exemple Déterminons la nature de $\sum \frac{1}{n^2+n}$
On a

$$\frac{1}{n^2+n} \underset{n\to+\infty}{\sim} \frac{1}{n^2}$$

Or $\sum 1/n^2$ converge et $1/n^2 \geqslant 0$ donc $\sum 1/(n^2+n)$ converge.

Exemple Déterminons la nature de $\sum \frac{1}{n+\sqrt{n}}$
On a

$$\frac{1}{n+\sqrt{n}} \underset{n\to+\infty}{\sim} \frac{1}{n}$$

Or $\sum 1/n$ diverge et $1/n \geqslant 0$ donc $\sum 1/(n+\sqrt{n})$ diverge.

Remarque Pour employer le résultat qui précède, il suffit seulement de vérifier la positivité de v_n, l'autre sera vraie (au moins à partir d'un certain rang) en vertu de l'équivalent.

Remarque Via passage à l'opposé, le résultat est aussi vrai pour les séries à termes négatifs.

Attention : La conversation de la nature d'une série par équivalence des termes n'est vraie que pour les séries à termes de signe constant.

9.3.3 Convergence absolue.

Définition

> Soit (u_n) une suite réelle ou complexe. On dit que la série $\sum u_n$ converge absolument si la série à termes positifs $\sum |u_n|$ converge.

Exemple La série $\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n^2}$ converge absolument (CVA)

En effet, $\sum_{n\geqslant 1} \frac{1}{n^2}$ converge.

Théorème

Si $\sum u_n$ converge absolument alors celle-ci converge et

$$\left|\sum_{n=0}^{+\infty} u_n\right| \leqslant \sum_{n=0}^{+\infty} |u_n|$$

dém. :
Cas (u_n) est une suite réelle à termes positifs : il n'y a rien à démontrer.
Cas (u_n) est une suite réelle. On introduit u_n^+ et u_n^- définis par

$$u_n^+ = \max(u_n, 0) \text{ et } u_n^- = \max(-u_n, 0)$$

On a

$$\forall n \in \mathbb{N}, u_n = u_n^+ - u_n^- \text{ et } |u_n| = u_n^+ + u_n^-$$

Puisque $0 \leqslant u_n^+, u_n^- \leqslant |u_n|$, on peut affirmer, par comparaison de séries à termes positifs, la convergence des séries $\sum u_n^+$ et $\sum u_n^-$ puis celle de $\sum u_n$ par différence de deux séries convergentes.
Cas (u_n) est une suite complexe. On introduit $\text{Re}(u_n)$ et $\text{Im}(u_n)$.
On a $|\text{Re}(u_n)|, |\text{Im}(u_n)| \leqslant |u_n|$ donc les séries réelles $\sum \text{Re}(u_n)$ et $\sum \text{Im}(u_n)$ convergent puis la série complexe $\sum u_n$ converge aussi.
□

*Bilan :*Pour une série réelle ou complexe :

$$\text{CVA} \Rightarrow \text{CV}$$

Pour une série à termes positifs :

$$\text{CVA} \Leftrightarrow \text{CV}$$

Remarque Plus généralement, pour une série à termes de signe constant à partir d'un certain rang, il y a aussi équivalence.

Attention : Il se peut que la série $\sum u_n$ converge alors que $\sum |u_n|$ diverge.

Définition

Une série convergente, mais non absolument convergente, est dite semi-convergente.

Exemple La série $\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}$ est semi-convergente.

9.3.4 Convergence par comparaison à une série positive

Théorème

Soit $\sum u_n$ une série numérique et $\sum v_n$ une série à termes positifs.
Si $u_n = O(v_n)$ et si $\sum v_n$ converge alors $\sum u_n$ converge absolument (et donc converge).

dém. :
Il existe $M \in \mathbb{R}$ et $N \in \mathbb{N}$ vérifiant
$$\forall n \geqslant N, |u_n| \leqslant M v_n$$
Quitte à modifier les premiers termes des séries (ce qui ne change pas la nature de celle-ci), on peut supposer la majoration vraie pour tout $n \in \mathbb{N}$. Or $\sum M v_n$ converge et $M v_n \geqslant 0$ donc, par comparaison de séries à termes positifs, $\sum |u_n|$ converge.
□

Corollaire

Si $u_n = o(v_n)$ et si $\sum v_n$ converge avec $v_n \geqslant 0$ alors $\sum u_n$ converge absolument et donc converge

Attention : Ces énoncés sont faux sans l'hypothèse $v_n \geqslant 0$.
Il est essentiel de comparer à une série à termes positifs !

Exemple Déterminons la nature de la série $\sum \frac{\sin n}{n^2}$.
On a
$$\left|\frac{\sin n}{n^2}\right| \leqslant \frac{1}{n^2}$$
donc
$$\frac{\sin n}{n^2} = O\left(\frac{1}{n^2}\right)$$
Or $\sum \frac{1}{n^2}$ converge et $\frac{1}{n^2} \geqslant 0$ donc, par domination, $\sum \frac{\sin n}{n^2}$ converge absolument et donc converge.

9.3.5 Séries et règles de référence

9.3.5.1 Séries de Riemann

Soit $\alpha \in \mathbb{R}$.

Théorème

La série à termes positifs $\sum_{n \geqslant 1} \frac{1}{n^\alpha}$ converge si, et seulement si, $\alpha > 1$.

dém. :
Cas $\alpha \leqslant 1$
Puisque pour tout $n \geqslant 1$,
$$\frac{1}{n^\alpha} \geqslant \frac{1}{n}$$

Puisque la série $\sum \frac{1}{n}$ diverge, on obtient par comparaison de séries à termes positifs que la série $\sum \frac{1}{n^\alpha}$ diverge.
Cas $\alpha > 1$
$\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ est une série à terme positifs. Nous allons montrer qu'elle converge en observant que ses sommes partielles sont majorées. Puisque la fonction $x \mapsto 1/x^\alpha$ est décroissante sur $]0,+\infty[$, on a pour tout $k \geqslant 2$

$$\frac{1}{k^\alpha} \leqslant \int_{k-1}^{k} \frac{\mathrm{d}t}{t^\alpha}$$

et alors

$$\sum_{k=2}^{n} \frac{1}{k^\alpha} \leqslant \int_1^n \frac{\mathrm{d}t}{t^\alpha} = \left[\frac{-1}{\alpha-1}\frac{1}{t^\alpha}\right]_1^n = \frac{1}{\alpha-1}\left(1-\frac{1}{n^{\alpha-1}}\right)$$

puis

$$S_n = 1 + \sum_{k=2}^{n} \frac{1}{k^\alpha} \leqslant 1 + \frac{1}{\alpha-1} = M$$

Par conséquent la série $\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ converge car c'est une série à termes positifs aux sommes partielles majorées.
□

Exemple $\sum \frac{1}{n^2}$ et $\sum \frac{1}{n^{1,001}}$ convergent alors que $\sum \frac{1}{n}$ et $\sum \frac{1}{\sqrt{n}}$ divergent.

Remarque Puisqu'il s'agit d'une série à termes positifs, il est possible de comparer à $\sum 1/n^\alpha$ pour étudier la nature d'une série numérique.

9.3.5.2 Règles de Riemann

Exemple Nature de $\sum_{n\geqslant 0} \frac{(-1)^n}{n^2-n+1}$

On a $\frac{(-1)^n}{n^2-n+1} \xrightarrow[n\to+\infty]{} 0$ mais ce n'est pas décisif.
Cependant

$$\left|\frac{(-1)^n}{n^2-n+1}\right| \underset{n\to+\infty}{\sim} \frac{1}{n^2}$$

Or $\sum \frac{1}{n^2}$ converge et $\frac{1}{n^2} \geqslant 0$ donc $\sum_{n\geqslant 0} \frac{(-1)^n}{n^2-n+1}$ converge.

Exemple Nature de $\sum_{n\geqslant 1} \left(\tan\frac{1}{n} - \frac{1}{n}\right)$

On sait

$$\tan u \underset{u\to 0}{=} u + \frac{1}{3}u^3 + o(u^3)$$

et donc

$$\tan\frac{1}{n} - \frac{1}{n} \underset{n\to+\infty}{\sim} \frac{1}{3n^3}$$

or $\sum \frac{1}{3n^3}$ converge et $\frac{1}{3n^3} \geqslant 0$ donc $\sum \tan\frac{1}{n} - \frac{1}{n}$ converge.

Exemple Nature de $\sum\limits_{n\geqslant 0} \frac{n+1}{n^2+1}$

On a

$$\frac{n+1}{n^2+1} \underset{n\to+\infty}{\sim} \frac{1}{n}$$

Or $\sum \frac{1}{n}$ diverge et $\frac{1}{n} \geqslant 0$ donc $\sum \frac{n+1}{n^2+1}$ diverge.

Exemple Nature de $\sum\limits_{n\geqslant 0} \mathrm{e}^{-n}$

On a

$$n^2\mathrm{e}^{-n} \xrightarrow[n\to+\infty]{} 0$$

donc

$$\mathrm{e}^{-n} \underset{n\to+\infty}{=} o\left(\frac{1}{n^2}\right)$$

or $\sum \frac{1}{n^2}$ converge et $\frac{1}{n^2} \geqslant 0$ donc $\sum \mathrm{e}^{-n}$ converge absolument et donc converge.

Exemple Nature de $\sum\limits_{n\geqslant 1} \frac{\ln(n)}{n^2+1}$

On a

$$n^{3/2}\frac{\ln(n)}{n^2+1} \underset{n\to+\infty}{\sim} \frac{\ln(n)}{\sqrt{n}} \to 0$$

donc

$$\frac{\ln(n)}{n^2+1} \underset{n\to+\infty}{=} o\left(\frac{1}{n^{3/2}}\right)$$

Or $\sum \frac{1}{n^{3/2}}$ converge et $\frac{1}{n^{3/2}} \geqslant 0$ donc $\sum \frac{\ln(n)}{n^2+1}$ converge absolument puis converge.

Exemple Nature de $\sum\limits_{n\geqslant 1} \frac{1}{\ln(n)}$

On a

$$n \times \frac{1}{\ln n} \xrightarrow[n\to+\infty]{} +\infty$$

donc pour n assez grand,

$$n \times \frac{1}{\ln n} \geqslant 1$$

puis

$$\frac{1}{\ln n} \geqslant \frac{1}{n}$$

Puisque $\sum \frac{1}{n}$ diverge et $\frac{1}{n} \geqslant 0$ la série $\sum_{n \geqslant 1} \frac{1}{\ln n}$ diverge.

Exemple Nature de $\sum_{n \geqslant 1} \frac{1}{d_n^2}$
avec d_n le nombre de diviseurs positifs de n.
Pour p nombre premier $d_p = 2$.
Puisqu'il y a une infinité de nombre premiers, $(1/d_n^2)$ ne tend pas vers 0 et donc la série diverge grossièrement.

*Bilan :*Les idées récurrentes :
- Si (u_n) ne tend pas vers 0 alors $\sum u_n$ diverge grossièrement ;
- Si $u_n \sim C/n^\alpha$ (avec $C \neq 0$) alors

$$\sum u_n \text{ converge si, et seulement si, } \alpha > 1\,;$$

- Si on détermine $\alpha > 1$ tel que $n^\alpha u_n \to 0$ alors $u_n = o\left(1/n^\alpha\right)$ et donc $\sum u_n$ converge absolument ;
- Si $nu_n \to \ell \neq 0$ alors $\sum u_n$ diverge.

9.3.5.3 Séries géométriques

Théorème

Soit $q \in \mathbb{C}$.
Si $|q| \geqslant 1$ alors $\sum q^n$ diverge grossièrement.
Si $|q| < 1$ alors $\sum q^n$ converge absolument et

$$\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}$$

dém. :
Cas $|q| \geqslant 1$:
On a $|q^n| = |q|^n \geqslant 1$ donc la suite (q^n) ne tend par vers 0. Il y a divergence grossière.
Cas $|q| < 1$:

$$\sum_{k=0}^{n} |q|^k = \frac{1-|q|^{n+1}}{1-|q|} \to \frac{1}{1-|q|}$$

donc $\sum q^n$ converge absolument.
De plus

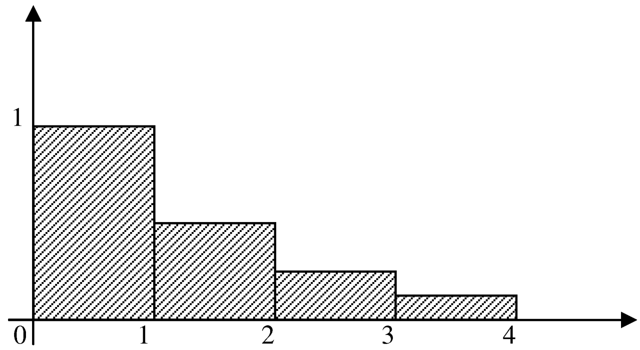
$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \to \frac{1}{1-q}$$

donc

$$\sum_{k=0}^{+\infty} q^k = \frac{1}{1-q}$$

□

Exemple $\displaystyle\sum_{n=0}^{+\infty} \frac{1}{2^n} = 1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n} + \cdots = 2.$



Exemple Pour $|x| < 1$,

$$\sum_{k=0}^{+\infty} (-1)^k x^{2k} = \frac{1}{1+x^2}$$

Exemple Pour $|z| < 1$,

$$\sum_{n=0}^{+\infty} (-1)^n z^n = \frac{1}{1+z}$$

9.3.5.4 Règle de d'Alembert

Théorème

Soit $\sum u_n$ une série à termes non nuls.
On suppose

$$\left|\frac{u_{n+1}}{u_n}\right| \to \ell \in \mathbb{R}^+ \cup \{+\infty\}$$

Si $\ell > 1$ alors $\sum u_n$ diverge grossièrement.
Si $\ell < 1$ alors $\sum u_n$ est absolument convergente.
Si $\ell = 1$ alors on ne peut rien conclure.

dém. :
Cas $\ell > 1$:
A partir d'un certain rang n_0

$$|u_{n+1}/u_n| \geqslant 1$$

et donc la suite $(|u_n|)_{n\geqslant n_0}$ est croissante. Elle ne peut alors converger vers 0 que si elle est constante égale à 0 ce qui est exclu.
Cas $\ell < 1$:
Soit $\varepsilon > 0$ (qu'on fixera par la suite). A partir d'un certain rang n_0,

$$||u_{n+1}/u_n| - \ell| \leqslant \varepsilon$$

et donc

$$|u_{n+1}/u_n| \leqslant \ell + \varepsilon$$

Par récurrence

$$|u_n| \leqslant (\ell+\varepsilon)^{n-n_0} |u_{n_0}| = M(\ell+\varepsilon)^n$$

avec $M = (\ell+\varepsilon)^{-n_0} |u_{n_0}|$. En choisissant initialement $\varepsilon > 0$ pour que $q = \ell + \varepsilon \in [0, 1[$, on a

$$u_n = O(q^n) \text{ avec } q^n \geqslant 0 \text{ et } \sum q^n \text{ converge}$$

On en déduit que $\sum u_n$ converge absolument et donc converge.
Cas $\ell = 1$:
Considérons $u_n = 1/n^\alpha$ avec $\alpha \in \mathbb{R}$.
On a

$$\left|\frac{u_{n+1}}{u_n}\right| \to 1$$

alors que $\sum u_n$ converge si, et seulement si, $\alpha > 1$.
□

Remarque C'est un critère grossier réservé aux suites dont le terme général comporte un produit (terme géométrique, factoriel,...) induisant la nature de la série.

Exemple Nature de $\sum\limits_{n\geqslant 0} u_n$ avec $u_n = 1/\binom{2n}{n}$
On a $u_n = \dfrac{(n!)^2}{(2n)!} > 0$ et

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{u_{n+1}}{u_n} = \frac{(n+1)^2}{(2n+1)(2n+2)} \to \frac{1}{4} < 1$$

donc $\sum\limits_{n\geqslant 0} u_n$ converge absolument puis converge.

9.4 Autres méthodes d'obtention de convergence

9.4.1 Séries alternées

Définition

Une suite réelle (u_n) est dite alternée si

$$\forall n \in \mathbb{N}, u_n = (-1)^n |u_n| \text{ ou } \forall n \in \mathbb{N}, u_n = (-1)^{n+1} |u_n|$$

Une série réelle $\sum u_n$ est dite alternée si la suite (u_n) l'est.

Exemple Les séries $\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}$ et $\sum_{n\geqslant 1} \ln\left(1+\frac{(-1)^{n-1}}{n}\right)$ sont alternées.

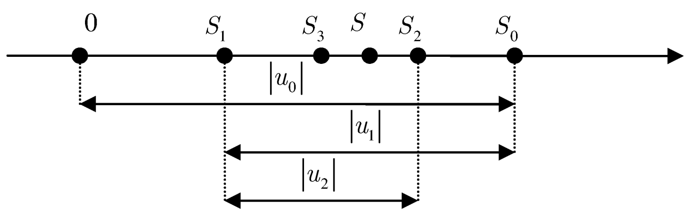
Théorème

Soit $\sum u_n$ une série alternée.
Si la suite $(|u_n|)_{n\geqslant 0}$ décroît vers 0 alors la série $\sum u_n$ est convergente.
De plus, son reste $R_n = \sum_{k=n+1}^{+\infty} u_k$ vérifie :
- R_n est du signe de u_{n+1} ;
- $|R_n| \leqslant |u_{n+1}|$.

dém. :
Quitte à considérer $(-u_n)$, on peut supposer

$$\forall n \in \mathbb{N}, u_n = (-1)^n |u_n|$$

Posons $S_n = \sum_{k=0}^{n} u_k$.



Nous allons établir l'adjacence des suites (S_{2n}) et (S_{2n+1}).

$$S_{2n+2} - S_{2n} = u_{2n+2} + u_{2n+1} = |u_{2n+2}| - |u_{2n+1}| \leqslant 0$$

Ainsi (S_{2n}) est décroissante.

$$S_{2n+3} - S_{2n+1} = u_{2n+3} + u_{2n+2} = -|u_{2n+3}| + |u_{2n+2}| \geqslant 0$$

Ainsi (S_{2n+1}) est croissante.
Enfin

$$S_{2n+1} - S_{2n} = u_{2n+1} \to 0$$

donc les deux suites sont adjacentes.
Par conséquent, elles convergent vers une même limite S.
Ainsi $\sum u_n$ converge et sa somme S est encadrée par les sommes partielles consécutives.
Considérons maintenant le reste

$$R_n = S - S_n$$

$R_{2n} = S - S_{2n}$. Or $S_{2n+1} \leqslant S \leqslant S_{2n}$ donc $R_{2n} \in [u_{2n+1}, 0]$.
$R_{2n+1} = S - S_{2n+1}$. Or $S_{2n+1} \leqslant S \leqslant S_{2n+2}$ donc $R_{2n+1} \in [0, u_{2n+2}]$
□

Corollaire

Le signe de la somme est celui de son premier terme.

dém. :
La somme S de la série est encadrée par $S_0 = u_0$ et $S_1 = u_0 + u_1$. Or $|u_1| \leqslant |u_0|$ donc $u_0 + u_1$ est du signe de u_0 et donc S aussi.
□

Exemple Déterminons la nature de

$$\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{\sqrt{n}}$$

C'est une série alternée.
$\left|\frac{(-1)^{n-1}}{\sqrt{n}}\right| = \frac{1}{\sqrt{n}}$ décroît vers 0 donc $\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{\sqrt{n}}$ converge.

Exemple Déterminons la nature de $\sum_{n\geqslant 2} \frac{(-1)^n}{n^3+1}$
1ère méthode :
C'est une série alternée et $\left|\frac{(-1)^n}{n^3+1}\right| = \frac{1}{n^3+1}$ décroît vers 0 donc $\sum_{n\geqslant 2} \frac{(-1)^n}{n^3+1}$ converge.
2ème méthode :
$\frac{(-1)^n}{n^3+1} = O\left(\frac{1}{n^3}\right)$ et $\sum \frac{1}{n^3}$ converge avec $\frac{1}{n^3} \geqslant 0$ donc $\sum_{n\geqslant 2} \frac{(-1)^n}{n^3+1}$ converge absolument.

9.4.2 Exploitation d'un DA à deux termes

Exemple Déterminons la nature de $\sum_{n\geqslant 1} \frac{(-1)^n}{n+(-1)^{n-1}}$
La série est alternée, mais son terme ne décroît pas en valeur absolue :

n	1	2	3	4	5
$\lvert u_n\rvert$	$1/2$	1	$1/4$	$1/3$	$1/6$

Pour déterminer sa nature, on forme un développement asymptotique à deux termes

$$\frac{(-1)^n}{n+(-1)^{n-1}} = \frac{(-1)^n}{n} + \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)$$

D'une part, la série alternée $\sum \frac{(-1)^n}{n}$ converge en vertu du critère spécial.
D'autre part, les séries $\sum \frac{1}{n^2}$ et $\sum o\left(\frac{1}{n^2}\right)$ convergent absolument.
Par somme, on peut conclure la convergence de la série étudiée

Exemple Déterminons la nature de $\sum_{n\geqslant 1} \ln\left(1+\frac{(-1)^{n-1}}{\sqrt{n}}\right)$
On écrit
$$\ln\left(1+\frac{(-1)^{n-1}}{\sqrt{n}}\right) = \frac{(-1)^{n-1}}{\sqrt{n}} - \frac{1}{2n} + o\left(\frac{1}{n}\right)$$
La série alternée $\sum \frac{(-1)^{n-1}}{\sqrt{n}}$ converge en vertu du critère spécial.
Mais
$$\frac{1}{2n} + o\left(\frac{1}{n}\right) \sim \frac{1}{2n}, \frac{1}{2n} \geqslant 0 \text{ et } \sum \frac{1}{2n} \text{ diverge}$$
donc par comparaison à une série à termes positifs, $\sum \frac{1}{2n} + o\left(\frac{1}{n}\right)$ diverge.
Finalement, par somme, la série $\sum_{n\geqslant 2} \frac{(-1)^n}{\sqrt{n}+(-1)^n}$ diverge.

Remarque Ici
$$\ln\left(1+\frac{(-1)^{n-1}}{\sqrt{n}}\right) \sim \frac{(-1)^{n-1}}{\sqrt{n}}$$
alors que
$$\sum \ln\left(1+\frac{(-1)^{n-1}}{\sqrt{n}}\right) \text{ diverge et } \sum \frac{(-1)^{n-1}}{\sqrt{n}} \text{ converge}$$
Cet exemple illustre que la conservation de la nature d'une série par équivalence des termes est incorrecte si la série n'est pas de signe constant.

9.4.3 Transformation d'Abel

Exemple Déterminons la nature de $\sum_{n\geqslant 1} \frac{\sin(n)}{n}$
On introduit $S_n = \sum_{k=0}^{n} \sin(k)$ de sorte que $\sin(n) = S_n - S_{n-1}$
$$\sum_{n=1}^{N} \frac{\sin(n)}{n} = \sum_{n=1}^{N} \frac{S_n - S_{n-1}}{n} = \sum_{n=1}^{N} \frac{S_n}{n} - \sum_{n=1}^{N} \frac{S_{n-1}}{n}$$
Par translation d'indice,
$$\sum_{n=1}^{N} \frac{\sin(n)}{n} = \sum_{n=1}^{N} \frac{S_n}{n} - \sum_{n=0}^{N-1} \frac{S_n}{n+1}$$

puis

$$\sum_{n=1}^{N}\frac{\sin(n)}{n}=\sum_{n=1}^{N}\frac{S_n}{n(n+1)}-S_0+\frac{S_{N+1}}{N+1}$$

Montrons que (S_n) est bornée.

$$S_n=\sum_{k=0}^{n}\sin(k)=\operatorname{Im}\left(\sum_{k=0}^{n}\mathrm{e}^{ik}\right)=\operatorname{Im}\left(\frac{\mathrm{e}^{i(n+1)}-1}{\mathrm{e}^{i}-1}\right)$$

donc

$$|S_n|\leqslant\left|\frac{1-\mathrm{e}^{i(n+1)}}{1-\mathrm{e}^{i}}\right|\leqslant\frac{2}{|1-\mathrm{e}^{i}|}$$

Puisque (S_n) est bornée, $\frac{S_{n+1}}{n+1}\to 0$ et $\frac{S_n}{n(n+1)}=O\left(\frac{1}{n^2}\right)$ donc $\sum\frac{S_n}{n(n+1)}$ converge absolument et sa somme partielle $\sum_{n=1}^{N}\frac{S_n}{n(n+1)}$ converge quand $n\to+\infty$.

Par opération, on en déduit que la suite de terme général $\sum_{n=1}^{N}\frac{\sin n}{n}$ converge quand $n\to+\infty$ et donc la série $\sum_{n\geqslant 1}\frac{\sin n}{n}$ converge.

On peut aussi montrer que

$$\sum_{n=1}^{+\infty}\frac{\sin n}{n}=\frac{\pi-1}{2}$$

mais c'est une autre histoire...

9.5 Applications

9.5.1 Lien suite-série

Théorème

La suite (u_n) et la série $\sum(u_{n+1}-u_n)$ sont de même nature.

dém. :

On a $S_n=\sum_{k=0}^{n}(u_{k+1}-u_k)=u_{n+1}-u_0$ donc la suite (S_n) converge si, et seulement si, (u_n) converge.

□

Exemple Montrons que la suite de terme général $u_n=\sum_{k=1}^{n}\frac{1}{\sqrt{k}}-2\sqrt{n}$ converge.

Etudions la série $\sum(u_{n+1}-u_n)$.

On a

$$u_{n+1}-u_n=\frac{1}{\sqrt{n+1}}-2\sqrt{n+1}+2\sqrt{n}$$

puis

$$u_{n+1}-u_n=\frac{1}{\sqrt{n}}\frac{1}{\sqrt{1+\frac{1}{n}}}-2\sqrt{n}\sqrt{1+\frac{1}{n}}+2\sqrt{n}$$

Ainsi

$$u_{n+1}-u_n=\frac{1}{\sqrt{n}}\left(1+O\left(1/n\right)\right)-2\sqrt{n}\left(1+\frac{1}{2n}+O\left(\frac{1}{n^2}\right)\right)+2\sqrt{n}=O\left(\frac{1}{n^{3/2}}\right)$$

La série $\sum\left(u_{n+1}-u_n\right)$ est absolument convergente donc converge puis (u_n) converge.

Exemple Soit (u_n) définie par $u_0=1$ et $u_n=\dfrac{2n}{2n+1}u_{n-1}$ pour $n\geqslant 1$

Montrons qu'il existe $A>0$, tel que $u_n\sim\dfrac{A}{\sqrt{n}}$.

On veut montrer que $v_n=\sqrt{n}u_n$ converge vers un réel >0.

Etudions la série $\sum\left(\ln v_n-\ln v_{n-1}\right)$.

$$\ln v_n-\ln v_{n-1}=\frac{1}{2}\ln\left(\frac{n}{n-1}\right)+\ln\left(\frac{2n}{2n+1}\right)=-\frac{1}{2}\ln\left(1-\frac{1}{n}\right)-\ln\left(1+\frac{1}{2n}\right)=O\left(\frac{1}{n^2}\right)$$

Ainsi $\sum\left(\ln v_n-\ln v_{n-1}\right)$ est absolument convergente donc la suite $(\ln v_n)$ converge.

En posant ℓ sa limite, $v_n\to\mathrm{e}^{\ell}=A>0$ et $u_n\sim\dfrac{A}{\sqrt{n}}$.

9.5.2 La constante d'Euler

Proposition

> La suite de terme général $u_n=\displaystyle\sum_{k=1}^{n}\frac{1}{k}-\ln n$ est convergente.

dém. :

Nous allons étudier la nature de la série de terme général $u_{n+1}-u_n$.

On a

$$u_{n+1}-u_n=\frac{1}{n+1}-\ln\left(1+\frac{1}{n}\right)=\frac{1}{n+1}-\frac{1}{n}+O\left(\frac{1}{n^2}\right)=O\left(\frac{1}{n^2}\right)$$

donc la série de terme général $u_{n+1}-u_n$ est absolument convergente donc convergente.

□

Définition

> On pose $\gamma=\displaystyle\lim_{n\to+\infty}\left(\sum_{k=1}^{n}\frac{1}{k}-\ln n\right)$ appelée constante d'Euler.
> On a $\gamma=0,577$ à 10^{-3} près.

Théorème

> $$\sum_{k=1}^{n}\frac{1}{k}=\ln n+\gamma+o(1)$$

dém. :
Puisque $u_n \to \gamma$ on peut écrire $u_n = \gamma + o(1)$ donc $\displaystyle\sum_{k=1}^{n} \frac{1}{k} - \ln n = \gamma + o(1)$

Cor : $\displaystyle\sum_{k=1}^{n} \frac{1}{k} \sim \ln n$

□

Exemple Calculons $\displaystyle\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}$
On peut affirmer que cette série alternée converge en vertu du critère spécial.

$$S_{2n} = \sum_{k=1}^{2n} \frac{(-1)^{k-1}}{k} = \left(1 + \frac{1}{3} + \cdots + \frac{1}{2n-1}\right) - \left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right)$$

donc

$$S_{2n} = \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{2n-1} + \frac{1}{2n}\right) - 2\left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right)$$

puis

$$S_{2n} = \sum_{k=1}^{2n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k} = \ln(2n) + \gamma - \ln n - \gamma + o(1) = \ln 2 + o(1)$$

Par suite

$$\sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} = \ln 2$$

9.5.3 Produit infini

Pour étudier l'existence de $\displaystyle\lim_{n\to+\infty} \prod_{k=0}^{n} u_k$, on passe au logarithme si le contexte le permet

Exemple Etudions l'existence de $\displaystyle\lim_{n\to+\infty} \prod_{k=1}^{n} \left(1 + \frac{(-1)^{k-1}}{k}\right)$

Pour tout $n \geqslant 1$, $1 + \dfrac{(-1)^{n-1}}{n} > 0$ donc

$$\ln \prod_{k=1}^{n} \left(1 + \frac{(-1)^{k-1}}{k}\right) = \sum_{k=1}^{n} \ln\left(1 + \frac{(-1)^{k-1}}{k}\right)$$

or

$$\ln\left(1 + \frac{(-1)^{k-1}}{k}\right) = \frac{(-1)^{k-1}}{k} - \frac{1}{2}\frac{1}{k^2} + o\left(\frac{1}{k^2}\right)$$

$\displaystyle\sum_{k\geqslant 1} \frac{(-1)^{k-1}}{k}$ est convergente et

$$\frac{1}{2}\frac{1}{k^2} + o\left(\frac{1}{k^2}\right) \sim \frac{1}{2}\frac{1}{k^2}$$

Par équivalence de série à termes positifs, la série $\sum \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)$ converge et donc

$$\sum_{k=1}^{n} \ln\left(1 + \frac{(-1)^{k-1}}{k}\right)$$

converge quand $n \to +\infty$. En posant ℓ sa limite, on a

$$\prod_{k=1}^{n} \left(1 + \frac{(-1)^{k-1}}{k}\right) \xrightarrow[n\to+\infty]{} \mathrm{e}^{\ell} > 0$$

Exemple Soit $\alpha, x \in \mathbb{R}$ avec $|\alpha| < 1$.
Etudions l'existence de la limite de $P_n(x) = \prod_{k=1}^{n} \left(1 - \alpha^k x\right)$ quand $n \to +\infty$.
Les premiers facteurs du produit ne sont pas nécessairement strictement positifs, mais puisque $1 - \alpha^k x \xrightarrow[k\to+\infty]{} 1$, il existe $N \in \mathbb{N}$ tel que

$$\forall k \geqslant N, 1 - \alpha^k x > 0$$

Pour $n \geqslant N$, on peut écrire

$$P_n(x) = P_{N-1}(x) \prod_{k=N}^{n} \left(1 - \alpha^k x\right)$$

Or

$$\ln\left[\prod_{k=N}^{n} \left(1 - \alpha^k x\right)\right] = \sum_{k=N}^{n} \ln\left(1 - \alpha^k x\right)$$

et

$$\ln\left(1 - \alpha^k x\right) \sim -\alpha^k x \text{ car } \alpha^k x \to 0$$

Puisque $|\alpha| < 1$, la série géométrique $\sum \alpha^n$ converge et, par équivalence de série à termes de signe constant, la série $\sum \ln\left(1 - \alpha^k x\right)$ converge. Ainsi

$$\sum_{k=N}^{n} \ln\left(1 - \alpha^k x\right) \xrightarrow[n\to+\infty]{} \ell$$

puis

$$P_n(x) \xrightarrow[n\to+\infty]{} P_N(x)\mathrm{e}^{\ell}$$

9.5.4 Musculation : séries de Bertrand

Théorème

Soit $(\alpha, \beta) \in \mathbb{R}^2$. On a

$$\sum \frac{1}{n^{\alpha}(\ln n)^{\beta}} \text{ converge si, et seulement si, } \alpha > 1 \text{ ou } (\alpha = 1 \text{ et } \beta > 1)$$

dém. :
Cas $\alpha < 1$:
On a

$$n \times \frac{1}{n^\alpha (\ln n)^\beta} = \frac{n^{1-\alpha}}{(\ln n)^\beta} \to +\infty$$

donc, à partir d'un certain rang,

$$\frac{1}{n^\alpha (\ln n)^\beta} \geqslant \frac{1}{n}$$

Or la série $\sum 1/n$ diverge et $1/n \geqslant 0$ donc la série étudiée diverge.
Cas $\alpha > 1$:
On peut introduire $\rho \in]1, \alpha[$ et on a

$$n^\rho \times \frac{1}{n^\alpha (\ln n)^\beta} = \frac{1}{n^{\alpha-\rho} (\ln n)^\beta} \xrightarrow[n \to +\infty]{} 0$$

donc la série étudiée est de terme général négligeable devant $1/n^\rho$ avec $\rho > 1$. Cette série est donc convergente.
Cas $\alpha = 1$ et $\beta \neq 1$:
Par le théorème des accroissement finis

$$\frac{1}{(\ln(n+1))^{\beta-1}} - \frac{1}{(\ln n)^{\beta-1}} \underset{n \to +\infty}{\sim} \frac{\beta - 1}{n (\ln n)^\beta}$$

et donc la série étudiée converge si, et seulement si, la suite $\left(1/(\ln n)^{\beta-1}\right)$ converge i.e. $\beta > 1$.
Cas $\alpha = 1$ et $\beta = 1$:
On exploite

$$\ln(\ln(n+1)) - \ln(\ln(n)) \underset{n \to +\infty}{\sim} \frac{1}{n \ln(n)}$$

pour conclure que la série étudiée diverge.
□

Chapitre 10

Fonctions réelles

10.1 Limite et continuité

I désigne un intervalle de $\mathbb{R}$.

10.1.1 Définitions quantifiées

10.1.1.1 Limite en $a \in \mathbb{R}$

Soit a un élément de I ou une extrémité finie de I.

Définition

On dit que $f : I \to \mathbb{R}$ tend vers $\ell \in \mathbb{R}$ en a si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in I, (|x - a| \leqslant \alpha \Rightarrow |f(x) - \ell| \leqslant \varepsilon)$$

On note alors $f \underset{a}{\to} \ell$ ou $f(x) \xrightarrow[x \to a]{} \ell$.

Remarque Cette définition peut être transformée en une définition équivalente en remplaçant :
- $|x - a| \leqslant \alpha$ par $|x - a| < \alpha$;
- $|f(x) - \ell| \leqslant \varepsilon$ par $|f(x) - \ell| < \varepsilon$.

Définition

On dit que $f : I \to \mathbb{R}$ tend vers $+\infty$ en a si

$$\forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in I, (|x - a| \leqslant \alpha \Rightarrow f(x) \geqslant M)$$

On note alors $f \underset{a}{\to} +\infty$ ou $f(x) \xrightarrow[x \to a]{} +\infty$.

Remarque Sous réserve d'existence, on définit aussi la limite à droite de f en a comme étant la limite en a de la restriction $f|_{I \cap]a,+\infty[}$.

10.1.1.2 Limite en $+\infty$

On suppose l'intervalle I non majoré.

Définition

On dit que $f : I \to \mathbb{R}$ tend vers $\ell \in \mathbb{R}$ en $+\infty$ si

$$\forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in I, (x \geqslant A \Rightarrow |f(x) - \ell| \leqslant \varepsilon)$$

Définition

On dit que $f : I \to \mathbb{R}$ tend vers $+\infty$ en $+\infty$ si

$$\forall M \in \mathbb{R}, \exists A \in \mathbb{R}, \forall x \in I, (x \geqslant A \Rightarrow f(x) \geqslant M)$$

Remarque Dans les cas « simples » une limite s'obtient :
- par opérations, quitte à lever des indéterminations par transformation d'écriture ;
- par comparaison, mais cela nécessite d'avoir parfois l'intuition de la limite à obtenir.

Exemple Etudions la limite quand $x \to +\infty$ de $x - \ln x$.
Quand $x \to +\infty$,

$$x - \ln x = x\left(1 - \frac{\ln x}{x}\right) \to +\infty$$

car par limite de référence $\dfrac{\ln x}{x} \to 0$.

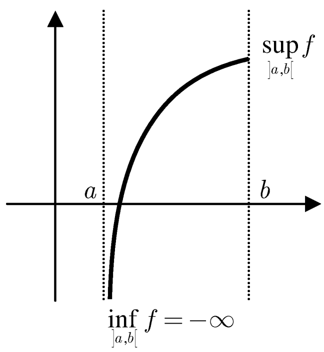
Attention : Ne pas rédiger $\lim \ldots = \lim \ldots = \ldots$.

10.1.1.3 Théorème de la limite monotone

Théorème

Soit $a < b \in \bar{\mathbb{R}}$. Si $f :]a,b[\to \mathbb{R}$ est monotone alors f admet des limites en a^+ et b^- qui sont

$$\inf_{]a,b[} f \text{ et } \sup_{]a,b[} f$$



Remarque Cet outil permet, entre autres, de calculer le sup et l'inf d'une fonction réelle à partir de son tableau de variation.

10.1.2 Continuité

Remarque Si $f : I \to \mathbb{R}$ admet une limite en $a \in I$ celle-ci est nécessairement égale à $f(a)$.

Définition

Une fonction $f : I \to \mathbb{R}$ est dite continue en $a \in I$ si $f(x) \xrightarrow[x \to a]{} f(a)$.
Une fonction $f : I \to \mathbb{R}$ est dite continue si elle l'est en tout $a \in I$.

Remarque Usuellement, la continuité d'une fonction s'obtient par argument d'opérations sur les fonctions continues.

Exemple Si $f, g : I \to \mathbb{R}$ sont continues alors la fonction $\sup(f, g) : x \mapsto \max(f(x), g(x))$ l'est aussi.
En effet, on remarque

$$\max(a, b) = \frac{1}{2}(a + b + |a - b|)$$

donc

$$\sup(f, g) = \frac{1}{2}(f + g + |f - g|)$$

est continue par opérations sur les fonctions continues.
En particulier, si $f : I \to \mathbb{R}$ est continue alors les fonctions $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0)$ le sont aussi.

Exemple Etudions la continuité de $f : \mathbb{R} \to \mathbb{R}$ définie par

$$f(x) = \begin{cases} \mathrm{e}^{-1/x} & \text{si } x > 0 \\ 0 & \text{si } x \leqslant 0 \end{cases}$$

Soit $a \in \mathbb{R}$.
Cas $a < 0$:
Au voisinage de a, $f(x) = 0$ et donc f est continue en a.
Cas $a > 0$:
Au voisinage de a, $f(x) = \mathrm{e}^{-1/x}$ et donc f est continue en a.
Cas $a = 0$.
Quand $x \to 0^+$, $f(x) = \mathrm{e}^{-1/x} \to 0 = f(0)$ et quand $x \to 0^-$, $f(x) = 0 \to 0 = f(0)$.
Ainsi f est aussi continue en 0 et finalement f est continue sur $\mathbb{R}$.

10.1.3 Théorème des valeurs intermédiaires

Théorème

> L'image d'un intervalle par une fonction continue est un intervalle.
> En particulier, une fonction continue prend toutes les valeurs comprises entre deux valeurs déjà prises.

Exemple Soit $f : [a, b] \to \mathbb{R}$ continue. On suppose

$$\forall x \in [a, b] , f(x) \in [a, b]$$

Montrons qu'il existe $x \in [a, b]$ tel que $f(x) = x$.
On introduit $\varphi(x) = f(x) - x$.
La fonction φ est continue par opérations sur les fonctions continues.
$\varphi(a) = f(a) - a \geqslant 0$ car $f(a) \in [a, b]$ et $\varphi(b) = f(b) - b \leqslant 0$ car $f(b) \in [a, b]$.
Par le théorème des valeurs intermédiaires, φ s'annule ce qui établit

$$\exists x \in [a, b] , f(x) = x$$

10.1.4 Théorème de la borne atteinte

Théorème

> Toute fonction continue sur un segment $[a, b]$ admet un minimum et un maximum.
> On dit qu'elle est bornée et atteint ses bornes.

Exemple Soit $f : [0, +\infty[\to \mathbb{R}$ continue. On suppose que $\ell = \lim\limits_{+\infty} f$ existe dans $\mathbb{R}$.
Montrons que f est bornée.
Pour $\varepsilon = 1$, il existe $A \in \mathbb{R}^+$ tel que pour tout $x \geqslant A$, $|f(x) - \ell| \leqslant 1$ et donc

$$|f(x)| \leqslant 1 + |\ell|$$

Ainsi f est bornée sur $[A, +\infty[$.
Sur $[0, A]$, f est continue sur un segment donc bornée.
Au final, la fonction f est bornée sur $\mathbb{R}^+$.

10.1.5 Théorème de la bijection continue strictement monotone

Théorème

> Si $f : I \to \mathbb{R}$ est continue et strictement monotone alors f réalise une bijection de I vers un intervalle J dont les extrémités sont les limites de f aux extrémités de I.
> De plus $f^{-1} : J \to I$ est continue, de même stricte monotonie que f.

Remarque Inversement, si $f : I \to J$ est une bijection continue, celle-ci est nécessairement strictement monotone et sa bijection réciproque est continue.

Exemple Etudions les bijections induites par $f : x \in \mathbb{R}^+ \mapsto x - 2\sqrt{x}$.
f est continue sur $\mathbb{R}^+$, dérivable sur $]0, +\infty[$ et $f'(x) = 1 - 1/\sqrt{x}$.

x	0		1		$+\infty$
$f'(x)$	$\Vert$	$-$	0	$+$	
$f(x)$	0	$\searrow$	-1	$\nearrow$	$+\infty$

Considérons $\varphi = f_{\restriction [1,+\infty[}$.
$\varphi'(x) > 0$ sauf pour $x = 1$ donc réalise une bijection de $[1, +\infty[$ vers $[-1, +\infty[$.

	1		$+\infty$
φ	1	$\nearrow$	$+\infty$

,

	-1		$+\infty$
φ^{-1}	1	$\nearrow$	$+\infty$

Considérons $\psi = f_{\restriction [0,1]}$.
$\psi'(x) < 0$ sauf pour $x = 0$ ou 1 donc ψ réalise une bijection de $[0, 1]$ vers $[-1, 0]$.

	1		0
ψ	0	$\searrow$	-1

,

	-1		0
ψ^{-1}	1	$\searrow$	0

10.2 Dérivation

I et J désignent des intervalles contenant chacun au moins deux points.

10.2.1 Nombre dérivé

Définition

On dit que $f : I \to \mathbb{R}$ est dérivable en $a \in I$ si le taux d'accroissement

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

admet une limite finie quand $h \to 0$ (avec $h \neq 0$). Cette limite est notée $f'(a)$.

Définition

On dit que $f : I \to \mathbb{R}$ est dérivable si elle est dérivable en tout $a \in I$; on peut alors introduire sa fonction dérivée

$$f' : I \to \mathbb{K}$$

Définition

On dit que $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^1$ si f est dérivable et si de surcroît sa dérivée est continue.

10.2.2 Théorème de Rolle

Théorème

Soit $a < b \in \mathbb{R}$, $f : [a, b] \to \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$.
Si $f(a) = f(b)$ alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.

dém. :
f est continue sur le segment $[a, b]$ donc f admet des extremums en $c, d \in [a, b]$

$$\forall x \in [a, b], f(c) \leqslant f(x) \leqslant f(d)$$

Si $f(c) = f(d)$ alors f est constante.
Sinon, l'un au moins des extremums de f n'est ni en a, ni en b et la fonction f' s'y annule
□

Exemple Soit $f : I \to \mathbb{R}$ une fonction n fois dérivable.
On suppose que f s'annule au moins $n+1$ fois. Montrons qu'il existe $c \in I$ tel que $f^{(n)}(c) = 0$.

dém. :
Introduisons $a_0 < a_1 < \ldots < a_n$ les valeurs d'annulation de f ordonnées.
Pour $i \in [\![1, n]\!]$, f est continue sur $[a_{i-1}, a_i]$, dérivable sur $]a_{i-1}, a_i[$ et $f(a_{i-1}) = f(a_i)$ donc par le théorème de Rolle, il existe $b_i \in]a_{i-1}, a_i[$ tel que $f'(b_i) = 0$.
Puisque

$$a_0 < b_1 < a_1 < b_2 < \ldots < b_n < a_n$$

les $b_1, \ldots, b_n$ sont deux à deux distincts. Ainsi f' s'annule n fois au moins.
En itérant ce processus, f'' s'annule $n-1$ fois au moins,..., $f^{(n)}$ s'annule 1 fois au moins.
□

Exemple Soit $U_n(X) = \left((X^2-1)^n\right)^{(n)}$. Montrons que U_n possède exactement n racines distinctes, toutes dans $]-1, 1[$.
Posons

$$P_n(X) = (X^2-1)^n = (X-1)^n(X+1)^n$$

1 et -1 sont racines de multiplicité n de P_n.
1 et -1 sont donc racines de $P_n, P_n', \ldots, P_n^{(n-1)}$.
En appliquant successivement le théorème de Rolle avec appui sur 1 et -1, on montre que pour tout $k \in [\![1, n]\!]$, $P_n^{(k)}$ admet au moins k racines dans $]-1, 1[$.
En particulier $U_n = P_n^{(n)}$ admet au moins $n = \deg U_n$ racines dans $]-1, 1[$. On en déduit que celles-ci sont simples et qu'il n'y en a pas d'autres.

10.2.3 Théorème des accroissements finis

Théorème

> Soit $a < b \in \mathbb{R}$, $f : [a, b] \to \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$.
> Il existe $c \in]a, b[$ tel que
>
> $$f(b) - f(a) = f'(c)(b-a)$$

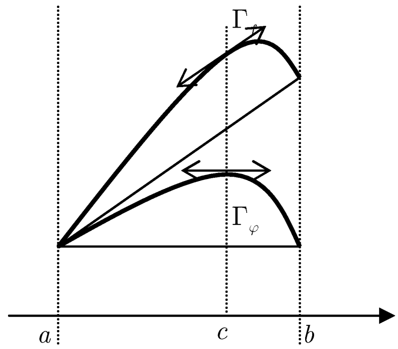
dém. :
Posons $K \in \mathbb{R}$ tel que

$$f(b) - f(a) = K(b-a)$$

i.e. K déterminé par

$$K = \frac{f(b) - f(a)}{b-a}$$

et introduisons $\varphi : x \mapsto f(x) - K(x-a)$.
φ est continue sur $[a, b]$, dérivable sur $]a, b[$ et $\varphi(a) = f(a) = \varphi(b)$.
Par application théorème de Rolle, il existe $c \in]a, b[$ vérifiant $\varphi'(c) = 0$ i.e. $f'(c) = K$.



□

Exemple Soit $f : [a, b] \to \mathbb{R}$ de classe $\mathcal{C}^2$ et g la fonction affine prenant les mêmes valeurs que f en a et b.
Montrons

$$\forall x_0 \in]a, b[\,, \exists c \in]a, b[\,, f(x_0) - g(x_0) = \frac{(x_0 - a)(x_0 - b)}{2} f''(c)$$

Cette identité est intéressant car elle permet de mesurer l'erreur commise lorsqu'on remplace $f(x)$ par $g(x)$ (comme dans la méthode d'intégration des trapèzes).

$$g(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a)$$

Posons $K \in \mathbb{R}$ tel que

$$f(x_0) = g(x_0) + \frac{(x_0 - a)(x_0 - b)}{2} K$$

i.e.

$$K = 2\frac{f(x_0) - g(x_0)}{(x_0 - a)(x_0 - b)}$$

Considérons la fonction

$$\varphi : x \mapsto f(x) - g(x) - \frac{(x - a)(x - b)}{2} K$$

La fonction φ est de classe $\mathcal{C}^2$ et s'annule en x_0, a, b.
Par application du théorème de Rolle, il existe $c \in]a, b[$ vérifiant $\varphi''(c) = 0$ i.e. $f''(c) = K$.

10.2.4 Inégalité des accroissements finis

Théorème

Soit $f : I \to \mathbb{R}$ dérivable et $M \in \mathbb{R}^+$. On a équivalence entre :
(i) $\forall x \in I, |f'(x)| \leqslant M$;
(ii) f est M lipschitzienne i.e.

$$\forall x, y \in I, |f(y) - f(x)| \leqslant M\,|y - x|$$

Exemple Si $f : [a, b] \to \mathbb{R}$ est de classe $\mathcal{C}^1$ alors f est lipschitzienne.
En effet, la fonction $|f'|$ est continue sur un segment donc bornée.

10.2.5 Théorème de la limite de la dérivée

Théorème

Soit $f : I \to \mathbb{R}$ et $a \in I$.
On suppose f continue sur I et dérivable sur $I \setminus \{a\}$.
Si $f'(x) \xrightarrow[x \to a, x \neq a]{} \ell \in \mathbb{R}$ alors f est dérivable en a et $f'(a) = \ell$.
Si $f'(x) \xrightarrow[x \to a, x \neq a]{} +\infty$ alors f n'est pas dérivable en a, mais présente une tangente verticale en a.

dém. :
Supposons $f'(x) \xrightarrow[x \to a, x \neq a]{} \ell \in \bar{\mathbb{R}}$.
Pour $h \neq 0$, on étudie le taux d'accroissement

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

Par le théorème des accroissements finis, il existe c_h compris entre a et $a + h$ tel que

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = f'(c_h)$$

Quand $h \to 0$ (avec $h \neq 0$), par encadrement $c_h \to a$ et par composition de limites

$$\frac{1}{h}\left(f(a+h) - f(a)\right) \to \ell$$

□

Corollaire

Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^k$ sur $I \setminus \{a\}$.
Si $f^{(i)}(x)$ possède une limite finie quand $x \to a$ pour chaque $i \in \{0, \ldots, k\}$ alors f admet un prolongement de classe $\mathcal{C}^k$ sur I.

Exemple Soit $f : \mathbb{R}^\star \to \mathbb{R}$ définie par $f(x) = \dfrac{\sin x}{x}$
Montrer que f se prolonge en une fonction de classe $\mathcal{C}^1$.

$$f(x) = \frac{\sin x}{x} \underset{x \to 0}{\sim} \frac{x}{x} \to 1$$

$$f'(x) = \frac{x \cos x - \sin x}{x^2} \underset{x \to 0}{=} \frac{o(x^2)}{x^2} \to 0$$

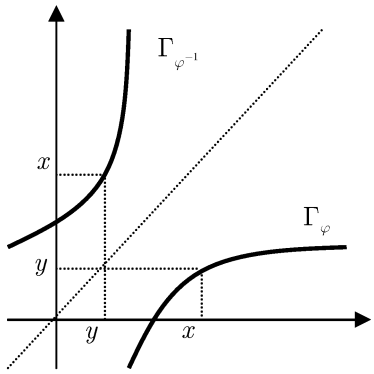
On peut donc prolonger f une fonction de classe $\mathcal{C}^1$ sur $\mathbb{R}$ en posant $f(0) = 1$.

10.2.6 Dérivation de bijection réciproque

Théorème

Soit $\varphi : I \to J$ une bijection continue et $x \in I$.
Si φ est dérivable en x et si $\varphi'(x) \neq 0$ alors φ^{-1} est dérivable en $y = \varphi(x)$ et

$$(\varphi^{-1})'(y) = \frac{1}{\varphi'(x)}$$



Corollaire

Si φ est dérivable et si φ' ne s'annule pas alors φ^{-1} est dérivable et

$$(\varphi^{-1})' = \frac{1}{\varphi' \circ \varphi^{-1}}$$

Remarque Cette formule de dérivation peut être retrouvée en dérivant la relation

$$\varphi \circ \varphi^{-1} = \text{Id}$$

Corollaire

Si φ est de classe $\mathcal{C}^n$ et si φ' ne s'annule pas alors φ^{-1} est de classe $\mathcal{C}^n$.

Exemple C'est ce résultat qui a fourni les dérivées suivantes

$$\frac{\mathrm{d}}{\mathrm{d}x}(\arcsin x) = \frac{1}{\sqrt{1-x^2}} \text{ et } \frac{\mathrm{d}}{\mathrm{d}x}(\arctan x) = \frac{1}{1+x^2}$$

Exemple Etudions la bijection réciproque de $f : \mathbb{R}^+ \to \mathbb{R}$ définie par $f(x) = \sqrt{x} + x + 1$.
f réalise une bijection de $\mathbb{R}^+$ sur $[1, +\infty[$ car c'est une fonction continue, strictement croissante (par opérations sur de telles fonctions) vérifiant $f(0) = 1$ et $\lim\limits_{+\infty} f = +\infty$.

La fonction f est dérivable sur $\mathbb{R}^{+\star}$ et

$$\forall x > 0, f'(x) = \frac{1}{2\sqrt{x}} + 1 \neq 0$$

Par le théorème précédent, on peut affirmer que son application réciproque f^{-1} est dérivable sur

$$f(\mathbb{R}^{+\star}) =]1, +\infty[$$

Etude de la dérivabilité en 1.
Quand $h \to 0$ (avec $h \neq 0$),

$$\frac{1}{h}\left(f^{-1}(1+h) - f^{-1}(1)\right) = \frac{1}{h} f^{-1}(1+h) \underset{x=f^{-1}(1+h)}{=} \frac{x}{f(x)-1} = \frac{x}{\sqrt{x}+x} \sim \frac{x}{\sqrt{x}} = \sqrt{x} \to 0$$

Ainsi f^{-1} est dérivable en 1 et $(f^{-1})'(1) = 0$.
Cela pouvait être attendu car la fonction f admet une tangente verticale en 0.

10.3 Intégration

I désigne un intervalle de $\mathbb{R}$ contenant au moins deux points.

10.3.1 Intégrale

Définition

Une fonction $f : [a, b] \to \mathbb{R}$ est dite continue par morceaux s'il existe un découpage

$$a_0 = a < a_1 < \cdots < a_n = b$$

vérifiant, pour tout $i \in \{1, \ldots, n\}$:
- f est continue sur $]a_{i-1}, a_i[$;
- f admet des limites finies en a_{i-1}^+ et a_i^-.
Une fonction $f : I \to \mathbb{R}$ est dite continue par morceaux si elle l'est sur tout segment $[a, b]$ inclus dans I.

Définition

Pour $f : I \to \mathbb{R}$ continue par morceaux et $a, b \in I$, il a été donné en première année un sens à l'intégrale

$$\int_a^b f(t)\,\mathrm{d}t$$

Exemple Calculons $\displaystyle\int_0^1 \frac{t+1}{t^2+t+1}\,\mathrm{d}t$.
On a

$$\frac{\mathrm{d}}{\mathrm{d}t}(t^2+t+1) = 2t+1$$

donc on décompose

$$\int_0^1 \frac{t+1}{t^2+t+1}\,\mathrm{d}t = \frac{1}{2}\int_0^1 \frac{2t+1}{t^2+t+1}\,\mathrm{d}t + \frac{1}{2}\int_0^1 \frac{1}{t^2+t+1}\,\mathrm{d}t$$

Or
$$\int_0^1 \frac{2t+1}{t^2+t+1}\,\mathrm{d}t = \left[\ln\left|t^2+t+1\right|\right]_0^1 = \ln 3$$
et
$$\int_0^1 \frac{1}{t^2+t+1}\,\mathrm{d}t = \int_0^1 \frac{1}{(t+1/2)^2+3/4}\,\mathrm{d}t$$
Sachant
$$\int \frac{\mathrm{d}u}{u^2+a^2} = \frac{1}{a}\arctan\frac{u}{a}$$
avec ici $u = t+1/2$ et $a = \sqrt{3}/2$, on obtient directement
$$\int_0^1 \frac{1}{t^2+t+1}\,\mathrm{d}t = \left[\frac{2}{\sqrt{3}}\arctan\frac{2t+1}{\sqrt{3}}\right]_0^1 = \frac{\pi}{3\sqrt{3}}$$
Finalement
$$\int_0^1 \frac{t+1}{t^2+t+1}\,\mathrm{d}t = \frac{1}{2}\ln 3 + \frac{\pi}{6\sqrt{3}}$$

Exemple Calculons $\displaystyle\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x$.
On réalise le changement de variable $x = \sin t$.
$\mathrm{d}x = \cos t\,\mathrm{d}t$, pour $t=0$, $x=0$ et pour $t=\pi/2$, $x=1$.
$$\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x = \int_0^{\pi/2} \sqrt{1-\sin^2 t}\cos t\,\mathrm{d}t = \int_0^{\pi/2} \cos^2 t\,\mathrm{d}t$$
Or $\cos 2a = 2\cos^2 a - 1$ donc
$$\cos^2 t = \frac{1}{2}(1+\cos 2t)$$
puis
$$\int_0^1 \sqrt{1-x^2}\,\mathrm{d}x = \left[\frac{t}{2}+\frac{\sin 2t}{4}\right]_0^{\pi/2} = \frac{\pi}{4}$$

10.3.2 Calcul des intégrales de Wallis

Exemple Calculons $I_n = \displaystyle\int_0^{\pi/2} \sin^n(t)\,\mathrm{d}t$
(ou encore $\displaystyle\int_0^{\pi/2} \cos^n(u)\,\mathrm{d}u$ via $u = \pi/2 - t$).
Pour $n \geqslant 2$,
$$I_n = \int_0^{\pi/2} \sin t.\sin^{n-1}(t)\,\mathrm{d}t$$
Par intégration par parties,
$$I_n = \left[-\cos t.\sin^{n-1} t\right]_0^{\pi/2} + (n-1)\int_0^{\pi/2} \cos^2(t)\sin^{n-2}(t)\,\mathrm{d}t$$

Or
$$\left[-\cos t.\sin^{n-1} t\right]_0^{\pi/2} = 0$$
et
$$\int_0^{\pi/2} \cos^2(t)\sin^{n-2}(t)\,\mathrm{d}t = \int_0^{\pi/2} (1-\sin^2(t))\sin^{n-2}(t)\,\mathrm{d}t = I_n - I_{n-2}$$
donc
$$I_n = (n-1)(I_n - I_{n-2})$$
puis enfin
$$I_n = \frac{n-1}{n} I_{n-2}$$
Par cette relation de récurrence, il est possible d'exprimer I_n en fonction de I_1 ou de I_0 selon la parité de n.
Cas n impair : $n = 2p+1$.
$$I_{2p+1} = \frac{2p}{2p+1} I_{2p-1} = \frac{2p}{2p+1}\frac{2p-2}{2p-1} I_{2p-3} = \dots$$
A terme
$$I_{2p+1} = \frac{2p}{2p+1}\frac{2p-2}{2p-1}\cdots\frac{2}{3} I_1$$
Or
$$2p(2p-2)\dots 2 = 2^p p! \text{ et } (2p+1)(2p-1)\dots 3 = \frac{(2p+1)!}{2^p p!}$$
De plus
$$I_1 = \int_0^{\pi/2} \sin(t)\,\mathrm{d}t = 1$$
donc
$$I_{2p+1} = \frac{(2^p p!)^2}{(2p+1)!}$$
Cas n pair : $n = 2p$. De façon analogue
$$I_{2p} = \frac{(2p)!}{(2^p p!)^2}\frac{\pi}{2}$$

10.3.3 Intégrale fonction de sa borne supérieure

Théorème

Si $f : I \to \mathbb{R}$ est continue, pour $a \in I$, l'application
$$x \mapsto \int_a^x f(t)\,\mathrm{d}t$$
est l'unique primitive de f s'annulant en a.

Remarque On a donc la formule de dérivation
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x)$$

Remarque On ne peut pas exprimer les primitives des fonctions suivantes à l'aide des fonctions usuelles

$$t \mapsto \mathrm{e}^{-t^2}, t \mapsto \frac{\sin t}{t}, t \mapsto \frac{\cos t}{t}, t \mapsto \frac{\mathrm{e}^t}{t}, t \mapsto \frac{1}{\ln t}, \ldots$$

Cependant celles-ci existent car toute fonction continue sur un intervalle y admet des primitives en vertu du résultat précédent.

Corollaire

Si $f : I \to \mathbb{R}$ est continue et si F est une primitive de f alors

$$\forall a, b \in I,\ \int_a^b f = [F]_a^b$$

Proposition

Soit $a < b$ et $f : [a, b] \to \mathbb{R}$

Si f est continue et si $\int_a^b f(t)\,\mathrm{d}t = 0$ alors f s'annule.

dém. :

En introduisant F une primitive de f, la relation $\int_a^b f(t)\,\mathrm{d}t = 0$ donne $F(a) = F(b)$ et le théorème de Rolle permet de conclure que $F' = f$ s'annule.

□

Proposition

Soit $a < b$ et $f : [a, b] \to \mathbb{R}$.

Si f est continue, $f \geqslant 0$ et si $\int_a^b f(t)\,\mathrm{d}t = 0$ alors $f = \tilde{0}$.

dém. :

On introduit F une primitive de f. Puisque $F' = f \geqslant 0$, on a F croissante et $\int_a^b f(t)\,\mathrm{d}t = 0$ donne $F(a) = F(b)$ et donc F est constante. On en déduit que $f = F' = 0$.

□

Exemple Etudions sur $]1, +\infty[$ la fonction $\varphi : x \mapsto \int_x^{x^2} \frac{\mathrm{d}t}{\ln t}$.

Définition :

La fonction $t \mapsto \frac{1}{\ln t}$ est définie et continue par morceaux sur $]1, +\infty[$ et

$$\forall x > 1, x, x^2 \in \,]1, +\infty[$$

Par suite $\varphi(x)$ est bien définie pour tout $x > 1$.

Variation :

Puisque $t \mapsto \frac{1}{\ln t}$ est continue sur $]1, +\infty[$, elle y admet une primitive de F et alors

$$\varphi(x) = F(x^2) - F(x)$$

Puisque F est de classe $\mathcal{C}^1$, φ l'est aussi et

$$\varphi'(x) = 2xF'(x) - F'(x) = \frac{x-1}{\ln x} \geqslant 0$$

Ainsi φ est croissante.
Limite en $+\infty$:
Quand $x \to +\infty$. Pour $t \in \left[x, x^2\right]$,

$$\frac{1}{2\ln x} \leqslant \frac{1}{\ln t} \leqslant \frac{1}{\ln x}$$

En intégrant,

$$\frac{1}{2}\frac{x^2 - x}{\ln x} \leqslant \varphi(x) \leqslant \frac{x^2 - x}{\ln x}$$

Or

$$\frac{x^2 - x}{\ln x} \sim \frac{x^2}{\ln x} \to +\infty$$

donc $\varphi(x) \to +\infty$.
Limite en 1^+ :
Quand $x \to 1^+$. Pour $t \in \left[x, x^2\right]$,

$$\frac{x}{t\ln t} \leqslant \frac{1}{\ln t} = \frac{t}{t\ln t} \leqslant \frac{x^2}{t\ln t}$$

En intégrant

$$x\int_x^{x^2} \frac{\mathrm{d}t}{t\ln t} \leqslant \varphi(x) \leqslant x^2\int_x^{x^2} \frac{\mathrm{d}t}{t\ln t}$$

Or

$$\int_x^{x^2} \frac{\mathrm{d}t}{t\ln t} = \left[\ln|\ln t|\right]_x^{x^2} = \ln 2$$

donc $\varphi(x) \to \ln 2$.
Finalement, on obtient le tableau de variation suivant

x	1		$+\infty$
$\varphi(x)$	$\ln 2$	$\nearrow$	$+\infty$

10.3.4 Formules de Taylor

10.3.4.1 Avec reste intégrale

Remarque On peut exprimer $f : I \to \mathbb{R}$ de classe $\mathcal{C}^1$ par sa dérivée avec la formule

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t$$

On peut généraliser :

Théorème

Soit $f : I \to \mathbb{R}$ et $a \in I$.
Si f est de classe $\mathcal{C}^{n+1}$ alors pour tout $x \in I$

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t$$

Remarque Par le changement de variable $x = a + \lambda(x-a)$, le reste intégrale se réécrit

$$(x-a)^{n+1} \int_0^1 \frac{(1-\lambda)^n}{n!} f^{(n+1)}(a+\lambda(x-a))\,\mathrm{d}u$$

Cette écriture révèle l'ordre de grandeur du reste intégrale...

10.3.4.2 Inégalité de Taylor-Lagrange

Remarque L'inégalité des accroissements finis donne

$$\forall x \in I, |f'(x)| \leqslant M \Rightarrow \forall a, x \in I, |f(x) - f(a)| \leqslant M\,|x-a|$$

On généralise :

Théorème

Soit $f : I \to \mathbb{K}$ et $M \in \mathbb{R}^+$.
Si f est de classe $\mathcal{C}^{n+1}$ et si

$$\forall x \in I, \left|f^{(n+1)}(x)\right| \leqslant M$$

alors pour chaque $a, x \in I$

$$\left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leqslant \frac{|x-a|^{n+1}}{(n+1)!} M$$

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^2$ telle que f et f'' soient bornées. On pose $M_0 = \sup |f|$ et $M_2 = \sup |f''|$.
Montrons que f' est bornée et

$$M_1 = \sup |f'| \leqslant 2\sqrt{M_0 M_2}$$

Par l'inégalité de Taylor-Lagrange

$$|f(a+h) - f(a) - hf'(a)| \leqslant \frac{h^2 M_2}{2}$$

On en déduit

$$|hf'(a)| \leqslant 2M_0 + \frac{h^2 M_2}{2}$$

Pour $h > 0$, cela conduit à

$$|f'(a)| \leqslant \frac{2M_0}{h} + \frac{h^2 M_2}{2}$$

La fonction f' est donc bornée et

$$M_1 \leqslant \frac{2M_0}{h} + \frac{h^2 M_2}{2}$$

Cette dernière relation vaut pour tout $h > 0$, il s'agit ensuite de trouver l'optimal. C'est $h = 2\sqrt{M_0/M_2}$ et l'on obtient

$$M_1 \leqslant 2\sqrt{M_0 M_2}$$

10.3.4.3 Formule de Taylor Young

Remarque Lorsqu'une fonction f est dérivable en a, on peut exprimer un développement limité à l'ordre 1

$$f(x) \underset{x\to a}{=} f(a) + f'(a)(x-a) + o\left(x-a\right)$$

Théorème

Si $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^n$ alors f admet un développement limité à l'ordre n en tout $a \in I$ de la forme

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o\left((x-a)^n\right)$$

10.3.4.4 Développements limités usuels

$$\frac{1}{1-u} = 1 + u + u^2 + \cdots + u^n + o(u^n) = \sum_{k=0}^{n} u^k + o(u^n)$$

$$\frac{1}{1+u} = 1 - u + u^2 + \cdots + (-1)^n u^n + o(u^n) = \sum_{k=0}^{n} (-1)^k u^k + o(u^n)$$

$$\ln(1+u) = u - \frac{1}{2}u^2 + \cdots + \frac{(-1)^{n-1}}{n}u^n + o(u^n) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}u^k + o(u^n)$$

$$\mathrm{e}^u = 1 + u + \frac{1}{2}u^2 + \frac{1}{6}u^3 + \cdots + \frac{1}{n!}u^n + o(u^n) = \sum_{k=0}^{n} \frac{1}{k!}u^k + o(u^n)$$

$$(1+u)^\alpha = 1 + \alpha u + \frac{\alpha(\alpha-1)}{2!}u^2 + \cdots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}u^n + o(u^n)$$

$$\cos u = 1 - \frac{1}{2}u^2 + \frac{1}{24}u^4 + \cdots + \frac{(-1)^n}{(2n)!}u^{2n} + o(u^{2n+1}) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!}u^{2k} + o(u^{2n+1})$$

$$\sin u = u - \frac{1}{6}u^3 + \frac{1}{120}u^5 + \cdots + \frac{(-1)^n}{(2n+1)!}u^{2n+1} + o(u^{2n+2}) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!}u^{2k+1} + o(u^{2n+2})$$

$$\mathrm{ch} u = 1 + \frac{1}{2}u^2 + \frac{1}{24}u^4 + \cdots + \frac{1}{(2n)!}u^{2n} + o(u^{2n+1}) = \sum_{k=0}^{n} \frac{1}{(2k)!}u^{2k} + o(u^{2n+1})$$

$\mathrm{sh}u = u + \dfrac{1}{6}u^3 + \cdots + \dfrac{1}{(2n+1)!}u^{2n+1} + o(u^{2n+2}) = \displaystyle\sum_{k=0}^{n} \dfrac{1}{(2k+1)!}u^{2k+1} + o(u^{2n+2})$

$\tan u = u + \dfrac{1}{3}u^3 + o(u^3)$

$\arctan u = u - \dfrac{1}{3}u^3 + \cdots + \dfrac{(-1)^n}{2n+1}u^{2n+1} + o(u^{2n+1}) = \displaystyle\sum_{k=0}^{n} \dfrac{(-1)^k}{2k+1}u^{2k+1} + o(u^{2n+1})$

10.4 Fonctions convexes

E désigne un $\mathbb{R}$-espace vectoriel.

10.4.1 Barycentre

Soit $(u_i)_{i\in I}$ une famille finie de vecteurs de E et $(\lambda_i)_{i\in I}$ une famille de coefficients réels avec

$$\sum_{i\in I} \lambda_i \neq 0$$

Définition

On appelle barycentre de la famille $(u_i)_{i\in I}$ affectés des coefficients $(\lambda_i)_{i\in I}$ le vecteur v de E déterminé par

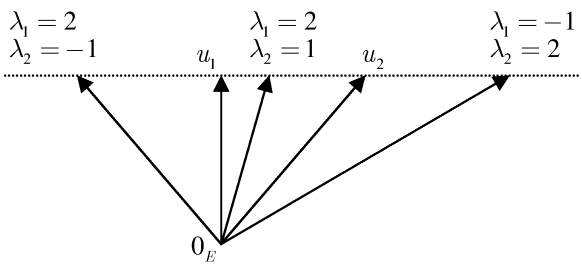
$$v = \frac{1}{\sum\limits_{i\in I} \lambda_i} \sum_{i\in I} \lambda_i u_i$$

On dit encore que v est le barycentre de la famille de vecteurs massiques $((u_i, \lambda_i))_{i\in I}$.

Remarque Dans le plan ou l'espace géométrique muni d'un repère d'origine O, on peut identifier point M et vecteur $\overrightarrow{OM}$.
On définit alors le centre de gravité (ou centre de masse) des points $A_1, \ldots, A_n$ affectés de masses $m_1, \ldots, m_n$ comme étant le point G tel que le vecteur $\overrightarrow{OG}$ est le barycentre de la famille de vecteurs $\left(\overrightarrow{OA_1}, \ldots, \overrightarrow{OA_n}\right)$ affectés des coefficients $(m_1, \ldots, m_n)$.
On peut montrer que ce centre de gravité ne dépend pas du choix du repère initial.

Exemple Le barycentre des u_1 et u_2 affectés des coefficients 1 et 1 correspond au vecteur milieu de u_1 et u_2.

Exemple



Remarque Les barycentres de deux vecteurs u_1, u_2 figurent sur la droite $u_1 + \text{Vect}(u_2 - u_1)$.

Définition

On appelle isobarycentre d'une famille de vecteurs $(u_1, \ldots, u_n)$ le barycentre v affecté de coefficients égaux à 1

$$v = \frac{1}{n}(u_1 + \cdots + u_n)$$

Proposition

Le barycentre est inchangé si :
a) on retire de la famille les vecteurs affectés d'un coefficient nul ;
b) on permute les vecteurs et les coefficients de la famille ;
c) on multiplie chaque coefficient par un scalaire non nul.

Remarque En exploitant un facteur de dilatation, tout barycentre peut être ramené à celui d'une famille dont la somme des coefficients vaut 1.

Théorème

On suppose $I = I_1 \cup I_2$ avec

$$I_1 \cap I_2 = \emptyset, \mu_1 = \sum_{i \in I_1} \lambda_i \neq 0 \text{ et } \mu_2 = \sum_{i \in I_2} \lambda_i \neq 0$$

Si v_1 et v_2 sont les barycentres des familles $((u_i, \lambda_i))_{i \in I_1}$ et $((u_i, \lambda_i))_{i \in I_2}$ alors le barycentre v de la famille $((u_i, \lambda_i))_{i \in I}$ est aussi le barycentre de la famille $((v_1, \mu_1), (v_2, \mu_2))$.

dém. :
On a $v_1 = \dfrac{1}{\mu_1} \displaystyle\sum_{i \in I_1} \lambda_i u_i$ et $v_2 = \dfrac{1}{\mu_2} \displaystyle\sum_{i \in I_2} \lambda_i u_i$ donc

$$\frac{1}{\mu_1 + \mu_2}(\mu_1 v_1 + \mu_2 v_2) = \frac{1}{\mu_1 + \mu_2} \sum_{i \in I_1 \cup I_2} \lambda_i u_i = \frac{1}{\sum\limits_{i \in I} \lambda_i} \sum_{i \in I} \lambda_i u_i$$

□

Remarque On peut calculer le barycentre d'une famille de plusieurs vecteurs en regroupant ceux-ci par paquets et se ramener à des situations où l'on ne considère que des familles de deux vecteurs.

10.4.2 Parties convexes

Définition

Soit $a, b \in E$. On appelle segment d'extrémités a et b l'ensemble $[a, b]$ constitué des barycentres des vecteurs a et b affectés de coefficients positifs :

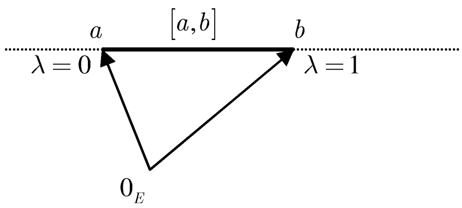
$$[a, b] = \{\lambda_1 a + \lambda_2 b / \lambda_1, \lambda_2 \geqslant 0, (\lambda_1, \lambda_2) \neq (0, 0)\}$$

En se ramenant à une somme de coefficients égale à 1

$$[a, b] = \{(1 - \lambda)a + \lambda b / \lambda \in [0, 1]\}$$

Remarque On peut aussi comprendre le segment $[a, b]$ comme obtenu par le paramétrage

$$[a, b] = \{a + \lambda(b - a) / \lambda \in [0, 1]\}$$

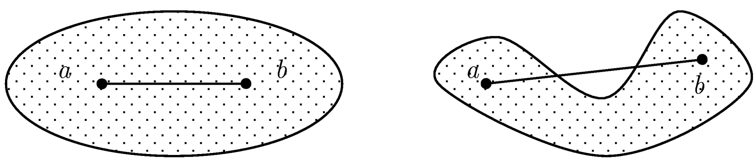


Définition

Une partie X de E est dite convexe si

$$\forall a, b \in X, [a, b] \subset X$$

Exemple



convexe non convexe

Exemple $\emptyset$ et E sont des parties convexes.

Exemple Les segments, les sous-espaces vectoriels et les sous-espaces affines sont des parties convexes.

Théorème

Soit X une partie de E. On a équivalence entre :
(i) X est une partie convexe ;
(ii) X contient tous les barycentres de ses vecteurs affectés de coefficients positifs.

dém. :
(ii) $\Rightarrow$ (i) Supposons (ii). Pour tout $a, b \in X$, la partie X contient le segment $[a, b]$ car celui-ci est constitué des barycentres de a et b affectés de coefficients positifs.
(i) $\Rightarrow$ (ii) Supposons X convexe et montrons par récurrence sur $n \geqslant 1$ que X contient les barycentres des familles de n éléments de X affectés de coefficients positifs.
Cas $n = 1$: il n'y a rien à démontrer.
Cas $n = 2$: on retrouve la définition de la convexité.
Supposons la propriété vraie au rang $n \geqslant 2$.
Soit v le barycentre de $((u_i, \lambda_i))_{1 \leqslant i \leqslant n+1}$ avec $u_i \in X$ et $\lambda_i \geqslant 0$.
On peut supposer les λ_i strictement positifs, sinon le problème est immédiatement résolu par l'hypothèse de récurrence. Considérons ensuite a le barycentre de la sous famille $((u_i, \lambda_i))_{1 \leqslant i \leqslant n}$. Par hypothèse de récurrence $a \in X$. Par associativité, v est barycentre de a et u_{n+1} affectés de coefficients positifs et donc

$$v \in [a, u_{n+1}] \subset X$$

Récurrence établie.
□

Remarque De manière semblable, on peut définir la notion de partie convexe du plan et de l'espace géométrique.

10.4.3 Fonction convexe, fonction concave

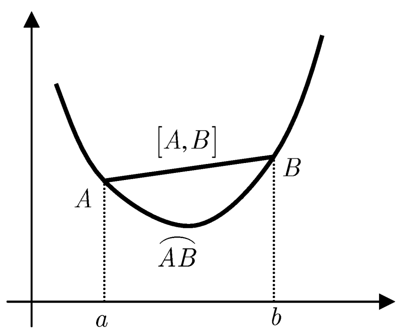
Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ est convexe si elle vérifie

$$\forall a, b \in I, \forall \lambda \in [0, 1], f((1-\lambda)a + \lambda b) \leqslant (1-\lambda) f(a) + \lambda f(b)$$

Proposition

Une fonction est convexe si ses arcs sont en dessous des cordes associées



dém. :
Pour $a, b \in I$, notons $A = (a, f(a))$ et $B = (b, f(b))$ les points du graphe de f d'abscisses a et b.
La corde d'extrémités A et B le segment $[A, B]$.

$$[A, B] = \{(1-\lambda)(a, f(a)) + \lambda(b, f(b))/\lambda \in [0,1]\}$$

soit encore

$$[A, B] = \{((1-\lambda)a + \lambda b, (1-\lambda)f(a) + \lambda f(b))\,/\lambda \in [0,1]\}$$

L'arc associé est $\overset{\frown}{AB}$ formé des points de Γ_f d'abscisses comprises entre a et b.

$$\overset{\frown}{AB} = \{(t, f(t))/t \in [a,b]\}$$

soit encore en écrivant $t = (1-\lambda)a + \lambda b$ avec $\lambda \in [0,1]$

$$\overset{\frown}{AB} = \{((1-\lambda)a + \lambda b, f\left((1-\lambda)a + \lambda b\right))\,/\lambda \in [0,1]\}$$

L'inégalité de convexité signifie alors que, pour une même abscisse, l'ordonnée du point de la corde est supérieure à celle du point de l'arc.

Ainsi, pour une fonction convexe, l'arc $\overset{\frown}{AB}$ est en dessous de la corde $[A, B]$.
□

Exemple Les fonctions affines $x \mapsto \alpha x + \beta$ sont convexes.
Pour ces fonctions, l'inégalité de convexité est en fait une égalité.

Exemple La fonction $|\,.\,|$ est convexe.
En effet,

$$\forall a, b \in \mathbb{R}, |\lambda a + (1-\lambda)b| \leqslant |\lambda|\,|a| + |1-\lambda|\,|b| = \lambda\,|a| + (1-\lambda)\,|b|$$

Définition

On dit qu'une fonction $f : I \to \mathbb{R}$ est concave si elle vérifie

$$\forall a, b \in I, \forall \lambda \in [0,1]\,, f\left((1-\lambda)a + \lambda b\right) \geqslant (1-\lambda)f(a) + \lambda f(b)$$

Remarque Pour une fonction concave, l'arc est au dessus de la corde.

Exemple Les fonctions affines sont concaves.

Proposition

Pour $f : I \to \mathbb{R}$, on a équivalence entre :
(i) f est concave ;
(ii) $-f$ est convexe.

dém. :
Par passage à l'opposé l'inégalité de convexité est renversée.
□

Remarque Par passage à l'opposé et renversement d'inégalité, les résultats qui suivent présentés pour les fonctions convexes se transposent aux fonctions concaves.

10.4.4 Caractérisation géométrique de la convexité

10.4.4.1 Épigraphe

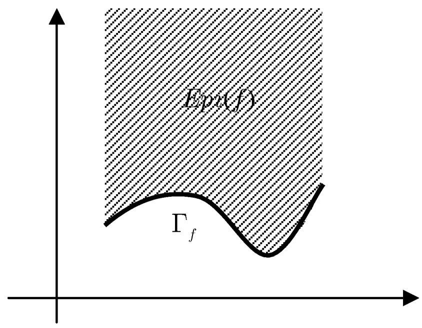
Définition

On appelle graphe d'une fonction $f : I \to \mathbb{R}$ l'ensemble

$$\Gamma_f = \left\{(x, y) \in \mathbb{R}^2 / x \in I \text{ et } f(x) = y\right\}$$

On appelle épigraphe d'une fonction $f : I \to \mathbb{R}$ l'ensemble

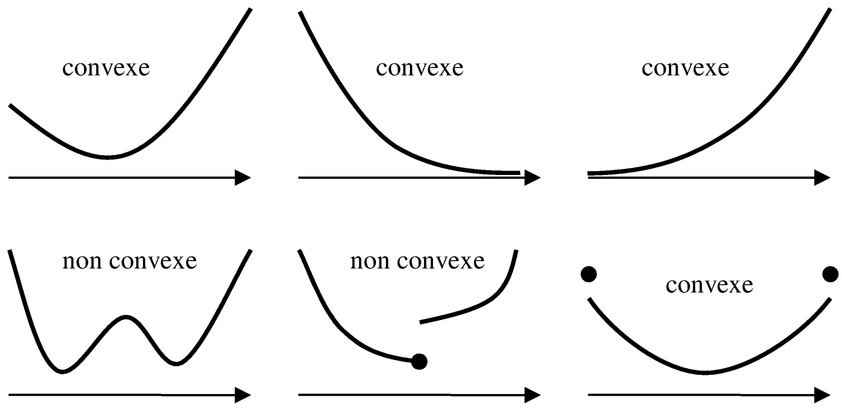
$$Epi(f) = \left\{(x, y) \in \mathbb{R}^2 / x \in I \text{ et } f(x) \leqslant y\right\}$$



Théorème

Pour $f : I \to \mathbb{R}$, on a équivalence entre :
(i) la fonction f est convexe ;
(ii) l'épigraphe de f est convexe.

dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe.
Soit A et B des points de l'épigraphe de f et A', B' les points du graphe de f de mêmes abscisses. Le segment $[A, B]$ est au dessus du segment $[A', B']$ lui même au dessus de l'arc $\overset{\frown}{A'B'}$. On en déduit que le segment $[A, B]$ est inclus dans l'épigraphe de f.
(ii) $\Rightarrow$ (i) Supposons l'épigraphe de f convexe.
Les cordes du graphe de f sont incluses dans l'épigraphe de f et sont donc au dessus des arcs. On en déduit que la fonction f est convexe.
Ex :



□

10.4.4.2 Inégalité des pentes

Définition

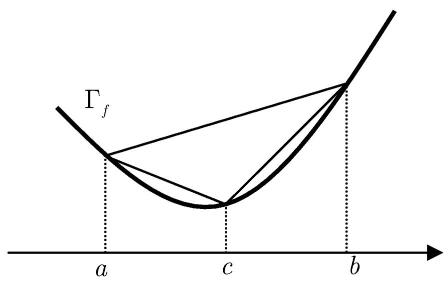
Pour $f : I \to \mathbb{R}$ et $a \neq b$ éléments de I, on note

$$\tau(a,b) = \frac{f(b)-f(a)}{b-a}$$

la pente (ou coefficient directeur) de la droite joignant les points d'abscisses a et b du graphe de f.

Théorème

Soit $f : I \to \mathbb{R}$. On a équivalence entre :
(i) f est convexe ;
(ii) $\forall a,b,c \in I$, $a < c < b \Rightarrow \tau(a,c) \leqslant \tau(a,b) \leqslant \tau(c,b)$;
(iii) $\forall a,b,c \in I$, $a < c < b \Rightarrow \tau(a,c) \leqslant \tau(c,b)$



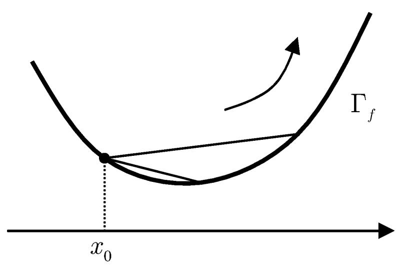
dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe
Soit $a,b,c \in I$ tels que $a < c < b$. On peut écrire $c = (1-\lambda)a + \lambda b$ avec

$$\lambda = \frac{c-a}{b-a} \in \left]0,1\right[$$

Par convexité
$$f(c) = f\left((1-\lambda)a + \lambda b\right) \leqslant (1-\lambda)f(a) + \lambda f(b)$$
donc
$$f(c) - f(a) \leqslant \lambda(f(b) - f(a)) = \frac{c-a}{b-a}(f(b) - f(a))$$
d'où $\tau(a,c) \leqslant \tau(a,b)$.
Aussi
$$f(b) - f(c) \geqslant (1-\lambda)(f(b) - f(a)) = \frac{b-c}{b-a}(f(b) - f(a))$$
ce qui fournit $\tau(a,b) \leqslant \tau(b,c)$.
(ii) $\Rightarrow$ (iii) C'est entendu
(iii) $\Rightarrow$ (i) Supposons (iii)
Soit $a, b \in I$ et $\lambda \in [0,1]$. Montrons
$$f\left((1-\lambda)a + \lambda b\right) \leqslant (1-\lambda)f(a) + \lambda f(b)$$
Si $a = b$: ok
Si $a \neq b$, quitte à échanger a et b d'une part, et λ et $1-\lambda$ d'autre part, on peut supposer $a < b$.
Si $\lambda = 0$ ou $\lambda = 1$: ok
Si $\lambda \in \left]0,1\right[$, posons $c = (1-\lambda)a + \lambda b$. Puisque $a < c < b$, on a $\tau(a,c) \leqslant \tau(c,b)$ ce qui donne
$$f(c) - f(a) \leqslant \frac{c-a}{b-c}(f(b) - f(c)) \text{ avec } \frac{c-a}{b-c} = \frac{c-a}{b-a}\frac{b-a}{b-c} = \frac{\lambda}{1-\lambda}$$
puis
$$f(c) \leqslant (1-\lambda)f(a) + \lambda f(b)$$
Ainsi f est convexe.
□

Corollaire

Si $f : I \to \mathbb{R}$ est convexe alors, pour chaque $x_0 \in I$, la fonction $x \mapsto \tau(x_0, x)$ est croissante



10.4.5 Fonctions convexes dérivables

Théorème

Soit $f : I \to \mathbb{R}$ dérivable. On a équivalence entre :
(i) f est convexe ;
(ii) f' est croissante.

dém. :
(i) $\Rightarrow$ (ii) Supposons f convexe. Soit $a, b \in I$ tels que $a < b$ et $x \in]a, b[$.
On a

$$\tau(a, x) \leqslant \tau(a, b) \leqslant \tau(b, x)$$

Quand $x \to a^+$, on obtient $f'(a) \leqslant \tau(a, b)$. Quand $x \to b^-$, on obtient $\tau(a, b) \leqslant f'(b)$.
Ainsi $f'(a) \leqslant f'(b)$ et f' est une fonction croissante.
(ii) $\Rightarrow$ (i) Supposons f' croissante.
Soit $a, b, c \in I$ tels que $a < c < b$.
Par le théorème des accroissements finis, il existe $\alpha \in]a, c[$ tel que $\tau(a, c) = f'(\alpha)$ et il existe $\beta \in]c, b[$ tel que $\tau(c, b) = f'(\beta)$. Puisque $\alpha \leqslant \beta$, on obtient

$$\tau(a, c) \leqslant \tau(c, b)$$

On peut alors conclure que f est convexe en vertu du théorème d'inégalité des pentes.
□

Corollaire

Soit $f : I \to \mathbb{R}$ deux fois dérivable. On a équivalence entre :
(i) f est convexe ;
(ii) $f'' \geqslant 0$.

dém. :
La monotonie de f' est donnée par le signe de f''.
□

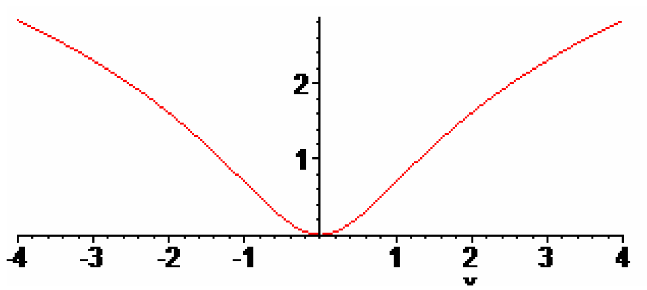
Exemple Les fonctions $x \mapsto x^2$, $x \mapsto \mathrm{e}^x$, $x \mapsto \mathrm{ch} x$ sont convexes.
En effet, ces fonctions sont de dérivées secondes positives.

Exemple La fonction $x \mapsto \ln x$ est une concave.
En effet, sa dérivée seconde négative.

Exemple Etudions la convexité de la fonction $f : x \mapsto \ln(1 + x^2)$ définie sur $\mathbb{R}$.
La fonction f est deux fois dérivable,

$$f'(x) = \frac{2x}{1 + x^2} \text{ et } f''(x) = 2\frac{1 - x^2}{(1 + x^2)^2}$$

du signe de $1 - x^2$
On en déduit que f est convexe sur $[-1, 1]$ et concave sur $]-\infty, -1]$ et sur $[1, +\infty[$.
Il y a inflexion aux points d'abscisse 1 et -1.



Notons que nous ne dirons pas que f est concave sur la réunion $]-\infty,-1]\cup[1,+\infty[$ car la notion de convexité d'une fonction réelle n'a de sens que pour une fonction définie sur un intervalle.

10.4.6 Inégalités de convexité

10.4.6.1 Position relative d'une courbe et de ses tangentes

Théorème

Si $f : I \to \mathbb{R}$ dérivable est convexe alors son graphe Γ_f est au dessus de chacune de ses tangentes.

dém. :
Soit $a \in I$. L'équation de la tangente T en a est

$$y = f'(a)(x-a) + f(a)$$

Considérons la fonction $g : I \to \mathbb{R}$ définie par

$$g(x) = f(x) - (f'(a)(x-a) + f(a))$$

Par opérations, la fonction g est dérivable et $g'(x) = f'(x) - f'(a)$.
La croissance de f' donne le signe de g' et on en déduit que g admet un minimum en a avec $g(a) = 0$.
Par suite, pour tout $x \in I$, $g(x) \geqslant 0$ puis l'inégalité

$$f(x) \geqslant f'(a)(x-a) + f(a)$$

□

Corollaire

Si $f : I \to \mathbb{R}$ dérivable est concave alors son graphe Γ_f est en dessous de chacune de ses tangentes.

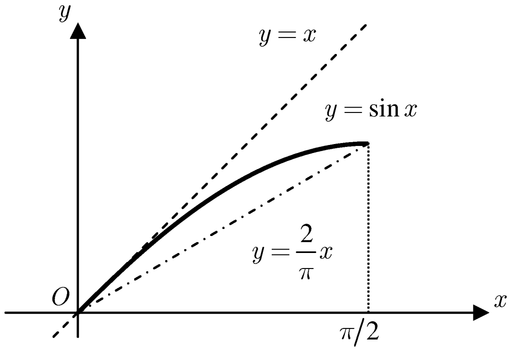
dém. :
Il suffit de considérer la fonction $-f$ qui est convexe.
□

10.4.6.2 Inégalités de convexité classiques

Exemple $\forall x \in \mathbb{R}, \mathrm{e}^x \geqslant 1 + x$
En effet, la fonction $x \mapsto \mathrm{e}^x$ est convexe, en positionnant son graphe par rapport à sa tangente en 0, on obtient la propriété.

Exemple $\forall x > -1, \ln(1+x) \leqslant x$
Puisque la fonction $x \mapsto \ln(1+x)$ est concave, il suffit de positionner son graphe par rapport à sa tangente en 0.

Exemple $\forall x \in [0,\pi/2]\,, \dfrac{2}{\pi}x \leqslant \sin x \leqslant x$



La fonction $x \mapsto \sin x$ est concave sur $[0, \pi/2]$, en positionnant son graphe par rapport à sa tangente en 0 et par rapport à sa corde joignant les points d'abscisse 0 et $\pi/2$, on obtient l'encadrement proposé.

10.4.6.3 Inégalité de Jensen

Théorème

Soit $f : I \to \mathbb{R}$ une fonction convexe et $n \in \mathbb{N}^\star$. On a

$$\forall a_1, \dots, a_n \in I, f(\lambda_1 a_1 + \dots + \lambda_n a_n) \leqslant \lambda_1 f(a_1) + \dots + \lambda_n f(a_n)$$

pour toute famille $\lambda_1, \dots, \lambda_n$ de réels positifs de somme 1.

dém. :
Posons $A_i = (a_i, f(a_i))$ points de l'épigraphe de f.
Puisque f est convexe, son épigraphe l'est aussi et celui-ci contient barycentre de la famille $((A_i, \lambda_i))_{1 \leqslant i \leqslant n}$.
Celui-ci est le couple

$$\left(\sum_{i=1}^n \lambda_i a_i, \sum_{i=1}^n \lambda_i f(a_i)\right)$$

et donc

$$f\left(\sum_{i=1}^n \lambda_i a_i\right) \leqslant \sum_{i=1}^n \lambda_i f(a_i)$$

□

Corollaire

Pour $f : I \to \mathbb{R}$ convexe, on a

$$\forall a_1, \dots, a_n \in I, f\left(\frac{a_1 + \dots + a_n}{n}\right) \leqslant \frac{1}{n}\left(f(a_1) + \dots + f(a_n)\right)$$

dém. :
Il suffit de prendre $\lambda_1 = \ldots = \lambda_n = 1/n$
□

Exemple Montrons

$$\forall a_1, \dots, a_n \in \mathbb{R}^+, \sqrt[n]{a_1 \dots a_n} \leqslant \frac{a_1 + \dots + a_n}{n}$$

Si l'un des a_i est nul, c'est immédiat.
Sinon, exploitons la concavité de $x \mapsto \ln x$.
Pour tout $a_1, \dots, a_n > 0$,

$$\frac{1}{n}\left(\ln a_1 + \cdots + \ln a_n\right) \leqslant \ln\left(\frac{a_1 + \cdots + a_n}{n}\right)$$

donc

$$\ln \sqrt[n]{a_1 \dots a_n} \leqslant \ln\left(\frac{a_1 + \cdots + a_n}{n}\right)$$

puis en composant avec la fonction exponentielle qui est croissante, on obtient l'inégalité voulue.

10.4.7 Musculation : dérivabilité et continuité des fonctions convexes

Théorème

Si $f : I \to \mathbb{R}$ est convexe alors en tout point $x_0 \in I$ qui n'est pas extrémité de I, f est dérivable à droite et à gauche avec

$$f'_g(x_0) \leqslant f'_d(x_0)$$

dém. :
Soit $a \in I$ tel que $a < x_0$.
L'application restreinte $\tau_{x_0} :]x_0, +\infty[\cap I$ est croissante et minorée par $\tau(a, x_0)$, cette application converge donc en x_0^+. Ainsi f est dérivable à droite en x_0 et

$$f'_d(x_0) \geqslant \tau(a, x_0)$$

L'application restreinte $\tau_{x_0} :]-\infty, x_0[\cap I$ est croissante et majorée, en vertu de l'étude précédente, par $f'_d(x_0)$. Cette application converge donc en x_0^- et f est dérivable à gauche en x_0 avec $f'_g(x_0) \leqslant f'_d(x_0)$.
□

Corollaire

Si $f : I \to \mathbb{R}$ est convexe alors f est continue en tout point intérieur à l'intervalle I.

dém. :
Car continue à droite et à gauche par dérivabilité à droite et à gauche.

□

Chapitre 11

Intégration sur un intervalle quelconque

On sait intégrer sur les segments $[a, b]$ et on souhaite étendre la notion à tout intervalle et ainsi donner un sens entre autre à

$$\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t \text{ et } \int_0^1 \frac{\mathrm{d}t}{\sqrt{t}}$$

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

11.1 Intégration sur $[a, +\infty[$

Soit $a \in \mathbb{R}$.

11.1.1 Convergence

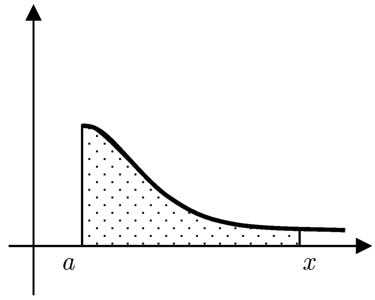
Définition

Soit $f : [a, +\infty[\to \mathbb{K}$ continue par morceaux.
On dit que l'intégrale de f sur $[a, +\infty[$ converge si l'intégrale partielle $\displaystyle\int_a^x f(t)\,\mathrm{d}t$ converge quand $x \to +\infty$.
On pose alors

$$\int_a^{+\infty} f(t)\,\mathrm{d}t \underset{\text{déf}}{=} \lim_{x\to+\infty} \int_a^x f(t)\,\mathrm{d}t$$

Cette intégrale s'écrit aussi $\displaystyle\int_a^{+\infty} f$ s'il n'est pas utile de préciser une variable d'intégration (qui par ailleurs est muette) ou encore $\displaystyle\int_{[a,+\infty[} f(t)\,\mathrm{d}t$.

Remarque L'intégrale converge si, et seulement si, l'aire hachurée converge quand $x \to +\infty$



Attention : Par essence, une intégrale impropre est une limite, pour la manipuler il faut préalablement en justifier l'existence.

Exemple Etude de $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t$.
La fonction $t \mapsto \mathrm{e}^{-t}$ est définie et continue par morceaux sur $[0, +\infty[$

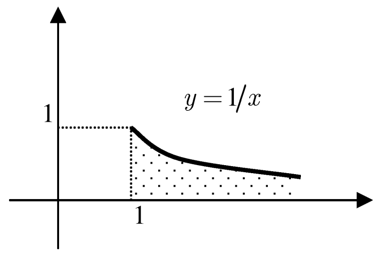
$$\int_0^x \mathrm{e}^{-t}\,\mathrm{d}t = 1 - \mathrm{e}^{-x} \xrightarrow[x\to+\infty]{} 1$$

donc $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t$ converge et

$$\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t = 1$$

Exemple Etude de $\displaystyle\int_0^{+\infty} 1\,\mathrm{d}t$.
La fonction $t \mapsto 1$ est définie et continue par morceaux sur $[0, +\infty[$.
$\displaystyle\int_0^x 1\,\mathrm{d}t = x \xrightarrow[x\to+\infty]{} +\infty$ donc $\displaystyle\int_0^{+\infty} 1\,\mathrm{d}t$ diverge.

Exemple Etude de $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$
La fonction $t \mapsto 1/t$ est définie et continue par morceaux sur $[1, +\infty[$
$\displaystyle\int_1^x \frac{\mathrm{d}t}{t} = \ln x \xrightarrow[x\to+\infty]{} +\infty$ donc $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ diverge.



Pour la fonction inverse, il y a trop d'espace entre la courbe et l'axe des abscisses pour que l'intégrale converge, la fonction inverse converge trop lentement vers 0 en $+\infty$.

11.1.2 Reste d'une intégrale convergente

Soit $f : [a, +\infty[\to \mathbb{K}$ continue par morceaux.

Théorème

Pour tout $b \in [a, +\infty[$, on a équivalence entre :
(i) $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}t$ converge ;
(ii) $\displaystyle\int_b^{+\infty} f(t)\,\mathrm{d}t$ converge.

dém. :
On a

$$\int_a^x f = \int_a^b f + \int_b^x f$$

donc une intégrale partielle converge si, et seulement si, l'autre converge aussi.
□

Corollaire

On ne change pas la nature d'une intégrale sur $[a, +\infty[$ en modifiant les valeurs de la fonction intégrée sur $[a, c]$. La nature de $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}t$ ne dépend que du comportement de f au voisinage de $+\infty$.

Définition

Si $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}t$ converge alors on peut introduire l'intégrale $\displaystyle\int_x^{+\infty} f(t)\,\mathrm{d}t$ pour tout $x \in [a, +\infty[$.
La fonction ainsi définie s'appelle le reste de l'intégrale convergente.

Théorème

Si $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}t$ converge alors pour tout $x \geqslant a$

$$\int_a^{+\infty} f(t)\,\mathrm{d}t = \int_a^x f(t)\,\mathrm{d}t + \int_x^{+\infty} f(t)\,\mathrm{d}t$$

De plus

$$\int_x^{+\infty} f(t)\,\mathrm{d}t \xrightarrow[x\to+\infty]{} 0$$

dém. :
Soit $x \in [a, +\infty[$ fixé. On introduit $y \in [x, +\infty[$ et on a

$$\int_a^y f(t)\,\mathrm{d}t = \int_a^x f(t)\,\mathrm{d}t + \int_x^y f(t)\,\mathrm{d}t$$

Quand $y \to +\infty$, on obtient

$$\int_a^{+\infty} f(t)\,\mathrm{d}t = \int_a^x f(t)\,\mathrm{d}t + \int_x^{+\infty} f(t)\,\mathrm{d}t$$

De plus

$$\int_x^{+\infty} f(t)\,\mathrm{d}t = \int_a^{+\infty} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t \xrightarrow[x\to+\infty]{} 0$$

□

11.1.3 Cas des fonctions continues

Soit $f : [a,+\infty[\to \mathbb{K}$ une fonction continue de primitive F.

Théorème

On a équivalence entre :
(i) $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}t$ converge ;
(ii) $F(x)$ converge quand $x \to +\infty$.
De plus, on a alors

$$\int_a^{+\infty} f(t)\,\mathrm{d}t = \lim_{x\to+\infty} F(x) - F(a) \underset{\text{déf}}{=} [F(x)]_a^{+\infty}$$

Exemple Etude de $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{t^2+1}$
L'intégrale converge car $\arctan t$ est primitive de l'intégrande et converge en $+\infty$.
De plus

$$\int_0^{+\infty} \frac{\mathrm{d}t}{t^2+1} = [\arctan t]_0^{+\infty} = \frac{\pi}{2}$$

Proposition

Si f est continue et si $\displaystyle\int_a^{+\infty} f$ converge alors

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_x^{+\infty} f\right) = -f(x)$$

dém. :
Introduisons une primitive F de f. Puisque l'intégrale converge, F admet une limite en $+\infty$ et on peut écrire

$$\int_x^{+\infty} f = \lim_{+\infty} F - F(x)$$

La fonction $x \mapsto \displaystyle\int_x^{+\infty} f$ est alors de classe $\mathcal{C}^1$ et

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_x^{+\infty} f\right) = -F'(x) = -f(x)$$

□

11.1.4 Propriétés

11.1.4.1 Linéarité

Théorème

Soit $f, g : [a, +\infty[\to \mathbb{K}$ continues par morceaux et $\lambda \in \mathbb{K}$.
Si les intégrales $\displaystyle\int_a^{+\infty} f$ et $\displaystyle\int_a^{+\infty} g$ convergent alors $\displaystyle\int_a^{+\infty} f + g$ et $\displaystyle\int_a^{+\infty} \lambda f$ convergent avec

$$\int_a^{+\infty} f + g = \int_a^{+\infty} f + \int_a^{+\infty} g \text{ et } \int_a^{+\infty} \lambda f = \lambda \int_a^{+\infty} f$$

Corollaire

L'ensemble constitué des fonctions continues par morceaux de $[a, +\infty[$ vers $\mathbb{K}$ dont l'intégrale converge définit un sous-espace vectoriel de $\mathcal{C}^0_{pm}([a, +\infty[\,, \mathbb{K})$.
L'application $f \mapsto \displaystyle\int_a^{+\infty} f$ y définit une forme linéaire.

Exemple Si $\displaystyle\int_a^{+\infty} f + g$ et $\displaystyle\int_a^{+\infty} f$ convergent alors $\displaystyle\int_a^{+\infty} g$ converge.
En effet, on peut écrire

$$g = (f + g) + (-1)g$$

Attention : Pour exploiter la relation $\displaystyle\int_a^{+\infty} f + g = \int_a^{+\infty} f + \int_a^{+\infty} g$, il faut préalablement justifier la convergence d'au moins deux des intégrales engagées !
Ceci empêche d'écrire des aberrations telles

$$\int_0^{+\infty} 0 \, \mathrm{d}t = \int_0^{+\infty} 1 \, \mathrm{d}t + \int_0^{+\infty} (-1) \, \mathrm{d}t$$

ou, un peu moins grossièrement

$$\int_1^{+\infty} \frac{\mathrm{d}t}{t(t+1)} = \int_1^{+\infty} \frac{\mathrm{d}t}{t} - \int_1^{+\infty} \frac{\mathrm{d}t}{t+1}$$

Exemple Si $\displaystyle\int_a^{+\infty} f$ converge et $\displaystyle\int_a^{+\infty} g$ diverge alors $\displaystyle\int_a^{+\infty} f$ diverge.

Attention : Si $\displaystyle\int_a^{+\infty} f$ et $\displaystyle\int_a^{+\infty} g$ divergent alors on ne peut rien dire sur la nature de $\displaystyle\int_a^{+\infty} f + g$.

11.1.4.2 Positivité

Théorème

Soit $f : [a, +\infty[\to \mathbb{R}$ continue par morceaux.
Si $\displaystyle\int_a^{+\infty} f$ converge et si $f \geqslant 0$ alors $\displaystyle\int_a^{+\infty} f \geqslant 0$

dém. :
En tant qu'intégrale bien ordonnée d'une fonction positive, pour tout $x \geqslant a$, on a

$$\int_a^x f \geqslant 0$$

A la limite quand $x \to +\infty$, on obtient $\displaystyle\int_a^{+\infty} f \geqslant 0$.
□

Corollaire

Soit $f, g : [a, +\infty[\to \mathbb{R}$ continues par morceaux
Si $\displaystyle\int_a^{+\infty} f$ et $\displaystyle\int_a^{+\infty} g$ convergent et si $f \leqslant g$ alors

$$\int_a^{+\infty} f \leqslant \int_a^{+\infty} g$$

dém. :
Avec convergence, on a

$$\int_a^{+\infty} g - \int_a^{+\infty} f = \int_a^{+\infty} g - f \geqslant 0$$

□

Théorème

Soit $f : [a, +\infty[\to \mathbb{R}$ continue.
Si $f \geqslant 0$ et si $\displaystyle\int_a^{+\infty} f$ converge avec $\displaystyle\int_a^{+\infty} f = 0$ alors f est la fonction nulle.

dém. :
Introduisons F une primitive de f. La fonction F est croissante et puisque l'intégrale de f converge et vaut 0, on a $F(a) = \lim\limits_{+\infty} F$. On en déduit que F est constante et donc $f = F' = 0$.
□

11.1.4.3 Conjugaison

Théorème

Soit $f : [a, +\infty[\to \mathbb{C}$ continue par morceaux.
Si $\displaystyle\int_a^{+\infty} f$ converge alors $\displaystyle\int_a^{+\infty} \bar{f}$ convergent et alors

$$\int_a^{+\infty} \bar{f} = \overline{\int_a^{+\infty} f}$$

dém. :
Par conjugaison de limites.
□

Corollaire

On a équivalence entre :
(i) $\int_a^{+\infty} f$ converge ;
(ii) $\int_a^{+\infty} \mathrm{Re} f$ et $\int_a^{+\infty} \mathrm{Im} f$ convergent.
De plus, on a alors

$$\int_a^{+\infty} f = \int_a^{+\infty} \mathrm{Re} f + i.\int_a^{+\infty} \mathrm{Im} f$$

Exemple Calcul de $\int_0^{+\infty} \cos(\omega t)\mathrm{e}^{-t}\,\mathrm{d}t$ et $\int_0^{+\infty} \sin(\omega t)\mathrm{e}^{-t}\,\mathrm{d}t$.
Introduisons $\int_0^{+\infty} \mathrm{e}^{(i\omega-1)t}\,\mathrm{d}t$.

$$\int_0^x \mathrm{e}^{i\omega t}\mathrm{e}^{-t}\,\mathrm{d}t = \int_0^x \mathrm{e}^{(i\omega-1)t}\,\mathrm{d}t = \left[\frac{\mathrm{e}^{(i\omega-1)t}}{i\omega-1}\right]_0^x \underset{x\to+\infty}{\to} \frac{1}{1-i\omega} = \frac{1+i\omega}{1+\omega^2}$$

On en déduit

$$\int_0^{+\infty} \cos(\omega t)\mathrm{e}^{-t}\,\mathrm{d}t = \frac{1}{1+\omega^2} \text{ et } \int_0^{+\infty} \sin(t)\mathrm{e}^{-t}\,\mathrm{d}t = \frac{\omega}{1+\omega^2}$$

11.2 Intégrabilité sur $[a,+\infty[$

Soit $a \in \mathbb{R}$.

11.2.1 Cas des fonctions positives

Théorème

Soit $f : [a,+\infty[\to \mathbb{R}$ continue par morceaux. Si f est positive on a équivalence entre :
(i) $\int_a^{+\infty} f$ converge ;
(ii) $\exists M \in \mathbb{R}^+, \forall x \in [a,+\infty[, \int_a^x f(t)\,\mathrm{d}t \leqslant M$.

dém. :
Puisque f est positive, pour tout $x \leqslant y \in [a,+\infty[$, on a

$$\int_a^x f(t)\,\mathrm{d}t \leqslant \int_a^y f(t)\,\mathrm{d}t$$

L'intégrale partielle $\int_a^x f(t)\,\mathrm{d}t$ définit donc une fonction croissante de x. Si celle-ci est majorée alors elle converge quand $x \to +\infty$ et la réciproque est vraie.
□

Remarque Au contraire, si $\int_a^{+\infty} f$ diverge avec $f \geqslant 0$ alors

$$\int_a^x f \xrightarrow[x \to +\infty]{} +\infty$$

11.2.2 Comparaison de fonctions positives

Théorème

Soit $f, g : [a, +\infty[\to \mathbb{R}$ continues par morceaux telles que $0 \leqslant f \leqslant g$.
Si $\int_a^{+\infty} g$ converge alors $\int_a^{+\infty} f$ aussi.
Si $\int_a^{+\infty} f$ diverge alors $\int_a^{+\infty} g$ aussi.

dém. :
Soit $x \in [a, +\infty[$. Puisque $f \leqslant g$, on a

$$\int_a^x f(t)\,\mathrm{d}t \leqslant \int_a^x g(t)\,\mathrm{d}t \leqslant \int_a^{+\infty} g(t)\,\mathrm{d}t$$

La fonction f est positive et ses intégrales partielles sont majorées, l'intégrale de f sur $[a, +\infty[$ est donc convergente.
□

Exemple Nature de $\int_0^{+\infty} \frac{\mathrm{e}^{-t}}{t+1}\,\mathrm{d}t$.
La fonction $f : t \mapsto \frac{\mathrm{e}^{-t}}{t+1}$ est définie et continue par morceaux sur $[0, +\infty[$.
Pour $t \geqslant 0$, on a $0 \leqslant f(t) \leqslant \mathrm{e}^{-t}$.
Or $\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t$ converge donc, par comparaison de fonctions positives, $\int_0^{+\infty} \frac{\mathrm{e}^{-t}}{t+1}\,\mathrm{d}t$ converge.

Exemple Nature de $\int_1^{+\infty} \frac{\ln(1+t)}{t}\,\mathrm{d}t$.
La fonction $f : t \mapsto \frac{\ln(1+t)}{t}$ est définie et continue par morceaux sur $[1, +\infty[$.
Pour $t \geqslant 1$, on a $f(t) \geqslant \frac{\ln 2}{t} \geqslant 0$.
Or $\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ diverge donc, par comparaison de fonctions positives, $\int_1^{+\infty} \frac{\ln(1+t)}{t}\,\mathrm{d}t$ diverge.

Théorème

Soit $f, g : [a, +\infty[\to \mathbb{R}^+$ continues par morceaux.
Si $f(t) \underset{t \to +\infty}{\sim} g(t)$ alors les intégrales $\int_a^{+\infty} f$ et $\int_a^{+\infty} g$ ont même nature.

dém. :
Pour t assez grand, on a la comparaison

$$\frac{1}{2}g(t) \leqslant f(t) \leqslant 2g(t)$$

qui est décisive !
□

11.2.3 Intégrabilité

Soit $f : [a, +\infty[\to \mathbb{K}$ continue par morceaux.

Définition

On dit que f est intégrable sur $[a, +\infty[$ si l'intégrale $\displaystyle\int_a^{+\infty} |f|$ converge.
On dit aussi que l'intégrale $\displaystyle\int_a^{+\infty} f$ est absolument convergente.

Remarque Si f est positive, il est équivalent de dire que f est intégrable sur $[a, +\infty[$ que de dire que son intégrale de f converge.

Exemple Intégrabilité de $t \mapsto \dfrac{\cos(t)}{1+t^2}$ sur $[0, +\infty[$.
On

$$0 \leqslant \left|\frac{\cos(t)}{1+t^2}\right| \leqslant \frac{1}{1+t^2}$$

Or il y a convergence de $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{1+t^2}$ donc, par comparaison de fonctions positives, il y a convergence de l'intégrale $\displaystyle\int_0^{+\infty} \left|\frac{\cos(t)}{1+t^2}\right| \mathrm{d}t$.
Ainsi, la fonction $t \mapsto \dfrac{\cos(t)}{1+t^2}$ est intégrable sur $[0, +\infty[$.

Théorème

Si f est intégrable sur $[a, +\infty[$ alors $\displaystyle\int_a^{+\infty} f$ converge et $\displaystyle\left|\int_a^{+\infty} f\right| \leqslant \int_a^{+\infty} |f|$

dém. :
Cas f à valeurs positives
C'est immédiat compte tenu des résultats qui précède.
Cas f à valeurs réelles
On pose $f^+ = \sup(f, 0)$ et $f^- = \sup(-f, 0)$.
Les fonctions $f^+, f^- : I \to \mathbb{R}^+$ sont continues par morceaux et vérifient $f = f^+ - f^-$.
On a aussi $|f| = f^+ + f^-$ donc $0 \leqslant f^+, f^- \leqslant |f|$.
Par comparaison de fonctions positives, les intégrales $\displaystyle\int_a^{+\infty} f^+$ et $\displaystyle\int_a^{+\infty} f^-$ convergent puis, par opérations, l'intégrale $\displaystyle\int_a^{+\infty} f$ converge aussi.

Cas f à valeurs complexes
On écrit $f = \text{Re}f + i\text{Im}f$.
$\text{Re}f, \text{Im}f : I \to \mathbb{R}$ sont continues par morceaux.
Puisque $|\text{Re}f|, |\text{Im}f| \leqslant |f|$, on a, par comparaison de fonctions positives, les intégrales $\int_a^{+\infty} |\text{Re}f|$ et $\int_a^{+\infty} |\text{Im}f|$ convergent donc $\int_a^{+\infty} \text{Re}f$ et $\int_a^{+\infty} \text{Im}f$ convergent puis par opérations $\int_a^{+\infty} f$ converge aussi.
Enfin, pour tout $x \in [a,+\infty[$

$$\left|\int_a^x f\right| \leqslant \int_a^x |f|$$

donc à la limite quand $x \to +\infty$

$$\left|\int_a^{+\infty} f\right| \leqslant \int_a^{+\infty} |f|$$

□
*Bilan :*Pour une fonction réelle ou complexe

$$f \text{ intégrable } \Rightarrow \int_a^{+\infty} f \text{ converge}$$

Pour une fonction positive, $f = |f|$ donc

$$f \text{ intégrable } \Leftrightarrow \int_a^{+\infty} f \text{ converge}$$

Remarque Plus généralement, pour une fonction de signe constant, il y a aussi équivalence.
On peut encore approfondir : si f est de signe constant au voisinage de $+\infty$ alors l'intégrabilité de f sur $[a,+\infty[$ équivaut à la convergence de l'intégrale $\int_a^{+\infty} f$.

Attention : Il se peut que $\int_a^{+\infty} f$ converge alors que $\int_a^{+\infty} |f|$ diverge.
Ce phénomène se rencontre lorsque la convergence de l'intégrale provient d'une compensation entre aires positive et négative.

Définition

Si $\int_a^{+\infty} f$ converge alors que $\int_a^{+\infty} |f|$ diverge, on dit que l'intégrale $\int_a^{+\infty} f$ est semi-convergente.

Exemple Les intégrales $\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t$ et $\int_0^{+\infty} \cos(t^2)\,\mathrm{d}t$ sont des intégrales semi-convergentes fameuses.

11.2.4 Intégrabilité par comparaison

11.2.4.1 Domination

Théorème

Soit $f : [a, +\infty[\to \mathbb{K}$ et $\varphi : [a, +\infty[\to \mathbb{R}^+$ continues par morceaux.
Si

$$\forall t \in [a, +\infty[\,, |f(t)| \leqslant \varphi(t) \text{ avec } \varphi \text{ intégrable}$$

alors f est intégrable.

dém. :
L'intégrale $\displaystyle\int_a^{+\infty} \varphi(t)\,\mathrm{d}t$ converge et donc, par comparaison de fonctions positives, $\displaystyle\int_a^{+\infty} |f(t)|\,\mathrm{d}t$ converge.
Ainsi f est intégrable.
□

11.2.4.2 Comparaisons asymptotiques

Définition

Soit $f, g : [a, +\infty[\to \mathbb{K}$.
On dit que f est dominée par g au voisinage de $+\infty$ si

$$\exists M \in \mathbb{R}^+, \exists A \in [a, +\infty[\,, \forall t \geqslant A, |f(t)| \leqslant M\,|g(t)|$$

On écrit alors

$$f(t) \underset{t\to+\infty}{=} O\left(g(t)\right)$$

Remarque Il revient au même de dire qu'il est possible d'écrire au voisinage de $+\infty$

$$f(t) = b(t)g(t) \text{ avec } b \text{ une fonction bornée}$$

Définition

Soit $f, g : [a, +\infty[\to \mathbb{K}$.
On dit que f est négligeable devant g au voisinage de $+\infty$ si

$$\forall \varepsilon > 0^+, \exists A \in [a, +\infty[\,, \forall t \geqslant A, |f(t)| \leqslant \varepsilon\,|g(t)|$$

On écrit alors

$$f(t) \underset{t\to+\infty}{=} o\left(g(t)\right)$$

Remarque Il revient au même de dire qu'il est possible d'écrire au voisinage de $+\infty$

$$f(t) = \varepsilon(t)g(t) \text{ avec } \varepsilon(t) \xrightarrow[t\to+\infty]{} 0$$

Définition

Soit $f,g:[a,+\infty[\to\mathbb{K}$.
On dit que f est équivalente à g au voisinage de $+\infty$ si l'on peut écrire

$$f(t)\underset{t\to+\infty}{=}g(t)+o\left(g(t)\right)$$

On écrit alors

$$f(t)\underset{t\to+\infty}{\sim}g(t)$$

Remarque Il revient au même de dire qu'il est possible d'écrire au voisinage de $+\infty$

$$f(t)=u(t)g(t)\text{ avec }u(t)\xrightarrow[t\to+\infty]{}1$$

11.2.4.3 Intégrabilité par comparaison asymptotique

Théorème

Soit $f:[a,+\infty[\to\mathbb{K}$ et $g:[a,+\infty[\to\mathbb{R}^+$ continues par morceaux.
Si $f(t)\underset{t\to+\infty}{=}O\left(g(t)\right)$ et si g est intégrable sur $[a,+\infty[$ alors f est intégrable sur $[a,+\infty[$

dém. :
Il existe $A\in[a,+\infty[$ et $M\in\mathbb{R}^+$ vérifiant

$$\forall t\in[A,+\infty[\,,|f(t)|\leqslant Mg(t)$$

En considérant $\varphi(t)=Mg(t)$, on peut affirmer par domination qu'il y a convergence de $\displaystyle\int_A^{+\infty}|f|$ et donc de $\displaystyle\int_a^{+\infty}|f|$ qui n'en diffère que d'une constante.
□

Corollaire

Si $f(t)\underset{t\to+\infty}{=}o\left(g(t)\right)$ et si g est intégrable sur $[a,+\infty[$ alors f est intégrable sur $[a,+\infty[$

dém. :
$f(t)\underset{t\to+\infty}{=}o\left(g(t)\right)$ alors $f(t)\underset{t\to+\infty}{=}O\left(g(t)\right)$.
□

Corollaire

Si $f(t)\underset{t\to+\infty}{\sim}g(t)$ alors l'intégrabilité de f sur $[a,+\infty[$ équivaut à celle de g.

dém. :
si $f(t)\underset{t\to+\infty}{\sim}g(t)$ alors $f(t)\underset{t\to+\infty}{=}O\left(g(t)\right)$ et aussi $g(t)\underset{t\to+\infty}{=}O\left(f(t)\right)$ de sorte que l'intégrabilité d'une fonction entraîne l'intégrabilité de l'autre.
□

Attention : Ces énoncés sont faux en terme de convergence d'intégrale. Il est indispensable de s'exprimer en terme d'intégrabilité. Cependant, on peut énoncer le théorème d'équivalence suivant :

11.2.5 Intégrales de Riemann

Soit $\alpha \in \mathbb{R}$.

Théorème

$$\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha} \text{ converge si, et seulement si, } \alpha > 1$$

dém. :
La fonction $t \mapsto 1/t^\alpha$ est définie et continue par morceaux sur $[1, +\infty[$.
Pour $\alpha \leqslant 1$,

$$\int_1^x \frac{\mathrm{d}t}{t^\alpha} \geqslant \int_1^x \frac{\mathrm{d}t}{t} = \ln x \xrightarrow[x \to +\infty]{} +\infty$$

et donc $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ diverge.
Pour $\alpha > 1$,

$$\int_1^x \frac{\mathrm{d}t}{t^\alpha} = \left[-\frac{1}{\alpha - 1}\frac{1}{t^{\alpha-1}}\right]_1^x \xrightarrow[x \to +\infty]{} \frac{1}{\alpha - 1}$$

et donc $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ converge.
□

Exemple $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^2}$ et $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^{1,00001}}$ convergent alors que $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ et $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{\sqrt{t}}$ divergent.

Corollaire

La fonction $t \mapsto 1/t^\alpha$ est intégrable sur $[1, +\infty[$ si, et seulement si, $\alpha > 1$.

11.2.6 En pratique

Exemple Nature de $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{t^4 + 1}$.
La fonction $f : t \mapsto 1/(t^4 + 1)$ est définie et continue par morceaux sur $[0, +\infty[$.
Quand $t \to +\infty$,
$f(t) \to 0$, on ne peut rien en conclure

$$f(t) \underset{t \to +\infty}{\sim} \frac{1}{t^4}$$

Or $t \mapsto 1/t^4$ est intégrable sur $[1, +\infty[$ (car $4 > 1$) donc f est intégrable sur $[1, +\infty[$, puis sur $[0, +\infty[$.
Ainsi, l'intégrale $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{t^4 + 1}$ est convergente.

Exemple Nature de $\displaystyle\int_0^{+\infty} \frac{t + 1}{t^2 + 1}\,\mathrm{d}t$.
La fonction $f : t \mapsto (t + 1)/(t^2 + 1)$ est définie et continue par morceaux sur $[0, +\infty[$.

On a

$$\frac{t+1}{t^2+1} \underset{t\to+\infty}{\sim} \frac{1}{t}$$

Or l'intégrale $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ diverge donc, par équivalence de fonctions positives, l'intégrale $\displaystyle\int_1^{+\infty} \frac{t+1}{t^2+1}\,\mathrm{d}t$ diverge.

On en déduit la divergence de $\displaystyle\int_0^{+\infty} \frac{t+1}{t^2+1}\,\mathrm{d}t$.

Exemple Nature de $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$.

La fonction $f : t \mapsto \mathrm{e}^{-t^2}$ est définie et continue par morceaux sur $[0,+\infty[$.

Quand $t \to +\infty$, $f(t) \to 0$ mais ce n'est en rien décisif. Cependant $t^2 f(t) \xrightarrow[t\to+\infty]{} 0$ donc

$$f(t) \underset{t\to+\infty}{=} o\left(\frac{1}{t^2}\right)$$

Or $t \mapsto 1/t^2$ est intégrable sur $[1,+\infty[$ (car $2>1$) donc f est intégrable sur $[0,+\infty[$.

L'intégrale $\displaystyle\int_0^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$ converge.

Exemple Nature de $\displaystyle\int_0^{+\infty} \frac{\cos(t)}{1+t^2}\,\mathrm{d}t$.

La fonction $f : t \mapsto \cos(t)/(1+t^2)$ est définie et continue par morceaux sur $[0,+\infty[$.

On a

$$f(t) \underset{t\to+\infty}{\sim} \frac{\cos(t)}{t^2}$$

donc

$$t^{3/2} f(t) \underset{t\to+\infty}{\sim} \frac{\cos(t)}{\sqrt{t}} \to 0$$

Ainsi $f(t) \underset{t\to+\infty}{=} o(1/t^{3/2})$ et on peut conclure que f est intégrable sur $[0,+\infty[$ et $\displaystyle\int_0^{+\infty} \frac{\cos t}{1+t^2}\,\mathrm{d}t$ converge.

Exemple Nature de $\displaystyle\int_1^{+\infty} \frac{1}{\ln(t+1)}\,\mathrm{d}t$.

La fonction $f : t \mapsto 1/\ln(t+1)$ est définie et continue par morceaux sur $[1,+\infty[$.

On a

$$t f(t) \xrightarrow[t\to+\infty]{} +\infty$$

Il existe $A \in [1,+\infty[$ tel que pour $t \geqslant A$, $tf(t) \geqslant 1$ et donc $f(t) \geqslant 1/t$.

Or $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t}$ diverge, donc par comparaison de fonctions positives (et moyennant un découpage des intégrales en A) on peut conclure que l'intégrale $\displaystyle\int_1^{+\infty} \frac{1}{\ln(t+1)}\,\mathrm{d}t$ diverge.

*Bilan :*Pour $f : [a, +\infty[\to \mathbb{K}$ continue par morceaux :
- Si $f(t) \underset{t\to+\infty}{\sim} C/t^\alpha$ (avec $C \neq 0$) quand $t \to +\infty$ alors

$$f \text{ est intégrable sur } [a, +\infty[\text{ si, et seulement si, } \alpha > 1\,;$$

- Si on détermine $\alpha > 1$ tel que $t^\alpha f(t) \xrightarrow[t\to+\infty]{} 0$ quand $t \to +\infty$ alors f est intégrable sur $[a, +\infty[$;
- Si $tf(t) \xrightarrow[t\to+\infty]{} \ell \neq 0$ alors l'intégrale de f sur $[a, +\infty[$ diverge.

11.2.7 Intégrabilité et limite en $+\infty$

Théorème

Soit $f : [a, +\infty[\to \mathbb{K}$ continue par morceaux.
Si $f(t) \to \ell \neq 0$ alors l'intégrale de f sur $[a, +\infty[$ diverge.

dém. :
Cas $\mathbb{K} = \mathbb{R}$
Quitte à considérer $-f$, on peut supposer $\ell > 0$. Puisque f tend vers ℓ en $+\infty$, il existe $A \in [a, +\infty[$ vérifiant

$$\forall t \geqslant A, f(t) \geqslant \ell/2$$

et alors

$$\int_a^x f(t)\,\mathrm{d}t = \int_a^A f(t)\,\mathrm{d}t + \int_A^x f(t)\,\mathrm{d}t$$

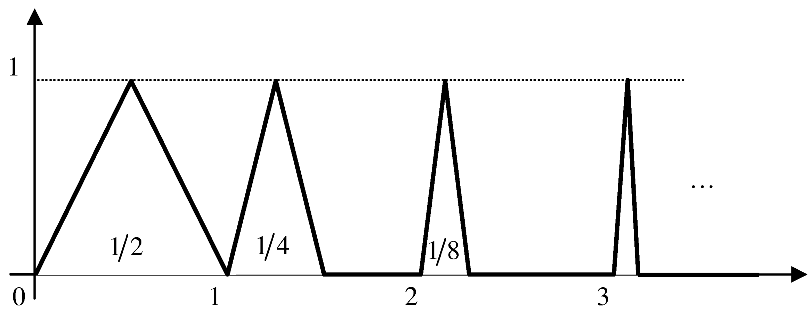
et donc

$$\int_a^x f(t)\,\mathrm{d}t \geqslant C^{te} + \frac{\ell}{2}(x-a) \xrightarrow[x\to+\infty]{} +\infty$$

Ainsi l'intégrale de $\displaystyle\int_a^{+\infty} f(t)\,\mathrm{d}t$ diverge (et donc f n'est pas intégrable)
Cas $\mathbb{K} = \mathbb{C}$
On raisonne par parties réelle ou imaginaire sachant que l'une des deux fonctions ne tend pas vers 0 en $+\infty$.
□

Attention : Etonnamment, la condition $f(t) \xrightarrow[t\to+\infty]{} 0$ n'est pas une condition nécessaire d'intégrabilité.

Exemple Soit $f : [0, +\infty[\to \mathbb{R}$ la fonction continue définie par



f est intégrable mais n'est pas de limite nulle en $+\infty$.
En effet, la fonction f est positive et

$$\int_0^x f(t)\,\mathrm{d}t \leqslant \int_0^{\lfloor x\rfloor+1} f(t)\,\mathrm{d}t = \sum_{n=1}^{\lfloor x\rfloor+1} \frac{1}{2^n} \leqslant \sum_{n=1}^{+\infty} \frac{1}{2^n} = 1$$

Les intégrales partielles de f sont majorées et donc $\displaystyle\int_0^{+\infty} f(t)\,\mathrm{d}t$ converge.
Aussi, f ne tend pas vers 0 en $+\infty$ car

$$f\left(n + \frac{1}{2^{n+1}}\right) = 1 \xrightarrow[n\to+\infty]{} 1$$

11.3 Extension à un intervalle quelconque

11.3.1 Intégration sur un intervalle semi ouvert

11.3.1.1 Intégration sur $[a, b[$

Soit $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$ avec $a < b$.

Définition

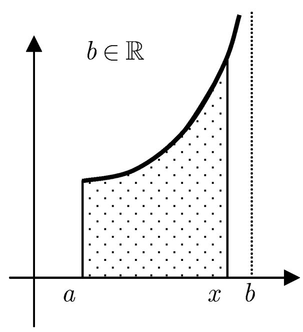
Soit $f : [a, b[\to \mathbb{K}$ continue par morceaux. On dit que l'intégrale de f sur $[a, b[$ converge si l'intégrale partielle $\displaystyle\int_a^x f(t)\,\mathrm{d}t$ converge quand $x \to b^-$.
On pose alors

$$\int_a^b f(t)\,\mathrm{d}t \underset{\text{déf}}{=} \lim_{x\to b^-} \int_a^x f(t)\,\mathrm{d}t$$

encore notée $\displaystyle\int_{[a,b[} f(t)\,\mathrm{d}t$.
On peut aussi introduire le reste d'intégrale convergente

$$\int_x^b f(t)\,\mathrm{d}t \xrightarrow[x\to b^-]{} 0$$

Remarque L'intégrale converge si, et seulement si, l'aire hachurée converge quand $x \to b^-$



Exemple Etude de $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^2}}$.
Il s'agit d'une intégrale impropre en la borne 1 (i.e. d'une intégrale sur $[0,1[$)
Puisque

$$\int_0^x \frac{\mathrm{d}t}{\sqrt{1-t^2}} = [\arcsin t]_0^x = \arcsin x \xrightarrow[x\to 1^-]{} \frac{\pi}{2}$$

Ainsi l'intégrale impropre $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^2}}$ converge et vaut $\dfrac{\pi}{2}$.
On peut aussi procéder à un calcul plus immédiat assurant directement la convergence

$$\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^2}} = \int_{[0,1[} \frac{\mathrm{d}t}{\sqrt{1-t^2}} = [\arcsin t]_0^{1^-} = \frac{\pi}{2}$$

11.3.1.2 Intégration sur $]a,b]$

Soit $a \in \mathbb{R}\cup\{-\infty\}$ et $b \in \mathbb{R}$ avec $a < b$.

Définition

Soit $f :]a,b] \to \mathbb{K}$ continue par morceaux avec $a \in \mathbb{R}\cup\{-\infty\}$ et $b \in \mathbb{R}$.
On dit que l'intégrale de f sur $]a,b]$ converge si l'intégrale partielle $\displaystyle\int_x^b f(t)\,\mathrm{d}t$ converge quand $x \to a^+$.
On pose alors

$$\int_{]a,b]} f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t = \lim_{x\to a^+} \int_x^b f(t)\,\mathrm{d}t$$

et on peut introduire le reste d'intégrale convergente

$$\int_a^x f(t)\,\mathrm{d}t \xrightarrow[x\to a^+]{} 0$$

Exemple Etude de $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{t}}$.
L'intégrale est impropre en la borne 0.
La fonction $t \mapsto 1/\sqrt{t}$ est définie et continue par morceaux sur $]0,1]$.

$$\int_x^1 \frac{\mathrm{d}t}{\sqrt{t}} = \left[2\sqrt{t}\right]_x^1 = 2 - 2\sqrt{x} \xrightarrow[x\to 0^+]{} 2$$

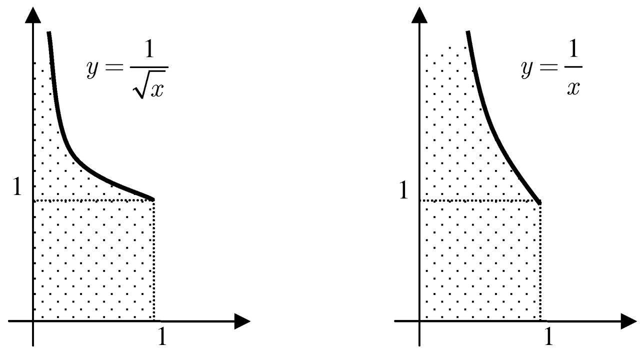
donc l'intégrale impropre $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{t}}$ converge et

$$\int_0^1 \frac{\mathrm{d}t}{\sqrt{t}} = 2$$

Exemple Etude de $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t}$.
L'intégrale est impropre en la borne 0.
La fonction $t \mapsto 1/t$ est définie et continue par morceaux sur $]0,1]$

$$\int_x^1 \frac{\mathrm{d}t}{t} = -\ln x \xrightarrow[x\to 0^+]{} +\infty$$

donc l'intégrale impropre $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t}$ diverge.
Pour la fonction inverse, il y a trop d'espace entre la courbe et l'axe des ordonnées pour que l'intégrale converge, cette fonction tend trop rapidement vers $+\infty$ en 0^+.



11.3.1.3 Lien avec une éventuelle intégration sur $[a,b]$

La notation $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ peut être ambiguë dans le cas où f est définie et continue par morceaux sur $[a,b]$. Cependant, il n'en est rien en vertu du résultat suivant.

Proposition

> Si $f : [a,b] \to \mathbb{K}$ continue par morceaux alors
>
> $$\int_a^x f(t)\,\mathrm{d}t \xrightarrow[x\to b^-]{} \int_a^b f(t)\,\mathrm{d}t$$
>
> où l'intégrale limite est comprise au sens de l'intégration sur un segment

dém. :
La fonction f est continue par morceaux sur le segment $[a,b]$, elle y est donc bornée par un certain $M \in \mathbb{R}^+$. On a alors

$$\left|\int_a^b f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\right| = \left|\int_x^b f(t)\,\mathrm{d}t\right| \leqslant \int_x^b |f(t)|\,\mathrm{d}t$$

puis

$$\left|\int_a^b f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\right| \leqslant \int_x^b M\,\mathrm{d}t = M(b-x) \xrightarrow[x\to+\infty]{} 0$$

□

Définition

Lorsqu'une fonction f est définie et continue par morceaux sur un segment $[a, b]$ on dit encore que son intégrale converge et l'on a

$$\int_{[a,b]} f(t)\,\mathrm{d}t = \int_{[a,b[} f(t)\,\mathrm{d}t = \int_{]a,b]} f(t)\,\mathrm{d}t$$

Cette valeur commune est celle désignée par

$$\int_a^b f(t)\,\mathrm{d}t$$

Remarque Soit $f : [a, b[\to \mathbb{K}$ continue. Si $f(t) \xrightarrow[t\to b^-]{} \ell \in \mathbb{K}$ alors on peut prolonger f par continuité en b. La fonction ainsi obtenue étant alors continue sur $[a, b]$, on peut affirmer que l'intégrale sur $[a, b[$ converge et vaut l'intégrale sur $[a, b]$. On dit alors que l'intégrale est faussement impropre en b.

Exemple Etude de $\displaystyle\int_0^{\pi/2} \frac{\sin t}{t}\,\mathrm{d}t$.

L'intégrale converge car faussement impropre en 0 puisque

$$\frac{\sin t}{t} \xrightarrow[t\to 0^+]{} 1$$

11.3.2 Intégrale sur un intervalle ouvert

Définition

Soit $f :]a, b[\to \mathbb{K}$ continue par morceaux avec $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in \mathbb{R} \cup \{+\infty\}$.
On dit que l'intégrale de f sur $]a, b[$ converge si, pour $c \in]a, b[$, les intégrales de f sur $]a, c]$ et sur $[c, b[$ convergent. On pose alors

$$\int_{]a,b[} f \underset{\text{déf}}{=} \int_{]a,c]} f + \int_{[c,b[} f$$

ou encore

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$$

Remarque Ni la notion, ni la valeur de l'intégrale ne dépendent du choix de $c \in]a, b[$.

Remarque Si $f : [a, b[\to \mathbb{K}$ est continue par morceaux, la convergence et la valeur des intégrales $\displaystyle\int_{]a,b[} f(t)\,\mathrm{d}t$ et $\displaystyle\int_{[a,b[} f(t)\,\mathrm{d}t$ sont les mêmes et encore une fois la notation $\displaystyle\int_a^b f(t)\,\mathrm{d}t$ ne crée pas d'ambiguïté.

Exemple Etude de $\displaystyle\int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{1+t^2}$.
La fonction $t \mapsto \dfrac{1}{1+t^2}$ est définie et continue par morceaux sur $\mathbb{R}$.
$\mathbb{R} = \left]-\infty, 0\right] \cup \left[0, +\infty\right[$
$\displaystyle\int_0^x \frac{\mathrm{d}t}{1+t^2} = \arctan x \xrightarrow[x\to+\infty]{} \frac{\pi}{2}$ donc $\displaystyle\int_{[0,+\infty[} \frac{\mathrm{d}t}{1+t^2}$ converge et vaut $\dfrac{\pi}{2}$.
$\displaystyle\int_x^0 \frac{\mathrm{d}t}{1+t^2} = -\arctan(x) \xrightarrow[x\to-\infty]{} \frac{\pi}{2}$ donc $\displaystyle\int_{]-\infty,0]} \frac{\mathrm{d}t}{1+t^2}$ converge et vaut $\dfrac{\pi}{2}$.
Par suite $\displaystyle\int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{1+t^2}$ converge et

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{1+t^2} = \pi$$

Exemple Etude de $\displaystyle\int_{\mathbb{R}} t\,\mathrm{d}t$.
La fonction $t \mapsto t$ est définie et continue par morceaux sur $\mathbb{R}$.
$\mathbb{R} = \left]-\infty, 0\right] \cup \left[0, +\infty\right[$
$\displaystyle\int_0^x t\,\mathrm{d}t = \frac{1}{2}x^2 \xrightarrow[x\to+\infty]{} +\infty$ donc $\displaystyle\int_{[0,+\infty[} t\,\mathrm{d}t$ diverge puis $\displaystyle\int_{\mathbb{R}} t\,\mathrm{d}t$ aussi.

Attention : Ici $\displaystyle\int_{-x}^{x} t\,\mathrm{d}t = 0 \xrightarrow[x\to+\infty]{} 0$. On n'aurait pu vouloir poser $\displaystyle\int_{\mathbb{R}} t\,\mathrm{d}t = 0$ mais cela n'est pas conforme à la définition.
En fait, on peut aussi remarquer $\displaystyle\int_{-x}^{x+1} t\,\mathrm{d}t = x + \frac{1}{2} \xrightarrow[x\to+\infty]{} +\infty$ et cette fois-ci $\displaystyle\int_{\mathbb{R}} t\,\mathrm{d}t$ n'a plus de sens.
Pour cette raison, la convergence d'une l'intégrale sur $\left]a, b\right[$ s'étudie en la coupant en deux et non en étudiant conjointement les deux bornes.

11.3.3 Propriétés

Les propriétés calculatoires de linéarité, de positivité et de conjugaison présentées pour les intégrales sur $\left[a, +\infty\right[$ restent vraies pour une intégration sur un intervalle I quelconque et se démontrent par des procédés analogues.

Théorème

L'ensemble des fonctions continues par morceaux de I vers $\mathbb{K}$ dont l'intégrale converge est un sous-espace vectoriel de l'espace $\mathcal{C}^0_{pm}(I, \mathbb{K})$ et l'application $f \mapsto \displaystyle\int_I f(t)\,\mathrm{d}t$ y définit une forme linéaire.

Théorème

Pour $f : I \to \mathbb{R}$ continue par morceaux.
Si $f \geqslant 0$ alors $\displaystyle\int_I f(t)\,\mathrm{d}t \geqslant 0$.

Théorème

Pour $f : I \to \mathbb{R}$ continue
Si $f \geqslant 0$ et si $\int_I f(t)\,\mathrm{d}t = 0$ alors f est la fonction nulle.

11.3.4 Relation de Chasles

Soit $f : I \to \mathbb{C}$ est continue par morceaux telle que $\int_I f$ converge.
Pour $a < b \in \bar{\mathbb{R}}$ des éléments ou des extrémités de I, la théorie qui précède permet de donner un sens à

$$\int_a^b f(t)\,\mathrm{d}t$$

en tant qu'intégrale convergente de f sur $[a, b]$, $]a, b]$, $[a, b[$ ou $]a, b[$ selon les possibilités. Si plusieurs interprétations sont possibles, celles-ci se correspondent. On pose encore

$$\int_b^a f(t)\,\mathrm{d}t = -\int_a^b f(t)\,\mathrm{d}t \text{ et } \int_a^a f(t)\,\mathrm{d}t = 0$$

On peut alors énoncer le résultat suivant

Théorème

Soit $f : I \to \mathbb{C}$ continue par morceaux telle que $\int_I f$ converge.
Pour tous a, b, c éléments ou extrémités de I, on a

$$\int_a^b f(t)\,\mathrm{d}t = \int_a^c f(t)\,\mathrm{d}t + \int_c^b f(t)\,\mathrm{d}t$$

avec convergence des intégrales engagées.

dém. :
Il suffit d'étudier tous les cas de figures possibles...

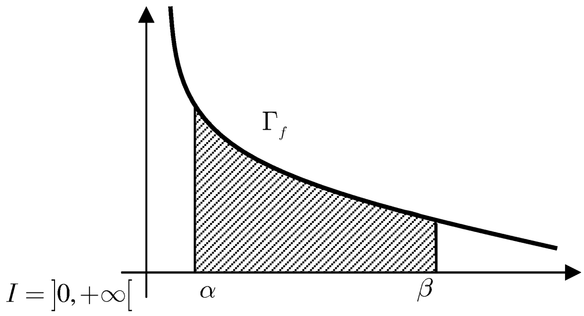
□

11.4 Intégrabilité sur un intervalle quelconque

11.4.1 Cas des fonctions positives

Théorème

Soit $f : I \to \mathbb{R}$ continue par morceaux et positive.
On a équivalence entre :
(i) $\displaystyle\int_I f$ converge ;
(ii) $\exists M \in \mathbb{R}, \forall [\alpha, \beta] \subset I, \displaystyle\int_\alpha^\beta f \leqslant M$.



dém. :
Notons $a < b \in \bar{\mathbb{R}}$ les extrémités de I.
(i) $\Rightarrow$ (ii) Supposons que $\displaystyle\int_I f = \int_a^b f$ converge. Pour tout $[\alpha, \beta] \subset I$,

$$\int_a^b f = \int_a^\alpha f + \int_\alpha^\beta f + \int_\beta^b f \geqslant \int_\alpha^\beta f$$

(ii) $\Rightarrow$ (i) Supposons (ii)
Cas $I = [a, b[$: l'intégrale partielle $\displaystyle\int_a^x f$ est croissante sur $[a, b[$ et majorée par M donc converge en b^-.
Ainsi $\displaystyle\int_{[a,b[} f$ converge.
Cas $I =]a, b]$: c'est analogue
Cas $I =]a, b[$: on découpe l'intervalle en $c \in]a, b[$.
□

Corollaire

Soit $f, g : I \to \mathbb{R}$ continues par morceaux telles que $0 \leqslant f \leqslant g$.
Si $\displaystyle\int_I g$ converge alors $\displaystyle\int_I f$ aussi.
Si $\displaystyle\int_I f$ diverge alors $\displaystyle\int_I g$ aussi.

11.4.2 Intégrabilité

Définition

On dit qu'une fonction $f : I \to \mathbb{K}$ continue par morceaux est intégrable sur I si l'intégrale $\int_I |f(t)|\,\mathrm{d}t$ converge. On dit encore que l'intégrale $\int_I f(t)\,\mathrm{d}t$ est absolument convergente.

Exemple Si $f : [a,b] \to \mathbb{K}$ est continue par morceaux alors f est intégrable sur $[a,b]$ mais aussi sur $]a,b]$, $[b,a[$ et $]a,b[$.

Théorème

Si $f : I \to \mathbb{K}$ continue par morceaux est intégrable alors l'intégrale $\int_I f$ converge et

$$\left|\int_I f\right| \leqslant \int_I |f|$$

dém. :
Cas f à valeurs positives : C'est immédiat par définition.
Cas f à valeurs réelles :
On pose $f^+ = \sup(f,0)$ et $f^- = \sup(-f,0)$.
Les fonctions $f^+, f^- : I \to \mathbb{R}^+$ sont continues par morceaux et vérifient $f = f^+ - f^-$.
On a aussi $|f| = f^+ + f^-$ donc $0 \leqslant f^+, f^- \leqslant |f|$.
Par comparaison de fonctions positives, les intégrales $\int_I f^+$ et $\int_I f^-$ convergent puis, par opérations, l'intégrale $\int_I f$ converge aussi.
Cas f à valeurs complexes
On écrit $f = \mathrm{Re}f + i\mathrm{Im}f$.
$\mathrm{Re}f, \mathrm{Im}f : I \to \mathbb{R}$ sont continues par morceaux.
Puisque $|\mathrm{Re}f|, |\mathrm{Im}f| \leqslant |f|$, on a, par comparaison de fonctions positives, $\int_I |\mathrm{Re}f|$ et $\int_I |\mathrm{Im}f|$ convergent donc $\int_I \mathrm{Re}f$ et $\int_I \mathrm{Im}f$ convergent puis par opérations $\int_I f$ aussi.
Démontrons maintenant l'inégalité

$$\left|\int_I f\right| \leqslant \int_I |f|$$

Notons $a < b \in \bar{\mathbb{R}}$ les extrémités de I.
Posons $c \in]a,b[$
Pour $x \in]a,c]$ et $y \in [c,b[$,

$$\left|\int_x^y f\right| \leqslant \int_x^y |f|$$

donne

$$\left|\int_x^c f + \int_c^y f\right| \leqslant \int_x^c |f| + \int_c^y |f|$$

A la limite quand $x \to a^+$

$$\left|\int_{]a,c]} f + \int_c^y f\right| \leqslant \int_{]a,c]} |f| + \int_c^y |f|$$

puis quand $y \to b^-$, on obtient

$$\left|\int_{]a,c]} f + \int_{[c,b[} f\right| \leqslant \int_{]a,c]} |f| + \int_{[c,b[} |f|$$

ce qui donne

$$\left|\int_I f\right| \leqslant \int_I |f|$$

□

*Bilan :*Pour une fonction réelle ou complexe

$$f \text{ intégrable } \Rightarrow \int_I f \text{ converge}$$

Pour une fonction positive, $f = |f|$ donc

$$f \text{ intégrable } \Leftrightarrow \int_I f \text{ convergence}$$

Plus généralement, pour une fonction de signe constant, il y a équivalence.

Attention : Il se peut que $\displaystyle\int_I f$ converge et $\displaystyle\int_I |f|$ diverge. Dans ce cas, on dit que l'intégrale $\displaystyle\int_I f$ est semi-convergente.

11.4.3 Opérations

11.4.3.1 Sur les fonctions

Théorème

Soit $f, g : I \to \mathbb{K}$ continues par morceaux et $\lambda, \mu \in \mathbb{K}$.
Si f et g sont intégrables alors $\lambda f + \mu g$ l'est aussi.

dém. :
On a

$$|\lambda f + \mu g| \leqslant |\lambda|\,|f| + |\mu|\,|g|$$

Or $\displaystyle\int_I |f(t)|\,\mathrm{d}t$ et $\displaystyle\int_I |g(t)|\,\mathrm{d}t$ convergent donc, par opérations $\displaystyle\int_I |\lambda|\,|f(t)| + |\mu|\,|g(t)|\,\mathrm{d}t$ converge.
Par comparaison de fonctions positives, $\displaystyle\int_I |\lambda f + \mu g|$ converge et donc $\lambda f + \mu g$ est intégrable.

□

Corollaire

L'ensemble $L^1(I, \mathbb{K})$ formé des fonctions de I vers $\mathbb{K}$ continues par morceaux et intégrable et un sous-espace vectoriel de l'espace $\mathcal{C}^0_{pm}(I, \mathbb{K})$ et l'application $f \mapsto \displaystyle\int_I f(t)\,\mathrm{d}t$ définit une forme linéaire sur $L^1(I, \mathbb{K})$.

Remarque En revanche, on ne peut rien dire quant au produit de deux fonctions intégrables.

Par exemple $1/\sqrt{t}$ est intégrable sur $]0,1]$ alors que $\dfrac{1}{t}=\dfrac{1}{\sqrt{t}}\times\dfrac{1}{\sqrt{t}}$ ne l'est pas.
Cependant, si f^2 et g^2 sont intégrables sur I alors le produit fg l'est aussi car

$$|fg|\leqslant\frac{1}{2}\left(|f|^2+|g|^2\right)$$

11.4.3.2 Sur l'intervalle

Proposition

> Soit $f : I\to\mathbb{K}$ continue par morceaux et J un intervalle inclus dans I.
> Si f est intégrable sur I alors f est intégrable sur J.

dém. :
Pour tout $[\alpha,\beta]\subset J$, on a

$$\int_\alpha^\beta |f(t)|\,\mathrm{d}t\leqslant\int_I |f(t)|\,\mathrm{d}t=M$$

et donc $\displaystyle\int_J |f(t)|\,\mathrm{d}t$ converge.
□

Proposition

> Soit $f :]a,b[\to\mathbb{K}$ continue par morceaux.
> f est intégrable sur $]a,b[$ si, et seulement si, f est intégrable sur $]a,c]$ et sur $[c,b[$.

dém. :
Car par définition

$$\int_{]a,b[}|f(t)|\,\mathrm{d}t \text{ converge si, et seulement si, } \int_{]a,c]}|f(t)|\,\mathrm{d}t \text{ et } \int_{[c,b[}|f(t)|\,\mathrm{d}t \text{ convergent}$$

□

11.4.4 Intégrabilité par comparaison

11.4.4.1 Domination

Théorème

> Soit $f : I\to\mathbb{K}$ et $\varphi : I\to\mathbb{R}^+$ continues par morceaux.
> Si
> $$\forall t\in I, |f(t)|\leqslant\varphi(t) \text{ avec } \varphi \text{ intégrable}$$
> alors f est intégrable.

dém. :
Par comparaison de fonctions positives, on obtient la convergence de $\displaystyle\int_I |f(t)|\,\mathrm{d}t$.
□

Exemple Si I est un intervalle borné et si $f : I\to\mathbb{K}$ est continue par morceaux et bornée alors f est intégrable sur I.

11.4.4.2 Comparaison asymptotique

Théorème

Soit $f, g : [a, b[\to \mathbb{C}$ continues par morceaux avec $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$.
Si $f(t) \underset{t \to b^-}{=} O\left(g(t)\right)$ et si g est intégrable alors f est intégrable.

Corollaire

Si $f(t) \underset{t \to b^-}{=} o\left(g(t)\right)$ avec g intégrable alors f l'est aussi.
Si $f(t) \underset{t \to b^-}{\sim} g(t)$ alors f est intégrable si, et seulement si, g l'est.

Remarque On peut énoncer des résultats analogues pour une étude d'intégrabilité sur $]a, b]$.

Exemple Soit $f, g : [a, b[\to \mathbb{R}^+$ continues par morceaux.
Si $f(t) \underset{t \to b^-}{\sim} g(t)$ alors les intégrales $\displaystyle\int_{[a,b[} f(t)\,\mathrm{d}t$ et $\displaystyle\int_{[a,b[} g(t)\,\mathrm{d}t$ ont même nature.

11.4.5 Intégrales de Riemann

11.4.5.1 Au voisinage de l'infini

Rappelons le résultat suivant.

Théorème

$$\int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha} \text{ converge si, et seulement si, } \alpha > 1$$

Par considération de symétrie, on a aussi

Théorème

$$\int_{-\infty}^{-1} \frac{\mathrm{d}t}{|t|^\alpha} \text{ converge si, et seulement si, } \alpha > 1$$

11.4.5.2 Au voisinage d'une extrémité finie

Théorème

Soit $a < b$ deux réels et $\alpha \in \mathbb{R}$

$$\int_a^b \frac{\mathrm{d}t}{(t-a)^\alpha} \text{ converge si, et seulement si, } \alpha < 1$$

dém. :
Etude de $\displaystyle\int_a^b \frac{\mathrm{d}t}{(t-a)^\alpha}$.

L'intégrale est impropre en a.
Cas $\alpha = 1$

$$\int_x^b \frac{\mathrm{d}t}{t-a} = [\ln(t-a)]_x^b = \ln(b-a) - \ln(x-a) \xrightarrow[x\to a^+]{} +\infty$$

et donc l'intégrale $\displaystyle\int_a^b \frac{\mathrm{d}t}{t-a}$ diverge.
Cas $\alpha \neq 1$
On a

$$\int_x^b \frac{\mathrm{d}t}{(t-a)^\alpha} = \left[\frac{1}{1-\alpha}\frac{1}{(t-a)^{\alpha-1}}\right]_x^b \xrightarrow[x\to a^+]{} \begin{cases} \dfrac{(b-a)^{1-\alpha}}{1-\alpha} & \text{si } \alpha < 1 \\ +\infty & \text{si } \alpha > 1 \end{cases}$$

et donc l'intégrale $\displaystyle\int_a^b \frac{\mathrm{d}t}{(t-a)^\alpha}$ converge si, et seulement si, $\alpha < 1$.
□

Exemple $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{t}}, \int_0^1 \frac{\mathrm{d}t}{t^{0,999}}$ convergent.
$\displaystyle\int_0^1 \frac{\mathrm{d}t}{t}$ et $\displaystyle\int_0^1 \frac{\mathrm{d}t}{t^2}$ divergent.

Exemple Pour $\lambda \in \mathbb{R}$, $\displaystyle\int_0^1 t^\lambda\,\mathrm{d}t = \int_0^1 \frac{\mathrm{d}t}{t^{-\lambda}}$ converge si, et seulement si, $\lambda > -1$.

Exemple L'intégrale $\displaystyle\int_0^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$ diverge pour toute valeur du réel α.

Théorème
Soit $a < b$ deux réels et $\alpha \in \mathbb{R}$

$$\int_a^b \frac{\mathrm{d}t}{(b-t)^\alpha} \text{ converge si, et seulement si, } \alpha < 1$$

dém. :
C'est une configuration symétrique de la précédente.
□

Exemple $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t}}$ converge alors que $\displaystyle\int_0^1 \frac{\mathrm{d}t}{1-t}$ diverge.

11.4.6 En pratique

11.4.6.1 Intégrabilité sur $[a, +\infty[$ ou $]-\infty, a]$

Les démarches d'intégrabilité déjà vu sur $[a, +\infty[$ se transposent à $]-\infty, a]$ en écrivant $|t|^\alpha$ au lieu de t^α lorsque l'exposant α est non entier.

Exemple Nature de $\displaystyle\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t$.

La fonction $f : t \mapsto \mathrm{e}^{-t^2}$ est définie et continue par morceaux sur $]-\infty, +\infty[$.
On a $t^2 f(t) \xrightarrow[t\to+\infty]{} 0$ donc f est intégrable sur $[0, +\infty[$
On a $t^2 f(t) \xrightarrow[t\to-\infty]{} 0$ donc f est intégrable sur $]-\infty, 0]$.
Finalement f est intégrable sur $\mathbb{R}$.

11.4.6.2 Intégrabilité sur $]0, a]$

Exemple Nature de $\displaystyle\int_0^{+\infty} \frac{t}{\mathrm{e}^t - 1}\,\mathrm{d}t$.La fonction $f : t \mapsto t/(\mathrm{e}^t - 1)$ est définie et continue par morceaux sur $]0, +\infty[$.
On a

$$\frac{t}{\mathrm{e}^t - 1} \underset{t\to0^+}{\sim} \frac{t}{t} \to 1$$

La fonction est prolongeable par continuité et l'intégrale est faussement impropre en 0.
On a aussi

$$t^2 \times \frac{t}{\mathrm{e}^t - 1} \underset{t\to+\infty}{\sim} t^3\mathrm{e}^{-t} \xrightarrow[t\to+\infty]{} 0$$

et donc f est intégrable sur $[1, +\infty[$.
Finalement, f est intégrable sur $]0, +\infty[$.

Exemple Nature de $\displaystyle\int_0^1 \frac{\cos t}{\sqrt{t}}\,\mathrm{d}t$.

La fonction $f : t \mapsto \cos(t)/\sqrt{t}$ est définie et continue par morceaux sur $]0, 1]$.
On a $f(t) \xrightarrow[t\to0^+]{} +\infty$ mais ce n'est en rien décisif.
Cependant

$$f(t) \underset{t\to0^+}{\sim} 1/\sqrt{t}$$

or $t \to 1/\sqrt{t}$ est intégrable sur $]0, 1]$ ($\alpha = 1/2 < 1$) donc f est intégrable sur $]0, 1]$ et $\displaystyle\int_0^1 \frac{\cos t}{\sqrt{t}}\,\mathrm{d}t$ converge.

Exemple Nature de $\displaystyle\int_0^1 \ln t\,\mathrm{d}t$.

La fonction $f : t \mapsto \ln t$ est définie et continue par morceaux sur $]0, 1]$.
$\sqrt{t} f(t) \xrightarrow[t\to+\infty]{} 0$ donc $f(t) \underset{t\to+\infty}{=} o\left(1/\sqrt{t}\right)$.
Or $t \to 1/\sqrt{t}$ est intégrable sur $]0, 1]$ ($\alpha = 1/2 < 1$) donc f est intégrable sur $]0, 1]$ et $\displaystyle\int_0^1 \ln t\,\mathrm{d}t$ converge.

Exemple Nature de $\displaystyle\int_0^1 \frac{\ln t}{t}\,\mathrm{d}t$.
La fonction $f : t \mapsto \ln(t)/t$ est définie et continue par morceaux sur $]0,1]$.
Quand $t \to 0^+$, $tf(t) \to -\infty$.
Il existe $a > 0$ tel que sur $]0,a]$, $f(t) \leqslant -1/t \leqslant 0$.
Par comparaison de fonctions négatives, l'intégrale $\displaystyle\int_0^1 \frac{\ln t}{t}\,\mathrm{d}t$ diverge.

*Bilan :*Pour $f :]0,a] \to \mathbb{C}$ continue par morceaux :
- si $f(t) \to \ell \in \mathbb{C}$ alors f est intégrable sur $]0,a]$;
- si $f(t) \underset{t\to 0^+}{\sim} C/t^\alpha$ alors f est intégrable sur $]0,a]$ si, et seulement si, $\alpha < 1$;
- s'il existe $\alpha < 1$ vérifiant $t^\alpha f(t) \xrightarrow[t\to 0^+]{} 0$ alors f est intégrable sur $]0,a]$;
- si $tf(t) \xrightarrow[t\to 0^+]{} \ell \neq 0$ alors l'intégrale de f sur $]0,a]$ diverge.

11.4.6.3 Intégration $]a,b]$ ou $[a,b[$

On transpose les démarches ci-dessus. Il pourra être pertinent de se ramener en 0 par translation/symétrie de la variable pour mieux percevoir les ordres de grandeur.

Exemple Nature de $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^3}}$.
La fonction $f : t \mapsto 1/\sqrt{1-t^3}$ est définie et continue par morceaux sur $[0,1[$.
Quand $t \to 1^-$, $t = 1-h$ avec $h \to 0^+$.

$$f(t) \sim \frac{1}{\sqrt{3h}} = \frac{1/\sqrt{3}}{\sqrt{1-t}}$$

Or $t \mapsto 1/\sqrt{1-t}$ est intégrable sur $[0,1[$ donc f aussi et $\displaystyle\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^3}}$ converge.

Exemple Nature de $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^2-1}$.
La fonction $f : t \mapsto 1/(t^2-1)$ est définie et continue par morceaux par morceaux sur $]1,+\infty[$.
Quand $t \to 1^+$, $t = 1+h$ avec $h \to 0^+$.

$$f(t) \sim \frac{1}{2h} = \frac{1}{2(t-1)}$$

Or $t \mapsto \dfrac{1}{t-1}$ n'est pas intégrable sur $]1,2]$ donc f non plus.
A fortiori, f n'est pas intégrable sur $]1,+\infty[$.
Puisque f est de signe constant, on peut affirmer que l'intégrale $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t^2-1}$ diverge.

Exemple Etude de $\displaystyle\int_0^1 \frac{t-1}{\ln t}\,\mathrm{d}t$.

La fonction $f : t \mapsto (t-1)/\ln t$ est définie et continue par morceaux sur $]0,0[=]0,1/2] \cup [1/2,1[$.
D'une part

$$f(t) \xrightarrow[t\to 0^+]{} 0$$

donc f est intégrable sur $]0,1/2]$.
D'autre part

$$f(t) \xrightarrow[t\to 1^-]{} 1$$

donc f est intégrable sur $[1/2,1[$.
Finalement f est intégrable sur $]0,1[$ et $\displaystyle\int_0^1 \frac{t-1}{\ln t}\,\mathrm{d}t$ converge.
Elle vaut $\ln 2$, mais c'est une longue histoire...

11.5 Calcul d'intégrales impropres

11.5.1 Par les intégrales partielles ou détermination de primitive

Pour justifier l'existence tout en calculant $\displaystyle\int_a^b f(t)\,\mathrm{d}t = \int_{[a,b[} f(t)\,\mathrm{d}t$ on peut

- calculer l'intégrale partielle $\displaystyle\int_a^x f(t)\,\mathrm{d}t$ puis passer à la limite quand $x \to b^-$,

- introduire une primitive F de f (supposée continue) et exploiter $\displaystyle\int_{[a,b[} f(t)\,\mathrm{d}t = [F]_a^{b^-}$.

Pour $\displaystyle\int_a^b f(t)\,\mathrm{d}t = \int_{]a,b[} f(t)\,\mathrm{d}t$ on peut

- calculer $\displaystyle\int_x^y f(t)\,\mathrm{d}t$ puis passer à la limite quand $x \to a^+$ et $y \to b^-$,

- introduire une primitive F de f et exploiter $\displaystyle\int_{]a,b[} f(t)\,\mathrm{d}t = [F]_{a^+}^{b^-}$.

Exemple Calcul de $\displaystyle\int_1^{+\infty} \frac{\mathrm{d}t}{t(t+1)}$.
On peut justifier l'existence a priori de l'intégrale par l'argument d'intégrabilité

$$\frac{1}{t(t+1)} \underset{t\to+\infty}{\sim} \frac{1}{t^2}$$

Ce qui suit va aussi justifier l'existence tout en donnant la valeur
On calcule l'intégrale grâce à la décomposition en éléments simples

$$\frac{1}{t(t+1)} = \frac{1}{t} - \frac{1}{t+1}$$

1ère méthode :

$$\int_1^x \frac{\mathrm{d}t}{t(t+1)} = \int_1^x \frac{\mathrm{d}t}{t} - \int_1^x \frac{\mathrm{d}t}{t+1} = \ln x - \ln(x+1) + \ln 2 \xrightarrow[x\to+\infty]{} \ln 2$$

2ème méthode :

$$\int_1^{+\infty} \frac{\mathrm{d}t}{t(t+1)} = \int_1^{+\infty} \left(\frac{1}{t} - \frac{1}{t+1}\right)\mathrm{d}t = \left[\ln\frac{t}{t+1}\right]_1^{+\infty} = -\ln\frac{1}{2} = \ln 2$$

Le calcul direct par primitive est souvent plus rapide, mais permet moins de liberté qu'un calcul mené par les intégrales partielles.

11.5.2 Changement de variable

Théorème

Soit $\varphi :]a,b[\to]\alpha,\beta[$ une bijection de classe $\mathcal{C}^1$ croissante et $f :]\alpha,\beta[\to \mathbb{K}$ une fonction continue par morceaux. On a équivalence entre :

(i) $\displaystyle\int_\alpha^\beta f(u)\,\mathrm{d}u$ converge ;

(ii) $\displaystyle\int_a^b f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t$ converge.

De plus, si tel est le cas

$$\int_a^b f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t = \int_\alpha^\beta f(u)\,\mathrm{d}u$$

dém. :

(i) $\Rightarrow$ (ii) Supposons la convergence de $\displaystyle\int_\alpha^\beta f(u)\,\mathrm{d}u$.

Soit $c \in]a,b[$ et $\gamma = \varphi(c)$. Pour $x \in [c,b[$, on a

$$\int_c^x f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t \underset{u=\varphi(t)}{=} \int_\gamma^{\varphi(x)} f(u)\,\mathrm{d}u$$

Puisque φ est une bijection croissante

$$\varphi(x) \xrightarrow[x\to b^-]{} \beta$$

et donc

$$\int_c^x f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t \xrightarrow[x\to b^-]{} \int_\gamma^\beta f(u)\,\mathrm{d}u$$

L'intégrale $\displaystyle\int_c^b f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t$ converge et vaut $\displaystyle\int_\gamma^\beta f(u)\,\mathrm{d}u$.

De même, l'intégrale $\displaystyle\int_a^c f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t$ converge et vaut $\displaystyle\int_\alpha^\gamma f(u)\,\mathrm{d}u$.

Finalement $\displaystyle\int_a^b f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t$ converge et vaut $\displaystyle\int_\alpha^\beta f(u)\,\mathrm{d}u$.

(ii) $\Rightarrow$ (i) Même démarche en exploitant φ^{-1}.

□

Remarque Si $\varphi :]a,b[\to]\alpha,\beta[$ est une bijection de classe $\mathcal{C}^1$ décroissante, on a un résultat analogue avec

$$\int_a^b f(\varphi(t))\,\varphi'(t)\,\mathrm{d}t = \int_\beta^\alpha f(u)\,\mathrm{d}u$$

Remarque En appliquant aussi ce résultat avec $|f|$ et en exploitant que φ' est de signe constant, on obtient aussi

$$u \mapsto f\left(\varphi(u)\right)\varphi'(u) \text{ intégrable sur }]a,b[\text{ si, et seulement si, } u \mapsto f(u) \text{ est intégrable sur }]\alpha,\beta[$$

Exemple Calcul de $\displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-\sqrt{t}}}{\sqrt{t}}\,\mathrm{d}t$.
La fonction $f : t \mapsto \mathrm{e}^{-\sqrt{t}}/\sqrt{t}$ est définie et continue par morceaux sur $]0,+\infty[$.
Réalisons le changement de variable $u = \sqrt{t}$
La fonction $t \mapsto \sqrt{t}$ réalise une bijection de classe $\mathcal{C}^1$ de $]0,+\infty[$ vers $]0,+\infty[$

$$u = \sqrt{t},\, t = u^2,\, \mathrm{d}t = 2u\,\mathrm{d}u$$

et donc

$$\int_0^{+\infty} \frac{\mathrm{e}^{-\sqrt{t}}}{\sqrt{t}}\,\mathrm{d}t = \int_0^{+\infty} 2\mathrm{e}^{-u}\,\mathrm{d}u$$

Puisque l'intégrale obtenue par le changement de variable est connue convergente, il en est de même de l'intégrale initiale et donc

$$\int_0^{+\infty} \frac{\mathrm{e}^{-\sqrt{t}}}{\sqrt{t}}\,\mathrm{d}t = \left[-2\mathrm{e}^{-u}\right]_0^{+\infty} = 2$$

11.5.3 Intégration par parties

Théorème

Soit I un intervalle d'extrémités $a < b \in \bar{\mathbb{R}}$ et $u, v : I \to \mathbb{K}$ de classe $\mathcal{C}^1$.
Si le produit uv converge en a^+ et b^- alors les intégrales

$$\int_a^b u'(t)v(t)\,\mathrm{d}t \text{ et } \int_a^b u(t)v'(t)\,\mathrm{d}t$$

ont même nature et en cas de convergence

$$\int_a^b u'(t)v(t)\,\mathrm{d}t = [uv]_{a^+}^{b^-} - \int_a^b u(t)v'(t)$$

dém. :
La fonction uv est de classe $\mathcal{C}^1$ avec $(uv)' = u'v + uv'$.
Si uv converge en a^+ et b^- alors, il y a convergence de l'intégrale

$$\int_a^b u'(t)v(t) + u(t)v'(t)\,\mathrm{d}t$$

et

$$\int_a^b u'(t)v(t) + u(t)v'(t)\,\mathrm{d}t = [uv]_a^b$$

Si l'une des intégrales

$$\int_a^b u'(t)v(t)\,\mathrm{d}t \text{ ou } \int_a^b u(t)v'(t)\,\mathrm{d}t$$

alors, par opérations, l'autre aussi et

$$\int_a^b u'(t)v(t)\,\mathrm{d}t + \int_a^b u(t)v'(t)\,\mathrm{d}t = [uv]_a^b$$

□

Exemple Soit $n \in \mathbb{N}$. Calcul de $I_n = \displaystyle\int_0^{+\infty} t^n \mathrm{e}^{-t}\,\mathrm{d}t$.
$f_n : t \mapsto t^n \mathrm{e}^{-t}$ est définie et continue par morceaux sur $[0, +\infty[$.
Quand $t \to +\infty$, $t^2 f_n(t) \to 0$ donc l'intégrale définissant I_n converge.
Posons $u'(t) = \mathrm{e}^{-t}$ et $v(t) = t^n$ avec $u(t) = -\mathrm{e}^{-t}$ et $v'(t) = nt^{n-1}$.
Les fonctions u et v sont de classe $\mathcal{C}^1$ et uv possède des limites finies en 0 et $+\infty$.
Par intégration par parties

$$I_n = \int_0^{+\infty} t^n \mathrm{e}^{-t}\,\mathrm{d}t = \left[-t^n \mathrm{e}^{-t}\right]_0^{+\infty} - \int_0^{+\infty} nt^{n-1}(-\mathrm{e}^{-t})\,\mathrm{d}t$$

avec convergence de l'intégrale introduite en second membre.
Ainsi $I_n = nI_{n-1}$ puis, sachant $I_0 = \displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t = 1$, on conclut

$$I_n = n!$$

Exemple Calcul de $\displaystyle\int_0^1 \frac{\ln(t)}{(1+t)^2}\,\mathrm{d}t$.
$f : t \mapsto \ln(t)/(1+t)^2$ est définie et continue par morceaux sur $]0, 1]$.

$$\sqrt{t}f(t) \underset{t\to 0^+}{\sim} \sqrt{t}\ln(t) \to 0$$

donc f intégrable sur $]0, 1]$ et donc l'intégrale $\displaystyle\int_0^1 \frac{\ln(t)}{(1+t)^2}\,\mathrm{d}t$ converge.
Posons $u'(t) = 1/(1+t)^2$ et $v(t) = \ln(t)$ avec $u(t) = -1/(1+t)$ et $v'(t) = 1/t$.
Les fonctions u et v sont de classe $\mathcal{C}^1$ mais le produit uv ne possède pas une limite finie en 0.
On ne peut procéder à cette intégration par parties... Il y a cependant deux solutions
1ère méthode : on réalise l'intégration par parties sur les intégrales partielles
Pour $x \in]0, 1]$

$$\int_x^1 \frac{\ln t}{(1+t)^2}\,\mathrm{d}t = \left[-\frac{\ln t}{1+t}\right]_x^1 + \int_x^1 \frac{\mathrm{d}t}{t(t+1)}$$

et donc

$$\int_x^1 \frac{\ln t}{(1+t)^2}\,\mathrm{d}t = \frac{\ln x}{1+x} + [\ln t - \ln(t+1)]_x^1$$

puis

$$\int_x^1 \frac{\ln t}{(1+t)^2}\,\mathrm{d}t = -\frac{x\ln x}{1+x} + \ln(1+x) - \ln 2 \underset{x\to 0}{\to} -\ln 2$$

donc
$$\int_0^1 \frac{\ln t}{(1+t)^2}\,\mathrm{d}t = -\ln 2$$
2ème méthode : on choisit $u(t) = t/(1+t)$ qui est aussi convenable et qui s'annulant en 0, permet d'avoir le produit uv convergeant en 0
$$\int_0^1 \frac{\ln t}{(1+t)^2}\,\mathrm{d}t = \left[\frac{t\ln t}{1+t}\right]_0^1 - \int_0^1 \frac{\mathrm{d}t}{t+1} = -\ln 2$$

11.6 Musculation

11.6.1 Intégrales de Bertrand

Théorème

Pour $\alpha,\beta \in \mathbb{R}$, $\displaystyle\int_{\mathrm{e}}^{+\infty} \frac{\mathrm{d}t}{t^\alpha (\ln t)^\beta}$ converge si, et seulement si, $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$).

dém. :
La fonction $f : t \mapsto 1/t^\alpha(\ln t)^\beta$ est définie, continue et positive sur $[\mathrm{e}, +\infty[$.
Cas $\alpha < 1$
$$tf(t) = \frac{t^{1-\alpha}}{(\ln t)^\beta} \xrightarrow[t\to+\infty]{} +\infty$$
donc pour t assez grand
$$f(t) \geqslant 1/t \geqslant 0$$
Or $\displaystyle\int_{\mathrm{e}}^{+\infty} \frac{\mathrm{d}t}{t}$ diverge donc par comparaison de fonctions positives, $\displaystyle\int_{\mathrm{e}}^{+\infty} \frac{\mathrm{d}t}{t^\alpha(\ln t)^\beta}$ diverge.
Cas $\alpha > 1$:
Sous cas inutile : $\beta > 0$
On a
$$t^\alpha f(t) \xrightarrow[t\to+\infty]{} 0$$
donc f est intégrable sur $[\mathrm{e}, +\infty[$ car $f(t) = o(1/t^\alpha)$ avec $\alpha > 1$.
Sous cas général :
On introduit $m \in \,]1, \alpha[$, on a
$$t^m f(t) = \frac{1}{t^{\alpha-m}(\ln t)^\beta} \xrightarrow[t\to+\infty]{} 0$$
donc f est intégrable sur $[\mathrm{e}, +\infty[$ car $f(t) = o(1/t^m)$ avec $m > 1$.
Cas $\alpha = 1$
$$\int_{\mathrm{e}}^{x} \frac{\mathrm{d}t}{t(\ln t)^\beta} \underset{u=\ln t}{=} \int_1^{\ln x} \frac{\mathrm{d}u}{u^\beta}$$
converge quand $x \to +\infty$ si, et seulement si, $\beta > 1$.
□

11.6.2 L'intégrale de Dirichlet

Proposition

L'intégrale $\displaystyle\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t$ converge.

dém. :
La fonction $t \mapsto \dfrac{\sin t}{t}$ est définie et continue par morceaux sur $]0,+\infty[$.
Cette fonction se prolonge par continuité en 0 donc $\displaystyle\int_{]0,1]} \frac{\sin t}{t}\,\mathrm{d}t$ converge.
Etudions $\displaystyle\int_{[1,+\infty[} \frac{\sin t}{t}\,\mathrm{d}t$
Soit $A \geqslant 1$. Par intégration par parties

$$\int_1^A \frac{\sin t}{t}\,\mathrm{d}t = \left[\frac{-\cos t}{t}\right]_1^A - \int_1^A \frac{\cos t}{t^2}\,\mathrm{d}t$$

Quand $A \to +\infty$,

$$\frac{\cos A}{A} \to 0 \text{ et } \int_1^A \frac{\cos t}{t^2}\,\mathrm{d}t \xrightarrow[A\to+\infty]{} \int_1^{+\infty} \frac{\cos t}{t^2}\,\mathrm{d}t$$

car cette dernière intégrale converge puisque

$$\frac{\cos t}{t^2} \underset{t\to+\infty}{=} O\left(\frac{1}{t^2}\right)$$

□

Remarque Par une intégration par parties judicieuse, on montre

$$\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t = \int_0^{+\infty} \frac{1-\cos t}{t^2}\,\mathrm{d}t$$

En exploitant $1 - \cos t = 2\sin^2(t/2)$ et le changement de variable $u = t/2$

$$\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t = \int_0^{+\infty} \frac{\sin^2 u}{u^2}\,\mathrm{d}u$$

Proposition

La fonction $t \mapsto \dfrac{\sin t}{t}$ n'est pas intégrable sur $]0,+\infty[$

dém. :
Montrons que $\displaystyle\int_0^{+\infty} \left|\frac{\sin t}{t}\right|\mathrm{d}t$ diverge, le problème se posant en $+\infty$.

$$\int_0^{n\pi} \frac{|\sin t|}{t}\,\mathrm{d}t = \sum_{k=1}^n \int_{(k-1)\pi}^{k\pi} \frac{|\sin t|}{t}\,\mathrm{d}t = \sum_{k=1}^n \int_0^{\pi} \frac{\sin u}{u+(k-1)\pi}\,\mathrm{d}u$$

Or

$$\int_0^{\pi} \frac{\sin u}{u+(k-1)\pi}\,\mathrm{d}u \geqslant \int_0^{\pi} \frac{\sin u}{k\pi}\,\mathrm{d}u = \frac{2}{k\pi}$$

donc

$$\int_0^{n\pi} \left|\frac{\sin t}{t}\right| \mathrm{d}t \geqslant \frac{2}{\pi}\sum_{k=1}^{n} \frac{1}{k} \to +\infty$$

□

Remarque On peut montrer que $\displaystyle\int_0^{+\infty} \frac{\sin t}{t}\,\mathrm{d}t = \frac{\pi}{2}$ mais c'est une longue histoire…

Chapitre 12

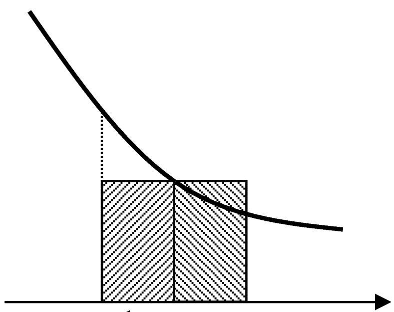
Comportement asymptotique de sommes et d'intégrales

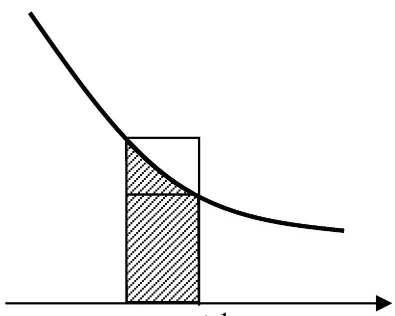
$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

12.1 Comparaison série intégrale

12.1.1 Principe

Cas f décroissante :





On a

$\int_n^{n+1} f(t)\,\mathrm{d}t \leqslant f(n) \leqslant \int_{n-1}^n f(t)\,\mathrm{d}t$ et $f(n+1) \leqslant \int_n^{n+1} f(t)\,\mathrm{d}t \leqslant f(n)$

Cas f croissante :

$\int_{n-1}^n f(t)\,\mathrm{d}t \leqslant f(n) \leqslant \int_n^{n+1} f(t)\,\mathrm{d}t$ et $f(n) \leqslant \int_n^{n+1} f(t)\,\mathrm{d}t \leqslant f(n+1)$

Théorème

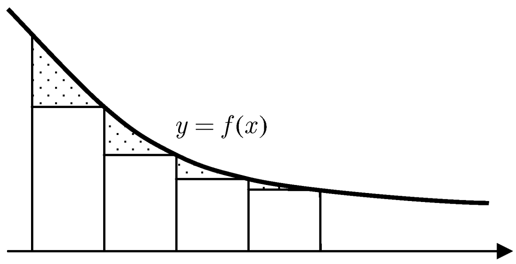
> Soit $f : [0, +\infty[\to \mathbb{R}$ continue par morceaux, décroissante et positive.
> La série de terme général $w_n = \int_{n-1}^n f(t)\,\mathrm{d}t - f(n)$ est convergente.

dém. :

Puisque f est décroissante, on a

$f(n) \leqslant \int_{n-1}^{n} f(t)\,\mathrm{d}t \leqslant f(n-1)$
et donc
$0 \leqslant w_n \leqslant f(n-1) - f(n)$
La nature de $\sum (f(n-1) - f(n))$ est celle de la suite $(f(n))$.
Or la fonction f est décroissante et minorée, elle converge donc en $+\infty$ et par conséquent, la suite $(f(n))$ aussi. Ainsi la série $\sum (f(n-1) - f(n))$ converge et, par comparaison de séries à termes positifs, la série de terme général w_n est convergente.
□

Remarque Cet énoncé signifie qu'il y a convergence des portions d'aire hachurée dans la figure ci-dessous



Corollaire

> Sous les hypothèses qui précédent, la série $\sum f(n)$ et l'intégrale impropre $\int_0^{+\infty} f(t)\,\mathrm{d}t$ sont de même nature.

dém. :
Puisque $\sum w_n$ converge, $\sum_{n\geqslant 1} f(n)$ et $\sum_{n\geqslant 1} \int_{n-1}^{n} f(t)\,\mathrm{d}t$ sont de même nature. Or
$\sum_{k=1}^{n} \int_{k-1}^{k} f(t)\,\mathrm{d}t = \int_0^n f(t)\,\mathrm{d}t$
Si l'intégrale $\int_0^{+\infty} f(t)\,\mathrm{d}t$ converge alors la série $\sum_{n\geqslant 1} f(n)$ converge.
Si l'intégrale $\int_0^{+\infty} f(t)\,\mathrm{d}t$ diverge alors, puisque f est positive $\int_0^x f(t)\,\mathrm{d}t \xrightarrow[x\to+\infty]{} +\infty$ et donc $\sum_{n\geqslant 1} f(n)$ diverge.
□

Exemple Pour $\alpha > 0$, la fonction $t \mapsto 1/t^\alpha$ est décroissante et l'on retrouve

$$\sum \frac{1}{n^\alpha} \text{ converge si, et seulement si, } \int_1^{+\infty} \frac{\mathrm{d}t}{t^\alpha} \text{ converge}$$

Remarque On peut aussi faire le lien entre la convergence des séries de Bertrand et celle des intégrales de Bertrand.

12.1.2 Reste d'une série de Riemann convergente

Pour $\alpha > 1$, la série $\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ est convergente.
Donnons un équivalent de son reste de rang n.
La fonction $t \mapsto \frac{1}{t^\alpha}$ est décroissante sur $]0, +\infty[$.
Pour $k \geqslant 2$,
$$\int_k^{k+1} \frac{\mathrm{d}t}{t^\alpha} \leqslant \frac{1}{k^\alpha} \leqslant \int_{k-1}^k \frac{\mathrm{d}t}{t^\alpha}$$
donc
$$\int_{n+1}^{N+1} \frac{\mathrm{d}t}{t^\alpha} \leqslant \sum_{k=n+1}^{N} \frac{1}{k^\alpha} \leqslant \int_n^N \frac{\mathrm{d}t}{t^\alpha}$$
Quand $N \to +\infty$,
$$\int_{n+1}^{+\infty} \frac{\mathrm{d}t}{t^\alpha} \leqslant \sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \leqslant \int_n^{+\infty} \frac{\mathrm{d}t}{t^\alpha}$$
avec convergence des intégrales engagées.
Or
$$\int_n^{+\infty} \frac{\mathrm{d}t}{t^\alpha} = \frac{1}{\alpha - 1}\frac{1}{n^{\alpha-1}} \text{ et } \int_{n+1}^{+\infty} \frac{\mathrm{d}t}{t^\alpha} \sim \frac{1}{\alpha-1}\frac{1}{n^{\alpha-1}}$$
donc par encadrement
$$\sum_{k=n+1}^{+\infty} \frac{1}{k^\alpha} \sim \frac{1}{\alpha-1}\frac{1}{n^{\alpha-1}}$$
Exemple En particulier
$$\sum_{k=n+1}^{+\infty} \frac{1}{k^2} \sim \frac{1}{n}$$

12.1.3 Sommes partielles d'une série de Riemann divergente

Pour $\alpha \leqslant 1$, la série $\sum_{n\geqslant 1} \frac{1}{n^\alpha}$ est divergente.
Donnons un équivalent de sa somme partielle de rang n.
Cas $\alpha = 1$.
On sait déjà :
$$\sum_{k=1}^{n} \frac{1}{k} = \ln n + \gamma + o(1)$$
Cas $0 < \alpha < 1$.
La fonction $t \mapsto \frac{1}{t^\alpha}$ est décroissante sur $]0, +\infty[$.
$$\int_k^{k+1} \frac{\mathrm{d}t}{t^\alpha} \leqslant \frac{1}{k^\alpha} \leqslant \int_{k-1}^k \frac{\mathrm{d}t}{t^\alpha}$$
En sommant
$$\int_1^{n+1} \frac{\mathrm{d}t}{t^\alpha} \leqslant \sum_{k=1}^{n} \frac{1}{k^\alpha} \leqslant \int_0^n \frac{\mathrm{d}t}{t^\alpha}$$
(avec convergence de l'intégrale de droite).
Or

$\displaystyle\int_0^n \frac{\mathrm{d}t}{t^\alpha} = \frac{n^{1-\alpha}}{1-\alpha}$ et $\displaystyle\int_1^{n+1} \frac{\mathrm{d}t}{t^\alpha} \sim \frac{n^{1-\alpha}}{1-\alpha}$
donc par comparaison
$\displaystyle\sum_{k=1}^n \frac{1}{k^\alpha} \sim \frac{n^{1-\alpha}}{1-\alpha}$
Cas $\alpha \leqslant 0$.
On écrit $\alpha = -\beta$ (avec $\beta \geqslant 0$) et on étudie
$\displaystyle\sum_{k=1}^n \frac{1}{k^\alpha} = \sum_{k=1}^n k^\beta$
La fonction $x \mapsto x^\beta$ est croissante sur $[0, +\infty[$.
$\displaystyle\int_{k-1}^k t^\beta \,\mathrm{d}t \leqslant k^\beta \leqslant \int_k^{k+1} t^\beta \,\mathrm{d}t$
En sommant
$\displaystyle\int_0^n t^\beta \,\mathrm{d}t \leqslant \sum_{k=1}^n k^\beta \leqslant \int_1^{n+1} t^\beta \,\mathrm{d}t$
Or
$\displaystyle\int_0^n t^\beta \,\mathrm{d}t = \frac{n^{\beta+1}}{\beta+1}$ et $\displaystyle\int_1^{n+1} t^\beta \,\mathrm{d}t \sim \frac{n^{\beta+1}}{\beta+1}$
donc par encadrement
$\displaystyle\sum_{k=1}^n k^\beta \sim \frac{n^{\beta+1}}{\beta+1}$ i.e. $\displaystyle\sum_{k=1}^n \frac{1}{k^\alpha} \sim \frac{n^{1-\alpha}}{1-\alpha}$

Exemple En particulier

$$\sum_{k=1}^n \frac{1}{\sqrt{k}} \sim 2\sqrt{n}$$

12.2 Sommation des relations de comparaison

12.2.1 Cas de la convergence

Théorème

Soit $\sum u_n$ une série numérique et $\sum v_n$ une série à termes positifs convergente.
Si $u_n = o(v_n)$ alors la série $\sum u_n$ converge et

$$\sum_{k=n+1}^{+\infty} u_k = o\left(\sum_{k=n+1}^{+\infty} v_k\right)$$

Si $u_n = O(v_n)$ alors la série $\sum u_n$ converge et

$$\sum_{k=n+1}^{+\infty} u_k = O\left(\sum_{k=n+1}^{+\infty} v_k\right)$$

Si $u_n \sim v_n$ alors la série $\sum u_n$ converge et

$$\sum_{k=n+1}^{+\infty} u_k \sim \sum_{k=n+1}^{+\infty} v_k$$

dém. :
Cas $u_n = o(v_n)$.
Par comparaison, la série $\sum u_n$ est absolument convergente.
Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que
$\forall n \geqslant N, |u_n| \leqslant \varepsilon |v_n| = \varepsilon v_n$
Pour $k \geqslant n+1$, $|u_k| \leqslant \varepsilon v_k$ puis en sommant
$$\left|\sum_{k=n+1}^{+\infty} u_k\right| \leqslant \sum_{k=n+1}^{+\infty} |u_k| \leqslant \varepsilon \sum_{k=n+1}^{+\infty} v_k$$
Ainsi
$$\sum_{k=n+1}^{+\infty} u_k = o\left(\sum_{k=n+1}^{+\infty} v_k\right)$$
Cas $u_n = O(v_n)$: démarche analogue sachant
$\exists M \in \mathbb{R}^+, \exists N \in \mathbb{N}, \forall n \geqslant N, |u_n| \leqslant M v_n$
Cas $u_n \sim v_n$.
Par équivalence de séries à termes positifs, la série $\sum u_n$ converge.
On a
$u_n = v_n + o(v_n) = v_n + w_n$ avec $w_n = o(v_n)$
donc
$$\sum_{k=n+1}^{+\infty} u_k = \sum_{k=n+1}^{+\infty} v_k + \sum_{k=n+1}^{+\infty} w_k = \sum_{k=n+1}^{+\infty} v_k + o\left(\sum_{k=n+1}^{+\infty} v_k\right) \sim \sum_{k=n+1}^{+\infty} v_k$$
□

Attention : La suite (v_n) de référence doit être positive ou, pour le moins, positive à partir d'un certain rang.

Exemple Déterminons un équivalent simple de

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^2+1}$$

On a $\frac{1}{k^2+1} \sim \frac{1}{k^2}$ et $\sum_{k\geqslant 1} \frac{1}{k^2}$ est une série à termes positifs convergente donc

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^2+1} \sim \sum_{k=n+1}^{+\infty} \frac{1}{k^2} \sim \frac{1}{n}$$

12.2.2 Cas de la divergence

Théorème

Soit $\sum u_n$ une série numérique et $\sum v_n$ une série à termes positifs divergente.
Si $u_n \underset{n\to+\infty}{=} o(v_n)$ alors

$$\sum_{k=0}^{n} u_k \underset{n\to+\infty}{=} o\left(\sum_{k=0}^{n} v_k\right)$$

Si $u_n \underset{n\to+\infty}{=} O(v_n)$ alors

$$\sum_{k=0}^{n} u_k \underset{n\to+\infty}{=} O\left(\sum_{k=0}^{n} v_k\right)$$

Si $u_n \underset{n\to+\infty}{\sim} v_n$ alors

$$\sum_{k=0}^{n} u_k \underset{n\to+\infty}{\sim} \sum_{k=0}^{n} v_k$$

dém. :
Remarquons que $\sum_{k=0}^{n} v_k \xrightarrow[n\to+\infty]{} +\infty$ car $\sum v_n$ est une série à termes positifs divergente.
Cas $u_n \underset{n\to+\infty}{=} o(v_n)$.
Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ vérifiant
$\forall n \geqslant N, |u_n| \leqslant \varepsilon\, |v_n| = \varepsilon v_n$
Pour $n \geqslant N$,
$\left|\sum_{k=0}^{n} u_k\right| \leqslant \left|\sum_{k=0}^{N-1} u_k\right| + \left|\sum_{k=N}^{n} u_k\right| \leqslant \left|\sum_{k=0}^{N-1} u_k\right| + \varepsilon \sum_{k=N}^{n} v_k$
Or, puisque $\sum_{k=0}^{n} v_k \to +\infty$, il existe $N' \in \mathbb{N}$ tel que

$\forall n \geqslant N', \left|\sum_{k=0}^{N-1} u_k\right| \leqslant \varepsilon \sum_{k=0}^{n} v_k$

Pour $n \geqslant \max(N, N')$, on obtient

$\left|\sum_{k=0}^{n} u_k\right| \leqslant 2\varepsilon \sum_{k=0}^{n} v_k$

Ainsi

$\sum_{k=0}^{n} u_k \underset{n\to+\infty}{=} o\left(\sum_{k=0}^{n} v_k\right)$

Cas $u_n \underset{n\to+\infty}{=} O(v_n)$: semblable.

Cas $u_n \underset{n\to+\infty}{\sim} v_n$: on écrit $u_n \underset{n\to+\infty}{=} v_n + o(v_n)$.

□

Attention : La suite (v_n) de référence doit être positive ou, pour le moins, positive à partir d'un certain rang.

Exemple Etudions

$$\sum_{k=1}^{n} \frac{1}{k+\sqrt{k}}$$

On a

$$\frac{1}{n+\sqrt{n}} \underset{n\to+\infty}{\sim} \frac{1}{n}$$

Or $\sum \frac{1}{n}$ est une série à termes positifs divergente donc

$$\sum_{k=1}^{n} \frac{1}{k+\sqrt{k}} \underset{n\to+\infty}{\sim} \sum_{k=1}^{n} \frac{1}{k} \sim \ln n$$

12.2.3 Théorème de Césaro

Soit $(u_n)_{n\geqslant 1}$ une suite numérique convergeant vers ℓ. On peut écrire

$$u_n \underset{n\to+\infty}{=} \ell + o(1) = \ell + \varepsilon_n \text{ avec } \varepsilon_n = o(1)$$

et alors

$$\frac{1}{n}(u_1 + \cdots + u_n) = \ell + \frac{1}{n}(\varepsilon_1 + \cdots + \varepsilon_n)$$

Puisque $\varepsilon_n = o(1)$ avec $\sum_{n\geqslant 0} 1$ est une série à termes positifs divergente

$$\sum_{k=1}^{n} \varepsilon_k = o\left(\sum_{k=1}^{n} 1\right) = o(n)$$

Ainsi

$$\frac{1}{n}(u_1 + \cdots + u_n) = \ell + \frac{1}{n}o(n) = \ell + o(1) \to \ell$$

Exemple Considérons la suite (u_n) donnée par

$$u_0 \in \left]0,\pi\right[\text{ et } \forall n \in \mathbb{N}, u_{n+1} = \sin(u_n)$$

La suite (u_n) est bien définie et à valeurs dans $\left]0,\pi\right[$

$$\forall x \in \left]0,\pi\right[, \sin(x) \in \left]0,1\right] \subset \left]0,\pi\right[$$

La suite (u_n) est décroissante car

$$u_{n+1} = \sin(u_n) \leqslant u_n$$

La suite (u_n) est donc converge et sa limite ℓ vérifie

$$\sin(\ell) = \ell$$

Cette limite est $\ell = 0$. Déterminons maintenant un équivalent de (u_n).
On a

$$\frac{1}{u_{n+1}^2} - \frac{1}{u_n^2} = \frac{(u_n - u_{n+1})(u_n + u_{n+1})}{(u_n u_{n+1})^2} \sim \frac{\frac{1}{3}u_n^4}{u_n^4} \to \frac{1}{3}$$

Donc par le théorème de Cesaro

$$\frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{1}{u_{k+1}^2} - \frac{1}{u_k^2}\right) = \frac{1}{n}\left(\frac{1}{u_n^2} - \frac{1}{u_0^2}\right) \to \frac{1}{3}$$

et on en déduit

$$u_n \sim \sqrt{\frac{3}{n}}$$

12.2.4 Musculation développement asymptotique à trois termes de H_n

Etudions

$$H_n = \sum_{k=1}^{n} \frac{1}{k}$$

On a déjà vu

$$H_n \underset{n\to+\infty}{=} \ln n + \gamma + o(1)$$

Approfondissons ce développement asymptotique. Posons

$$\varepsilon_n = \sum_{k=1}^{n} \frac{1}{k} - \ln n - \gamma$$

Nous allons exprimer ε_n comme le reste d'une série convergente.

$$\ln n = \sum_{k=2}^{n} \ln k - \ln(k-1) = \sum_{k=2}^{n} -\ln\left(1 - \frac{1}{k}\right)$$

donc

$$\varepsilon_n = 1 + \sum_{k=2}^{n}\left(\frac{1}{k} + \ln\left(1 - \frac{1}{k}\right)\right) - \gamma$$

Puisque $\varepsilon_n \to 0$, on a

$$\sum_{k=2}^{+\infty} \frac{1}{k} + \ln\left(1 - \frac{1}{k}\right) = \gamma - 1$$

puis

$$\varepsilon_n = \sum_{k=2}^{n} \left(\frac{1}{k} + \ln\left(1 - \frac{1}{k}\right)\right) - \sum_{k=2}^{+\infty} \left(\frac{1}{k} + \ln\left(1 - \frac{1}{k}\right)\right) = -\sum_{k=n+1}^{+\infty} \left(\frac{1}{k} + \ln\left(1 - \frac{1}{k}\right)\right)$$

Or

$$\frac{1}{n} + \ln\left(1 - \frac{1}{n}\right) \underset{n\to+\infty}{\sim} -\frac{1}{2n^2}$$

et $\sum \frac{1}{n^2}$ est une série à termes positifs convergente donc

$$\sum_{k=n+1}^{+\infty} \left(\frac{1}{k} + \ln\left(1 - \frac{1}{k}\right)\right) \underset{n\to+\infty}{\sim} -\frac{1}{2} \sum_{k=n+1}^{+\infty} \frac{1}{k^2} = -\frac{1}{2n}$$

puis enfin $\varepsilon_n \sim 1/2n$. Finalement

$$\sum_{k=1}^{n} \frac{1}{k} \underset{n\to+\infty}{=} \ln n + \gamma + \frac{1}{2n} + o\left(\frac{1}{n}\right)$$

12.3 Intégration des relations de comparaison

12.3.1 Cas de la convergence sur $[a, +\infty[$

Théorème

Soit $f : [a, +\infty[\to \mathbb{K}$ et $g : [a, +\infty[\to \mathbb{R}^+$ continues par morceaux.
On suppose que g est intégrable.
Si $f(t) \underset{t\to+\infty}{=} o\left(g(t)\right)$ alors f est intégrable et

$$\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} o\left(\int_x^{+\infty} g(t)\,\mathrm{d}t\right)$$

Si $f(t) \underset{t\to+\infty}{=} O\left(g(t)\right)$ alors f est intégrable et

$$\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} O\left(\int_x^{+\infty} g(t)\,\mathrm{d}t\right)$$

Si $f(t) \underset{t\to+\infty}{\sim} g(t)$ alors f est intégrable et

$$\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} \int_x^{+\infty} g(t)\,\mathrm{d}t$$

dém. :
Dans les trois cas, la fonction f est évidemment intégrable

Cas $f(t) \underset{t\to+\infty}{=} o\left(g(t)\right)$.
Soit $\varepsilon > 0$. Il existe $A \in [a, +\infty[$ tel que
$\forall t \in [A, +\infty[\,, |f(t)| \leqslant \varepsilon\, |g(t)| \leqslant \varepsilon g(t)$
et alors, pour $x \geqslant A$

$$\left|\int_x^{+\infty} f(t)\,\mathrm{d}t\right| \leqslant \int_x^{+\infty} |f(t)|\,\mathrm{d}t \leqslant \int_x^{+\infty} \varepsilon g(t)\,\mathrm{d}t = \varepsilon \int_x^{+\infty} g(t)\,\mathrm{d}t$$

Ainsi

$$\int_x^{+\infty} f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} o\left(\int_x^{+\infty} g(t)\,\mathrm{d}t\right)$$

Cas $f(t) \underset{t\to+\infty}{=} O\left(g(t)\right)$. Démarche analogue avec

$$\exists A \in [a, +\infty[\,, \exists M \in \mathbb{R}^+, \forall t \in [A, +\infty[\,, |f(t)| \leqslant Mg(t)$$

Cas $f(t) \underset{t\to+\infty}{\sim} g(t)$. On peut écrire

$$f(t) \underset{t\to+\infty}{=} g(t) + o\left(g(t)\right)$$

puis, avec convergence des intégrales écrites

$$\int_x^{+\infty} f(t)\,\mathrm{d}t = \int_x^{+\infty} g(t)\,\mathrm{d}t + o\left(\int_x^{+\infty} g(t)\,\mathrm{d}t\right) \underset{x\to+\infty}{\sim} \int_x^{+\infty} g(t)\,\mathrm{d}t$$

□

Attention : La fonction de référence g est positive, ou pour le moins, au voisinage de $+\infty$.

Exemple Déterminons un équivalent quand $x \to +\infty$ de

$$\int_x^{+\infty} \frac{\mathrm{d}t}{t^3+1}$$

Puisque

$$\frac{1}{t^3+1} \underset{t\to+\infty}{\sim} \frac{1}{t^3} \text{ avec } \frac{1}{t^3} \geqslant 0 \text{ et intégrable sur } [1, +\infty[$$

Par intégration de relation de comparaison, on obtient

$$\int_x^{+\infty} \frac{\mathrm{d}t}{t^3+1} \underset{x\to+\infty}{\sim} \int_x^{+\infty} \frac{\mathrm{d}t}{t^3} = \left[-\frac{1}{2t^2}\right]_x^{+\infty} = \frac{1}{2x^2}$$

Exemple Déterminons un équivalent quand $x \to +\infty$ du terme

$$\int_x^{+\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t$$

L'intégrale étudiée est convergente puisque $t^2\mathrm{e}^{-t}/t \xrightarrow[t\to+\infty]{} 0$.

Procédons à une intégration par parties avec $u(t) = -\mathrm{e}^{-t}$ et $v(t) = 1/t$.
Les fonctions u et v sont de classe $\mathcal{C}^1$ et le produit uv converge en $+\infty$. On a donc

$$\int_x^{+\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t = \frac{\mathrm{e}^{-x}}{x} - \int_x^{+\infty} \frac{\mathrm{e}^{-t}}{t^2}\,\mathrm{d}t$$

Or

$$\frac{\mathrm{e}^{-t}}{t^2} \underset{t\to+\infty}{=} o\left(\frac{\mathrm{e}^{-t}}{t}\right)$$

donc, par intégration de relation de comparaison

$$\int_x^{+\infty} \frac{\mathrm{e}^{-t}}{t^2}\,\mathrm{d}t = o\left(\int_x^{+\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t\right)$$

et finalement

$$\int_x^{+\infty} \frac{\mathrm{e}^{-t}}{t}\,\mathrm{d}t \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^{-x}}{x}$$

12.3.2 Cas de la divergence sur $[a,+\infty[$

Théorème

Soit $f : [a,+\infty[\to \mathbb{K}$ et $g : [a,+\infty[\to \mathbb{R}^+$ continues par morceaux.
On suppose que g n'est pas intégrable.
Si $f(t) \underset{t\to+\infty}{=} o\left(g(t)\right)$ alors

$$\int_a^x f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} o\left(\int_a^x g(t)\,\mathrm{d}t\right)$$

Si $f(t) \underset{t\to+\infty}{=} O\left(g(t)\right)$ alors

$$\int_a^x f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} O\left(\int_a^x g(t)\,\mathrm{d}t\right)$$

Si $f(t) \underset{t\to+\infty}{\sim} g(t)$ alors

$$\int_a^x f(t)\,\mathrm{d}t \underset{x\to+\infty}{\sim} \int_a^x g(t)\,\mathrm{d}t$$

dém. :
Puisque la fonction g est positive, mais non intégrable, on a
$\displaystyle\int_a^x g(t)\,\mathrm{d}t \xrightarrow[x\to+\infty]{} +\infty$
Cas $f(t) \underset{t\to+\infty}{=} o\left(g(t)\right)$.
Soit $\varepsilon > 0$. Il existe $A \in [a,+\infty[$ tel que
$\forall t \in [A,+\infty[\,, |f(t)| \leqslant \varepsilon\,|g(t)| \leqslant \varepsilon g(t)$
et alors, pour $x \geqslant A$

$$\left|\int_a^x f(t)\,\mathrm{d}t\right| \leqslant \int_a^x |f(t)|\,\mathrm{d}t \leqslant \int_a^A |f(t)|\,\mathrm{d}t + \varepsilon \int_A^x g(t)\,\mathrm{d}t$$

Puisque le terme $\int_a^A |f(t)|\ \mathrm{d}t$ est constant et que $\int_a^x g(t)\,\mathrm{d}t$ tend vers l'infini, il existe $A' \in [a, +\infty[$ tel que

$$\forall x \geqslant A',\ \int_a^A |f(t)|\ \mathrm{d}t \leqslant \varepsilon \int_a^x |g(t)|\ \mathrm{d}t$$

et alors, pour tout $x \geqslant \max(A, A')$

$$\left|\int_a^x f(t)\,\mathrm{d}t\right| \leqslant \varepsilon \int_a^x g(t)\,\mathrm{d}t + \varepsilon \int_A^x g(t)\,\mathrm{d}t \leqslant 2\varepsilon \int_a^x g(t)\,\mathrm{d}t$$

Ainsi

$$\int_a^x f(t)\,\mathrm{d}t \underset{x\to+\infty}{=} o\left(\int_a^x g(t)\,\mathrm{d}t\right)$$

Cas $f(t) \underset{t\to+\infty}{=} O\left(g(t)\right)$. Démarche analogue.
Cas $f(t) \underset{t\to+\infty}{\sim} g(t)$. On peut écrire

$$f(t) \underset{t\to+\infty}{=} g(t) + o\left(g(t)\right)$$

puis

$$\int_a^x f(t)\,\mathrm{d}t = \int_a^x g(t)\,\mathrm{d}t + o\left(\int_a^x g(t)\,\mathrm{d}t\right) \underset{x\to+\infty}{\sim} \int_a^x g(t)\,\mathrm{d}t$$

□

Exemple Soit $f : [0, +\infty[\to \mathbb{R}$ continue admettant une limite ℓ en $+\infty$.
On peut écrire $f(t) \underset{t\to+\infty}{=} \ell + o(1)$ et donc, par intégration de relation de comparaison

$$\int_0^x f(t)\,\mathrm{d}t = \ell x + o(x)$$

Exemple Déterminons un équivalent quand $x \to +\infty$ du terme

$$\int_1^x \frac{\ln t}{t+1}\,\mathrm{d}t$$

On a

$$\frac{\ln t}{t+1} \underset{t\to+\infty}{\sim} \frac{\ln t}{t} \text{ avec } \frac{\ln t}{t} \geqslant 0 \text{ et non intégrable sur } [1, +\infty[$$

Par intégration de relation de comparaison, on obtient

$$\int_1^x \frac{\ln t}{t+1}\,\mathrm{d}t \sim \int_1^x \frac{\ln t}{t}\,\mathrm{d}t = \frac{1}{2}\left(\ln x\right)^2$$

12.3.3 Enoncé général

Théorème

Soit $a < b$ avec $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$.
Soit $f : [a, b[\to \mathbb{K}$ et $g : [a, b[\to \mathbb{R}^+$ continues par morceaux vérifiant

$$f(x) \underset{x \to b^-}{=} o\left(g(x)\right)$$

Si g est intégrable sur $[a, b[$ alors f aussi

$$\int_x^b f(t)\,\mathrm{d}t \underset{x \to +\infty}{=} o\left(\int_x^b g(t)\,\mathrm{d}t\right)$$

Si g n'est pas intégrale sur $[a, b[$ alors

$$\int_a^x f(t)\,\mathrm{d}t \underset{x \to +\infty}{=} o\left(\int_a^x g(t)\,\mathrm{d}t\right)$$

dém. :
Analogue aux précédentes.
□

Remarque Cet énoncé se transpose aux situations $f(x) \underset{x \to b^-}{=} O\left(g(x)\right)$ et $f(x) \underset{x \to b^-}{\sim} g(x)$.
Cet énoncé se transpose aux intégrales sur $]a, b]$.

Exemple On retrouve la formule permettant d'intégrer les développements limités

$$\int_a^x o\left((t-a)^n\right)\,\mathrm{d}t = o\left((x-a)^{n+1}\right)$$

Exemple Si $f : [0, 1] \to \mathbb{R}$ est continue alors $f(t) \underset{t \to 0^+}{=} f(0) + o(1)$ et donc

$$\int_0^x f(t)\,\mathrm{d}t \underset{x \to 0^+}{=} f(0)x + o(x)$$

Exemple Déterminons un équivalent quand $x \to 0^+$ de

$$\int_x^1 \frac{\mathrm{e}^t}{t}\,\mathrm{d}t$$

On a

$$\frac{\mathrm{e}^t}{t} \underset{t \to 0^+}{\sim} \frac{1}{t} \text{ et } t \mapsto \frac{1}{t} \text{ est positive et non intégrable sur }]0, 1]$$

donc, par intégration de relation de comparaison

$$\int_x^1 \frac{\mathrm{e}^t}{t}\,\mathrm{d}t \sim \int_x^1 \frac{\mathrm{d}t}{t} = \ln x$$

12.3.4 Musculation

Soit $f : [0, +\infty[\to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ ne s'annulant pas et vérifiant

$$\frac{xf'(x)}{f(x)} \xrightarrow[x\to+\infty]{} \alpha \neq 1$$

Etudions l'existence de

$$\int_0^{+\infty} f(t)\,\mathrm{d}t$$

On a

$$\frac{f'(x)}{f(x)} \underset{x\to+\infty}{\sim} \frac{\alpha}{x} + o\left(\frac{1}{x}\right)$$

Par intégration de relation de comparaison

$$\ln\left(f(x)\right) = \ln(x^\alpha) + o(\ln x) \underset{x\to+\infty}{\sim} \ln(x^\alpha)$$

On ne peut cependant pas aller jusqu'à affirmer $f(x) \underset{x\to+\infty}{\sim} x^\alpha$... mais l'on va néanmoins déterminer la nature de l'intégrale $\displaystyle\int_0^{+\infty} f(t)\,\mathrm{d}t$.

Cas $\alpha < 1$. On a

$\ln(xf(x)) = \ln(x) + \ln(f(x)) \underset{x\to+\infty}{\sim} (1-\alpha)\ln x \xrightarrow[x\to+\infty]{} +\infty$

Ainsi $xf(x) \xrightarrow[x\to+\infty]{} +\infty$ et donc

$$\int_0^{+\infty} f(t)\,\mathrm{d}t \text{ diverge}$$

Cas $\alpha > 1$. On introduit $\rho \in \,]1, \alpha[$

$$\ln(x^\rho f(x)) = \rho\ln(x) + \ln(f(x)) \underset{x\to+\infty}{\sim} (\rho-\alpha)\ln x \xrightarrow[x\to+\infty]{} -\infty$$

et donc

$$f(x) \underset{x\to+\infty}{=} o\left(\frac{1}{x^\rho}\right)$$

ce qui assure que f est intégrable sur $[0, +\infty[$.

Chapitre 13

Familles sommables

13.1 Ensembles dénombrables

13.1.1 Définition

Définition

Un ensemble est dit dénombrable s'il est en bijection avec $\mathbb{N}$ (dans un sens ou dans l'autre).

Exemple $\mathbb{N}^\star$ est dénombrable.
Il suffit de considérer la bijection $s : \mathbb{N} \to \mathbb{N}^\star$ donnée par $s(n) = n + 1$.

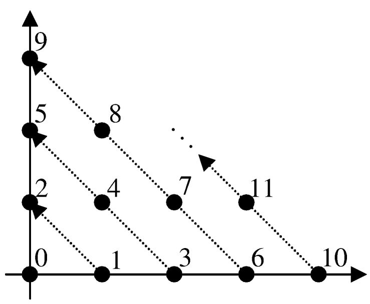
Exemple $\mathbb{Z}$ est dénombrable.
Il suffit de considérer la bijection $\delta : \mathbb{N} \to \mathbb{Z}$ donnée par

$$\delta(n) = \begin{cases} n/2 & \text{si } n \text{ est pair} \\ -(n+1)/2 & \text{sinon} \end{cases}$$

pour laquelle

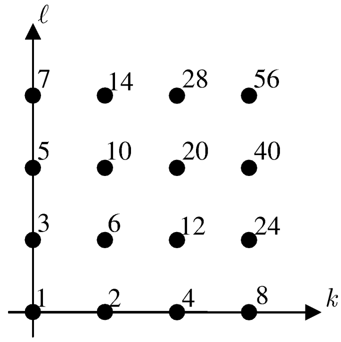
n	0	1	2	3	4	5	$\cdots$
$\delta(n)$	0	-1	1	-2	2	-3	$\cdots$

Exemple $\mathbb{N}^2$ est dénombrable.
Il suffit de considérer la bijection $\pi : \mathbb{N}^2 \to \mathbb{N}$ numérotant les éléments de $\mathbb{N}^2$ comme illustré ci-dessous



On peut aussi construire une bijection de $\mathbb{N}^2$ vers $\mathbb{N}^\star$ en posant

$$\varphi(k,\ell) = 2^k(2\ell+1)$$



Remarque Dire qu'un ensemble est dénombrable signifie qu'il est possible de numéroter de façon exhaustive ses éléments.

Définition

Si E est un ensemble dénombrable et si $\varphi : \mathbb{N} \to E$ est une application bijective, on dit que la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = \varphi(n)$ est une énumération des éléments de E.

13.1.2 Propriétés

Théorème

Toute partie infinie de $\mathbb{N}$ est dénombrable.

dém. :
Soit F une partie infinie de $\mathbb{N}$. Considérons la suite (u_n) définie par récurrence en posant

$$u_0 = \min F \text{ et } \forall n \in \mathbb{N}, u_{n+1} = \min\left(F \setminus \{u_0,\dots,u_n\}\right)$$

La suite $(u_n)_{n\in\mathbb{N}}$ est constituée d'éléments de F et est strictement croissante. De plus, tout élément de F figure dans cette suite. Considérons en effet $x \in F$. Puisque la suite (u_n) tend vers $+\infty$, il existe $N \in \mathbb{N}$ tels que $x < u_{N+1}$ et donc $x \notin F\setminus\{u_0,\dots,u_N\}$. Or $x \in F$ donc $x \in \{u_0,\dots,u_N\}$.
La fonction $\varphi : \mathbb{N} \to F$ définie par $\varphi(n) = u_n$ réalise alors une bijection de $\mathbb{N}$ vers F.
□

Théorème

Un ensemble est fini ou dénombrable si, et seulement si, il est en bijection avec une partie de $\mathbb{N}$.

dém. :
($\Rightarrow$) Si un ensemble est fini de cardinal n alors il est en bijection avec $[\![1,n]\!]$ (comprendre $\emptyset$, quand $n=0$). Si un ensemble est dénombrable, il est par définition en bijection avec $\mathbb{N}$.
($\Leftarrow$) Soit E un ensemble en bijection avec une partie F de $\mathbb{N}$ via une application $\varphi : E \to F$.
Si l'ensemble E est fini, le problème est résolu.

Si l'ensemble E est infini alors F est une partie infinie de $\mathbb{N}$ et il existe alors une bijection de $\psi : \mathbb{N} \to F$. L'application $\varphi^{-1} \circ \psi$ est alors une bijection de $\mathbb{N}$ vers E. L'ensemble E est dans ce cas dénombrable.
□

Définition

Un ensemble est dit au plus dénombrable s'il est fini ou bien dénombrable i.e. s'il est en bijection avec une partie de $\mathbb{N}$.

13.1.3 Opérations

13.1.3.1 Inclusion

Théorème

Toute partie d'un ensemble dénombrable est au plus dénombrable.

dém. :
Car par restriction en bijection avec une partie de $\mathbb{N}$.
□

Corollaire

S'il existe une injection d'un ensemble E dans un ensemble dénombrable alors E est dénombrable.

dém. :
Soit $\varphi : E \to F$ injective avec F dénombrable. Par l'application φ, E est en bijection avec $\varphi(E)$ qui est une partie de F donc φ est en bijection avec une partie au plus dénombrable.
□

13.1.3.2 Produit cartésien

Théorème

Si E et F sont des ensembles dénombrables alors $E \times F$ est dénombrable.

dém. :
Soit $\varphi : E \to \mathbb{N}$, $\psi : F \mapsto \mathbb{N}$ et $\pi : \mathbb{N}^2 \mapsto \mathbb{N}$ bijectives. L'application $(x, y) \mapsto \pi(\varphi(x), \psi(y))$ est une bijection de $E \times F$ vers $\mathbb{N}$.
□

Corollaire

Si $E_1, \ldots, E_n$ sont des ensembles au plus dénombrables alors $E_1 \times \ldots \times E_n$ est dénombrable.

dém. :
Par récurrence sur $n \in \mathbb{N}^\star$.
Cas $n = 1$: ok
Supposons la propriété établie au rang $n \geqslant 1$.
Soit $E_1, \ldots, E_n, E_{n+1}$ dénombrables.
Par hypothèse de récurrence $E = E_1 \times \ldots \times E_n$ est dénombrable et donc, par le théorème $E \times E_{n+1}$ est dénombrable. Or $E \times E_{n+1}$ n'est autre que $E_1 \times \ldots \times E_n \times E_{n+1}$.
Récurrence établie.
□

Exemple $\mathbb{Q}$ est une partie dénombrable.
En effet, on peut construire une injection de $\mathbb{Q}$ dans $\mathbb{Z} \times \mathbb{N}$ par l'application

$$r = p/q \mapsto (p, q)$$

en notant p/q le représentant irréductible du nombre rationnel r

Or l'ensemble $\mathbb{Z} \times \mathbb{N}$ est dénombrable et $\mathbb{Q}$ est alors dénombrable car c'est un ensemble infini en bijection avec une partie d'un ensemble dénombrable.

Remarque En revanche, l'ensemble $\mathbb{R}$ n'est pas dénombrable ni $\wp(\mathbb{N})$ ou $\{0,1\}^{\mathbb{N}}$.

13.1.3.3 Réunion

Théorème

Soit $(E_i)_{i \in I}$ une famille d'ensembles.
Si chaque E_i est au plus dénombrable et que l'ensemble d'indexation I est aussi dénombrable alors la réunion $\bigcup_{i \in I} E_i$ est au plus dénombrable.

dém. :
Cette démonstration est hors programme.
Entrapercevons cependant le résultat dans le cas d'une réunion dénombrable d'ensembles dénombrables.
On peut introduire $\varphi_i : \mathbb{N} \to E_i$ bijective pour chaque $i \in I$ et $\psi : \mathbb{N} \to I$ bijective. Considérons alors l'application

$$f : \mathbb{N}^2 \to \bigcup_{i \in I} E_i$$

définie par $f(k, \ell) = \varphi_k(\ell)$. Celle-ci est une surjection de $\mathbb{N}^2$ sur $\bigcup\limits_{i \in I} E_i$.
Pour chaque $x \in \bigcup_{i \in I} E_i$, l'ensemble des antécédents $f^{-1}(\{x\})$ est non vide ce qui permet de définir une injection de $\bigcup_{i \in I} E_i$ dans $\mathbb{N}^2$.
□

13.2 Familles sommables

Si $(u_i)_{i \in I}$ est une famille finie de réels ou de complexes, on sait donner un sens à la somme de ses termes

$$\sum_{i \in I} u_i$$

La notion de famille sommable vise à étendre aux familles infinies dénombrables cette notion.
Contrairement aux séries, la sommation ne sera pas ordonnée, le résultat du calcul sera indépendant de la manière dont il est organisé.
I désigne un ensemble au plus dénombrable (I fini, $I = \mathbb{N}$, $I = \mathbb{Z}$, $I = \mathbb{N}^2$,. . .)

13.2.1 Familles à termes positifs

Définition

On dit qu'une famille $(u_i)_{i\in I}$ de réels positifs est sommable s'il existe un réel M tel que

$$\forall F \text{ fini} \subset J, \sum_{i\in F} u_i \leqslant M$$

Si tel est le cas, on pose

$$\sum_{i\in I} u_i = \sup_{F \text{ finie}\subset I} \sum_{i\in F} u_i$$

Sinon, on pose

$$\sum_{i\in I} u_i = +\infty$$

Exemple On suppose I fini. La famille $(u_i)_{i\in I}$ est assurément sommable et $\sum\limits_{i\in I} u_i$ désigne à nouveau la somme de ses termes.

Exemple On dit que la famille $(u_i)_{i\in I}$ est à support fini si son support $J = \{i \in I / u_i \neq 0\}$ est fini.
Si la famille $(u_i)_{i\in I}$ est à support fini alors celle-ci est sommable.
En effet, pour toute partie F finie $\subset I$,

$$\sum_{i\in F} u_i \leqslant \sum_{i\in F\cup J} u_i = \sum_{i\in J} u_i = M$$

De plus $\sum\limits_{i\in I} u_i = \sum\limits_{i\in J} u_j$ car ici le majorant est un maximum.

Exemple La famille de réels positifs $(u_n)_{n\in\mathbb{N}}$ est sommable si, et seulement si, la série $\sum u_n$ converge.
De plus, on a alors

$$\sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$$

En effet, si la famille (u_n) est sommable alors $\sum u_n$ converge car il s'agit d'une série à termes positifs aux sommes partielles majorées. De plus

$$\sum_{n=0}^{+\infty} u_n = \lim_{N\to+\infty} \sum_{n=0}^{N} u_n \leqslant \sum_{n\in\mathbb{N}} u_n$$

Inversement, si la série $\sum u_n$ converge alors pour toute partie F finie $\subset I$, il existe $N \in \mathbb{N}$ tel que $F \subset [\![0, N]\!]$ et donc

$$\sum_{n\in F} u_n \leqslant \sum_{n=0}^{N} u_n \leqslant \sum_{n=0}^{+\infty} u_n$$

La famille $(u_n)_{n\in\mathbb{N}}$ est alors sommable et

$$\sum_{n\in\mathbb{N}} u_n \leqslant \sum_{n=0}^{+\infty} u_n$$

Exemple Soit $q \in [0,1[$ et $u_n = q^{|n|}$ pour $n \in \mathbb{Z}$. La famille $(u_n)_{n\in\mathbb{N}}$ est sommable.
En effet, pour toute partie F finie $\subset \mathbb{Z}$, il existe $N \in \mathbb{N}$ tel que $F \subset [\![-N, N]\!]$ et alors

$$\sum_{i\in F} u_i \leqslant \sum_{n=-N}^{N} q^{|n|} = 1 + 2\sum_{n=1}^{N} q^n = 1 + 2q\frac{1-q^N}{1-q} \leqslant \frac{1+q}{1-q}$$

De plus, on a

$$\sum_{n\in\mathbb{Z}} q^{|n|} = \frac{1+q}{1-q}$$

car

$$\forall F \text{ finie} \subset I,\ \sum_{i\in F} u_i \leqslant \frac{1+q}{1-q} \text{ et } \sum_{n=-|N|}^{N} q^{|n|} \xrightarrow[N\to+\infty]{} \frac{1+q}{1-q}$$

Remarque Si $(u_i)_{i\in I}$ est sommable alors pour tout permutation $\sigma \in \mathcal{S}(I)$, la famille permutée $(u_{\sigma(i)})_{i\in I}$ l'est aussi et de même somme.
En effet, les sommes finies considérées pour étudier $(u_i)_{i\in I}$ et $(u_{\sigma(i)})_{i\in I}$ sont les mêmes.

13.2.2 Comparaison

Théorème

> Soit $(u_i)_{i\in I}$ et $(v_i)_{i\in I}$ deux familles de réels positifs indexées par I.
> Si $u_i \leqslant v_i$ pour tout $i \in I$ et si la famille $(v_i)_{i\in I}$ est sommable alors la famille $(u_i)_{i\in I}$ l'est aussi et
>
> $$\sum_{i\in I} u_i \leqslant \sum_{i\in I} v_i$$

dém. :
Pour toute partie finie F incluse dans I

$$\sum_{i\in F} u_i \leqslant \sum_{i\in F} v_i \leqslant \sum_{i\in I} v_i$$

□

Théorème

> Soit $(u_i)_{i\in I}$ une famille de réels positifs et $J \subset I$.
> Si la famille $(u_i)_{i\in I}$ est sommable alors la sous-famille $(u_i)_{i\in J}$ l'est aussi et
>
> $$\sum_{i\in J} u_i \leqslant \sum_{i\in I} u_i$$

dém. :
Pour toute partie finie F incluse dans J

$$\sum_{i\in F} u_i \leqslant \sum_{i\in I} u_i$$

□

13.2.3 Regroupement de la sommation

Soit $(u_i)_{i\in I}$ une famille de réels positifs indexée par un ensemble I dénombrable.

Théorème

On suppose $I = I_1 \cup I_2$ avec I_1, I_2 disjoints. On a équivalence entre
(i) $(u_i)_{i\in I}$ est sommable ;
(ii) $(u_i)_{i\in I_1}$ et $(u_i)_{i\in I_2}$ sont sommables.
De plus, on a alors

$$\sum_{i\in I} u_i = \sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i$$

dém. :
(i) $\Rightarrow$ (ii) Supposons $(u_i)_{i\in I}$ sommable. Puisque $I_1, I_2 \subset I$, les sous-familles $(u_i)_{i\in I_1}$ et $(u_i)_{i\in I_2}$ sont sommables. De plus, pour F_1 finie $\subset I_1$ et F_2 finie $\subset I_2$

$$\sum_{i\in F_1} u_i + \sum_{i\in F_2} u_i = \sum_{i\in F_1\cup F_2} u_i \leqslant \sum_{i\in I} u_i = M$$

donc

$$\sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i \leqslant \sum_{i\in I} u_i$$

(ii) $\Rightarrow$ (i) Supposons $(u_i)_{i\in I_1}$ et $(u_i)_{i\in I_2}$ sommables.
Pour F finie $\subset I$, on a

$$\sum_{i\in F} u_i = \sum_{i\in F\cap I_1} u_i + \sum_{i\in F\cap I_2} u_i \leqslant \sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i = M$$

donc $(u_i)_{i\in I}$ est sommable et

$$\sum_{i\in I} u_i \leqslant \sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i$$

□

Remarque Ce résultat s'étend évidemment à $I = I_1 \cup I_2 \cup \ldots \cup I_N$ avec $(I_j)_{1\leqslant j\leqslant N}$ deux à deux disjoints.

Exemple Soit $(u_n)_{n\in\mathbb{Z}}$ une famille de réels positifs.
La famille $(u_n)_{n\in\mathbb{Z}}$ est sommable si, et seulement si, les familles $(u_n)_{n\in\mathbb{N}^\star}$ et $(u_{-n})_{n\in\mathbb{N}^\star}$ le sont.
De plus, on a alors

$$\sum_{n\in\mathbb{Z}} u_n = u_0 + \sum_{n\in\mathbb{N}^\star} u_n + \sum_{n\in\mathbb{N}^\star} u_{-n}$$

13.2.4 Sommation par paquets

Théorème

Soit $(u_i)_{i\in I}$ une famille dénombrable de réels positifs et $(I_n)_{n\in\mathbb{N}}$ une famille de parties de I vérifiant

$$\forall n \neq m, I_n \cap I_m = \emptyset \text{ et } \bigcup_{n\in\mathbb{N}} I_n = I$$

On a équivalence entre :
(i) la famille $(u_i)_{i\in I}$ est sommable ;
(ii) chaque famille $(u_i)_{i\in I_n}$ est sommable et la série $\sum\left(\sum_{i\in I_n} u_i\right)$ converge.
De plus, si tel est le cas

$$\sum_{i\in I} u_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} u_i\right)$$

dém. :
Cette démonstration est hors programme.
(i) $\Rightarrow$ (ii) Supposons $(u_i)_{i\in I}$ sommable
Pour tout $n\in\mathbb{N}$, $I_n \subset I$ donc $(u_i)_{i\in I_n}$ est aussi sommable.
Pour tout $N\in\mathbb{N}$, considérons la partition finie de I réalisée à partir de $I_0,\ldots,I_N$ et $J = \bigcup_{n\geqslant N+1} I_n$.
On a

$$\sum_{n=0}^{N}\sum_{i\in I_n} u_i \leqslant \sum_{n=0}^{N}\sum_{i\in I_n} u_i + \sum_{i\in J} u_i = \sum_{i\in I} u_i$$

Puisque $\sum\left(\sum_{i\in I_n} u_i\right)$ est une série à termes positifs aux sommes partielles majorées, celle-ci converge et

$$\sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} u_i\right) \leqslant \sum_{i\in I} u_i$$

(ii) $\Rightarrow$ (i) Supposons (ii). Soit une partie F finie $\subset I$. Il existe $N\in\mathbb{N}$ tel que

$$F \subset \bigcup_{n=0}^{N} I_n$$

et alors

$$\sum_{i\in F} u_i = \sum_{n=0}^{N}\sum_{i\in F\cap I_n} u_i \leqslant \sum_{n=0}^{N}\sum_{i\in I_n} u_i \leqslant \sum_{n=0}^{+\infty}\sum_{i\in I_n} u_i = M$$

La famille $(u_i)_{i\in I}$ est donc sommable et

$$\sum_{i\in I} u_i \leqslant \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} u_i\right)$$

□

Exemple Soit $x\in[0,1[$. Montrons $\dfrac{x}{1-x} = \displaystyle\sum_{n=0}^{+\infty} \frac{x^{2^n}}{1-x^{2^{n+1}}}$.
La famille $(x^p)_{p\in\mathbb{N}^\star}$ est sommable. Pour $n\in\mathbb{N}$, considérons $I_n = \{2^n(2k+1)/k\in\mathbb{N}\}$.

Par sommation par paquets

$$\sum_{p=1}^{+\infty} x^p = \sum_{n=0}^{+\infty} \sum_{p\in I_n} x^p = \sum_{n=0}^{+\infty}\sum_{k=0}^{+\infty} x^{2^n(2k+1)}$$

et ainsi

$$\frac{x}{1-x} = \sum_{n=0}^{+\infty} \frac{x^{2^n}}{1-x^{2^{n+1}}}$$

Corollaire

Si $\varphi : \mathbb{N} \to I$ est une bijection alors on a équivalence entre :
(i) $(u_i)_{i\in I}$ est sommable ;
(ii) $\sum u_{\varphi(n)}$ converge.
De plus, si tel est le cas

$$\sum_{i\in I} u_i = \sum_{n=0}^{+\infty} u_{\varphi(n)}$$

dém. :
Il suffit de considérer la partition de I constituée de $I_n = \{\varphi(n)\}$.
□

Remarque En conséquence, après indexation des éléments de I, la sommabilité de la famille $(u_i)_{i\in I}$ se ramène à la convergence d'une série à termes positifs.

13.2.5 Extension aux familles réelles ou complexes

Soit $(u_i)_{i\in I}$ une famille de nombres réels ou complexes indexée par un ensemble I au plus dénombrable.

Définition

On dit que la famille $(u_i)_{i\in I}$ est sommable si la famille $(|u_i|)_{i\in I}$ l'est i.e. s'il existe un réel M tel que

$$\forall F \text{ fini} \subset J, \sum_{i\in F} |u_i| \leqslant M$$

Théorème

S'il existe une famille de réels positif $(v_i)_{i\in I}$ sommable vérifiant

$$\forall i \in I, |u_i| \leqslant v_i$$

alors la famille $(u_i)_{i\in I}$ est sommable

Définition

Soit $(u_i)_{i\in I}$ une famille sommable de réels. Pour tout $i \in I$, on introduit

$$u_i^+ = \max(u_i, 0) \text{ et } u_i^- = \max(-u_i, 0)$$

Les familles de réels positifs $(u_i^+)_{i\in I}$ et $(u_i^-)_{i\in I}$ étant sommables, on pose

$$\sum_{i\in I} u_i = \sum_{i\in I} u_i^+ - \sum_{i\in I} u_i^-$$

Définition

Soit $(u_i)_{i\in I}$ une famille sommable de complexes. Les familles de réels $(\mathrm{Re} u_i)_{i\in I}$ et $(\mathrm{Im} u_i)_{i\in I}$ étant sommables, on pose

$$\sum_{i\in I} u_i = \sum_{i\in I} \mathrm{Re}(u_i) + \mathrm{i}.\sum_{i\in I} \mathrm{Im}(u_i)$$

Exemple On suppose I fini. La famille $(u_i)_{i\in I}$ est assurément sommable et $\sum_{i\in I} u_i$ désigne à nouveau la somme de ses termes.

Exemple On suppose la famille $(u_i)_{i\in I}$ est à support fini et l'on introduit son support
$J = \{i \in I / u_i \neq 0\}$.
La famille $(u_i)_{i\in I}$ est assurément sommable et $\sum_{i\in I} u_i = \sum_{i\in J} u_j$

Exemple Une famille de réels ou de complexes $(u_n)_{n\in\mathbb{N}}$ est sommable si, et seulement si, la famille $(|u_n|)_{n\in\mathbb{N}}$ l'est . Ceci revient à affirmer la convergence de la série $\sum |u_n|$.
Ainsi, la sommabilité de $(u_n)_{n\in\mathbb{N}}$équivaut à la convergence absolument de $\sum u_n$.
De plus, on a alors

$$\sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$$

13.2.6 Sommation par paquets

Théorème

Soit $(u_i)_{i\in I}$ une famille dénombrable de réels positifs et $(I_n)_{n\in\mathbb{N}}$ une familles de parties de I vérifiant

$$\forall n \neq m, I_n \cap I_m = \emptyset \text{ et } \bigcup_{n\in\mathbb{N}} I_n = I$$

Si la famille $(u_i)_{i\in I}$ est sommable alors chaque famille $(u_i)_{i\in I_n}$ l'est aussi et la série $\sum\left(\sum_{i\in I_n} u_i\right)$ converge absolument.
De plus, on a alors

$$\sum_{i\in I} u_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} u_i\right)$$

dém. :
Cette démonstration est hors programme.
Puisque la famille $(u_i)_{i\in I}$ est sommable, la famille $(|u_i|)_{i\in I}$ l'est aussi et donc les familles $(|u_i|)_{i\in I_n}$ le sont encore et la série $\sum\left(\sum_{i\in I_n} |u_i|\right)$ converge. Ainsi les familles $(u_i)_{i\in I_n}$ sont sommables et la série $\sum\left(\sum_{i\in I_n} u_i\right)$ est absolument convergente car dans le cadre réel

$$\left|\sum_{i\in I_n} u_i\right| \leqslant \sum_{i\in I_n} u_i^+ + \sum_{i\in I_n} u_i^- \leqslant \sum_{i\in I_n} |u_i|$$

et dans le cadre complexe

$$\left|\sum_{i\in I_n} u_i\right| \leqslant \left|\sum_{i\in I_n} \mathrm{Re}(u_i)\right| + \left|\sum_{i\in I_n} \mathrm{Im}(u_i)\right| \leqslant 2\sum_{i\in I_n} |u_i|$$

Il reste à établir l'égalité

$$\sum_{i\in I} u_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} u_i\right)$$

Celle-ci est connue si tous les termes u_i sont réels positifs.
Celle-ci est encore vraie si tous les u_i sont réels en raisonnant par u_i^+ et u_i^-.
Celle-ci est aussi vraie si tous les u_i sont complexes en raisonnant par $\mathrm{Re}(u_i)$ et $\mathrm{Im}(u_i)$.
□

Corollaire

On suppose $I = I_1 \cup I_2$ avec I_1, I_2 disjoints.
Si $(u_i)_{i\in I}$ est sommable alors $(u_i)_{i\in I_1}$ et $(u_i)_{i\in I_2}$ sont sommables et

$$\sum_{i\in I} u_i = \sum_{i\in I_1} u_i + \sum_{i\in I_2} u_i$$

dém. :
Prendre $I_n = \emptyset$ pour $n \neq 1, 2$.
□

Corollaire

Soit $\varphi : \mathbb{N} \to I$ une bijection. Si la famille $(u_i)_{i\in I}$ est sommable alors la série $\sum u_{\varphi(n)}$ converge et

$$\sum_{i\in I} u_i = \sum_{n=0}^{+\infty} u_{\varphi(n)}$$

dém. :
On utilise $I_n = \{\varphi(n)\}$.
□

Remarque Pour utiliser ces résultats, il faut préalablement justifier la sommabilité de $(|u_i|)_{i\in I}$ ce qui pourra se faire en employant le résultat analogue connu pour les familles de réels positifs.

Exemple Considérons $u_n = (-1)^n/n$ et $I = \mathbb{N}^\star$.
La série $\sum u_n$ converge et cependant la famille $(u_n)_{n\in\mathbb{N}^\star}$ n'est pas sommable.
En effet, pour $I_1 = \{2p/p \in \mathbb{N}^\star\}$ et $I_2 = \{2p+1/p \in \mathbb{N}\}$, les familles $(u_n)_{n\in I_1}$ et $(u_n)_{n\in I_2}$ ne sont pas sommables.

13.2.7 Propriétés

13.2.7.1 Linéarité

Théorème

Soit $(u_i)_{i\in I}$ et $(v_i)_{i\in I}$ deux familles d'éléments de $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $\lambda, \mu \in \mathbb{K}$.
Si $(u_i)_{i\in I}$ et $(v_i)_{i\in I}$ sont sommables alors $(\lambda u_i + \mu v_i)_{i\in I}$ l'est aussi et

$$\sum_{i\in I} \lambda u_i + \mu v_i = \lambda \sum_{i\in I} u_i + \mu \sum_{i\in I} v_i$$

dém. :
Pour tout $i \in I$,

$$|\lambda u_i + \mu v_i| \leqslant |\lambda|\,|u_i| + |\mu|\,|v_i|$$

donc toute partie F finie $\subset I$,

$$\sum_{i\in F} |\lambda u_i + \mu v_i| \leqslant |\lambda| \sum_{i\in F} |u_i| + |\mu| \sum_{i\in F} |v_i| \leqslant |\lambda| \sum_{i\in I} |u_i| + |\mu| \sum_{i\in I} |v_i| = M$$

Ainsi $(\lambda u_i + \mu v_i)_{i\in I}$ est sommable.
De plus, si $\varphi : \mathbb{N} \to I$ est une bijection

$$\sum_{i\in I} \lambda u_i + \mu v_i = \sum_{n=0}^{+\infty} \lambda u_{\varphi(n)} + \mu v_{\varphi(n)}$$

Par linéarité des séries convergentes

$$\sum_{i\in I} \lambda u_i + \mu v_i = \lambda \sum_{n=0}^{+\infty} u_{\varphi(n)} + \mu \sum_{n=0}^{+\infty} v_{\varphi(n)} = \lambda \sum_{i\in I} u_i + \mu \sum_{i\in I} v_i$$

□

Corollaire

L'ensemble des familles $(u_i)_{i\in I}$ sommables est un sous-espace vectoriel de l'espace $\mathbb{K}^I$ des familles indexées sur I et l'application $(u_i)_{i\in I} \mapsto \sum_{i\in I} u_i$ y définit une forme linéaire.

13.2.7.2 Positivité

Théorème

Soit $(u_i)_{i\in I}$ une famille de réels positifs.
Si $(u_i)_{i\in I}$ est sommable alors $\sum_{i\in I} u_i \geqslant 0$.

dém. :
Par définition $\sum_{i\in I} u_i$ est la borne supérieure d'un ensemble de quantités positives.
□

Corollaire

Si $(u_i)_{i\in I}$ et $(v_i)_{i\in I}$ sont deux familles de réels sommables vérifiant

$$\forall i \in I, u_i \leqslant v_i$$

alors

$$\sum_{i\in I} u_i \leqslant \sum_{i\in I} v_i$$

dém. :
Il suffit de considérer la famille positive $(v_i - u_i)_{i\in I}$.
□

Théorème

Soit $(u_i)_{i\in I}$ une famille de réels positifs.
Si $(u_i)_{i\in I}$ est sommable et si $\sum_{i\in I} u_i = 0$ alors $u_i = 0$ pour tout $i \in I$.

dém. :
Pour tout $i \in I$, on a

$$0 \leqslant u_i \leqslant \sum_{i\in I} u_i = 0$$

car la somme est la borne supérieure de l'ensemble des sommes sur les parties finies F ; il suffit ici de considérer $F = \{i\}$.
□

13.2.7.3 Conjugaison

Théorème

Si $(u_i)_{i\in I}$ est une famille de complexes sommable alors $(\overline{u_i})_{i\in I}$ l'est aussi et

$$\sum_{i\in I} \overline{u_i} = \overline{\sum_{i\in I} u_i}$$

Corollaire

On a équivalence entre :
(i) la famille $(u_i)_{i\in I}$ est sommable ;
(ii) les familles $(\mathrm{Re}(u_i))_{i\in I}$ et $(\mathrm{Im}(u_i))_{i\in I}$ sont sommables.

dém. :
(i) $\Rightarrow$ (ii) via $\mathrm{Re}(u_i) = (u_i + \overline{u_i})/2$ et $\mathrm{Re}(u_i) = (u_i - \overline{u_i})/2\mathrm{i}$.
(ii) $\Rightarrow$ (i) via $u_i = \mathrm{Re}(u_i) + \mathrm{i}.\mathrm{Im}(u_i)$.
□

13.2.7.4 Inégalité triangulaire

Théorème

Si $(u_i)_{i\in I}$ est une famille de réels ou de complexes sommable alors

$$\left|\sum_{i\in I} u_i\right| \leqslant \sum_{i\in I} |u_i|$$

dém. :
Soit $\varphi : \mathbb{N} \to I$ est une bijection

$$\left|\sum_{i\in I} u_i\right| = \left|\sum_{n=0}^{+\infty} u_{\varphi(n)}\right| \leqslant \sum_{n=0}^{+\infty} \left|u_{\varphi(n)}\right| = \sum_{i\in I} |u_i|$$

□

13.3 Application à la réorganisation des sommes

13.3.1 Permutation des termes d'une série

Soit $\sum u_n$ une série et $\sigma \in \mathcal{S}(\mathbb{N})$. Que dire de la série $\sum u_{\sigma(n)}$?

Exemple Considérons la série $\displaystyle\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}$ de somme $S = \ln 2$ et permutons ses termes.

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k+1} - \frac{1}{2k+2} + \cdots$$

Permutons les termes de S de la manière suivante :

$$S = 1 - \left(\frac{1}{2} + \frac{1}{4}\right) + \frac{1}{3} - \left(\frac{1}{6} + \frac{1}{8}\right) + \cdots + \frac{1}{2k+1} - \left(\frac{1}{4k+2} + \frac{1}{4k+4}\right) + \cdots$$

on obtient

$$S = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \cdots - \frac{1}{4k+2} + \frac{1}{4k+4} + \cdots$$

puis

$$S = \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\right) = \frac{1}{2}S$$

Ainsi, on peut changer la somme d'une série en en permutant ses termes !

Théorème

Si $\sum u_n$ converge absolument alors pour tout $\sigma \in \mathcal{S}(\mathbb{N})$, la série permutée $\sum u_{\sigma(n)}$ converge absolument et

$$\sum_{n=0}^{+\infty} u_{\sigma(n)} = \sum_{n=0}^{+\infty} u_n$$

dém. :
Si $\sum u_n$ converge absolument alors $(u_n)_{n\in\mathbb{N}}$ est sommable et

$$\sum_{n=0}^{+\infty} u_n = \sum_{n\in\mathbb{N}} u_n$$

La famille permutée $(u_{\sigma(n)})_{n\in\mathbb{N}}$ est alors elle aussi sommable et

$$\sum_{n\in\mathbb{N}} u_{\sigma(n)} = \sum_{n\in\mathbb{N}} u_n$$

On en déduit que la série $\sum u_{\sigma(n)}$ converge absolument et donc

$$\sum_{n=0}^{+\infty} u_{\sigma(n)} = \sum_{n\in\mathbb{N}} u_{\sigma(n)} = \sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$$

□

Exemple Nature de $\displaystyle\sum_{n\geqslant 1} \frac{1}{n\sigma(n)}$ pour $\sigma \in \mathcal{S}(\mathbb{N}^\star)$.
Sachant

$$ab \leqslant \frac{1}{2}\left(a^2+b^2\right)$$

on a

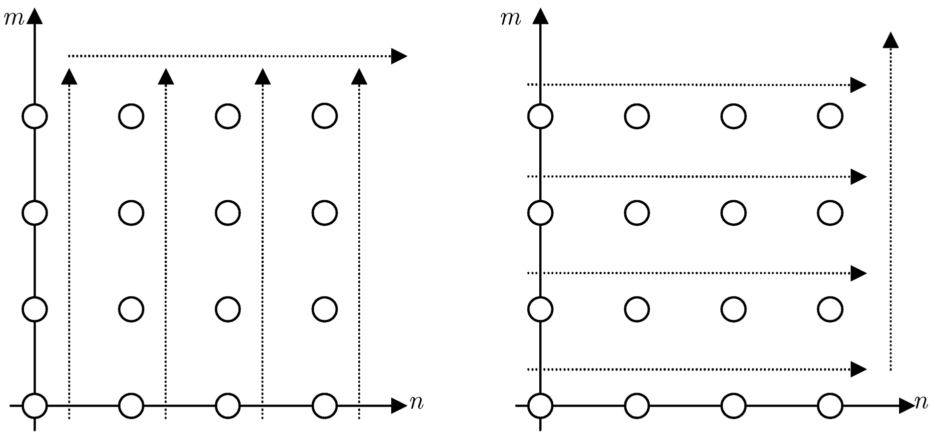
$$\frac{1}{n\sigma(n)} \leqslant \frac{1}{2}\left(\frac{1}{n^2}+\frac{1}{\sigma(n)^2}\right)$$

Or $\sum \frac{1}{n^2}$ converge absolument et donc $\sum \frac{1}{\sigma(n)^2}$ aussi.
Par comparaison de séries à termes positifs, on obtient la convergence de $\displaystyle\sum_{n\geqslant 1} \frac{1}{n\sigma(n)}$.

13.3.2 Sommes doubles

Soit $(u_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels ou de complexes. A-t-on

$$\sum_{m=0}^{+\infty}\sum_{n=0}^{+\infty} u_{m,n} = \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} u_{m,n} \text{ ?}$$



Théorème

Soit $(u_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels ou de complexes. On a équivalence entre
(i) la famille $(u_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable ;
(ii) pour tout $n\in\mathbb{N}$, la série $\sum\limits_m |u_{m,n}|$ converge et la série $\sum\limits_n \sum\limits_{m=0}^{+\infty} |u_{m,n}|$ converge.
De plus, on a alors

$$\sum_{(m,n)\in\mathbb{N}^2} u_{m,n} = \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} u_{m,n}$$

dém. :
On caractérise la sommabilité $(|u_{m,n}|)_{(m,n)\in\mathbb{N}^2}$ par le théorème de sommation par paquets avec $I_n = \mathbb{N}\times\{n\}$.
Une fois la sommabilité acquise, on calcule la somme par la même organisation par paquets.
□

Corollaire

On a alors

$$\sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} u_{m,n} = \sum_{m=0}^{+\infty}\sum_{n=0}^{+\infty} u_{m,n}$$

avec convergence des séries écrites.

dém. :
On calcule

$$\sum_{(m,n)\in\mathbb{N}^2} u_{m,n}$$

en procédant à deux sommations par paquets.
La première avec $I_n = \mathbb{N}\times\{n\}$, la seconde avec $J_m = \{m\}\times\mathbb{N}$.
□

Exemple Montrons

$$\sum_{m=1}^{+\infty}\sum_{n=m}^{+\infty}\frac{1}{n^3} = \sum_{n=1}^{+\infty}\frac{1}{n^2}$$

Commençons par interpréter le premier membre sous la forme

$$\sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty} u_{m,n}$$

Posons $u_{m,n} = \dfrac{1}{n^3}$ si $n \geqslant m$ et 0 sinon.

$\displaystyle\sum_{m\geqslant 1} |u_{m,n}|$ converge car $u_{m,n} = 0$ pour $m > n$.

$\displaystyle\sum_{m=1}^{+\infty} |u_{m,n}| = \sum_{m=1}^{n} \frac{1}{n^3} = \frac{1}{n^2}$ donc $\displaystyle\sum_{n\geqslant 1}\sum_{m=1}^{+\infty} |u_{m,n}|$ converge.

Ainsi, la famille $(u_{m,n})_{(m,n)\in(\mathbb{N}^\star)^2}$ est sommable et par le théorème de Fubini, on a l'égalité

$$\sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty} u_{m,n} = \sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty} u_{m,n}$$

avec convergence des séries engagées. On obtient ainsi

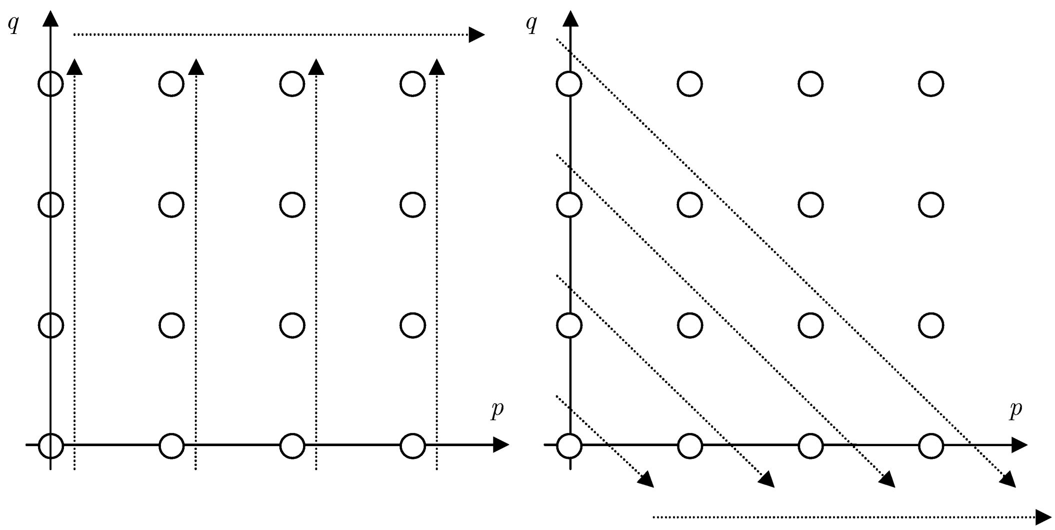
$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \sum_{m=1}^{+\infty}\sum_{n=m}^{+\infty} \frac{1}{n^3}$$

13.3.3 Produit de Cauchy

Soit $\sum u_m$ et $\sum v_n$ deux séries convergentes. On a

$$\left(\sum_{m=0}^{+\infty} u_m\right)\left(\sum_{n=0}^{+\infty} v_n\right) = \sum_{m=0}^{+\infty}\left(u_m \sum_{n=0}^{+\infty} v_n\right) = \sum_{m=0}^{+\infty}\sum_{n=0}^{+\infty} (u_m v_n)$$

qui se comprend $(u_0v_0+u_0v_1+u_0v_2+\cdots)+(u_1v_0+u_1v_1+u_1v_2+\cdots)+(u_2v_0+u_2v_1+u_2v_2+\cdots)+\cdots$.
Peut-on réorganiser la somme en $u_0v_0 + (u_0v_1 + u_1v_0) + (u_0v_2 + u_1v_1 + u_2v_0) + \cdots$?



Définition

On appelle produit de Cauchy des séries $\sum u_n$ et $\sum v_n$ la série de terme général

$$w_n = \sum_{k=0}^{n} u_k v_{n-k}$$

Théorème

Si $\sum u_m$ et $\sum v_n$ sont deux séries absolument convergentes alors la famille $(u_m v_n)_{(m,n)\in\mathbb{N}^2}$ est sommable et

$$\sum_{(m,n)\in\mathbb{N}^2} u_m v_n = \left(\sum_{m=0}^{+\infty} u_m\right)\left(\sum_{n=0}^{+\infty} u_n\right)$$

dém. :

Pour tout $n \in \mathbb{N}$, la série $\sum_m |u_m v_n|$ converge et la série $\sum_n \sum_{m=0}^{+\infty} |u_m v_n|$ converge donc la famille $(u_m v_n)_{(m,n)\in\mathbb{N}^2}$ est sommable.

□

Corollaire

Si $\sum u_m$ et $\sum v_n$ convergent absolument alors la série produit de Cauchy $\sum w_n$ converge absolument aussi et on a

$$\sum_{n=0}^{+\infty} w_n = \left(\sum_{m=0}^{+\infty} u_m\right)\left(\sum_{n=0}^{+\infty} v_n\right)$$

dém. :

On procède à une sommation par paquets avec

$$I_p = \left\{(m,n) \in \mathbb{N}^2 / m + n = p\right\}$$

sachant

$$\sum_{(m,n)\in I_p} u_m v_n = w_p$$

□

Exemple Soit $a \in \mathbb{C}$ tel que $|a| < 1$. Montrons $\dfrac{1}{(1-a)^2} = \displaystyle\sum_{n=0}^{+\infty} (n+1)a^n$

Par sommation géométrique

$$\frac{1}{(1-a)^2} = \left(\sum_{n=0}^{+\infty} a^n\right)\left(\sum_{n=0}^{+\infty} a^n\right)$$

Par produit de Cauchy de séries absolument convergentes,

$$\frac{1}{(1-a)^2} = \sum_{n=0}^{+\infty}\sum_{k=0}^{n}\left(a^k a^{n-k}\right) = \sum_{n=0}^{+\infty} (n+1)a^n$$

Exemple Pour $x \in \mathbb{R}$, on pose $f(x) = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$.
Vérifions
$$\forall x, y \in \mathbb{R}, f(x)f(y) = f(x+y)$$
On vérifie aisément l'absolue convergence de la série définissant $f(x)$ par application du critère d'Alembert. On a
$$f(x)f(y) = \left(\sum_{n=0}^{+\infty} \frac{1}{n!} x^n\right)\left(\sum_{n=0}^{+\infty} \frac{1}{n!} y^n\right)$$
Par produit de Cauchy de séries absolument convergentes
$$f(x)f(y) = \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!}$$
Or
$$\sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \frac{(x+y)^n}{n!}$$
donc
$$f(x)f(y) = \sum_{n=0}^{+\infty} \frac{(x+y)^n}{n!} = f(x+y)$$
On a établira ultérieurement que f n'est autre que la fonction exponentielle.

Chapitre 14

Espaces normés

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$ et E désigne un $\mathbb{K}$-espace vectoriel.

14.1 Norme

14.1.1 Définition

Définition

On appelle norme sur E toute application $N : E \to \mathbb{R}^+$ vérifiant :
1) $\forall x \in E$, $N(x) = 0 \Rightarrow x = 0_E$ [séparation].
2) $\forall \lambda \in \mathbb{K}$, $\forall x \in E$, $N(\lambda.x) = |\lambda|\, N(x)$ [homogénéité]
3) $\forall x, y \in E$, $N(x+y) \leqslant N(x) + N(y)$ [inégalité triangulaire].
On dit alors que le couple (E, N) est un espace normé.

Remarque Les normes sont usuellement notées $N(.)$, $\| \,.\, \|$ ou $|\,.\,|$, elles servent à définir la longueur d'un vecteur.

Exemple La valeur absolue sur $\mathbb{R}$ et le module sur $\mathbb{C}$ sont des normes.

Exemple Soit E un espace préhilbertien réel de produit scalaire noté $\langle ., . \rangle$.
La norme euclidienne associée à ce produit scalaire est une norme. Celle-ci est définie par

$$\forall x \in E, \|x\| = \sqrt{\langle x, x \rangle}$$

Exemple Si F est un sous-espace vectoriel d'un espace E normé par $\| \,.\, \|$ alors la restriction $\| \,.\, \| : F \to \mathbb{R}^+$ définit une norme sur F.

Proposition

Si $\| \,.\, \|$ est une norme sur E alors :
a) $\forall x \in E$, $\|x\| = 0 \Leftrightarrow x = 0_E$;
b) $\forall x \in E$, $\|-x\| = \|x\|$;
c) $\forall x, y \in E$, $|\|x\| - \|y\|| \leqslant \|x - y\|$ [inégalité triangulaire renversée].

dém. :
a) ($\Rightarrow$) par définition et ($\Leftarrow$) par homogénéité avec $\lambda = 0$.
b) par homogénéité avec $\lambda = -1$.
c) par l'inégalité triangulaire $\|x\| = \|x - y + y\| \leqslant \|x - y\| + \|y\|$ donc $\|x\| - \|y\| \leqslant \|x - y\|$ et par un raisonnement symétrique $\|y\| - \|x\| \leqslant \|x - y\|$.
□

Définition

Un vecteur x d'un espace E normé par $\| \,.\, \|$ est dit unitaire si $\|x\| = 1$.

Exemple Si $x \neq 0_E$ alors $u = \dfrac{1}{\|x\|}x$ est un vecteur unitaire colinéaire à x.

14.1.2 Normes usuelles sur $\mathbb{K}^n$

Pour $x = (x_1, \dots, x_n) \in \mathbb{K}^n$, on pose

$$\|x\|_1 \underset{\text{déf}}{=} |x_1| + \cdots + |x_n| = \sum_{k=1}^{n} |x_k|,\ \|x\|_2 \underset{\text{déf}}{=} \sqrt{|x_1|^2 + \cdots + |x_n|^2} = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2} \quad \text{et}$$

$$\|x\|_\infty \underset{\text{déf}}{=} \max\{|x_1|, \dots, |x_n|\} = \max_{1 \leqslant k \leqslant n} |x_k|$$

Théorème

$\| \,.\, \|_1$ définit une norme sur $\mathbb{K}^n$.

dém. :
$\| \,.\, \|_1 : \mathbb{K}^n \to \mathbb{R}^+$ est bien définie. Soit $x \in \mathbb{K}^n$. Si $\|x\|_1 = 0$ alors $\sum\limits_{k=1}^{n} |x_k| = 0$.
Par somme nulle de quantités positives $|x_1| = \ldots = |x_n| = 0$ et donc $x = 0_{\mathbb{K}^n}$.
Soit $\lambda \in \mathbb{K}$ et $x \in \mathbb{K}^n$

$$\|\lambda x\| = \sum_{k=1}^{n} |\lambda x_k| = \sum_{k=1}^{n} |\lambda| \, |x_k| = |\lambda| \sum_{k=1}^{n} |x_k| \, |\lambda x_n| = |\lambda| \, \|x\|_1$$

Soit $x, y \in \mathbb{K}^n$.

$$\|x + y\|_1 = \sum_{k=1}^{n} |x_k + y_k| \leqslant \sum_{k=1}^{n} (|x_k| + |y_k|) = \sum_{k=1}^{n} |x_k| + \sum_{k=1}^{n} |y_k| = \|x\|_1 + \|y\|_1$$

Finalement $\| \,.\, \|_1$ est une norme sur $\mathbb{K}^n$.
□

Théorème

$\| \,.\, \|_2$ définit une norme sur $\mathbb{K}^n$.

dém. :
$\| \,.\, \|_2 : \mathbb{K}^n \to \mathbb{R}^+$ est bien définie.
Soit $x \in \mathbb{K}^n$. Si $\|x\|_2 = 0$ alors $\sum\limits_{k=1}^{n} |x_k|^2 = 0$.

Par somme nulle de quantités positives $|x_1|^2 = \ldots = |x_n|^2 = 0$ et donc $x = 0_{\mathbb{K}^n}$.
Soit $\lambda \in \mathbb{K}$ et $x \in \mathbb{K}^n$

$$\|\lambda x\|_2 = \sqrt{\sum_{k=1}^{n} |\lambda x_k|^2} = \sqrt{|\lambda|^2 \sum_{k=1}^{n} |x_k|^2} = |\lambda| \, \|x\|_2$$

Soit $x, y \in \mathbb{K}^n$

$$\|x+y\|_2^2 = \sum_{k=1}^{n} |x_k + y_k|^2$$

Or $|x_k + y_k| \leqslant |x_k| + |y_k|$ donc

$$\|x+y\|_2^2 = \sum_{k=1}^{n} |x_k|^2 + 2\sum_{k=1}^{n} |x_k| \, |y_k| + \sum_{k=1}^{n} |y_k|^2$$

Rappelons l'inégalité de Cauchy-Schwarz

$$\forall a_k, b_k \in \mathbb{R}, \left|\sum_{k=1}^{n} a_k b_k\right| \leqslant \left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \left(\sum_{k=1}^{n} b_k^2\right)^{1/2}$$

On en déduit

$$\sum_{k=1}^{n} |x_k| \, |y_k| \leqslant \sqrt{\sum_{k=1}^{n} |x_k|^2} \sqrt{\sum_{k=1}^{n} |y_k|^2}$$

donc

$$\|x+y\|_2^2 \leqslant (\|x\|_2 + \|y\|_2)^2$$

puis

$$\|x+y\|_2 \leqslant \|x\|_2 + \|y\|_2$$

Finalement $\| \, . \, \|_2$ est une norme sur $\mathbb{K}^n$.
□

Théorème

$\| \, . \, \|_\infty$ définit une norme sur $\mathbb{K}^n$.

dém. :
$\| \, . \, \|_\infty : \mathbb{K}^n \to \mathbb{R}^+$ est bien définie
Soit $x \in \mathbb{K}^n$. Si $\|x\|_\infty = 0$ alors pour tout $1 \leqslant k \leqslant n$, $0 \leqslant |x_k| \leqslant \|x\|_\infty$ donc $|x_k| = 0$ et donc $x = 0_{\mathbb{K}^n}$.
Soit $\lambda \in \mathbb{K}$ et $x \in \mathbb{K}^n$

$$\|\lambda x\|_\infty = \max_{1 \leqslant k \leqslant n} |\lambda x_k| = |\lambda| \max_{1 \leqslant k \leqslant n} |x_k| = |\lambda| \, \|x\|_\infty$$

Soit $x, y \in \mathbb{K}^n$.

$$\|x+y\|_\infty = \max_{1 \leqslant k \leqslant n} |x_k + y_k| \leqslant \max_{1 \leqslant k \leqslant n} (|x_k| + |y_k|) \leqslant \max_{1 \leqslant k \leqslant n} |x_k| + \max_{1 \leqslant k \leqslant n} |y_k| = \|x\|_\infty + \|y\|_\infty$$

Finalement $\| \, . \, \|_\infty$ est une norme sur $\mathbb{K}^n$.
□

Remarque Plus généralement, pour $p \in [1, +\infty[$, on peut montrer que

$$\|x\|_p \underset{\text{déf}}{=} (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

définit une norme sur $\mathbb{K}^n$. De plus

$$\|x\|_\infty = \lim_{p \to +\infty} \|x\|_p$$

14.1.3 Distance associée

Soit $\| \, . \, \|$ une norme sur E.

Définition

On appelle distance associée à la norme $\| \, . \, \|$ sur E l'application $d : E \times E \to \mathbb{R}^+$ définie par

$$d(x, y) \underset{\text{déf}}{=} \|y - x\|$$

Exemple Sur $E = \mathbb{R}$ ou $\mathbb{C}$, $d(x, y) = |y - x|$ définit la distance associée à $| \, . \, |$.

Proposition

a) $\forall x, y \in E, d(x, y) = 0 \Leftrightarrow x = y$ [séparation] ;
b) $\forall x, y \in E, d(x, y) = d(y, x)$ [symétrie] ;
c) $\forall x, y, z \in E, d(x, z) \leqslant d(x, y) + d(y, z)$ [inégalité triangulaire] ;
d) $\forall x, y, z \in E, d(x + z, y + z) = d(x, y)$ [invariance par translation].

dém. :
a) $\|y - x\| = 0 \Leftrightarrow y - x = 0_E$.
b) $\|y - x\| = \|x - y\|$.
c) $\|z - x\| = \|(z - y) + (y - x)\| \leqslant \|z - y\| + \|y - x\|$.
d) $\|(y + z) - (x + z)\| = \|y - x\|$.
□

14.1.4 Boules

Soit $\| \, . \, \|$ une norme sur E.

Définition

Soit $a \in E$ et $r > 0$. On définit :
- la boule ouverte de centre a et de rayon r :

$$B(a,r) \underset{\text{déf}}{=} \{x \in E / \|x-a\| < r\}$$

- la boule fermée de centre a et de rayon r :

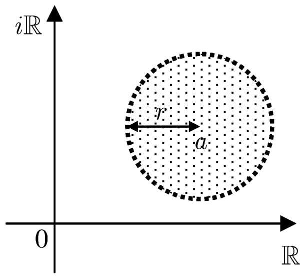
$$B_f(a,r) \underset{\text{déf}}{=} \{x \in E / \|x-a\| \leqslant r\}$$

- la sphère de centre a et de rayon r :

$$S(a,r) \underset{\text{déf}}{=} \{x \in E / \|x-a\| = r\}$$

Exemple Dans $(\mathbb{R}, |\,.\,|)$, $B(a,r) =]a-r, a+r[$, $B_f(a,r) = [a-r, a+r]$.

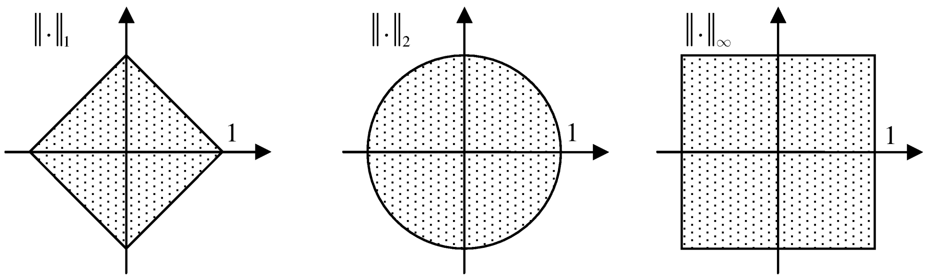
Exemple Dans $(\mathbb{C}, |\,.\,|)$, $B(a,r) = D(a,r) \underset{\text{déf}}{=} \{z \in \mathbb{C} / |z-a| < r\}$ est le disque ouvert de centre a et de rayon r.



Définition

Les boules de centre 0_E et de rayon 1, sont appelées boules unités.

Exemple Boules unités fermées sur $E = \mathbb{R}^2$



Proposition

$B(a,r) = a + rB(0_E, 1)$ et $B_f(a,r) = a + rB_f(0_E, 1)$.
Ainsi, les boules générales se déduisent des boules des boules unités par homothéties et translations.

dém. :
$a + rB(0_E, 1) = \{a + ru / \|u\| < 1\}$.
Si $x \in a + rB(0_E, 1)$ alors $\|x - a\| = \|ru\| = r\|u\| < r$ donc $x \in B(a,r)$.
Si $x \in B(a,r)$ alors pour $u = \dfrac{1}{r}(x-a)$, on a $x = a + ru$ avec $\|u\| < 1$.
□

Proposition

Les boules sont des parties convexes.

dém. :
Etudions $B(a,r)$.
Soit $x, y \in B(a,r)$. $[x,y] = \{(1-\lambda)x + \lambda y / \lambda \in [0,1]\}$.
Soit $z \in [x,y]$. On peut écrire $z = (1-\lambda)x + \lambda y$ avec $\lambda \in [0,1]$.
On a alors $\|z - a\| \leqslant \lambda\|x - a\| + (1-\lambda)\|y - a\| < \lambda r + (1-\lambda)r = r$ l'inégalité stricte étant maintenue car l'un au moins des deux facteurs λ ou $1-\lambda$ est strictement positif.
□

14.1.5 Bornitude

Soit $\|\,.\,\|$ une norme sur E.

Définition

Une partie A de E est dite bornée s'il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall x \in A, \|x\| \leqslant M$$

Exemple Les boules sont des parties bornées. En effet

$$\forall x \in B_f(a,r),\ \|x\| \leqslant \|a\| + \|x - a\| \leqslant \|a\| + r = M$$

Définition

Soit X un ensemble. On dit qu'une fonction vectorielle $f : X \to E$ est bornée lorsque son image l'est i.e.

$$\exists M \in \mathbb{R}^+, \forall x \in X, \|f(x)\| \leqslant M$$

Exemple La fonction $x \mapsto (2 + \cos x)\sin x$ est bornée.
En effet,

$$\forall x \in \mathbb{R}, |(2 + \cos x)\sin x| = |2 + \cos x|\,|\sin x| \leqslant 3$$

Il est plus aisé de raisonner ainsi que par les concepts de fonctions minorées et majorées.

Définition

Pour $X = \mathbb{N}$, une fonction au départ de $\mathbb{N}$ est communément appelée une suite. La définition qui précède se transpose donc aux suites de vecteurs et par conséquent une suite $(u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ est dite bornée si

$$\exists M \in \mathbb{R}^+, \forall n \in \mathbb{N}, \|u_n\| \leqslant M$$

Théorème

Soit $f, g : X \to E$ et $\lambda, \mu \in \mathbb{K}$.
Si f et g sont bornées alors $\lambda f + \mu g$ l'est aussi.

dém. :
Il existe $M, M' \in \mathbb{R}^+$ tels que

$$\forall x \in X, \|f(x)\| \leqslant M \text{ et } \|g(x)\| \leqslant M'$$

On a alors

$$\forall x \in X, \|\lambda f(x) + \mu g(x)\| \leqslant |\lambda| M + |\mu| M'$$

donc $\lambda f + \mu g$ est bornée.
□

Corollaire

L'ensemble $\mathcal{B}(X, E)$ des fonctions bornées de X vers E est un sous-espace vectoriel de l'espace $\mathcal{F}(X, E)$ des fonctions de X vers E .

14.2 Espaces normés usuels

14.2.1 Normes sur un espace de dimension finie

Théorème

Tout $\mathbb{K}$-espace vectoriel de dimension finie peut être muni d'une norme.

dém. :
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}$.
Cas $n = 0$.
$E = \{0_E\}$ est muni de la norme définie par $N(0_E) = 0$.
Cas $n \in \mathbb{N}^\star$.
Soit $e = (e_1, \dots, e_n)$ une base de E. Pour tout $x \in E$, il existe d'uniques $x_1, \dots, x_n$ vérifiant

$$x = x_1.e_1 + \cdots + x_n.e_n$$

Posons $\varphi_j : E \to \mathbb{K}$ l'application qui à x associe sa j-ème coordonnée dans la base e.
L'application φ_j est une forme linéaire sur E.
Considérons $\| \,.\, \|$ une norme sur $\mathbb{K}^n$ et posons, pour tout $x \in E$,

$$N(x) = \|(\varphi_1(x), \dots, \varphi_n(x))\|$$

L'application N est bien définie de E vers $\mathbb{R}^+$.
Si $N(x) = 0$ alors $(\varphi_1(x), \dots, \varphi_n(x)) = 0_{\mathbb{K}^n}$ et donc $x = 0_E$.
Soit $\lambda \in \mathbb{K}$ et $x \in E$.

$$N(\lambda.x) = \|(\varphi_1(\lambda.x), \dots, \varphi_n(\lambda.x))\| = \|\lambda.(\varphi_1(x), \dots, \varphi_n(x))\| = |\lambda| N(x)$$

Soit $x, y \in E$.

$$N(x+y) = \|(\varphi_1(x+y), \dots, \varphi_n(x+y))\| = \|(\varphi_1(x), \dots, \varphi_n(x)) + (\varphi_1(y), \dots, \varphi_n(y))\| \leqslant N(x)+N(y)$$

□

Définition

En choisissant sur $\mathbb{K}^n$, $\|\,.\,\| = \|\,.\,\|_1$, $\|\,.\,\|_2$, ou $\|\,.\,\|_\infty$, la norme N ci-dessus est notée $\|\,.\,\|_{1,e}$, $\|\,.\,\|_{2,e}$ ou $\|\,.\,\|_{\infty,e}$.

Exemple Soit $E = \mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{B} = (E_{i,j})$ sa base canonique.
Pour $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$

$$\|A\|_1 = \sum_{i=1}^{n}\sum_{j=1}^{p} |a_{i,j}|,\ \|A\|_2 = \left(\sum_{i=1}^{n}\sum_{j=1}^{p} |a_{i,j}|^2\right)^{1/2} \text{ et } \|A\|_\infty = \max_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant p}} |a_{i,j}|$$

14.2.2 Norme de la convergence uniforme

Soit X un ensemble non vide. Pour $f : X \to \mathbb{K}$ bornée, on pose

$$\|f\|_\infty \underset{\text{déf}}{=} \sup_{x\in X} |f(x)|$$

Cette borne supérieure existe car

$$\{|f(x)|\,/x \in X\} \text{ est une partie de } \mathbb{R} \text{ non vide et majorée}$$

Cette borne supérieure désigne le plus petit réel M vérifiant

$$\forall x \in X, |f(x)| \leqslant M$$

Théorème

$\|\,.\,\|_\infty$ définit une norme sur l'espace $\mathcal{B}(X, \mathbb{K})$.

dém. :
L'application $\|\,.\,\|_\infty$ est bien définie de $\mathcal{B}(X, \mathbb{K})$ vers $\mathbb{R}^+$.
Soit $f \in \mathcal{B}(X, \mathbb{K})$. Si $\|f\|_\infty = 0$ alors $\sup\{|f(x)|\,/x \in X\} = 0$ donc pour tout $x \in X$, $\|f(x)\| = 0$ puis $f = 0$.
Soit $\lambda \in \mathbb{K}$ et $f \in \mathcal{B}(X, \mathbb{K})$. Pour tout $x \in X$,

$$|\lambda f(x)| = |\lambda|\,|f(x)| \leqslant |\lambda|\,\|f\|_\infty$$

donc $\|\lambda f\|_\infty \leqslant |\lambda|\,\|f\|_\infty$. Pour $\lambda \neq 0$,

$$\|f\|_\infty = \left\|\frac{1}{\lambda}\lambda f\right\|_\infty \leqslant \left|\frac{1}{\lambda}\right|\|\lambda f\|_\infty$$

et donc $|\lambda| \|f\|_\infty \leqslant \|\lambda f\|_\infty$ puis l'égalité. Pour $\lambda = 0$, l'égalité est bien entendu aussi vérifiée.
Soit $f, g \in \mathcal{B}(X, \mathbb{K})$. Pour tout $x \in X$

$$|(f+g)(x)| = |f(x) + g(x)| \leqslant |f(x)| + |g(x)| \leqslant \|f\|_\infty + \|g\|_\infty$$

□

Corollaire

$\| \,.\, \|_\infty$ définit une norme sur l'espace $\mathcal{B}(\mathbb{N}, \mathbb{K})$ des suites bornées où

$$\|u\|_\infty = \sup_{n\in\mathbb{N}} |u_n|$$

dém. :
Il suffit de considérer $X = \mathbb{N}$.
□

14.2.3 Norme de la convergence en moyenne et en moyenne quadratique

Soit $a < b$ deux réels et $E = \mathcal{C}\left([a, b], \mathbb{K}\right)$ l'espace des fonctions continues de $[a, b]$ vers $\mathbb{K}$.
Cet espace est inclus dans celui des fonctions bornée de $[a, b]$ vers $\mathbb{K}$. On peut donc le munir de la norme induite

$$\|f\|_\infty = \sup_{t\in[a,b]} |f(t)|$$

et, de surcroît, la borne supérieure est ici un maximum en vertu du théorème de la borne atteinte.
Pour $f : [a, b] \to \mathbb{K}$ continue, on pose aussi

$$\|f\|_1 \underset{\text{déf}}{=} \int_a^b |f(t)| \,\mathrm{d}t \text{ et } \|f\|_2 \underset{\text{déf}}{=} \left(\int_a^b |f(t)|^2 \,\mathrm{d}t\right)^{1/2}$$

Théorème

$\| \,.\, \|_1$ définit une norme sur $\mathcal{C}\left([a, b], \mathbb{K}\right)$.

dém. :
L'application $\| \,.\, \|_1 : \mathcal{C}\left([a, b], \mathbb{K}\right) \to \mathbb{R}^+$ est bien définie.
Soit $f \in \mathcal{C}\left([a, b], \mathbb{K}\right)$. Si $\|f\|_1 = 0$ alors $\displaystyle\int_a^b |f(t)| \,\mathrm{d}t = 0$ or $|f|$ est continue et positive sur $[a, b]$ donc $f = \tilde{0}$.
Soit $\lambda \in \mathbb{K}$ et $f \in \mathcal{C}\left([a, b], \mathbb{K}\right)$.

$$\|\lambda.f\|_1 = \int_a^b |\lambda f(t)| \,\mathrm{d}t = \int_a^b |\lambda| \, |f(t)| \,\mathrm{d}t = |\lambda| \int_a^b |f(t)| \,\mathrm{d}t = |\lambda| \, \|f\|_1$$

Soit $f, g \in \mathcal{C}\left([a, b], \mathbb{K}\right)$.

$$\|f+g\|_1 = \int_a^b |f(t) + g(t)| \,\mathrm{d}t \leqslant \int_a^b |f(t)| + |g(t)| \,\mathrm{d}t = \|f\|_1 + \|g\|_1$$

□

Théorème

$\| \,.\, \|_2$ définit une norme sur $\mathcal{C}([a,b],\mathbb{K})$.

dém. :
L'application $\| \,.\, \|_2 : \mathcal{C}([a,b],\mathbb{K}) \to \mathbb{R}^+$ est bien définie.
Soit $f \in \mathcal{C}([a,b],\mathbb{K})$. Si $\|f\|_2 = 0$ alors $\int_a^b |f(t)|^2 \,\mathrm{d}t = 0$ or $|f|^2$ est continue et positive sur $[a,b]$ donc $f = \tilde{0}$.
Soit $\lambda \in \mathbb{K}$ et $f \in \mathcal{C}([a,b],\mathbb{K})$.

$$\|\lambda.f\|_2 = \left(\int_a^b |\lambda f(t)|^2 \,\mathrm{d}t\right)^{1/2} = \left(\int_a^b |\lambda|^2 |f(t)|^2 \,\mathrm{d}t\right)^{1/2} = |\lambda| \left(\int_a^b |f(t)|^2 \,\mathrm{d}t\right)^{1/2} = |\lambda| \, \|f\|_2$$

Soit $f, g \in \mathcal{C}([a,b],\mathbb{K})$.

$$\|f+g\|_2^2 = \int_a^b |f(t)+g(t)|^2 \,\mathrm{d}t \leqslant \int_a^b (|f(t)| + |g(t)|)^2 \,\mathrm{d}t$$

En développant

$$\|f+g\|_2^2 \leqslant \int_a^b |f(t)|^2 \,\mathrm{d}t + 2\int_a^b |f(t)|\,|g(t)| \,\mathrm{d}t + \int_a^b |g(t)|^2 \,\mathrm{d}t$$

Par l'inégalité de Cauchy-Schwarz

$$\left|\int_a^b |f(t)g(t)| \,\mathrm{d}t\right| \leqslant \left(\int_a^b |f(t)|^2 \,\mathrm{d}t\right)^{1/2} \left(\int_a^b |g(t)|^2 \,\mathrm{d}t\right)^{1/2}$$

donc

$$\|f+g\|_2^2 \leqslant \|f\|_2^2 + 2\,\|f\|_2\,\|g\|_2 + \|g\|_2^2$$

□

14.2.4 Produit d'espaces normés

Soit $(E_1, N_1), \ldots, (E_p, N_p)$ des espaces normés. Considérons le produit cartésien

$$E = E_1 \times \cdots \times E_p = \prod_{j=1}^{p} E_j$$

E est un $\mathbb{K}$-espace vectoriel dont les éléments x sont des tuples $(x_1, \ldots, x_p)$ avec

$$\forall 1 \leqslant j \leqslant p, x_j \in E_j$$

Le vecteur nul est le tuple nul

$$0_E = \left(0_{E_1}, \ldots, 0_{E_p}\right)$$

Les opérations sur E se déduisent de celles sur les espaces E_j

$$\lambda.(x_1, \ldots, x_p) = (\lambda.x_1, \ldots, \lambda.x_p) \text{ et } (x_1, \ldots, x_p) + (y_1, \ldots, y_p) = (x_1 + y_1, \ldots, x_p + y_p)$$

Pour $x = (x_1, \dots, x_p) \in E$, on pose

$$\|x\| = \max_{1 \leqslant j \leqslant p} N_j(x_j)$$

Théorème

$\| \, . \, \|$ définit une norme sur E.

dém. :
L'application $\| \, . \, \|$ est bien définie de E vers $\mathbb{R}^+$.
Soit $x = (x_1, \dots, x_p) \in E$. Si $\|x\| = 0$ alors

$$\forall j \in \{1, \dots, p\}, N_j(x_j) = 0$$

et donc

$$\forall j \in \{1, \dots, p\}, x_j = 0_{E_j}$$

On en déduit $x = 0_E$.
Soit $\lambda \in \mathbb{K}$ et $x = (x_1, \dots, x_p) \in E$

$$\|\lambda.x\| = \max_{1 \leqslant j \leqslant p} N_j(\lambda x_j) = \max_{1 \leqslant j \leqslant p} |\lambda| N_j(x_j) = |\lambda| \max_{1 \leqslant j \leqslant p} N_j(x_j) = |\lambda| \|x\|$$

Soit $x, y \in E$

$$\|x + y\| = \max_{1 \leqslant j \leqslant p} N_j(x_j + y_j) \leqslant \max_{1 \leqslant j \leqslant p} (N_j(x_j) + N_j(y_j)) \leqslant \max_{1 \leqslant j \leqslant p} N_j(x_j) + \max_{1 \leqslant j \leqslant p} N_j(y_j) = \|x\| + \|y\|$$

□

Définition

$(E, \| \, . \, \|)$ est appelé espace normé produit des espaces normés $(E_1, N_1), \dots, (E_p, N_p)$

14.2.5 Normes d'algèbres

Soit $(E, +, \times, .)$ une $\mathbb{K}$-algèbre.

Définition

On appelle norme d'algèbre sur E toute application $\| \, . \, \| : E \to \mathbb{R}^+$ vérifiant :
1) $\| \, . \, \|$ est une norme sur E ;
2) $\forall x, y \in E, \|xy\| \leqslant \|x\| \|y\|$ [sous-multiplicativité]
On dit alors que le couple $(E, \| \, . \, \|)$ est une algèbre normée.

Exemple La valeur absolue est une norme d'algèbre sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Exemple $\| \, . \, \|_\infty$ est une norme d'algèbre sur $\mathbb{K}^n$.

Exemple $\| \, . \, \|_\infty$ est une norme d'algèbre sur $\mathcal{B}(X, \mathbb{K})$.

Exemple Sur $\mathcal{M}_n(\mathbb{K})$, $\|\,.\,\|_\infty$ définie par

$$\|A\|_\infty = \max_{1\leqslant i,j\leqslant n} |a_{i,j}|$$

n'est pas une norme d'algèbre car on a seulement

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}), \|AB\|_\infty \leqslant n\,\|A\|_\infty\,\|B\|_\infty$$

Cependant, l'application $\|\,.\,\|$ définie par $\|A\| = n\,\|A\|_\infty$ est encore une norme sur $\mathcal{M}_n(\mathbb{K})$ et celle-ci vérifie

$$\|AB\| \leqslant \|A\|\,\|B\|$$

C'est une norme d'algèbre sur $\mathcal{M}_n(\mathbb{K})$.

14.3 Equivalence de normes

14.3.1 Comparaison de normes

Définition

On dit qu'une norme N_1 sur E est dominée par une norme N_2 lorsque

$$\exists \alpha > 0, \forall x \in E, N_1(x) \leqslant \alpha N_2(x)$$

Exemple Sur $\mathbb{K}^n$, comparons deux à deux les normes $\|\,.\,\|_1$, $\|\,.\,\|_2$ et $\|\,.\,\|_\infty$.
a) $\|x\|_\infty \leqslant \|x\|_1 \leqslant n\,\|x\|_\infty$.
En effet , $\|x\|_\infty = \max\limits_{1\leqslant k\leqslant n} |x_k| \leqslant \sum\limits_{k=1}^n |x_k| = \|x\|_1$ et $\|x\|_1 = \sum\limits_{k=1}^n |x_k| \leqslant \sum\limits_{k=1}^n \|x\|_\infty = n\,\|x\|_\infty$
b) $\|x\|_\infty \leqslant \|x\|_2 \leqslant \sqrt{n}\,\|x\|_\infty$.
En effet, $\|x\|_\infty = \max\limits_{1\leqslant k\leqslant n} |x_k|^2 \leqslant \sum\limits_{k=1}^n |x_k|^2 = \|x\|_2^2$ et $\|x\|_2^2 \leqslant \sum\limits_{k=1}^n \|x\|_\infty^2 = n\,\|x\|_\infty$.
c) $\|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{n}\,\|x\|_2$.
En effet, $\|x\|_2^2 = \sum\limits_{k=1}^n |x_k|^2 \leqslant \left(\sum\limits_{k=1}^n |x_k|\right)^2 = \|x\|_1^2$ et $\|x\|_1 \leqslant \sqrt{n}\,\|x\|_2$ par l'inégalité de Cauchy-Schwarz.

Exemple Sur $E = \mathcal{C}([0,1], \mathbb{K})$ comparons les normes $\|\,.\,\|_1$ et $\|\,.\,\|_\infty$.
On a $\|f\|_1 = \int_0^1 |f(t)|\,\mathrm{d}t \leqslant \int_0^1 \|f\|_\infty\,\mathrm{d}t = \|f\|_\infty$.
Ainsi $\|\,.\,\|_1$ est dominée par $\|\,.\,\|_\infty$.
Montrons qu'en revanche $\|\,.\,\|_\infty$ n'est pas dominée par $\|\,.\,\|_1$.
Pour cela considérons $f_n : t \mapsto t^n$.
On a

$$\|f_n\|_1 = \frac{1}{n+1} \text{ et } \|f_n\|_\infty = 1$$

Par l'absurde, supposons qu'il existe $\alpha > 0$ tel que

$$\forall f \in E, \|f\|_\infty \leqslant \alpha \|f\|_1$$

Appliquée en $f = f_n$, on obtient

$$1 \leqslant \frac{\alpha}{n+1} \xrightarrow[n\infty]{} 0$$

C'est absurde !

Remarque Sur $E = \mathcal{C}([a,b], \mathbb{K})$, on a aussi

$$\|f\|_1 \leqslant \sqrt{b-a}\,\|f\|_2 \text{ et } \|f\|_2 \leqslant \sqrt{b-a}\,\|f\|_\infty$$

Cependant $\|\,.\,\|_2$ n'est pas dominée par $\|\,.\,\|_1$, ni $\|\,.\,\|_\infty$ n'est pas dominée par $\|\,.\,\|_2$.

14.3.2 Normes équivalentes

Définition

Deux normes N_1 et N_2 sur un même espace E sont dites équivalentes lorsqu'elles se dominent mutuellement i.e.

$$\exists \alpha, \beta > 0, \forall x \in E, \alpha N_2(x) \leqslant N_1(x) \leqslant \beta N_2(x)$$

Proposition

L'équivalence de norme définit une relation d'équivalence sur l'ensemble des normes sur E.

Exemple Sur $\mathbb{K}^n$, les normes $\|\,.\,\|_1$, $\|\,.\,\|_2$ et $\|\,.\,\|_\infty$ sont équivalentes.

Théorème

Sur un $\mathbb{K}$-espace vectoriel de dimension finie, les normes sont deux à deux équivalentes.
(admis)

Exemple Sur l'espace de dimension infinie $E = \mathcal{C}([a,b], \mathbb{K})$, les normes $\|\,.\,\|_1$, $\|\,.\,\|_2$ et $\|\,.\,\|_\infty$ ne sont pas équivalentes.

Exemple Soit $E = \mathcal{C}^1([0,1], \mathbb{R})$. On y définit les normes

$$N(f) = |f(0)| + \|f'\|_\infty \text{ et } N'(f) = \|f\|_\infty + \|f'\|_\infty$$

Celles-ci sont équivalentes.
En effet, il est évident que $N(f) \leqslant N'(f)$ mais aussi, sachant

$$f(x) = f(0) + \int_0^x f'(t)\,\mathrm{d}t$$

on a

$$\|f\|_\infty \leqslant N(f)$$

puis

$$N'(f) \leqslant 2N(f)$$

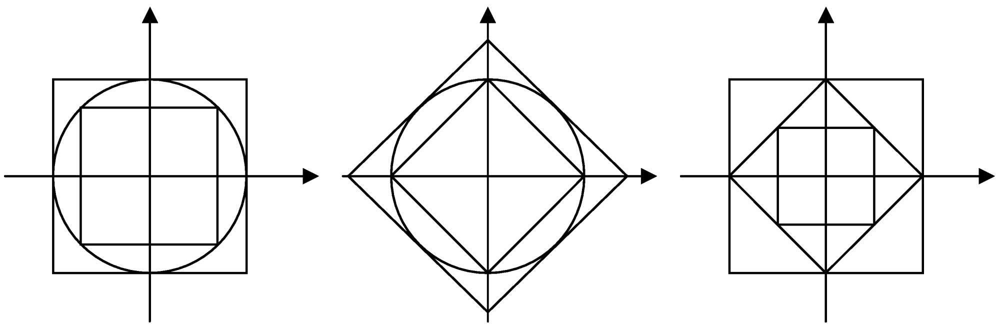
14.3.3 Encadrement des boules

Proposition

Si N_1 et N_2 sont deux normes équivalentes alors toute boule de centre a pour l'une des normes est incluse et contient des boules de même centre a pour l'autre norme.

dém. :
Supposons $\alpha N_2 \leqslant N_1 \leqslant \beta N_2$ et considérons $B = B_1(a, r)$.
On a $B_2(a, r/\beta) \subset B$ car $N_2(x-a) < r/\beta \Rightarrow N_1(x-a) < r$
et $B \subset B_2(a, r/\alpha)$ car $N_1(x-a) < r \Rightarrow N_2(x-a) < r/\alpha$.
□

Exemple Sur $\mathbb{R}^2$



$\|\cdot\|_\infty \leq \|\cdot\|_2 \leq \sqrt{2}\|\cdot\|_\infty$ $\quad$ $\frac{1}{\sqrt{2}}\|\cdot\|_1 \leq \|\cdot\|_2 \leq \|\cdot\|_1$ $\quad$ $\|\cdot\|_\infty \leq \|\cdot\|_1 \leq 2\|\cdot\|_\infty$

14.3.4 Notion invariante par passage à une norme équivalente

Définition

On dit qu'une notion est invariante par passage à une norme équivalente si, lorsqu'elle est vérifiée dans une espace normé (E, N_1), elle l'est encore dans l'espace normée (E, N_2) quand N_2 est équivalente à N_1.

Exemple La notion de partie bornée est invariante par passage à une norme équivalente.
En effet, une partie est bornée si, et seulement si, elle est incluse dans une boule de centre 0_E et cette notion n'est pas changée lorsqu'on passe à une norme équivalente.
De même pour la notion de suite ou de fonction bornée.

Exemple La notion de vecteur unitaire n'est pas invariante par passage à une norme équivalente.

Remarque Lorsque deux normes ne sont pas équivalentes, certaines propriétés peuvent être vraies pour une norme sans l'être pour l'autre.

Exemple Dans $E = \mathcal{C}([0,1], \mathbb{K})$, considérons la suite $(f_n)_{n\in\mathbb{N}}$ des fonctions $f_n : t \mapsto nt^n$.
Cette suite est bornée pour $\|\,.\,\|_1$, mais ne l'est pas pour $\|\,.\,\|_\infty$.
On a $\|f_n\|_1 = \dfrac{n}{n+1} \to 1$ donc la suite $(f_n)_{n\in\mathbb{N}}$ est bornée pour $\|\,.\,\|_1$.
En revanche $\|f_n\|_\infty = n \to +\infty$ donc la suite $(f_n)_{n\in\mathbb{N}}$ n'est pas bornée pour $\|\,.\,\|_\infty$.
Conclusion : on retrouve à nouveau que $\|\,.\,\|_1$ et $\|\,.\,\|_\infty$ ne sont pas équivalentes sur E.

14.4 Suites d'éléments d'un espace normé

On s'intéresse ici aux suites d'éléments d'un espace normé. L'étude s'appliquera aux suites numériques, aux suites d'éléments de $\mathbb{K}^n$, aux suites matricielles ou encore aux suites de fonctions...
$(E, \|\,.\,\|)$ désigne un espace normé.

14.4.1 Convergence

Définition

On dit qu'une suite $u = (u_n)_{n\in\mathbb{N}}$ d'éléments de E tend vers $\ell \in E$ si $\|u_n - \ell\| \to 0$ i.e. :

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow \|u_n - \ell\| \leqslant \varepsilon$$

On note alors $u_n \xrightarrow[n\to+\infty]{} \ell$ ou $u_n \xrightarrow[n\to+\infty]{\|\,.\,\|} \ell$.

Exemple Etudions $u_n = \left(\dfrac{\sin n}{n}, \dfrac{n+1}{n}\right) \in \mathbb{R}^2$.
Pour $\|\,.\,\| = \|\,.\,\|_1$,

$$\|u_n - (0,1)\| = \left|\frac{\sin n}{n}\right| + \left|\frac{n+1}{n} - 1\right| \to 0$$

donc $u_n \to (0,1)$.

Exemple Soit $\|\,.\,\|$ une norme d'algèbre sur $\mathcal{M}_p(\mathbb{K})$.
Si $\|A\| < 1$ alors $A^n \xrightarrow[n\to+\infty]{} O_p$.
En effet

$$\|A^n - O_p\| = \|A^n\| \leqslant \|A\|^n \to 0$$

Théorème

Si $u_n \to \ell$ et $u_n \to \ell'$ alors $\ell = \ell'$.

dém. :
$0 \leqslant \|\ell - \ell'\| \leqslant \|\ell - u_n\| + \|u_n - \ell'\| \to 0$ donc $\|\ell - \ell'\| = 0$ puis $\ell = \ell'$.
□

Définition

On dit qu'une suite $u = (u_n)_{n\in\mathbb{N}}$ d'éléments de E converge s'il existe $\ell \in E$ tel que $u_n \to \ell$.
Cet élément ℓ est alors unique, on l'appelle limite de u et on note

$$\ell = \lim u \text{ ou } \ell = \lim_{n\to+\infty} u_n$$

Remarque Si deux suites sont égales à partir d'un certain rang, elles ont même nature et même éventuelle limite : on ne modifie pas la limite d'une suite en modifiant la valeur d'un nombre fini de ses termes.

14.4.2 Opérations

Théorème

Si $u_n \to \ell$ alors $\|u_n\| \to \|\ell\|$.
Par conséquent toute suite convergente est bornée.

dém. :
Par l'inégalité triangulaire renversée

$$|\|u_n\| - \|\ell\|| \leqslant \|u_n - \ell\| \to 0$$

□

Théorème

Si $u_n \in E \to \ell$ et $v_n \in E \to \ell'$ alors $\lambda u_n + \mu v_n \to \lambda\ell + \mu\ell'$.
Si de plus E est une algèbre normée alors $u_n v_n \to \ell\ell'$.

dém. :
$\|\lambda u_n + \mu\ell - (\lambda\ell + \mu\ell')\| \leqslant |\lambda| \|u_n - \ell\| + |\mu| \|v_n - \ell'\| \to 0$.
$\|u_n v_n - \ell\ell'\| \leqslant \|u_n v_n - u_n\ell'\| + \|u_n\ell' - \ell\ell'\| \leqslant \|u_n\| \|v_n - \ell'\| + |\ell'| \|u_n - \ell\| \to 0$.
□

Théorème

Si $\alpha_n \in \mathbb{K} \to \alpha$ et $u_n \in E \to \ell$ alors $\alpha_n.u_n \to \alpha.\ell$.

dém. :
$\|\alpha_n.u_n - \alpha.\ell\| \leqslant \|\alpha_n.u_n - \alpha.u_n\| + \|\alpha.u_n - \alpha.\ell\| = |\alpha_n - \alpha| \|u_n\| + |\alpha| \|u_n - \ell\| \to 0$.
□

14.4.3 Effet d'un changement de norme

Théorème

Si N_1 est dominée par N_2 alors toute suite convergeant pour N_2 converge vers la même limite pour N_1.

dém. :
Car avec les notations qui précèdent

$$N_1(u_n - \ell) \leqslant \alpha N_2(u_n - \ell) \to 0$$

□

Corollaire

Deux normes équivalentes définissent les mêmes suites convergentes et celles-ci ont mêmes limites pour les deux normes.

Attention : Si N_1 et N_2 ne sont pas équivalentes, il se peut qu'une suite converge pour une norme et diverge pour l'autre voire qu'elle converge deux pour ces deux normes, mais vers des limites différentes !

Exemple $E = \mathcal{C}([0,1],\mathbb{R})$ muni de $\|\,.\,\|_1$ et $\|\,.\,\|_\infty$. Etudions la convergence de la suite des fonctions $f_n : t \mapsto t^n$ pour ces deux normes.
On a $\|f_n\|_1 = \dfrac{1}{n+1} \to 0$ donc $f_n \xrightarrow{\|\,.\,\|_1} \tilde{0}$.
Or $\|f_n\|_\infty = 1$ qui ne tend pas vers 0.
Conclusion : on retrouve à nouveau que $\|\,.\,\|_1$ et $\|\,.\,\|_\infty$ ne sont pas équivalentes sur E.

14.4.4 Convergence en dimension finie

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $p \in \mathbb{N}^\star$ muni d'une base $e = (e_1, \dots, e_p)$.
Soit $u = (u(n))_{n\in\mathbb{N}}$ une suite d'éléments de E. Pour tout $n \in \mathbb{N}$, on peut écrire

$$u(n) = u_1(n).e_1 + \cdots + u_p(n).e_p$$

Définition

Les suites scalaires $u_j = (u_j(n))_{n\in\mathbb{N}}$ sont appelées suites coordonnées (ou composantes) de la suite vectorielle u dans la base e .

Exemple Supposons $E = \mathbb{R}^2$ et $u_n = \big(n^2, 1/(n+1)\big)$.
Les suites coordonnées de u dans la base canonique de $\mathbb{R}^2$ sont $(n^2)_{n\in\mathbb{N}}$ et $(1/(n+1))_{n\in\mathbb{N}}$.

Théorème

On a équivalence entre :
(i) u converge ;
(ii) les suites $u_1, \dots, u_p$ convergent.
De plus, si tel est le cas,

$$\lim u = (\lim u_1).e_1 + \cdots + (\lim u_n).e_p$$

dém. :
Choisissons $\|\,.\,\| = \|\,.\,\|_{\infty,e}$.

(i) $\Rightarrow$ (ii) Supposons que la suite u converge vers $\ell = \ell_1 e_1 + \cdots + \ell_p e_p$.
Pour tout $j \in \{1, \ldots, p\}$,

$$|u_j(n) - \ell_j| \leqslant \|u(n) - \ell\| \to 0$$

donc $u_j(n) \to \ell_j$.
(ii) $\Rightarrow$ (i) Supposons que pour tout $j \in \{1, \ldots, p\}$, $u_j(n) \to \ell_j$. Considérons alors $\ell = \ell_1 e_1 + \cdots + \ell_p e_p$.
On a

$$\|u(n) - \ell\|_{\infty,e} = \max\{|u_1(n) - \ell_1|, \ldots, |u_p(n) - \ell_p|\} \leqslant \sum_{j=1}^{p} |u_j(n) - \ell_j| \to 0$$

donc $u \to \ell$.
□

Exemple Dans $\mathbb{R}^2$,

$$\left(n \sin \frac{1}{n}, \left(1 + \frac{1}{n}\right)^n\right) \xrightarrow[n \to +\infty]{} (1, \mathrm{e})$$

Exemple Dans $\mathcal{M}_{p,q}(\mathbb{K})$,

$$A_n \to A \Leftrightarrow \forall i \in \{1, \ldots, p\}, \forall j \in \{1, \ldots, q\}, [A_n]_{i,j} \to [A]_{i,j}$$

Exemple Dans $\mathcal{M}_p(\mathbb{K})$,

$$A_n \to A \text{ et } B_n \to B \Rightarrow A_n B_n \to AB$$

En effet

$$[A_n B_n]_{i,j} = \sum_{k=1}^{n} [A_n]_{i,k} [B_n]_{k,j} \xrightarrow[n \to +\infty]{} [AB]_{i,j}$$

Exemple Soit $A \in \mathcal{M}_p(\mathbb{K})$. On suppose $A^n \to B$. Montrons $B^2 = B$.
Par extraction $A^{2n} \to B$ et par ce qui précède

$$A^{2n} = A^n \times A^n \to B^2$$

Par unicité de la limite

$$B^2 = B$$

14.4.5 Convergence dans un espace produit

Soit $(E_1, N_1), \ldots, (E_p, N_p)$ des espaces normés et $E = E_1 \times \cdots \times E_p = \prod_{j=1}^{p} E_j$ muni de la norme

$$\|x\| = \max_{1 \leqslant j \leqslant p} N_j(x_j)$$

Soit $u = (u(n))$ une suite d'éléments de E. Pour tout $n \in \mathbb{N}$,

$$u(n) = (u_1(n), \ldots, u_p(n))$$

Définition

Les suites vectorielles $u_j = (u_j(n))$ sont appelées suites coordonnées de la suite u.

Exemple Supposons $E_1 = E_2 = \mathcal{M}_p(\mathbb{K})$ et pour $A \in \mathcal{M}_p(\mathbb{K})$ considérons

$$u_n = \left(A^n, \frac{1}{n+1}A\right)$$

Les suites coordonnées de u sont $(A^n)_{n\in\mathbb{N}}$ et $\left(\frac{1}{n+1}A\right)_{n\in\mathbb{N}}$.

Théorème

On a équivalence entre :
(i) u converge ;
(ii) les suites $u_1, \ldots, u_p$ convergent.
De plus, si tel est le cas

$$\lim u = (\lim u_1, \ldots, \lim u_p)$$

Exemple Si $A_n \to A$ et $B_n \to B$ dans $\mathcal{M}_p(\mathbb{K})$ alors $(A_n + B_n, A_nB_n) \to (A + B, AB)$ dans $\mathcal{M}_p(\mathbb{K})^2$.

14.4.6 Séries d'éléments d'un espace normé

Soit (u_n) une suite d'éléments de l'espace normé $(E, \|\,.\,\|)$.

14.4.6.1 Vocabulaire

Définition

On appelle série de terme général u_n la suite (S_n) définie par

$$S_n = \sum_{k=0}^{n} u_k$$

Cette série est notée $\sum u_n$ et le terme S_n est appelé somme partielle de rang n de cette série.

Exemple Les séries numériques sont un cas particulier.

Exemple Soit $A \in \mathcal{M}_p(\mathbb{K})$, $\sum A^n$ et $\sum \frac{1}{n!}A^n$ sont des séries matricielles.

Définition

On dit que la série $\sum u_n$ converge si la suite (S_n) converge.
Sa limite S est alors appelée somme de la série et est notée

$$\sum_{n=0}^{+\infty} u_n$$

On introduit aussi

$$R_n = \sum_{k=n+1}^{+\infty} u_k = S - S_n$$

appelé reste de rang n de la série.

14.4.6.2 Série absolument convergente

Définition

Une série $\sum u_n$ d'éléments de E est dite absolument convergente s'il y a convergence de la série numérique à termes positifs $\sum \|u_n\|$.

Théorème

Si l'espace E est de dimension finie, l'absolue convergence d'une série d'éléments de E entraîne sa convergence

dém. :
Introduisons $e = (e_1, \ldots, e_p)$ une base de E.
Soit $\sum u(n)$ une série d'éléments de E et $u_1, \ldots, u_p$ les suites coordonnées dans e de la suite u.

$$u(n) = u_1(n).e_1 + \cdots + u_p(n).e_p$$

Toutes les normes étant équivalentes sur E, il existe $\alpha > 0$ tel que

$$\| \,.\, \|_{\infty,e} \leqslant \alpha \| \,.\, \|$$

et alors, pour tout $j \in \{1, \ldots, p\}$

$$\forall n \in \mathbb{N}, |u_j(n)| \leqslant \|u\|_{\infty,e} \leqslant \alpha \|u\|$$

Par comparaison de séries à termes positifs, il y a convergence absolue, et donc convergence de $\sum_n u_j(n)$.
On en déduit la convergence de la série $\sum u(n)$ car sa suite de sommes partielles converge.
□

Exemple Soit $E = \mathcal{M}_p(\mathbb{K})$ muni d'une norme d'algèbre $\| \,.\, \|$.
Soit $A \in \mathcal{M}_p(\mathbb{K})$ vérifiant $\|A\| < 1$. Etudions $\sum A^n$.
La série matricielle $\sum A^n$ converge absolument car

$$\|A^n\| \leqslant \|A\|^n$$

et la série $\sum \|A\|^n$ converge puisque $\|A\| < 1$.

On peut alors introduire la matrice

$$B = \sum_{n=0}^{+\infty} A^n$$

Pour $N \in \mathbb{N}$, on a

$$(I_p - A)\sum_{n=0}^{N} A^k = I_p - A^{N+1} \xrightarrow[N\to+\infty]{} I_p$$

Or on a aussi

$$(I_p - A)\sum_{n=0}^{N} A^k \xrightarrow[N\to+\infty]{} (I_p - A)B$$

On en déduit

$$B = (I_p - A)^{-1}$$

14.4.7 Musculation

Théorème

Soit E un espace de dimension finie et $f : E \to E$ une application telle qu'il existe $k \in [0,1[$ vérifiant

$$\forall x, y \in E, \|f(x) - f(y)\| \leqslant k\,\|x - y\|$$

Montrons que f admet un unique point fixe.

dém. :
Unicité : si x et y sont deux points fixes de f alors

$$\|x - y\| = \|f(x) - f(y)\| \leqslant k\,\|x - y\|$$

Sachant $k \in [0,1[$, ceci entraîne $x = y$.
Existence : soit $x_0 \in E$ et $(x_n)_{n\in\mathbb{N}}$ donnée par $x_{n+1} = f(x_n)$. On vérifie par récurrence $\|x_{n+1} - x_n\| \leqslant k^n\,\|x_1 - x_0\|$. On en déduit que la série télescopique $\sum x_{n+1} - x_n$ converge absolument et donc la suite (x_n) converge. On peut alors introduire x_∞ sa limite. Puisque $\|f(x_\infty) - x_{n+1}\| \leqslant k\,\|x_\infty - x_n\| \to 0$, on obtient $x_{n+1} \to f(x_\infty)$ puis, par unicité de la limite, $f(x_\infty) = x_\infty$.
□

Chapitre 15

Suites et séries de fonctions numériques

Les fonctions étudiées sont à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.
I et J désignent des intervalles de $\mathbb{R}$ contenant au moins deux points.

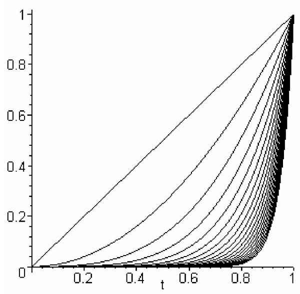
15.1 Suites de fonctions

15.1.1 Présentation

Définition

On appelle suite de fonctions de I vers $\mathbb{K}$ toute suite (u_n) d'éléments de $\mathcal{F}(I, \mathbb{K})$.

Exemple Considérons $u_n : [0,1] \to \mathbb{R}$ définie par $u_n(t) = t^n$.
$(u_n)_{n \in \mathbb{N}}$ est une suite de fonctions de $[0,1]$ vers $\mathbb{R}$.



15.1.2 Convergence simple

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$.

Définition

On dit que la suite de fonctions (u_n) converge simplement vers $u : I \to \mathbb{K}$ si

$$\forall t \in I, \quad u_n(t) \xrightarrow[n \to +\infty]{} u(t)$$

On note alors $u_n \xrightarrow[I]{CVS} u$.

Exemple Convergence simple de $(u_n)_{n\in\mathbb{N}}$ avec

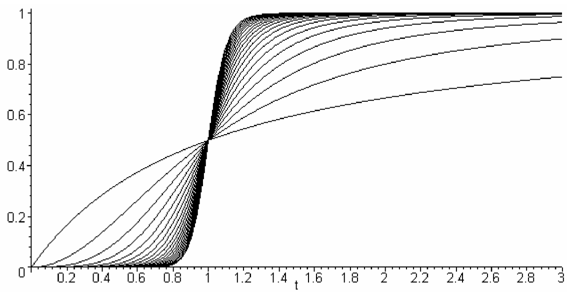
$$u_n(t) = t^n \text{ avec } t \in [0,1]$$

Soit $t \in [0,1]$.
Quand $n \to +\infty$
Si $t \in [0,1[$ alors $u_n(t) \to 0$.
Si $t = 1$ alors $u_n(t) \to 1$.
Par suite $u_n \xrightarrow{CVS} u$ avec

$$u : t \mapsto \begin{cases} 0 & \text{si } t \in [0,1[\\ 1 & \text{si } t = 1 \end{cases}$$

Exemple Convergence simple de $(u_n)_{n\in\mathbb{N}}$ avec

$$u_n(t) = \frac{t^n}{1+t^n} \text{ avec } t \in \mathbb{R}^+$$

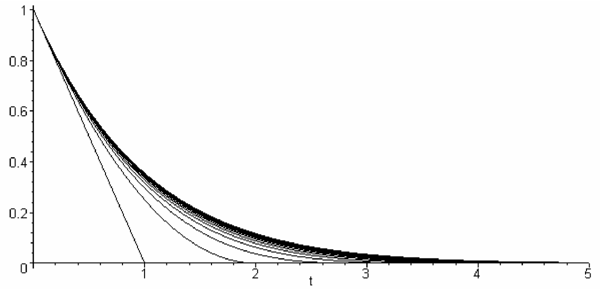


Soit $t \in \mathbb{R}^+$.
Quand $n \to +\infty$
Si $t \in [0,1[$ alors $u_n(t) \to 0$.
Si $t = 1$ alors $u_n(t) = 1/2 \to 1/2$.
Si $t \in]1,+\infty[$ alors $u_n(t) \to 1$.
Finalement $u_n \xrightarrow{CVS} u$ avec

$$u : t \mapsto \begin{cases} 0 & \text{si } t \in [0,1[\\ 1/2 & \text{si } t = 1 \\ 1 & \text{si } t =]1,+\infty[\end{cases}$$

Exemple Convergence simple de $(u_n)_{n\geqslant 1}$ avec

$$u_n(t) = \begin{cases} \left(1 - \dfrac{t}{n}\right)^n & \text{si } t \in [0,n[\\ 0 & \text{si } t \in [n,+\infty[\end{cases}$$



Soit $t \in \mathbb{R}^+$.
Quand $n \to +\infty$
Pour n assez grand, $t < n$ donc

$$u_n(t) = \left(1 - \frac{t}{n}\right)^n = \exp\left(n \ln(1 - t/n)\right) \to \mathrm{e}^{-t}$$

Ainsi $u_n \xrightarrow{CVS} u$ avec

$$u : t \mapsto \mathrm{e}^{-t}$$

Théorème

Si $u_n \xrightarrow[I]{CVS} u$ et $u_n \xrightarrow[I]{CVS} v$ alors $u = v$.

dém. :
Pour tout $t \in I$, on a $u_n(t) \to u(t)$ et $u_n(t) \to v(t)$ donc $u(t) = v(t)$.
□

Définition

Si $u_n \xrightarrow[I]{CVS} u$ alors on dit que u est la limite simple de la suite (u_n) et on note

$$u = \lim_{n \to +\infty} u_n$$

15.1.3 Propriétés de la limite simple

Proposition

Si $u_n \xrightarrow[I]{CVS} u$ et si chaque u_n est positive alors u est positive.

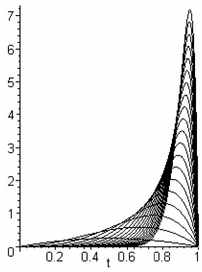
dém. :
Si toutes les fonctions u_n sont positives alors pour tout $t \in I$, $u(t) \geqslant 0$ par passage à la limite de l'inégalité$u_n(t) \geqslant 0$.
□

Proposition

Si $u_n \xrightarrow[I]{CVS} u$ et si chaque u_n est croissante alors u est croissante.

dém. :
Si toutes les fonctions u_n sont croissantes alors pour tout $x \leqslant y \in I$, $u(x) \leqslant u(y)$ par passage à la limite de l'inégalité$u_n(x) \leqslant u_n(y)$.
□
(!)$u_n \xrightarrow{CVS} u$ et chaque u_n continue n'implique pas u continue !
$u_n \xrightarrow{CVS} u$ n'implique pas $\int_I u_n(t)\,\mathrm{d}t \to \int_I u(t)\,\mathrm{d}t$!

Exemple Etudions $\int_0^1 u_n(t)\,\mathrm{d}t$ avec $u_n(t) = n^2 t^n (1-t)$



Soit $t \in [0,1]$.
Quand $n \to +\infty$.
Si $t \in [0,1[$ alors $u_n(t) \to 0$ par croissance comparée.
Si $t = 1$ alors $u_n(t) = 0 \to 0$.
Finalement $u_n \xrightarrow{CVS} \tilde{0}$.
Cependant
$$\int_0^1 u_n(t)\,\mathrm{d}t = n^2 \left(\int_0^1 t^n\,\mathrm{d}t - \int_0^1 t^{n+1}\,\mathrm{d}t \right) = \frac{n}{n+1} \to 1$$
En fait
$$u_n\left(1 - \frac{1}{n}\right) \sim \frac{n}{\mathrm{e}} \to +\infty\,!$$

15.1.4 Convergence uniforme

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$.

Définition

On dit que (u_n) converge uniformément vers $u : I \to \mathbb{K}$ si
$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow \forall t \in I, |u_n(t) - u(t)| \leqslant \varepsilon$$
On dit alors que u est limite uniforme de la suite (u_n) et on note
$$u_n \xrightarrow{CVU} u \text{ ou } u_n \xrightarrow[I]{CVU} u$$

Remarque Comparativement, dire que (u_n) converge simplement vers u signifie :
$$\forall t \in I, \forall \varepsilon > 0, \exists N > 0, \forall n \in \mathbb{N}, n \geqslant N \Rightarrow |u_n(t) - u(t)| \leqslant \varepsilon$$

Pour la convergence simple, le rang N est susceptible de dépendre de t alors que pour la convergence uniforme N doit convenir pour tout $t \in I$ (on dit qu'il est uniforme en t).

Remarque La convergence simple se comprend comme la convergence des fonctions « point par point » .
La convergence uniforme se comprend comme la convergence des fonctions « dans leur globalité » .

Théorème

Si $u_n \xrightarrow{CVU} u$ alors $u_n \xrightarrow{CVS} u$.
Ainsi, s'il y a convergence uniforme, c'est vers la limite simple de la suite de fonctions ; en particulier il y a unicité de la limite uniforme.

dém. :
Qui peut le plus, peut le moins.
□

Théorème

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$ convergeant simplement vers $u : I \to \mathbb{K}$.
S'il existe une suite réelle (α_n) vérifiant

$$\forall t \in I, |u_n(t) - u(t)| \leqslant \alpha_n \text{ et } \alpha_n \xrightarrow[n\to+\infty]{} 0$$

alors la convergence de la suite (u_n) est uniforme.

dém. :
Pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ vérifiant

$$\forall n \in \mathbb{N}, n \geqslant N \Rightarrow |\alpha_n| \leqslant \varepsilon$$

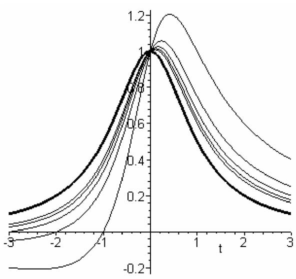
et alors

$$\forall n \in \mathbb{N}, n \geqslant N \Rightarrow \forall t \in I, |u_n(t) - u(t)| \leqslant \varepsilon$$

□

Exemple Convergence uniforme de $(u_n)_{n\geqslant 1}$ avec

$$u_n(t) = \frac{t+n}{n(1+t^2)} \text{ pour } t \in \mathbb{R}$$



Soit $t \in \mathbb{R}$.
Quand $n \to +\infty$,

$$u_n(t) \to \frac{1}{1+t^2}$$

Ainsi $u_n \xrightarrow{CVS} u$ avec

$$u : t \mapsto \frac{1}{1+t^2}$$

Etudions

$$u_n(t) - u(t) = \frac{1}{n}\frac{t}{1+t^2}$$

En vertu de l'inégalité

$$2\,|ab| \leqslant a^2 + b^2$$

on a

$$|u_n(t) - u(t)| \leqslant \frac{1}{2n} = \alpha_n$$

Puisque $\alpha_n \to 0$, on obtient finalement $u_n \xrightarrow{CVU} u$.

15.1.5 Convergence en norme uniforme

L'algèbre $\mathcal{B}(I, \mathbb{K})$ des fonctions bornées de I vers $\mathbb{K}$ est normée par

$$\|f\|_\infty = \sup_{t \in I} |f(t)|$$

Définition

La norme infinie $\|\,.\,\|_\infty$ est encore appelée norme uniforme et est parfois notée $\|\,.\,\|_{\infty,I}$.

Remarque On peut calculer exactement $\|f\|_\infty$ à partir du tableau de variation de f.

Théorème

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$.
On a équivalence entre :
(i) (u_n) converge uniformément vers une fonction $u : I \to \mathbb{K}$;
(ii) A partir d'un certain rang, les fonctions $u_n - u$ sont bornées et $\|u_n - u\|_{\infty,I} \to 0$.

dém. :
Ecrire

$$\forall t \in I, |u_n(t) - u(t)| \leqslant \varepsilon$$

équivaut à signifier

$$u_n - u \text{ bornée et } \|u_n - u\|_{\infty,I} \leqslant \varepsilon$$

□

Exemple Convergence uniforme de $(u_n)_{n\in\mathbb{N}}$ avec

$$u_n(t) = t^n \text{ pour } t \in [0,1]$$

$u_n \xrightarrow[{[0,1]}]{CVS} u$ avec

$$u(t) = \begin{cases} 0 & \text{si } t \in [0,1[\\ 1 & \text{si } t = 1 \end{cases}$$

Etudions $u_n - u$. On a

$$u_n(t) - u(t) = \begin{cases} t^n & \text{si } t \in [0,1[\\ 1 & \text{si } t = 1 \end{cases}$$

donc $\|u_n - u\|_\infty = 1$ qui ne tend pas vers 0 donc la suite de fonctions (u_n) ne converge pas uniformément.
Cependant, pour $a \in [0,1[$,

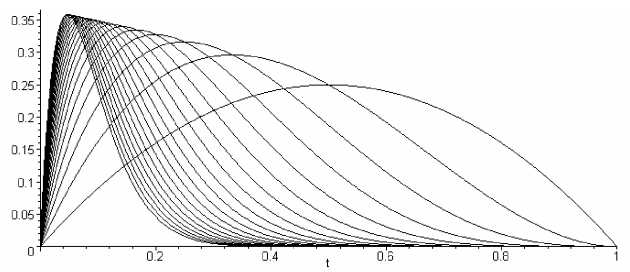
$$\|u_n - u\|_{\infty,[0,a]} = a^n \to 0$$

donc

$$u_n \xrightarrow[{[0,a]}]{CVU} \tilde{0}$$

Exemple Convergence uniforme de $(u_n)_{n\in\mathbb{N}}$ avec

$$u_n(t) = nt(1-t)^n \text{ pour } t \in [0,1]$$



Soit $t \in [0,1]$
Quand $n \to +\infty$
Si $t = 0$ alors $u_n(t) = 0 \to 0$.
Si $t \in]0,1]$ alors $u_n(t) \to 0$ par croissances comparées.
Finalement $u_n \xrightarrow{CVS} u = \tilde{0}$.
En étudiant les variations de $\delta_n(t) = u_n(t) - u(t)$ on obtient

t	0		$1/(n+1)$		1
$u_n(t) - u(t)$	0	$\nearrow$	$u_n(1/(n+1))$	$\searrow$	0

donc

$$\|u_n - u\|_\infty = u_n\left(\frac{1}{n+1}\right) = \frac{n}{n+1}\left(1 - \frac{1}{n+1}\right)^n \sim \frac{1}{\mathrm{e}}$$

Par conséquent la suite de fonctions (u_n) ne converge pas uniformément.
Cependant pour $a \in]0,1]$.

Pour n assez grand $\dfrac{1}{n+1} \leqslant a$, et puisque

t	0		$1/(n+1)$		a		1
$u_n(t) - u(t)$	0	$\nearrow$	$u_n(1/(n+1))$	$\searrow$	$u_n(a)$	$\searrow$	0

On obtient donc

$$\|u_n - u\|_\infty = u_n(a) \xrightarrow[n\to+\infty]{} 0$$

Ainsi $u_n \xrightarrow[[a,1]]{CVU} \tilde{0}$ pour tout $a \in]0,1]$.

15.2 Séries de fonctions

15.2.1 Présentation

Soit $(u_n)_{n\geqslant n_0}$ une suite de fonctions de I vers $\mathbb{K}$.

Définition

On appelle série de fonctions de terme général u_n la suite de fonctions $(S_n)_{n\geqslant n_0}$ avec

$$S_n = \sum_{k=n_0}^{n} u_k$$

Cette série de fonctions est notée $\sum\limits_{n\geqslant n_0} u_n$ et S_n est appelée somme partielle de rang n de celle-ci.

Remarque Dans la suite on supposera $n_0 = 0$ quitte à poser nulles les premières fonctions de la suite $(u_n)_{n\in\mathbb{N}}$. La série de fonctions est alors simplement notée $\sum u_n$.

Exemple Considérons $u_n : \mathbb{R} \to \mathbb{R}$ définie par $u_n(t) = t^n$.
La série de fonctions $\sum u_n$ est la suite de fonctions (S_n) avec

$$S_n(t) = \sum_{k=0}^{n} u_k(t) = \sum_{k=0}^{n} t^k = \begin{cases} \dfrac{1-t^{n+1}}{1-t} & \text{si } t \neq 1 \\ n+1 & \text{si } t = 1 \end{cases}$$

15.2.2 Convergence simple

Soit $\sum u_n$ une série de fonctions de I vers $\mathbb{K}$.

Définition

On dit que la série de fonctions $\sum u_n$ converge simplement si la suite (S_n) de ses sommes partielles converge simplement vers une certaine fonction S.
Cette fonction S est appelée somme de la série de fonctions et on note

$$S = \sum_{n=0}^{+\infty} u_n$$

Théorème

On a équivalence entre :
(i) la série de fonctions $\sum u_n$ converge simplement sur I ;
(ii) la série numérique $\sum u_n(t)$ converge pour chaque $t \in I$.
De plus, si tel est le cas

$$\left(\sum_{n=0}^{+\infty} u_n\right)(t) = \sum_{n=0}^{+\infty} u_n(t)$$

dém. :
(i) $\Leftrightarrow \forall t \in I$, $(S_n(t))$ converge.
Or $S_n(t) = \left(\sum_{k=0}^{n} u_k\right)(t) = \sum_{k=0}^{n} u_k(t)$ donc
(i) $\Leftrightarrow \forall t \in I, \sum u_n(t)$ converge.
De plus, on a alors

$$S(t) = \lim_{n \to +\infty} S_n(t) = \sum_{n=0}^{+\infty} u_n(t)$$

□

Définition

Si la série de fonctions $\sum u_n$ converge simplement, on peut introduire son reste de rang n

$$R_n = \sum_{k=n+1}^{+\infty} u_k : t \mapsto \sum_{k=n+1}^{+\infty} u_k(t)$$

Proposition

Si la série de fonctions $\sum u_n$ converge simplement alors sa somme S vérifie

$$S = S_n + R_n \text{ et } R_n \xrightarrow{CVS} \tilde{0}$$

dém. :

Pour tout $t \in I$,

$$S(t) = \left(\sum_{k=0}^{+\infty} u_k\right)(t) = \sum_{k=0}^{+\infty} u_k(t) = \sum_{k=0}^{n} u_k(t) + \sum_{k=n+1}^{+\infty} u_k(t) = S_n(t) + R_n(t)$$

De plus, pour tout $t \in I$, $R_n(t) \to 0$ car $R_n(t)$ est le reste d'une série numérique convergente.
□

Exemple Convergence simple de $\sum u_n$ avec

$$u_n : \mathbb{R} \to \mathbb{R} \text{ définie par } u_n(t) = t^n$$

Pour $t \in \mathbb{R}$, la série numérique $\sum u_n(t) = \sum t^n$ converge si, et seulement si, $t \in \left]-1,1\right[$.
Par conséquent, la série de fonctions $\sum u_n$ converge simplement sur $\left]-1,1\right[$.
Sa somme S est définie sur $\left]-1,1\right[$ et

$$S(t) = \sum_{n=0}^{+\infty} t^n = \frac{1}{1-t} \text{ pour } t \in \left]-1,1\right[$$

Exemple Convergence simple de $\sum_{n \geqslant 1} u_n$ avec

$$u_n : \mathbb{R} \to \mathbb{R} \text{ définie par } u_n(t) = 1/n^t$$

Pour $t \in \mathbb{R}$, $\sum u_n(t) = \sum 1/n^t$ converge si, et seulement si, $t > 1$.
Par conséquent, la série de fonctions $\sum u_n$ converge simplement sur $\left]1,+\infty\right[$.
Sa somme est définie sur $\left]1,+\infty\right[$, on la note ζ et cela définit la fonction zêta de Riemann

$$\zeta(t) = \sum_{n=1}^{+\infty} \frac{1}{n^t} \text{ pour } t \in \left]1,+\infty\right[$$

Remarque L'étude de la convergence simple de $\sum u_n$ fournit le domaine de définition de la fonction $\sum_{n=0}^{+\infty} u_n$.

15.2.3 Convergence uniforme

Soit $\sum u_n$ une série de fonctions de I vers $\mathbb{K}$.

Définition

> On dit que la série de fonctions $\sum u_n$ converge uniformément lorsque la suite (S_n) de ses sommes partielles converge uniformément.

Théorème

On a équivalence entre :
(i) la série de fonctions $\sum u_n$ converge uniformément sur I ;
(ii) la série de fonctions $\sum u_n$ converge simplement et $R_n \xrightarrow[I]{CVU} \tilde{0}$.

dém. :
(i) $\Leftrightarrow \exists S : I \to \mathbb{K}, S_n \xrightarrow{CVU} S$
$\Leftrightarrow \exists S : I \to \mathbb{K}, S_n \xrightarrow{CVS} S$ et $S_n - S \xrightarrow{CVU} \tilde{0}$
$\Leftrightarrow$ (ii)
□

Remarque Pour étudier la convergence uniforme de (R_n) vers la fonction nulle, on pourra :
- raisonner par majoration uniforme, c'est-à-dire déterminer (α_n) telle que

$$\forall t \in I, |R_n(t)| \leqslant \alpha_n \text{ avec } \alpha_n \to 0$$

- évaluer $\|R_n\|_\infty$ et étudier si $\|R_n\|_\infty \to 0$.

Exemple Convergence uniforme de $\sum_{n\geqslant 1} u_n$ avec

$$u_n(t) = \frac{(-1)^n}{n+t} \text{ pour } t \in \mathbb{R}^+$$

Pour $t \in \mathbb{R}^+$, la série $\sum \frac{(-1)^n}{n+t}$ est convergente en vertu du CSSA donc la série de fonctions $\sum u_n$ converge simplement sur $\mathbb{R}^+$.
La fonction $S : t \mapsto \sum_{n=1}^{+\infty} \frac{(-1)^n}{n+t}$ est donc définie sur $\mathbb{R}^+$
On a

$$R_n(t) = \sum_{k=n+1}^{+\infty} \frac{(-1)^k}{k+t}$$

Par le CSSA,

$$|R_n(t)| \leqslant \frac{1}{n+1+t} \leqslant \frac{1}{n+1} \xrightarrow[n\to+\infty]{} 0$$

Par majoration uniforme, on peut affirmer que $\sum u_n$ converge uniformément sur $\mathbb{R}^+$.

15.2.4 Convergence normale

Soit $\sum u_n$ une série de fonctions de I vers $\mathbb{K}$.

Définition

On dit que la série de fonctions $\sum u_n$ converge normalement lorsque :
- les fonctions u_n sont toutes bornées ;
- la série numérique $\sum \|u_n\|_\infty$ est convergente.

Théorème

Si la série de fonctions $\sum u_n$ converge normalement alors celle-ci converge uniformément et la convergence est absolue en tout point.

dém. :
Supposons la série de fonctions $\sum u_n$ normalement convergente sur I.
Pour tout $t \in I$, $|u_n(t)| \leqslant \|u_n\|_\infty$ donc par comparaison de séries à termes positifs, la série numérique $\sum u_n(t)$ est absolument convergente.
En particulier, cette série converge et donc la série de fonctions $\sum u_n$ converge simplement.
Aussi, pour tout $t \in I$,

$$|R_n(t)| \leqslant \sum_{k=n+1}^{+\infty} |u_k(t)| \leqslant \sum_{k=n+1}^{+\infty} \|u_k\|_\infty$$

donc

$$|R_n(t)| \leqslant \sum_{k=n+1}^{+\infty} \|u_k\|_\infty \to 0$$

Par majoration uniforme de limite nulle, on peut affirmer que la série de fonctions $\sum u_n$ converge uniformément.
□

Remarque $CVN \Rightarrow CVU \Rightarrow CVS$.
Les réciproques sont fausses.

Remarque Pour montrer qu'une série de fonctions $\sum u_n$ converge normalement sur I, il suffit de déterminer (α_n) telle que

$$\forall t \in I, |u_n(t)| \leqslant \alpha_n \text{ et } \sum \alpha_n \text{ converge}$$

Exemple Convergence uniforme de $\sum u_n$ avec

$$u_n(t) = \frac{\sin(nt)}{n^2+1} \text{ pour } t \in \mathbb{R}$$

On a

$$|u_n(t)| \leqslant \frac{1}{n^2+1}$$

Or $\sum \frac{1}{n^2+1}$ converge et donc, par majoration uniforme, la série de fonctions $\sum u_n$ converge normalement.
Par conséquent, $\sum u_n$ converge simplement et uniformément sur $\mathbb{R}$.

Exemple Convergence uniforme de $\sum_{n\geqslant 1} u_n$ avec

$$u_n(t) = \frac{1}{n} - \frac{1}{n+t} \text{ pour } t \in [0, +\infty[$$

Pour $t \in \mathbb{R}^+$, $\sum u_n(t) = \sum \left(\frac{1}{n} - \frac{1}{n+t}\right)$ avec

$$\frac{1}{n} - \frac{1}{n+t} = \frac{t}{n(n+t)} \underset{n\to+\infty}{\sim} \frac{t}{n^2}$$

Par équivalence de série à termes positifs, il y a convergence de $\sum u_n(t)$ et donc la série de fonctions converge simplement sur $\mathbb{R}^+$.
Etudions sa convergence normale. Puisque

t	0		$+\infty$
$u_n(t)$	0	$\nearrow$	$1/n$

u_n est bornée et $\|u_n\|_{\infty,\mathbb{R}^+} = 1/n$. Il n'y a pas convergence normale sur $\mathbb{R}^+$
Cependant pour $a \geqslant 0$, on a

$$\forall t \in [0,a], |u_n(t)| \leqslant u_n(a)$$

et puisque $\sum u_n(a)$ converge, il y a convergence normale (et donc uniforme) de la série de fonctions étudiée sur $[0,a]$ pour tout $a \in \mathbb{R}^+$.

Remarque En pratique la convergence uniforme d'une série de fonctions s'obtient le plus souvent :
- par convergence normale ;
- par $\|R_n\|_\infty \to 0$ via exploitation du critère spécial des séries alternées si cela est contextuel.

15.3 Continuité et limite

15.3.1 Continuité

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$.

Théorème

Si $u_n \xrightarrow{CVU} u$ et si chaque u_n est continue en $a \in I$ alors u est continue en a.

dém. :
Exploitons

$$|u(t) - u(a)| \leqslant |u(t) - u_n(t)| + |u_n(t) - u_n(a)| + |u_n(a) - u(a)|$$

Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que pour tout $n \in \mathbb{N}$,

$$n \geqslant N \Rightarrow \forall t \in I, |u_n(t) - u(t)| \leqslant \varepsilon$$

Fixons un tel $n \geqslant N$. La relation précédente donne

$$|u(t) - u(a)| \leqslant 2\varepsilon + |u_n(t) - u_n(a)|$$

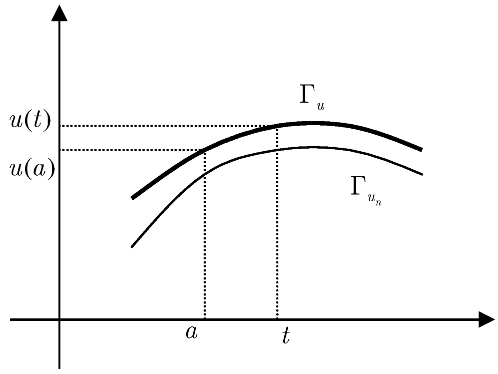
La fonction u_n étant continue en a, il existe $\alpha > 0$ tel que

$$\forall t \in I, |t-a| \leqslant \alpha \Rightarrow |u_n(t) - u_n(a)| \leqslant \varepsilon$$

En vertu de la relation initiale, on a alors

$$\forall t \in I, |t-a| \leqslant \alpha \Rightarrow |u(t)-u(a)| \leqslant 3\varepsilon$$

Ainsi, la fonction u est continue en a.



□

Corollaire

> La limite uniforme d'une suite de fonctions continues est continue.

Exemple Soit $u_n : [0,1] \to \mathbb{R}$ définie par $u_n(t) = t^n$.
La limite simple de (u_n) n'est pas continue alors que chaque u_n l'est : il n'y a pas convergence uniforme sur $[0,1]$!

Corollaire

> Si $\sum u_n$ est une série de fonctions continues uniformément convergente alors sa somme S est continue.

dém. :
$S_n = \sum_{k=0}^{n} u_k \xrightarrow{CVU} S$ et chaque S_n est continue donc S est continue.
□

Exemple Définition et continuité sur $[0,1]$ de la fonction

$$S : t \mapsto \sum_{n=0}^{+\infty} \frac{(-1)^n t^n}{2n+1}$$

Introduisons

$$u_n : t \in [0,1] \mapsto \frac{(-1)^n t^n}{2n+1}$$

Pour tout $t \in [0,1]$, la série numérique $\sum u_n(t)$ converge via CSSA.
Par suite la série de fonctions $\sum u_n$ converge simplement sur $[0,1]$ et donc S est définie sur $[0,1]$.
De plus, par le CSSA,

$$|R_n(t)| \leqslant \frac{t^{n+1}}{2n+3} \leqslant \frac{1}{2n+3} \xrightarrow[n\to+\infty]{} 0$$

Par majoration uniforme de limite nulle, on peut affirmer que la série de fonctions $\sum u_n$ converge uniformément sur $[0,1]$. Or chaque u_n est continue donc la somme S est continue sur $[0,1]$.

15.3.2 Continuité par convergence uniforme sur tout segment

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$.

Définition

On dit que la suite de fonctions (u_n) converge uniformément sur tout segment de I vers u : $I \to \mathbb{K}$ lorsque

$$\forall\,[a,b] \subset I, u_n \xrightarrow[{[a,b]}]{CVU} u$$

Proposition

Si tel est le cas, la suite (u_n) converge simplement vers u sur I.

dém. :
Pour $t \in I$, il existe $[a,b] \subset I$ tel que $t \in [a,b]$ et $u_n \xrightarrow[{[a,b]}]{CVU} u$ entraîne $u_n(t) \xrightarrow[n\to+\infty]{} u(t)$.
□

Exemple Si (u_n) converge uniformément sur I alors (u_n) converge a fortiori uniformément sur tout segment de I.

Attention : La réciproque est fausse : la convergence uniforme sur tout segment de I n'implique pas la convergence uniforme sur I.

Exemple Précédemment, pour $u_n(t) = \dfrac{1}{n} - \dfrac{1}{n+t}$, on a vu que $\sum u_n$ convergeait normalement sur $[0,a]$ pour tout $a > 0$ donc $\sum u_n$ converge uniformément sur tout segment de $[0,+\infty[$.

Théorème

Si (u_n) converge uniformément vers u sur tout segment de I et si chaque u_n est continue alors u est continue.

dém. :
Soit $t_0 \in I$.
Si t_0 n'est pas extrémité de I, il existe $\alpha > 0$ tels que $[t_0 - \alpha, t_0 + \alpha] \subset I$.
Par convergence uniforme de (u_n) sur le segment $[t_0 - \alpha, t_0 + \alpha]$, on peut affirmer que la fonction u est continue sur ce segment et en particulier la fonction u est continue en t_0.
Si t_0 est une extrémité de I : idem avec des segments $[t_0, t_0 + \alpha]$ ou $[t_0 - \alpha, t_0]$.
□

Corollaire

Si $\sum u_n$ est une série de fonctions continues convergeant uniformément sur tout segment de I alors sa somme est continue.

Exemple Définition et continuité sur $\mathbb{R}$ de la fonction $S : t \mapsto \displaystyle\sum_{n=0}^{+\infty} \frac{t^n}{(2n+1)!}$

Introduisons $u_n : \mathbb{R} \to \mathbb{R}$ définie par $u_n(t) = \dfrac{t^n}{(2n+1)!}$.
Pour $t \in \mathbb{R}$.

$$u_n(t) = \frac{1}{(2n+1)} \frac{1}{2n} \frac{t^n}{(2n-1)!} = o\left(\frac{1}{n^2}\right)$$

car par croissances comparées

$$\frac{t^n}{(2n-1)!} \to 0$$

La série numérique $\sum u_n(t)$ est absolument convergente et donc convergente.
Ainsi, la série $\sum u_n$ converge simplement sur $\mathbb{R}$ et donc S est définie sur $\mathbb{R}$.
Etudions la convergence uniforme via convergence normale.
La fonction u_n n'est pas bornée sur $\mathbb{R}$, il n'y a pas convergence normale sur $\mathbb{R}$.
Soit $a > 0$.
Sur $[-a, a]$,

$$|u_n(t)| \leqslant \frac{a^n}{(2n+1)!} = u_n(a)$$

Puisque la série numérique $\sum u_n(a)$ converge, on peut par majoration uniforme, affirmer que la série de fonctions $\sum u_n$ converge normalement, et donc uniformément, sur $[-a, a]$.
Puisque ceci vaut pour tout $a > 0$, on peut affirmer que $\sum u_n$ converge uniformément sur tout segment de $\mathbb{R}$, or chaque u_n est continue donc S est continue sur $\mathbb{R}$.

15.3.3 Limite et comportement asymptotique

Soit (u_n) une suite de fonctions de I vers $\mathbb{K}$ et a un point ou une extrémité éventuellement infinie de I.

Théorème

Si (u_n) converge uniformément sur I vers $u : I \to \mathbb{K}$ et si chaque u_n tend vers une limite finie ℓ_n en a alors la suite (ℓ_n) converge et

$$u(t) \xrightarrow[t \to a]{} \lim_{n \to +\infty} \ell_n$$

Autrement dit

$$\lim_{t \to a} \lim_{n \to +\infty} u_n(t) = \lim_{n \to +\infty} \lim_{t \to a} u_n(t)$$

dém. :
Commençons par établir que la suite (ℓ_n) est bornée.
Pour $\varepsilon = 1 > 0$, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geqslant N, \forall t \in I, |u_n(t) - u(t)| \leqslant 1$$

et donc

$$\forall n \geqslant N, \forall t \in I, |u_n(t) - u_N(t)| \leqslant 2$$

En passant à la limite quand $t \to a$, on obtient

$$|\ell_n - \ell_N| \leqslant 2$$

Ainsi, la suite (ℓ_n) est bornée.
Par le théorème de Bolzano Weierstrass, elle possède une suite extraite convergente $(\ell_{\varphi(n)})$ de limite ℓ_∞.
Montrons que u tend vers ℓ_∞ en a.
Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que

$$\forall n \geqslant N, \forall t \in I, |u_n(t) - u(t)| \leqslant \varepsilon$$

En particulier

$$\forall n \geqslant N, \forall t \in I, \left|u_{\varphi(n)}(t) - u(t)\right| \leqslant \varepsilon$$

Parallèlement, il existe $N' \in \mathbb{N}$ tel que

$$\forall n \geqslant N', \left|\ell_{\varphi(n)} - \ell_\infty\right| \leqslant \varepsilon$$

Considèrons, $n = \max(N, N')$. Puisque $u_{\varphi(n)} \underset{a}{\to} \ell_{\varphi(n)}$, on obtient au voisinage de a

$$\left|u_{\varphi(n)}(t) - \ell_{\varphi(n)}\right| \leqslant \varepsilon$$

puis

$$|u(t) - \ell_\infty| \leqslant 3\varepsilon$$

Ainsi u converge vers ℓ_∞ en a. Ceci détermine alors la valeur de ℓ_∞ de façon unique et puisque la suite (ℓ_n) est bornée et ne possède qu'une seule valeur d'adhérence, elle converge vers celle-ci.
□

Corollaire

Si $\sum u_n$ converge uniformément sur I et si chaque u_n tend vers une limite finie ℓ_n en a alors la série numérique $\sum \ell_n$ converge et

$$\sum_{n=0}^{+\infty} u_n(t) \xrightarrow[t \to a]{} \sum_{n=0}^{+\infty} \ell_n$$

Autrement dit

$$\lim_{t \to a} \sum_{n=0}^{+\infty} u_n(t) = \sum_{n=0}^{+\infty} \lim_{t \to a} u_n(t)$$

dém. :
$S_n = \sum_{k=0}^{n} u_n$ converge uniformément vers $S = \sum_{n=0}^{+\infty} u_n$ et $S_n \underset{a}{\to} \sum_{k=0}^{n} \lim_a u_k$ donc par le théorème de la double limite, la suite $\left(\sum_{k=0}^{n} \lim_a u_k\right)$ converge et $S \underset{a}{\to} \lim_{n \to +\infty} \sum_{k=0}^{n} \lim_a u_k$.
□

Exemple a) Définition et continuité de $S(x) = \sum_{n=1}^{+\infty} \frac{1}{n^2 + x^2}$ pour $x \in \mathbb{R}$.

b) Limite en $+\infty$.

c) Equivalent en $+\infty$.
a) Posons
$$u_n(x) = \frac{1}{n^2 + x^2}$$
Les fonctions u_n sont définies et continues sur $\mathbb{R}$ et
$$\forall x \in \mathbb{R}, |u_n(x)| \leqslant \frac{1}{n^2}$$
Puisque la série $\sum 1/n^2$ converge, la série de fonctions $\sum u_n$ converge normalement, et donc uniformément sur $\mathbb{R}$. On en déduit que S est définie et continue sur $\mathbb{R}$.
b) On a
$$\forall n \in \mathbb{N}^\star, \lim_{x \to +\infty} \frac{1}{n^2 + x^2} = 0$$
Puisqu'il y a convergence uniforme au voisinage de $+\infty$, on peut appliquer le théorème de la double limite et affirmer
$$\lim_{x \to +\infty} S(x) = \sum_{n=1}^{+\infty} 0 = 0$$
c) La fonction $t \mapsto 1/(t^2 + x^2)$ est décroissante et donc
$$\int_n^{n+1} \frac{\mathrm{d}t}{t^2 + x^2} \leqslant \frac{1}{n^2 + x^2} \leqslant \int_{n-1}^{n} \frac{\mathrm{d}t}{t^2 + x^2}$$
En sommant, on obtient
$$\int_1^{+\infty} \frac{\mathrm{d}t}{t^2 + x^2} \leqslant S(x) \leqslant \int_0^{+\infty} \frac{\mathrm{d}t}{t^2 + x^2}$$
Puisque
$$\int_0^{+\infty} \frac{\mathrm{d}t}{t^2 + x^2} = \left[\frac{1}{x} \arctan\left(\frac{t}{x}\right)\right]_0^{+\infty} = \frac{\pi}{2x} \text{ et } \int_1^{+\infty} \frac{\mathrm{d}t}{t^2 + x^2} = \frac{\pi/2 - \arctan(1/x)}{x} \sim \frac{\pi}{2x}$$
on obtient
$$S(x) \underset{x \to +\infty}{\sim} \frac{\pi}{2x}$$

Exemple a) Définition et continuité de $S(t) = \displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n}{nt + 1}$ pour $t > 0$.
b) Limite de S en $+\infty$.
c) Développement asymptotique à deux termes en $+\infty$.
a) Introduisons $u_n : t \in \mathbb{R}^{+\star} \mapsto \dfrac{(-1)^n}{nt + 1}$.
Pour $t > 0$, $\sum u_n(t)$ converge en vertu du CSSA.
$\sum u_n$ converge simplement sur $\mathbb{R}^{+\star}$ donc S est définie sur $\mathbb{R}^{+\star}$.
Par le critère spécial des séries alternées
$$|R_n(t)| \leqslant \frac{1}{(n+1)t + 1}$$

Pour $a > 0$, sur $[a, +\infty[$,

$$|R_n(t)| \leqslant \frac{1}{(n+1)a+1} \xrightarrow[n\to+\infty]{} 0$$

$\sum u_n$ converge uniformément sur $[a, +\infty[$ pour tout $a > 0$ donc $\sum u_n$ converge uniformément sur tout segment de $\mathbb{R}^{+\star}$. On en déduit que la fonction S est continue.
b) $\sum u_n$ converge uniformément sur $[a, +\infty[$ et

$$\lim_{+\infty} u_n = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{sinon} \end{cases}$$

Par le théorème de la double limite, la série $\sum \lim\limits_{+\infty} u_n$ converge et

$$\lim_{+\infty} S = \sum_{n=0}^{+\infty} \lim_{+\infty} u_n = 1 + 0 + 0 + \cdots = 1$$

c) On a déjà $S(t) \underset{t\to+\infty}{=} 1 + o(t)$. Déterminons un équivalent de $S(t) - 1$ quand $t \to +\infty$.
On a

$$S(t) - 1 = \sum_{n=1}^{+\infty} \frac{(-1)^n}{nt+1}$$

donc

$$t(S(t) - 1) = \sum_{n=1}^{+\infty} \frac{(-1)^n t}{nt+1}$$

Introduisons

$$v_n : t > 0 \mapsto \frac{(-1)^n t}{nt+1}$$

Le critère spécial des séries alternées s'applique à $\sum v_n(t)$ donc

$$|R_n(t)| \leqslant \frac{t}{(n+1)t+1} \leqslant \frac{1}{n+1} \xrightarrow[n\to+\infty]{} 0$$

$\sum v_n$ converge uniformément sur $\mathbb{R}^{+\star}$ et puisque

$$\lim_{+\infty} v_n = \frac{(-1)^n}{n}$$

le théorème de la double limite s'applique et la série $\sum \frac{(-1)^n}{n}$ est donc convergente avec

$$\lim_{t\to+\infty} t(S(t) - 1) = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n} = -\ln 2$$

On en déduit

$$S(t) - 1 \underset{t\to+\infty}{\sim} -\frac{\ln 2}{t}$$

15.4 Intégration et dérivation

15.4.1 Intégration sur un segment

Théorème

Soit (u_n) une suite de fonctions de $[a,b]$ vers $\mathbb{K}$.
Si (u_n) converge uniformément sur $[a,b]$ et si chaque u_n est continue alors la fonction $u = \lim\limits_{n\to+\infty} u_n$ est continue et la suite $\left(\int_a^b u_n(t)\,\mathrm{d}t\right)$ converge vers $\int_a^b u(t)\,\mathrm{d}t$.
Autrement dit

$$\lim_{n\to+\infty}\int_a^b u_n(t)\,\mathrm{d}t = \int_a^b u(t)\,\mathrm{d}t$$

dém. :
u est continue car limite uniforme d'une suite de fonctions continues, on peut donc introduire $\int_a^b u$.
Puisque

$$\left|\int_a^b u_n(t)\,\mathrm{d}t - \int_a^b u(t)\,\mathrm{d}t\right| \leqslant \int_a^b |u_n(t)-u(t)|\,\mathrm{d}t \leqslant (b-a)\,\|u_n-u\|_\infty \to 0$$

on a

$$\int_a^b u_n(t)\,\mathrm{d}t \to \int_a^b u(t)\,\mathrm{d}t$$

□

Corollaire

Soit $\sum u_n$ est une série de fonctions de $[a,b]$ vers $\mathbb{K}$
Si
1) chaque u_n est continue ;
2) $\sum u_n$ converge uniformément sur $[a,b]$;
alors sa somme $\sum\limits_{n=0}^{+\infty} u_n$ est continue et la série numérique $\sum \int_a^b u_n(t)\,\mathrm{d}t$ converge vers $\int_a^b \sum\limits_{n=0}^{+\infty} u_n(t)\,\mathrm{d}t$.
Autrement dit

$$\sum_{n=0}^{+\infty}\int_a^b u_n(t)\,\mathrm{d}t = \int_a^b \sum_{n=0}^{+\infty} u_n(t)\,\mathrm{d}t$$

dém. :
$S_n = \sum\limits_{k=0}^{n} u_k \xrightarrow{CVU} S = \sum\limits_{n=0}^{+\infty} u_n$ donc $\int_a^b S_n \to \int_a^b S$ i.e. $\sum\limits_{k=0}^{n}\int_a^b u_k \to \int_a^b \sum\limits_{n=0}^{+\infty} u_n$.
□

Exemple Calculons $\int_0^1 S(t)\,\mathrm{d}t$ avec $S(t) = \sum\limits_{n=1}^{+\infty}\left(\frac{1}{n} - \frac{1}{n+t}\right)$.

Introduisons $u_n : [0,1] \to \mathbb{R}$ définie par $u_n(t) = \dfrac{1}{n} - \dfrac{1}{n+t}$.
On a
$$\|u_n\|_\infty = \frac{1}{n(n+1)} = O\left(\frac{1}{n^2}\right)$$
La série de fonctions $\sum u_n$ converge normalement sur $[0,1]$ donc uniformément et
$$\int_0^1 S(t)\,\mathrm{d}t = \int_0^1 \sum_{n=0}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+t}\right)\,\mathrm{d}t = \sum_{n=1}^{+\infty} \left(\int_0^1 \frac{1}{n} - \frac{1}{n+t}\,\mathrm{d}t\right)$$
Or
$$\int_0^1 \frac{1}{n} - \frac{1}{(n+t)}\,\mathrm{d}t = \frac{1}{n} - \ln\frac{n+1}{n}$$
et
$$\sum_{k=1}^{n} \left(\frac{1}{k} - \ln\frac{k+1}{k}\right) = \sum_{k=1}^{n} \frac{1}{k} - \ln(n+1) \to \gamma$$
donc
$$\int_0^1 S(t)\,\mathrm{d}t = \gamma$$

Attention : Ces résultats ne valent que pour une intégration sur un segment !

Exemple Considérons $u_n : [0,+\infty[\to \mathbb{R}$ définie par $u_n(t) = \dfrac{1}{n}\mathrm{e}^{-t/n}$.
$\|u_n\|_\infty = \dfrac{1}{n} \to 0$ donc $u_n \xrightarrow[{[0,+\infty[}]{CVU} \tilde{0}$ alors que $\displaystyle\int_0^{+\infty} u_n(t)\,\mathrm{d}t = 1$ ne tend pas vers 0 !

15.4.2 Dérivation

Lemme

Soit (φ_n) une suite de fonctions continues de I vers $\mathbb{K}$ et $a \in I$.
On pose
$$\Phi_n(x) = \int_a^x \varphi_n(t)\,\mathrm{d}t$$
Si (φ_n) converge uniformément sur tout segment de I vers une fonction φ, alors la suite de fonctions (Φ_n) converge uniformément sur tout segment de I vers la fonction Φ avec
$$\Phi(x) = \int_a^x \varphi(t)\,\mathrm{d}t$$

dém. :
Notons que φ_n et φ sont continues ce qui permet d'introduire les intégrales définissant Φ_n et Φ.
Soit $[\alpha,\beta]$ un segment de I. Quitte à agrandir ce segment, on peut supposer que $a \in [\alpha,\beta]$.

Pour tout $x \in [\alpha, \beta]$
Cas $x \geqslant a$

$$|\Phi_n(x) - \Phi(x)| \leqslant \int_a^x |\varphi_n(t) - \varphi(t)| \,\mathrm{d}t \leqslant (x-a) \|\varphi_n - \varphi\|_{\infty,[\alpha,\beta]} \leqslant (\beta - \alpha) \|\varphi_n - \varphi\|_{\infty,[\alpha,\beta]}$$

Cas $x \leqslant a$
Idem.
Ainsi

$$\|\Phi_n - \Phi\|_{\infty,[\alpha,\beta]} \leqslant (\beta - \alpha) \|\varphi_n - \varphi\|_{\infty,[\alpha,\beta]} \to 0$$

□

Théorème

Soit (u_n) une suite de fonctions de classe $\mathcal{C}^1$ de I vers $\mathbb{K}$
Si (u_n) converge simplement sur I et si (u_n') converge uniformément sur tout segment de I ;
alors la fonction $u = \lim\limits_{n\to+\infty} u_n$ est de classe $\mathcal{C}^1$ et $u' = \lim\limits_{n\to+\infty} u_n'$.
Ainsi

$$\left(\lim_{n\to+\infty} u_n\right)' = \lim_{n\to+\infty} u_n'$$

De plus, la convergence de la suite (u_n) est uniforme sur tout segment de I.

dém. :
Posons $\varphi_n = u_n'$ et $\varphi = \lim u_n' = \lim \varphi_n$.
Soit $a \in I$ et Φ_n définie par

$$\Phi_n(x) = \int_a^x \varphi_n(t)\,\mathrm{d}t$$

Par le lemme, (Φ_n) converge uniformément sur tout segment de I vers Φ donnée par

$$\Phi(x) = \int_a^x \varphi(t)\,\mathrm{d}t$$

L'application Φ est de classe $\mathcal{C}^1$ avec $\Phi' = \varphi$
Parallèlement

$$\Phi_n(x) = \int_a^x u_n'(t)\,\mathrm{d}t = u_n(x) - u_n(a) \xrightarrow[n\to+\infty]{} u(x) - u(a)$$

pour tout $x \in I$.
Par unicité de limite,

$$\Phi(x) = u(x) - u(a)$$

puis

$$u(x) = \Phi(x) + u(a)$$

Par suite u est de classe $\mathcal{C}^1$ avec $u' = \varphi = \lim u_n'$.
De plus, soit $[\alpha, \beta] \subset I$.
On a

$$u_n(x) - u(x) = \Phi_n(x) - \Phi(x) + u_n(a) - u(a)$$

donc

$$\|u_n - u\|_{\infty,[\alpha,\beta]} \leqslant \|\Phi_n - \Phi\|_{\infty,[\alpha,\beta]} + |u_n(a) - u(a)|$$

or $\Phi_n \xrightarrow[{[\alpha,\beta]}]{CVU} \Phi$ et $u_n(a) \to u(a)$ donc

$$\|u_n - u\|_{\infty,[\alpha,\beta]} \to 0$$

Ainsi la convergence de (u_n) est uniforme sur $[\alpha, \beta]$.
□

Corollaire

Soit $\sum u_n$ une série de fonctions de classe $\mathcal{C}^1$ de I vers $\mathbb{K}$.
Si $\sum u_n$ converge simplement sur I et si $\sum u_n'$ converge uniformément sur tout segment de I
alors la somme $\sum_{n=0}^{+\infty} u_n$ est de classe $\mathcal{C}^1$ et

$$\left(\sum_{n=0}^{+\infty} u_n\right)' = \sum_{n=0}^{+\infty} u_n'$$

Attention : L'hypothèse de travail est « classe $\mathcal{C}^1$ » et non seulement « dérivable » !

Exemple Monotonie sur $]0, +\infty[$ de la fonction $S : t \mapsto \sum_{n=0}^{+\infty} \frac{(-1)^n}{n+t}$
Introduisons les fonctions $u_n :]0, +\infty[\to \mathbb{R}$ définies par

$$u_n(t) = \frac{(-1)^n}{n+t}$$

Soit $t > 0$. la série numérique $\sum u_n(t)$ converge en vertu du CSSA.
La série de fonctions $\sum u_n$ converge alors simplement sur $]0, +\infty[$ et sa somme S est donc bien définie sur $]0, +\infty[$.
u_n est de classe $\mathcal{C}^1$ et

$$u_n'(t) = \frac{(-1)^{n+1}}{(n+t)^2}$$

Soit $t > 0$. La série numérique $\sum u_n'(t)$ converge en vertu du CCSA
On a

$$|R_n(t)| \leqslant \frac{1}{(n+1+t)^2} \leqslant \frac{1}{(n+1)^2} \xrightarrow[n\to+\infty]{} 0$$

Ainsi la série de fonctions $\sum u_n'$ converge uniformément sur $]0, +\infty[$.
On peut alors affirmer que S est de classe $\mathcal{C}^1$ et

$$S'(t) = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}}{(t+n)^2}$$

Par le CSSA, $S'(t)$ est du signe de son premier terme $\frac{(-1)^{0+1}}{t^2} \leqslant 0$.

La fonction S est donc décroissante.
Pour compléter le tableau de variation de S, exploitons le CSSA pour encadrer S par deux sommes partielles consécutives :

$$\frac{1}{t} - \frac{1}{t+1} \leqslant S(t) \leqslant \frac{1}{t}$$

On peut alors affirmer $S \underset{+\infty}{\longrightarrow} 0$ et $S \underset{0^+}{\longrightarrow} +\infty$.

15.4.3 Dérivées d'ordres supérieurs

Théorème

Soit (u_n) une suite de fonctions de classe $\mathcal{C}^p$ de I vers $\mathbb{K}$.
Si les suites (u_n),…, $(u_n^{(p-1)})$ convergent simplement sur I et si la suite de fonctions $(u_n^{(p)})$ converge uniformément sur tout segment de I alors la fonction $u = \lim\limits_{n\to+\infty} u_n$ est de classe $\mathcal{C}^p$ et pour tout $k \in \{1,\dots,p\}$,

$$u^{(k)} = \lim_{n\to+\infty}\left(u_n^{(k)}\right)$$

dém. :
Par récurrence sur $p \in \mathbb{N}^\star$.
Pour $p = 1$: ok
Supposons la propriété vraie au rang p et étudions celle-ci au rang $p+1$.
Puisque $(u_n^{(p)})$ converge simplement et que $(u_n^{(p+1)})$ converge uniformément sur tout segment, on peut affirmer que $\lim\limits_{n\to+\infty} u_n^{(p)}$ est de classe $\mathcal{C}^1$ et

$$\left(\lim_{n\to+\infty} u_n^{(p)}\right)' = \lim_{n\to+\infty}\left(u_n^{(p+1)}\right)$$

De plus $(u_n^{(p)})$ converge uniformément sur tout segment.
Par l'hypothèse de récurrence, on a alors $\lim\limits_{n\to+\infty} u_n$ de classe $\mathcal{C}^p$ et pour tout $k \in \{1,\dots,p\}$,

$$\left(\lim_{n\to+\infty} u_n\right)^{(k)} = \lim_{n\to+\infty}\left(u_n^{(k)}\right)$$

En particulier

$$\left(\lim_{n\to+\infty} u_n\right)^{(p)} = \lim_{n\to+\infty}\left(u_n^{(p)}\right)$$

est une fonction de classe $\mathcal{C}^1$ et donc u est de classe $\mathcal{C}^{p+1}$ avec

$$\left(\lim_{n\to+\infty} u_n\right)^{(p+1)} = \lim_{n\to+\infty}\left(u_n^{(p)}\right)' = \lim_{n\to+\infty}\left(u_n^{(p+1)}\right)$$

Récurrence établie.
□

Corollaire

Soit $\sum u_n$ une série de fonctions de classe $\mathcal{C}^p$ de I vers $\mathbb{K}$.
Si les séries $\sum u_n, \dots, \sum u_n^{(p-1)}$ convergent simplement et si la série $\sum u_n^{(p)}$ converge uniformément sur tout segment de I alors la fonction $\sum_{n=0}^{+\infty} u_n$ est de classe $\mathcal{C}^p$ et, pour tout $k \in \{1, \dots, p\}$,

$$\left(\sum_{n=0}^{+\infty} u_n\right)^{(k)} = \sum_{n=0}^{+\infty} u_n^{(k)}$$

15.4.4 Application : l'exponentielle réelle

Exemple Pour $x \in \mathbb{R}$, posons

$$e(x) = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$$

a) Définition.
b) Dérivation.
a) On introduit $u_n(x) = x^n/n!$ définie sur $\mathbb{R}$.
Pour $x = 0$, la série $\sum u_n(0)$ est évidemment convergente et

$$e(0) = \sum_{n=0}^{+\infty} \frac{0^n}{n!} = 1 + 0 + \cdots = 1 \text{ car } 0^0 = 1$$

Pour $x \neq 0$,

$$\left|\frac{u_{n+1}(x)}{u_n(x)}\right| = \frac{|x|}{n+1} \xrightarrow[n\to+\infty]{} 0 < 1$$

Par application de la règle de d'Alembert, la série $\sum u_n(x)$ converge.
Ainsi, la série de fonctions $\sum u_n$ converge simplement sur $\mathbb{R}$.
Dérivation :
Les fonctions u_n sont de classe $\mathcal{C}^1$ et

$$u_n'(x) = \begin{cases} 0 & \text{si } n = 1 \\ x^{n-1}/(n-1)! & \text{si } n > 1 \end{cases}$$

Soit $a > 0$. Pour $x \in [-a, a]$ et $n \geqslant 1$,

$$|u_n'(x)| \leqslant \frac{a^{n-1}}{(n-1)!}$$

Or $\sum a^{n-1}/(n-1)!$ converge donc $\sum u_n'$ converge normalement sur $[-a, a]$.
Résumons : les u_n sont de classe $\mathcal{C}^1$, $\sum u_n$ converge simplement et $\sum u_n'$ converge uniformément sur tout segment. On en déduit que la fonction e est de classe $\mathcal{C}^1$ et

$$e'(x) = \sum_{n=1}^{+\infty} u_n'(x) = \sum_{n=1}^{+\infty} \frac{x^{n-1}}{(n-1)!} = e(x)$$

Ainsi, la fonction e est solution de l'équation différentielle $y' = y$ et vérifie $y(0) = 1$. On reconnaît l'exponentielle réelle.

Théorème

$$\forall x \in \mathbb{R}, \mathrm{e}^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$$

Exemple En particulier

$$\mathrm{e} = \sum_{n=0}^{+\infty} \frac{1}{n!}$$

15.4.5 Application : étude de la fonction zêta

Exemple Pour $s \in]1, +\infty[$, posons

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$$

a) Définition et classe $\mathcal{C}^\infty$.
b) Monotonie et convexité.
c) Etude en $+\infty$.
d) Etude en 1^+.
a) Posons $u_n : s \mapsto 1/n^s$ définie sur $]1, +\infty[$ pour $n \in \mathbb{N}^\star$.
La série de fonctions $\sum u_n$ converge simplement sur $]1, +\infty[$ et la fonction ζ est sa somme.
b) Les fonctions u_n sont de classe $\mathcal{C}^\infty$ sur $]1, +\infty[$ et

$$u_n^{(k)}(s) = \frac{(-\ln n)^k}{n^s}$$

Sur $[a, b] \subset]1, +\infty[$,

$$\forall s \in [a, b], \left|u_n^{(k)}(s)\right| \leqslant \frac{(\ln n)^k}{n^a}$$

Soit $\rho \in]1, a[$, on a

$$n^\rho \times \frac{(\ln n)^k}{n^a} \xrightarrow[n \to +\infty]{} 0$$

et il y a donc convergence de la série $\sum \frac{(\ln n)^k}{n^a}$.
Par majoration uniforme, la série de fonctions $\sum u_n^{(k)}$ converge normalement sur $[a, b]$.
Par convergence uniformément sur tout segment de $]1, +\infty[$, on peut affirmer que ζ est de classe $\mathcal{C}^\infty$ sur $]1, +\infty[$ et

$$\zeta^{(p)}(s) = \sum_{n=1}^{+\infty} \frac{(-\ln n)^p}{n^s}$$

b) Monotonie :

$$\zeta'(s) = \sum_{n=1}^{+\infty} \frac{-\ln n}{n^s} \leqslant 0$$

ζ est décroissante.

Convexité :

$$\zeta''(s) = \sum_{n=1}^{+\infty} \frac{(\ln n)^2}{n^s} \geqslant 0$$

ζ est convexe.
c) Limite en $+\infty$:

$$\lim_{s\to+\infty} \frac{1}{n^s} = \begin{cases} 0 & \text{si } n > 1 \\ 1 & \text{si } n = 1 \end{cases}$$

Pour appliquer le théorème de la double limite, observons la convergence uniforme au voisinage de $+\infty$.
Pour $s \geqslant 2$

$$|u_n(s)| \leqslant \frac{1}{n^2}$$

Or $\sum \frac{1}{n^2}$ converge normalement, donc $\sum u_n$ converge normalement et donc uniformément sur $[2, +\infty[$
Par le théorème de la double limite

$$\lim_{s\to+\infty} \zeta(s) = \sum_{n=1}^{+\infty} \lim_{s\to+\infty} \frac{1}{n^s} = 1 + 0 + 0 + \cdots = 1$$

Equivalent de $\zeta(s) - 1$ quand $s \to +\infty$:
On a

$$\zeta(s) - 1 = \frac{1}{2^s} + \sum_{n=3}^{+\infty} \frac{1}{n^s}$$

avec

$$0 \leqslant \sum_{n=3}^{+\infty} \frac{1}{n^s} \leqslant \int_2^{+\infty} \frac{\mathrm{d}t}{t^s} = \frac{1}{(s-1)} \frac{1}{2^s}$$

donc

$$\zeta(s) - 1 = \frac{1}{2^s} + o\left(\frac{1}{2^s}\right) \sim \frac{1}{2^s}$$

d) Limite en 1^+ :
Par monotonie, on peut affirmer que la fonction ζ admet une limite en 1^+.
Puisque

$$\zeta(s) \geqslant \sum_{k=1}^{n} \frac{1}{k^s}$$

à la limite

$$\lim_{s\to1^+} \zeta(s) \geqslant \sum_{k=1}^{n} \frac{1}{k}$$

Or ceci vaut pour tout n et on sait $\sum_{k=1}^{n} \frac{1}{k} \xrightarrow[n\to+\infty]{} +\infty$ donc

$$\lim_{s\to1^+} \zeta(s) = +\infty$$

Equivalent en 1^+ :

La fonction $t \mapsto \dfrac{1}{t^s}$ est décroissante donc

$$\int_n^{n+1} \frac{\mathrm{d}t}{t^s} \leqslant \frac{1}{n^s} \leqslant \int_{n-1}^{n} \frac{\mathrm{d}t}{t^s}$$

On en déduit

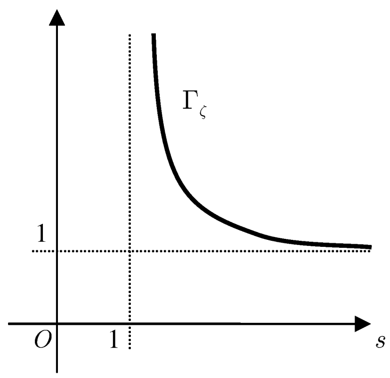
$$\int_1^{+\infty} \frac{\mathrm{d}t}{t^s} \leqslant \zeta(s) \leqslant 1 + \int_1^{+\infty} \frac{\mathrm{d}t}{t^s}$$

i.e.

$$\frac{1}{s-1} \leqslant \zeta(s) \leqslant 1 + \frac{1}{s-1}$$

Par suite

$$\zeta(s) \underset{s\to 1}{\sim} \frac{1}{s-1}$$



Chapitre 16

Topologie des espaces normés

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$
$(E, \| \, . \, \|)$ désigne un $\mathbb{K}$-espace vectoriel normé.
Les notions qui suivront ne seront pas modifiées lorsqu'on passe d'une norme à une norme équivalente.
En particulier, si l'espace E est de dimension finie, elles ne dépendent pas de la norme choisie.

16.1 Intérieur et adhérence

X désigne une partie de E.

16.1.1 Intérieur d'une partie

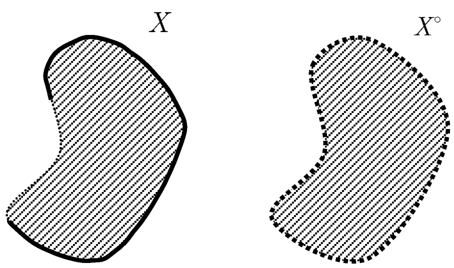
Définition

Un élément $a \in E$ est dit intérieur à une partie X si X est voisinage de a i.e.

$$\exists \alpha > 0, B(a, \alpha) \subset X$$

On note X° l'ensemble des éléments intérieurs à X appelé intérieur de X.

Exemple



Exemple Les éléments intérieurs à X sont éléments de X i.e. $X^\circ \subset X$.

Exemple Si $X \subset \mathbb{R}$ alors a est intérieur à X si, et seulement si,

$$\exists \alpha > 0,]a - \alpha, a + \alpha[\subset \mathbb{R}$$

Exemple L'intérieur d'un intervalle non vide est l'intervalle ouvert de mêmes extrémités.

Exemple L'intérieur du demi-plan complexe

$$P = \{z \in \mathbb{C}/\text{Im}(z) \geqslant 0\}$$

est

$$P^\circ = \{z \in \mathbb{C}/\text{Im}(z) > 0\}$$

Exemple L'intérieur d'une boule ouverte $B(a, r)$ est elle-même.
En effet, pour tout $x \in B(a, r)$, on vérifie $B(x, \alpha) \subset B(a, r)$ avec $\alpha = r - \|x - a\| > 0$.

16.1.2 Adhérence d'une partie

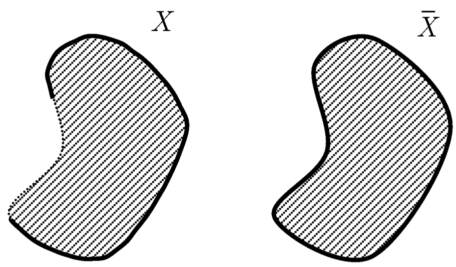
Définition

On dit qu'un élément a est adhérent à X si X intercepte tous les voisinages de a i.e. :

$$\forall \alpha > 0, B(a, \alpha) \cap X \neq \emptyset$$

On appelle adhérence de X l'ensemble noté $\bar{X}$ des éléments adhérents à X.

Exemple



Exemple Les éléments de X sont adhérents à X i.e. $X \subset \bar{X}$.

Exemple Si $X \subset \mathbb{R}$ alors a est adhérent à X si, et seulement si,

$$\forall \alpha > 0,]a - \alpha, a + \alpha[\cap X \neq \emptyset$$

Exemple L'adhérence d'un intervalle non vide est l'intervalle fermé de mêmes extrémités.

Exemple 0 est adhérent à $\mathbb{C}^\star$.

Proposition

On a

$$\mathcal{C}_E\bar{X} = (\mathcal{C}_E X)^\circ \text{ et } \mathcal{C}_E X^\circ = \overline{\mathcal{C}_E X}$$

dém. :

$$x \in \mathcal{C}_E\bar{X} \Leftrightarrow x \notin \bar{X} \Leftrightarrow \exists\alpha > 0, B(a,\alpha) \cap \bar{X} = \emptyset \Leftrightarrow \exists\alpha > 0, B(a,\alpha) \subset \mathcal{C}_E X \Leftrightarrow a \in (\mathcal{C}_E X)^\circ$$

L'autre égalité se déduit de la précédente par passage au complémentaire et substitution de $\mathcal{C}_E X$ à X.
□

16.1.3 Caractérisation séquentielle des points adhérents

Théorème

Soit X une partie non vide.
On a équivalence entre :
(i) a est adhérent à X ;
(ii) $\exists (x_n) \in X^{\mathbb{N}}, x_n \to a$;

dém. :
(i) $\Rightarrow$ (ii) Supposons que pour tout $\alpha > 0$, $B(a,\alpha) \cap X \neq \emptyset$.
Pour $n \in \mathbb{N}$ et $\alpha = \dfrac{1}{n+1} > 0$, l'ensemble $B\left(a, 1/(n+1)\right) \cap X$ est non vide.
Soit x_n un élément de celui-ci. En faisant varier n, cela définit une suite $(x_n) \in X^{\mathbb{N}}$ vérifiant

$$\|x_n - a\| \leqslant \frac{1}{n+1} \to 0$$

et donc $x_n \to a$.
(ii) $\Rightarrow$ (i) Supposons (ii). Pour tout $\alpha > 0$, il existe $N \in \mathbb{N}$ tel que

$$\forall n \geqslant N, \|x_n - a\| < \alpha$$

et donc $B(a,\alpha) \cap X \neq \emptyset$.
□

Exemple Si X est une partie non vide et majorée de $\mathbb{R}$ alors le réel $\sup X$ est adhérent à X.
En effet, il existe une suite d'éléments de X convergeant vers $\sup X$

Exemple La matrice nulle O_p est adhérente à $\mathrm{GL}_p(\mathbb{K})$.
En effet,

$$\frac{1}{n} I_p \in \mathrm{GL}_p(\mathbb{K}) \to O_p$$

Exemple L'adhérence d'une boule ouverte est la boule fermée de mêmes centre et rayon.
En effet, si $x \in \overline{B(a,r)}$ alors il existe $(x_n) \in B(a,r)^{\mathbb{N}}$ telle que $x_n \to x$ et l'inégalité $\|x_n - a\| < r$ donne à la limite $\|x - a\| \leqslant r$ donc $x \in B_f(a,r)$.
Inversement, si $x \in B_f(a,r)$ alors $x = \lim(x_n)$ avec

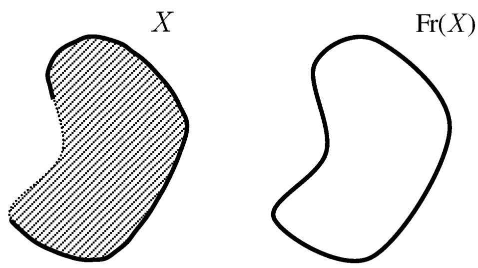
$$x_n = a + \frac{n}{n+1}(x-a) \in B(a,r)$$

16.1.4 Frontière

Définition

On appelle frontière d'une partie X de E l'ensemble Fr$(X) = \overline{X} \backslash X^\circ$.

Exemple



Exemple Dans $E = \mathbb{R}$, Fr $([a,b[) = [a,b] \setminus]a,b[= \{a,b\}$ et Fr$(\mathbb{Q}) = \overline{\mathbb{Q}} \backslash \mathbb{Q}^\circ = \mathbb{R}$.

Exemple Dans $E = \mathbb{C}$, la frontière du demi-plan $P = \{z \in \mathbb{C} / \mathrm{Im}(z) \geqslant 0\}$ est la droite réelle $\mathbb{R}$.

Exemple La frontière d'une boule (ouverte ou fermée) est la sphère de mêmes centre et rayon.

Proposition

Fr$(X) = \overline{X} \cap \overline{\mathcal{C}_E X} =$ Fr$(\mathcal{C}_E X)$.

dém. :
Fr$(X) = \bar{X} \backslash X^\circ = \bar{X} \cap \mathcal{C}_E(X^\circ) = \bar{X} \cap \overline{\mathcal{C}_E X}$.
□

Proposition

$\bar{X} = X \cup$ Fr(X) et $X^\circ = X \backslash$Fr(X).

16.2 Parties ouvertes et parties fermées

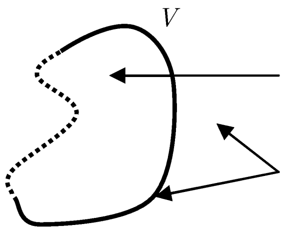
16.2.1 Voisinage

Définition

On appelle voisinage d'un élément a de E toute partie V de E vérifiant

$$\exists\alpha > 0, B(a, \alpha) \subset V$$

Exemple



Exemple Dans $E = \mathbb{R}$, une partie V de $\mathbb{R}$ est un voisinage de $a \in \mathbb{R}$ si, et seulement si,

$$\exists\alpha > 0,]a - \alpha, a + \alpha[\subset V$$

Proposition

Si V est un voisinage de a et W une partie de E contenant V alors W est un voisinage de a.

dém. :
Il existe $\alpha > 0$ tel que $B(a, \alpha) \subset V$ or $V \subset W$ donc $B(a, \alpha) \subset W$
□

Proposition

Si $V_1, \ldots, V_n$ sont des voisinages de a alors $V_1 \cap \ldots \cap V_n$ est un voisinage de a.

dém. :
Il existe $\alpha_1, \ldots, \alpha_n > 0$ tels que pour tout $i \in \{1, \ldots, n\}$, $B(a, \alpha_i) \subset V_i$.
Pour $\alpha = \min\{\alpha_1, \ldots, \alpha_n\} > 0$, $B(a, \alpha) \subset V_1 \cap \ldots \cap V_n$.
□

Remarque Ce résultat est faux pour une intersection infinie. Par exemple

$$\bigcap_{n \in \mathbb{N}^\star} [-1/n, 1/n] = \{0\}$$

est une intersection infinie de voisinage de 0 qui n'est pas un voisinage de 0.

16.2.2 Parties ouvertes

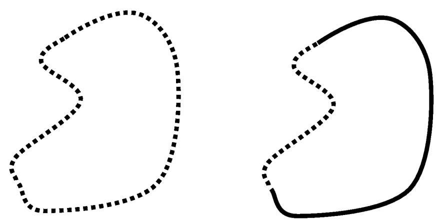
Définition

Une partie U de E est dite ouverte si elle est voisinage de chacun de ses points i.e. :

$$\forall a \in U, \exists \alpha > 0, B(a, \alpha) \subset U$$

On dit encore que U est un ouvert de E.

Exemple



Partie ouverte Partie non ouverte

Exemple Une partie U est ouverte si, et seulement si, $U^\circ = U$.
En particulier, on a alors $U \cap \mathrm{Fr}(U) = \emptyset$.

Exemple $\emptyset$ et E sont des parties ouvertes de E.

Exemple Dans $E = \mathbb{R}$, les intervalles ouverts $]a, b[\,,]a, +\infty[\,,]-\infty, a[$ sont des parties ouvertes.

Exemple Une boule ouverte $B(a, r)$ est une partie ouverte.
En effet, pour $x \in B(a, r)$ et $\alpha = r - \|x - a\| > 0$, on a $B(x, \alpha) \subset B(a, r)$.

Théorème

Une réunion (finie ou infinie) de parties ouvertes est une partie ouverte.

dém. :
Soit $(U_i)_{i \in I}$ une famille de parties ouvertes de E et $U = \bigcup_{i \in I} U_i$.
Soit $a \in U$, il existe $i \in I$ tel que $a \in U_i$. Puisque U_i est un ouvert, il existe $\alpha > 0$ tel que $B(a, \alpha) \subset U_i$ et donc $B(a, \alpha) \subset U$.
□

Exemple X° est la réunion des ouverts inclus dans X.
Par suite, X° est le plus grand ouvert inclus dans X.
En effet
Notons U la réunion des ouverts inclus dans X.
U est un ouvert inclus dans X et U contient tous les ouverts inclus dans X. Montrons $U = X^\circ$
U est un ouvert inclus dans X donc X est voisinage de chacun des points de U et donc $U \subset X^\circ$.
Inversement, si $a \in X^\circ$ il existe $\alpha > 0$ tel que $B(a,\alpha) \subset X$. $B(a,\alpha)$ est alors un ouvert inclus dans X donc $B(a,\alpha) \subset U$ puis $a \in U$. Ainsi $X^\circ \subset U$ puis $=$.

Exemple Soit $X \subset E$ et $\alpha > 0$. $X_\alpha = \bigcup_{a\in X} B(a,\alpha)$ est un ouvert de E contenant X.

Théorème

Une intersection finie de parties ouvertes est une partie ouverte.

dém. :
Soit $(U_i)_{1\leqslant i\leqslant n}$ une famille finie de parties ouvertes de E et $U = \bigcap_{i=1}^n U_i$.
Soit $a \in U$. Pour tout $i \in \{1,\dots,n\}$, il existe $\alpha_i > 0$ tel que $a \in U_i$. Pour $\alpha = \min\{\alpha_1,\dots,\alpha_n\} > 0$, on a pour tout $i \in \{1,\dots,n\}$, $B(a,\alpha) \subset B(a,\alpha_i) \subset U_i$ donc $B(a,\alpha) \subset U$.
□

Remarque Une intersection infinie de parties ouvertes peut ne pas être ouverte :
$\bigcap_{n\in\mathbb{N}^\star} \left]-1/n, 1/n\right[= \{0\}$
n'est pas une partie ouverte.

Proposition

Si $U_1,\dots,U_p$ sont des parties ouvertes des espaces normés $E_1,\dots,E_p$ alors $U = U_1\times\dots\times U_p$ est une partie ouverte de l'espace normé produit $E = E_1\times\dots\times E_p$.

dém. :
Commençons par préciser les boules de E.
Notons $N_1,\dots,N_p$ les normes sur $E_1,\dots,E_p$ et $\|\,.\,\|$ la norme sur E.
Pour $x = (x_1,\dots,x_p) \in E$, $\|x\| = \max_{1\leqslant j\leqslant p} N_j(x_j)$.
Soit $a = (a_1,\dots,a_p)$ et $r > 0$.

$$x \in B(a,r) \Leftrightarrow \forall j \in \{1,\dots,p\}, x_j \in B_j(a_j,r)$$

Ainsi

$$B(a,r) = \prod_{j=1}^p B_j(a_j,r)$$

Soit $U_1,\dots,U_p$ des parties ouvertes de E et $U = U_1\times\dots\times U_p$.
Soit $a = (a_1,\dots,a_p) \in U$. Pour tout $j \in \{1,\dots,p\}$, $a_j \in U_j$, or U_j est ouvert donc il existe $\alpha_j > 0$

tel que $B_j(a_j, \alpha_j) \subset U_j$. Considérons alors $\alpha = \min\{\alpha_1, \ldots, \alpha_p\} > 0$. Pour tout $j \in \{1, \ldots, p\}$, $B_j(a_j, \alpha) \subset U_j$ donc

$$B(a, \alpha) = \prod_{j=1}^{p} B_j(a_j, \alpha) \subset \prod_{j=1}^{p} U_j = U$$

□

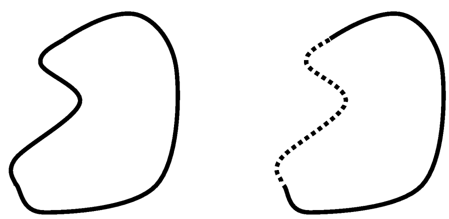
Exemple Dans $\mathbb{R}^2$, le produit cartésien de deux intervalles ouverts de $\mathbb{R}$ est un ouvert de $\mathbb{R}^2$.

16.2.3 Parties fermées

Définition

Une partie F de E est dite fermée si son complémentaire est une partie ouverte.
On dit encore que F est un fermé de E.

Exemple



Partie fermée Partie non fermée

Exemple Une partie F est fermée si, et seulement si, $\bar{F} = F$.
En particulier, on a alors $\text{Fr}(F) \subset F$.

Exemple E et $\emptyset$ sont des fermés.

Exemple Dans $E = \mathbb{R}$, les intervalles fermés $[a, b]$, $[a, +\infty[$, $]-\infty, a]$ sont des parties fermées de $\mathbb{R}$.

Théorème

Une intersection (finie ou infinie) de parties fermées est un fermé.
Une union finie de parties fermées est fermée.

dém. :
Par passage au complémentaire d'une union ou d'une intersection d'ouverts.
□

Exemple $\text{Fr}(X)$ est une partie fermée.
En effet, $\text{Fr}(X) = \bar{X} \cap \overline{\mathcal{C}_E X}$ est l'intersection de deux fermés.

Exemple $\bar{X}$ est l'intersection des fermés contenant X.
Par suite, $\bar{X}$ est le plus petit fermé contenant X.
En effet, notons F l'intersection de tous les fermes contenant X.
F est un fermé qui contient.
Si $a \notin \bar{X}$ alors il existe $\alpha > 0$ tel que $B(a, \alpha) \subset \mathcal{C}_E X$ i.e. $X \subset \mathcal{C}_E B(a, \alpha)$.
Or $C_E B(a, \alpha)$ est un fermé et donc $a \notin F$ car $a \notin \mathcal{C}_E B(a, \alpha)$.
Inversement, si $a \notin F$, puisque F est fermé, il existe $\alpha > 0$ tel que $B(a, \alpha) \subset \mathcal{C}_E F$ et donc $X \subset F \subset \mathcal{C}_E B(a, \alpha)$. On en déduit que $a \notin \bar{X}$.

Remarque Une union infinie de parties fermées peut ne pas être fermée : $\bigcup_{n \in \mathbb{N}^\star} [1/n, 1] =]0, 1]$

16.2.4 Caractérisation séquentielle des parties fermées

Théorème

> Soit F une partie de E. On a équivalence entre :
> (i) F est fermée ;
> (ii) $\forall (x_n) \in F^{\mathbb{N}}, x_n \to a \Rightarrow a \in F$
> On dit qu'une partie fermée contient les limites de ses suites convergentes.

dém. :
(i) $\Rightarrow$ (ii) Par contraposée.
Supposons qu'il existe $(x_n) \in F^{\mathbb{N}}$ telle que $x_n \to a$ et $a \notin F$.
Soit $\alpha > 0$. Pour n assez grand, $\|x_n - a\| < \alpha$ donc $x_n \in B(a, \alpha)$ et donc $B(a, \alpha) \cap F \neq \emptyset$.
Ainsi $a \in \mathcal{C}_E F$ et

$$\forall \alpha > 0, B(a, \alpha) \not\subset \mathcal{C}_E F$$

La partie $\mathcal{C}_E F$ n'est pas ouverte et donc F n'est pas fermée.
(ii) $\Rightarrow$ (i) Par contraposée.
Supposons F non fermée i.e. $\mathcal{C}_E F$ non ouvert.
Il existe $a \in \mathcal{C}_E F$ tel que

$$\forall \alpha > 0, B(a, \alpha) \cap F \neq \emptyset$$

Soit $n \in \mathbb{N}$. Pour $\alpha = 1/(n+1) > 0$, il existe $x_n \in B(a, 1/(n+1)) \cap F$.
En faisant varier n, ceci détermine une suite $(x_n) \in F^{\mathbb{N}}$ telle que $x_n \to a$ avec $a \notin F$.
□

Exemple Les singletons sont des parties fermées.

Exemple Les boules fermées sont des parties fermées.
En effet, si $(x_n) \in B_f(a, r)^{\mathbb{N}}$ converge vers ℓ alors pour tout $n \in \mathbb{N}$, $\|x_n - a\| \leqslant r$ donne à la limite $\|\ell - a\| \leqslant r$ et donc $\ell \in B_f(a, r)$. La caractérisation séquentielle des parties fermées permet alors de conclure.

Exemple Les sphères sont des parties fermées.

Proposition

Si $F_1, \ldots, F_p$ sont des parties fermées des espaces vectoriels normés $E_1, \ldots, E_p$ alors $F = F_1 \times \ldots \times F_p$ est une partie fermée de l'espace vectoriel normé produit $E = E_1 \times \cdots \times E_p$.

dém. :
Soit $(x(n)) \in F^{\mathbb{N}}$ une suite convergente de limite a.
On peut écrire $x(n) = (x_1(n), \ldots, x_p(n))$ avec $x_j(n) \to a_j$ où $a = (a_1, \ldots, a_p)$.
Pour tout $j \in \{1, \ldots, p\}$, $(x_j(n)) \in F_j^{\mathbb{N}}$, or F_j est fermée, donc $a_j \in F_j$ puis $a \in F$.
□

Exemple Dans $\mathbb{R}^2$, le produit cartésien de deux intervalles fermés de $\mathbb{R}$ est un fermé de $\mathbb{R}^2$.

16.3 Topologie et continuité

16.3.1 Topologie relative

Soit X une partie de E.

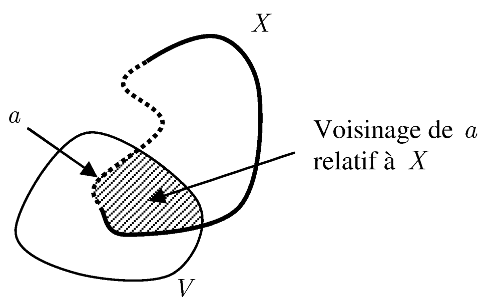
16.3.1.1 Voisinage relatif à X

Soit a un élément de E.

Définition

On appelle voisinage de a relatif à X, tout ensemble de la forme $V \cap X$ avec V voisinage de a.

Exemple



Exemple $[0,1]$ est un voisinage de 0 relatif à $\mathbb{R}^+$.
En effet, $[0,1] = [-1,1] \cap \mathbb{R}^+$.

Proposition

Soit A une partie de X. On a équivalence entre :
(i) A est un voisinage de a relatif à X ;
(ii) il existe $\alpha > 0$ tel que $B(a, \alpha) \cap X \subset A$.

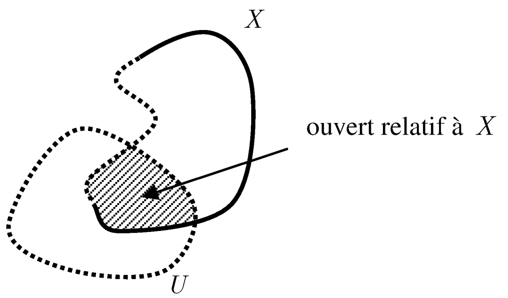
dém. :
(i) $\Rightarrow$ (ii) Si A est un voisinage de a relatif à X alors il existe V voisinage de a tel que $A = V \cap X$. Il existe $\alpha > 0$ tel que $B(a, \alpha) \subset V$ et alors $B(a, \alpha) \cap X \subset A$.
(ii) $\Rightarrow$ (i) Supposons qu'il existe $\alpha > 0$ tel que $B(a, \alpha) \cap X \subset A$. Pour $V = B(a, \alpha) \cup A$, V est un voisinage de A et $V \cap X = (B(a, \alpha) \cap X) \cup (A \cap X) = A$.
□

16.3.1.2 Ouvert relatif à X

Définition

On appelle ouvert relatif à X tout ensemble de la forme $U \cap X$ avec U ouvert de E.

Exemple



Exemple $[0, 1[$ est un ouvert relatif de $\mathbb{R}^+$.
En effet, $[0, 1[=]-1, 1[\cap \mathbb{R}^+$.

Exemple $\emptyset$ et X sont des ouverts relatifs à X.

Proposition

Soit A une partie de X. On a équivalence entre :
(i) A est un ouvert relatif à X ;
(ii) A est voisinage relatif à X de chacun de ses points.

dém. :
(i) $\Rightarrow$ (ii) Si A est un ouvert relatif à X alors $A = U \cap X$ avec U ouvert.
Pour tout $a \in A$, $a \in U$ or U est ouvert donc U est voisinage de a et $A = U \cap X$ est voisinage de a relatif à X.
(ii) $\Rightarrow$ (i) Supposons (ii)
Soit $a \in A$. A est un voisinage relatif à X de a donc il existe $\alpha_a > 0$ tel que $B(a, \alpha_a) \cap X \subset A$.
Posons alors

$$U = \bigcup_{a \in A} B(a, \alpha_a)$$

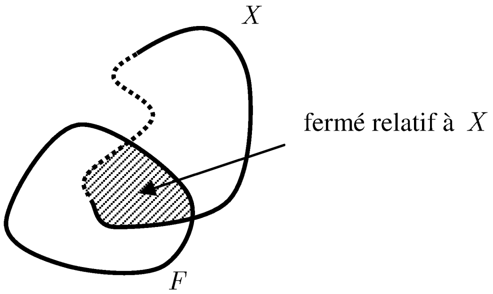
U est ouvert comme réunion d'ouverts et on vérifie $A = U \cap X$.
□

16.3.1.3 Fermé relatif à X

Définition

On appelle fermé relatif à X tout ensemble de la forme $F \cap X$ avec F fermé de E.

Exemple



Exemple $]0,1]$ est un fermé relatif de $]0,+\infty[$.
En effet, $]0,1] =]0,+\infty[\cap [0,1]$.

Exemple $\emptyset$ et X sont des fermés relatifs à X.

Théorème

Soit A une partie de X. On a équivalence entre :
(i) A est un fermé relatif à X ;
(ii) A contient les limites de ses suites convergeant dans X.
(iii) le complémentaire de A dans X est un ouvert relatif à X ;

dém. :
(i) $\Rightarrow$ (ii) Supposons $A = F \cap X$ avec F fermé. Si $(x_n) \in A^{\mathbb{N}}$ converge vers $x \in X$ alors puisque $(x_n) \in F^{\mathbb{N}}$, on a $x \in F$ donc $x \in F \cap X = A$.
(ii) $\Rightarrow$ (iii) Par contraposée. Supposons que le complémentaire de A dans X n'est pas un ouvert relatif à X. Il existe alors $a \in X \backslash A$ tel que $X \backslash A$ n'est pas voisinage relatif à X de a. Pour tout $\alpha > 0$, on a alors $B(a,\alpha) \cap X \not\subset X \backslash A$ et donc $B(a,\alpha) \cap A \neq \emptyset$. Cette propriété utilisée avec $\alpha = 1/(n+1)$ permet de construire une suite $(x_n) \in A^{\mathbb{N}}$ telle que $x_n \to a \in X \backslash A$.
(iii) $\Rightarrow$ (i) Si $X \backslash A = U \cap X$ avec U ouvert alors $A = X \cap F$ avec $F = \mathcal{C}_E U$ fermée.
□

16.3.2 Continuité et topologie

Théorème

Soit $f : X \subset E \to F$. On a équivalence entre :
(i) f est continue ;
(ii) l'image réciproque de chaque ouvert de F est un ouvert relatif à X ;
(iii) l'image réciproque de chaque fermé de F est un fermé relatif à X.

dém. :
(i) $\Rightarrow$ (ii) Supposons f continue et considérons V un ouvert de F. Pour tout $a \in f^{-1}(V)$, $f(a) \in V$ or V est ouvert et donc il existe $\varepsilon > 0$ tel que $B(f(a), \varepsilon) \subset V$. Par continuité de f en a, il existe $\alpha > 0$ vérifiant

$$\forall x \in X, \|x - a\|_E < \alpha \Rightarrow \|f(x) - f(a)\|_F < \varepsilon$$

et donc

$$\forall x \in B(a, \alpha) \cap X, f(x) \in B(f(a), \varepsilon) \subset V$$

et ainsi

$$B(a, \alpha) \cap X \subset f^{-1}(V)$$

Par suite $f^{-1}(V)$ est ouvert relatif à X car voisinage de chacun de ses points.
(ii) $\Rightarrow$ (i) Supposons (ii). Pour tout $a \in X$ et tout $\varepsilon > 0$ considérons l'ouvert $V = B(f(a), \varepsilon)$. Par hypothèse, $f^{-1}(V)$ est un ouvert relatif à X. Or $a \in f^{-1}(V)$ donc $f^{-1}(V)$ est un voisinage de a relatif à X et donc il existe $\alpha > 0$ tel que

$$B(a, \alpha) \cap X \subset f^{-1}\left(B(\ell, \varepsilon)\right)$$

On a alors

$$\forall x \in X,\ \|x - a\|_E < \alpha \Rightarrow \|f(x) - f(a)\|_F < \varepsilon$$

(ii) $\Leftrightarrow$ (iii) via $f^{-1}(C_F Y) = C_X f^{-1}(Y)$ pour $Y \subset F$.
□

Remarque Le résultat est faux en terme d'image directe

$$\sin(]0, \pi[) =]0, 1] \text{ et } \exp(\mathbb{R}^-) =]0, 1]$$

Corollaire

Pour $f : E \to F$ continue, l'image réciproque d'une partie ouverte (resp. fermée) de F est une partie ouverte de E (resp. fermée).

dém. :
Car un ouvert (resp. un fermé) relatif à E est un ouvert (resp. un fermé) de E.
□

Exemple $U = \left\{(x, y) \in \mathbb{R}^2 / x < y\right\}$ est un ouvert de $\mathbb{R}^2$.
En effet, considérons $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = y - x$.
$U = f^{-1}(\mathbb{R}^{+\star})$ or f est continue et $\mathbb{R}^{+\star}$ est ouvert donc U est un ouvert relatif à $\mathbb{R}^2$ i.e. un ouvert de $\mathbb{R}^2$.

Exemple Soit X continue et $\bar{X} = E$.
Les ensembles X, E et $\forall a \in E, \forall \alpha > 0, B(a, \alpha) \cap X \neq \emptyset$ sont fermés.
Les ensembles $\forall a \in E, \exists (x_n) \in X^{\mathbb{N}}, x_n \to a$, $E \subset \bar{X}$ et $\mathbb{Q}$ sont ouverts.

Exemple $\mathbb{R}$ est une partie ouverte de $\mathbb{R} \backslash \mathbb{Q}$.
En effet, $\mathbb{D}$ et $\det$ est continue et $\mathbb{R}$ est un ouvert.
De même, on obtient que $\mathrm{GL}_n(\mathbb{K})$ est un fermé.

Exemple $\mathcal{M}_n(\mathbb{K})$ est une partie fermée de $A \in \mathcal{M}_n(\mathbb{K})$.
En effet, $A_p = A + \frac{1}{p} I_n \to A$ avec A continue et p fermé.

16.4 Densité

16.4.1 Définition

Définition

Une partie X de E est dite dense si $\bar{X} = E$.

Théorème

On a équivalence entre :
(i) X est une partie dense de E ;
(ii) $\forall a \in E, \forall \alpha > 0, B(a, \alpha) \cap X \neq \emptyset$;
(iii) $\forall a \in E, \exists (x_n) \in X^{\mathbb{N}}, x_n \to a$.

dém. :
(ii) et (iii) signifient $E \subset \bar{X}$.
□

Exemple $\mathbb{Q}$ est une partie dense de $\mathbb{R}$.
En effet, tout réel est limite d'une suite de rationnels.

Exemple Aussi, $\mathbb{R} \backslash \mathbb{Q}$ et $\mathbb{D}$ sont des parties denses de $\mathbb{R}$.

Exemple $\mathrm{GL}_n(\mathbb{K})$ est une partie dense de $\mathcal{M}_n(\mathbb{K})$.
En effet, pour tout $A \in \mathcal{M}_n(\mathbb{K})$, on a $A_p = A + \frac{1}{p} I_n \to A$.
Or la matrice A n'a qu'un nombre fini de valeurs propres, donc pour p assez grand $f, g : E \to F$.

16.4.2 Continuité et densité

Théorème

Soit $f, g : E \to F$ continues.
Si f et g sont égales sur une partie X de E dense alors $f = g$.

dém. :
Soit $x \in X$. Il existe $(x_n) \in X^{\mathbb{N}}$ telle que $x_n \to x$. Or pour tout $n \in \mathbb{N}$ $f(x_n) = g(x_n)$ donc à la limite

$$f(x) = g(x)$$

□

Exemple Déterminons les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues vérifiant

$$\forall x, y \in \mathbb{R}, f(x + y) = f(x) + f(y)$$

Soit f solution.
On a $f(0+0) = f(0) + f(0)$ donc $f(0) = 0$.
On a $f(2a) = f(a+a) = f(a) + f(a) = 2f(a)$,...
Par récurrence, on montre

$$\forall a \in \mathbb{R}, \forall n \in \mathbb{N}, f(na) = nf(a)$$

Puisque $f(x) + f(-x) = f(0) = 0$ on a $f(-x) = -f(x)$.
Par suite

$$\forall a \in \mathbb{R}, \forall n \in \mathbb{Z}, f(na) = nf(a)$$

Soit $x = p/q \in \mathbb{Q}$ avec $p \in \mathbb{Z}$ et $q \in \mathbb{N}^\star$.
$f(x) = pf(1/q)$ et $f(1) = qf(1/q)$ donc $f(x) = \dfrac{p}{q}f(1) = \alpha x$ en posant $\alpha = f(1)$.
Les fonctions $x \mapsto f(x)$ et $x \mapsto \alpha x$ sont continues sur $\mathbb{R}$ et coïncident sur la partie $\mathbb{Q}$ dense dans $\mathbb{R}$, elles sont donc égales sur $\mathbb{R}$.

Exemple Montrons que

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}), \chi_{AB} = \chi_{BA}$$

Soit $\lambda \in \mathbb{K}$ et $B \in \mathcal{M}_n(\mathbb{K})$.
Pour $A \in \mathrm{GL}_n(\mathbb{K})$,

$$\chi_{AB}(\lambda) = \det(\lambda I_n - AB) = \det(A)\det(\lambda A^{-1} - B)$$

puis

$$\chi_{AB}(\lambda) = \det(\lambda A^{-1} - B)\det(A) = \det(\lambda I_n - BA) = \chi_{BA}(\lambda)$$

Les applications $A \mapsto \chi_{AB}(\lambda)$ et $A \mapsto \chi_{BA}(\lambda)$ sont continues sur $\mathcal{M}_n(\mathbb{K})$ et coïncident sur $\mathrm{GL}_n(\mathbb{K})$ partie dense de $\mathcal{M}_n(\mathbb{K})$, elles sont donc égales sur $\mathcal{M}_n(\mathbb{K})$.
Ainsi, pour tout $\lambda \in \mathbb{K}$, $\chi_{AB}(\lambda) = \chi_{BA}(\lambda)$ et donc $\chi_{AB} = \chi_{BA}$.

16.4.3 Approximations uniformes

Soit $a < b \in \mathbb{R}$.

16.4.3.1 Par des fonctions en escalier

Rappel :
On appelle subdivision d'un segment $[a, b]$ toute suite réelle finie $\sigma = (a_0, a_1, \ldots, a_n)$ avec

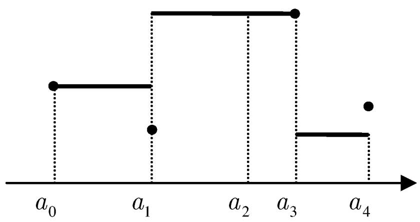
$$a_0 = a < a_1 < \ldots < a_{n-1} < a_n = b$$

Définition

Une fonction $\varphi : [a,b] \to \mathbb{K}$ est dite en escalier s'il existe une subdivision $\sigma = (a_0, a_1, \ldots, a_n)$ de $[a,b]$
vérifiant

$$\forall i \in \{1, \ldots, n\}, \varphi_{\restriction]a_{i-1},a_i[} \text{ est constante}$$

Une telle subdivision est alors dite adaptée à φ.



Théorème

Soit $f : [a,b] \to \mathbb{K}$ continue par morceaux.
Pour tout $\varepsilon > 0$, il existe une fonction en escalier $\varphi : [a,b] \to \mathbb{K}$ vérifiant

$$\forall t \in [a,b], |f(t) - \varphi(t)| \leqslant \varepsilon$$

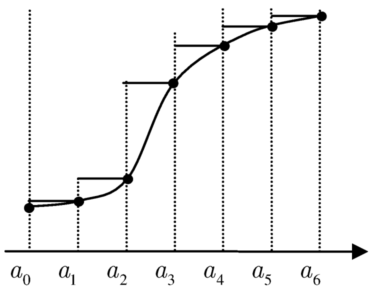
dém. :
Cas f continue sur $[a,b]$.
Soit $\varepsilon > 0$. Puisque f est continue sur le segment $[a,b]$, elle y est uniformément continue et donc il existe $\alpha > 0$ tel que

$$\forall s,t \in [a,b], |s-t| < \alpha \Rightarrow |f(s) - f(t)| \leqslant \varepsilon$$

Soit $n \in \mathbb{N}^\star$ tel que $(b-a)/n \leqslant \alpha$ et $\sigma = (a_0, \ldots, a_n)$ la subdivision de $[a,b]$ définie par

$$a_i = a + i\frac{b-a}{n}$$

Considérons $\varphi : [a,b] \to \mathbb{C}$ définie par $\varphi(t) = f(a_i)$ sur $]a_{i-1}, a_i]$ et $\varphi(a) = f(a)$.



La fonction φ est une fonction en escalier et pour tout $i \in \{1, \ldots, n\}$ et tout $t \in]a_{i-1}, a_i]$, on a

$$|t - a_i| \leqslant \frac{b-a}{n} \leqslant \alpha$$

et donc
$$|f(t) - \varphi(t)| \leqslant \varepsilon$$
Cas f continue par morceaux sur $[a, b]$.
Soit $\sigma = (a_0, \ldots, a_n)$ une subdivision de $[a, b]$ adaptée à f.
Pour tout $i \in \{1, \ldots, n\}$, on peut prolonger $f_{\restriction]a_{i-1}, a_i[}$ en une fonction continue f_i définie sur $[a_{i-1}, a_i]$.
La fonction f_i étant continue, il existe (φ_i) fonction en escalier telle que
$$\forall t \in [a_{i-1}, a_i]\,, |f_i(t) - \varphi_i(t)| \leqslant \varepsilon$$
Posons alors $\varphi : [a, b] \to E$ définie par
$$\varphi(a_i) = f(a_i) \text{ et } \varphi(t) = \varphi_i(t) \text{ si } t \in]a_{i-1}, a_i[$$
On a clairement par construction
$$\forall t \in [a, b]\,, |f(t) - \varphi(t)| \leqslant \varepsilon$$

□

Corollaire

L'ensemble $\mathcal{E}([a, b], \mathbb{K})$ des fonctions en escalier de $[a, b]$ vers $\mathbb{K}$ est une partie dense de l'espace $\mathcal{C}^0_{pm}([a, b], \mathbb{K})$ normé par $\|\,.\,\|_\infty$.
Toute fonction continue par morceaux est limite uniforme d'une suite de fonctions en escalier.

Exemple Montrons
$$\forall f \in \mathcal{C}^0_{pm}([a, b], \mathbb{K})\,, \lim_{n \to +\infty} \int_a^b f(t)\mathrm{e}^{int}\,\mathrm{d}t = 0$$
Cas f constante : C'est immédiat par calcul.
Cas f en escalier : C'est immédiat en découpant l'intégrale.
Cas f continue par morceaux :
Soit $\varepsilon > 0$. Il existe $\varphi : [a, b] \to \mathbb{K}$ en escalier vérifiant
$$\forall t \in [a, b]\,, |f(t) - \varphi(t)| \leqslant \varepsilon$$
et alors pour tout $n \in \mathbb{N}$
$$\int_a^b f(t)\mathrm{e}^{int}\,\mathrm{d}t = \int_a^b \varphi(t)\mathrm{e}^{int}\,\mathrm{d}t + \int_a^b (f(t) - \varphi(t))\,\mathrm{e}^{int}\,\mathrm{d}t$$
avec
$$\left|\int_a^b (f(t) - \varphi(t))\,\mathrm{e}^{int}\,\mathrm{d}t\right| \leqslant \int_a^b |f(t) - \varphi(t)|\,\mathrm{d}t \leqslant (b - a)\varepsilon$$
Or
$$\int_a^b \varphi(t)\mathrm{e}^{int}\,\mathrm{d}t \xrightarrow[n \to +\infty]{} 0$$
donc pour n assez grand
$$\left|\int_a^b f(t)\mathrm{e}^{int}\,\mathrm{d}t\right| \leqslant (b - a + 1)\varepsilon$$

16.4.3.2 Par des fonctions polynômes

On note $\mathcal{P}([a, b], \mathbb{K})$ l'espace des fonctions polynomiales de $[a, b]$ vers $\mathbb{K}$.

Théorème

Soit $f : [a, b] \to \mathbb{K}$ une fonction continue.
Pour tout $\varepsilon > 0$, il existe une fonction $\varphi : [a, b] \to \mathbb{K}$ polynomiale vérifiant

$$\forall t \in [a, b], |f(t) - \varphi(t)| \leqslant \varepsilon$$

Corollaire

$\mathcal{P}([a, b], \mathbb{K})$ est une partie dense de $\mathcal{C}([a, b], \mathbb{K})$ normé par $\|\,.\,\|_\infty$
Toute fonction continue sur $[a, b]$ est limite uniforme d'une suite de fonctions polynomiales.

Remarque Puisque $\|\,.\,\|_1$ et $\|\,.\,\|_2$ sont dominées par $\|\,.\,\|_\infty$, $\mathcal{P}([a, b], \mathbb{K})$ est encore une partie dense de $\mathcal{C}([a, b], \mathbb{K})$ normé par $\|\,.\,\|_1$ ou $\|\,.\,\|_2$.

Remarque Pour $k \in \mathbb{N}^\star \cup \{\infty\}$, on a $\mathcal{P}([a, b], \mathbb{K}) \subset \mathcal{C}^k([a, b], \mathbb{K})$.
Par conséquent, $\mathcal{C}^k([a, b], \mathbb{K})$ est aussi une partie dense de $\mathcal{C}([a, b], \mathbb{K})$ normé par $\|\,.\,\|_\infty$, $\|\,.\,\|_2$ ou $\|\,.\,\|_1$.

Exemple Soit $f \in \mathcal{C}([0, 1], \mathbb{R})$ vérifiant

$$\forall n \in \mathbb{N}, \int_0^1 t^n f(t)\,\mathrm{d}t = 0$$

Montrons que f est la fonction nulle.
Pour tout $P \in \mathbb{R}[X]$, on a par linéarité

$$\int_0^1 P(t)f(t)\,\mathrm{d}t = 0$$

Par le théorème de Weierstrass, il existe une suite de fonctions polynômes (φ_n) convergeant uniformément vers f sur $[a, b]$. On a alors

$$\left|\int_0^1 \varphi_n(t)f(t)\,\mathrm{d}t - \int_0^1 f^2(t)\,\mathrm{d}t\right| \leqslant \int_0^1 |\varphi_n(t) - f(t)|\,|f(t)|\,\mathrm{d}t$$

et donc

$$\left|\int_0^1 \varphi_n(t)f(t)\,\mathrm{d}t - \int_0^1 f^2(t)\,\mathrm{d}t\right| \leqslant \|\varphi_n - f\|_\infty \int_0^1 |f(t)|\,\mathrm{d}t \to 0$$

Ainsi

$$\int_0^1 \varphi_n(t)f(t)\,\mathrm{d}t \to \int_0^1 f^2(t)\,\mathrm{d}t$$

et puisque $\displaystyle\int_0^1 \varphi_n(t)f(t)\,\mathrm{d}t = 0$, on en déduit

$$\int_0^1 f^2(t)\,\mathrm{d}t = 0$$

Par nullité de l'intégrale d'une fonction continue et positive, on peut conclure $f = 0$.

16.4.4 Musculation : Sous-groupe de $(\mathbb{R}, +)$

Théorème

> Les sous-groupes de $(\mathbb{R}, +)$ sont de la forme $a\mathbb{Z}$ avec $a \in \mathbb{R}$ ou bien sont des parties denses de $\mathbb{R}$.

dém. :
Soit H un sous-groupe de $(\mathbb{R}, +)$.
Si $H = \{0\}$ alors $H = a\mathbb{Z}$ avec $a = 0$.
Sinon, il existe $h \in H$ tel que $h \neq 0$ et, quitte à considérer son opposé, on peut supposer $h > 0$.
Posons alors $a = \inf H^+$ avec $H^+ = \{h \in H / h > 0\}$.
Cette borne inférieure existe car H^+ est une partie de $\mathbb{R}$ non vide et minorée.
Cas $a > 0$:
Montrons $H = a\mathbb{Z}$.
Commençons par justifier $a \in H$.
Puisque $a = \inf H^+$, $2a$ n'est pas minorant de H^+ et donc il existe $b \in H^+$ tel que $a \leqslant b < 2a$.
Si $b > a$ alors $b - a > 0$ or, par opération dans le sous-groupe H, on a $b - a \in H$. Ainsi $b - a \in H^+$.
Cependant $b - a < a = \inf H^+$, c'est absurde.
On en déduit $b = a$ et, puisque $b \in H^+$, on obtient $a \in H$.
Sachant $a \in H$, on peut affirmer $a\mathbb{Z} = \langle a \rangle \subset H$.
Inversement, soit $x \in H$.
Par division euclidienne, on peut écrire $x = aq + r$ avec $a \in \mathbb{Z}$ et $r \in [0, a[$.
Notons que $r = x - aq \in H$ car $x \in H$ et $aq \in a\mathbb{Z} \subset H$.
Si $r > 0$ alors $r \in H^+$. Or $r < a = \inf H^+$. C'est absurde.
On en déduit $r = 0$ puis $x = aq \in a\mathbb{Z}$.
Par double inclusion, on obtient $H = a\mathbb{Z}$.
Cas $a = 0$:
Montrons que H est dense dans $\mathbb{R}$.
Soit $x \in \mathbb{R}$ et $\varepsilon > 0$.
Puisque $\inf H^+ = 0$, il existe $h \in H^+$ tel que $0 < h < \varepsilon$.
Posons alors $n = \lfloor x/h \rfloor \in \mathbb{Z}$.
On a $x/h - 1 < n \leqslant x/h$ donc $x - h < nh \leqslant x$ puis $nh \in]x - \varepsilon, x]$.
Or $nh \in H$ donc on peut affirmer $H \cap]x - \varepsilon, x + \varepsilon[\neq \emptyset$.
□

Exemple Montrons que $\{\cos(n) / n \in \mathbb{N}\}$ est dense dans $[-1, 1]$.
Considérons $H = \mathbb{Z} + 2\pi\mathbb{Z}$.
H est un sous-groupe de $(\mathbb{R}, +)$.
S'il est de la forme $a\mathbb{Z}$ avec $a \in \mathbb{R}$ alors, puisque $\mathbb{Z} \subset H = a\mathbb{Z}$, on a $a \in \mathbb{Q}$.
De plus, puisque $2\pi\mathbb{Z} \subset H = a\mathbb{Z}$, on a aussi $\pi \in a\mathbb{Q}$.
On en déduit que π est rationnel.
C'est absurde.
On peut donc affirmer que $H = \mathbb{Z} + 2\pi\mathbb{Z}$ est un sous-groupe dense dans $\mathbb{R}$.
Considérons alors $x \in [-1, 1]$ et $\theta = \arccos x \in [0, \pi] \subset \mathbb{R}$.
Il existe une suite d'éléments de H convergeant vers θ et donc il existe deux suites d'entiers (a_n) et (b_n) telles que $a_n + 2\pi b_n \to \theta$.
On a alors $\cos(|a_n|) = \cos(a_n + b_n) \to \cos\theta = x$.

Chapitre 17

Continuité d'une fonction vectorielle

E et F désignent des $\mathbb{K}$-espaces vectoriels normés par $\| \,.\, \|_E$ et $\| \,.\, \|_F$. Les notions qui vont suivre sont inchangées lorsqu'on passe d'une norme à une norme équivalente. En particulier, elles ne dépendent pas du choix de la norme lorsque les espaces sont de dimensions finies.
X désigne une partie de E.
On s'intéresse ici aux applications $f : X \subset E \to F$. En pratique, l'étude s'appliquera :
- aux fonctions numériques d'une ou plusieurs variables réelles ;
- aux fonctions d'une variable complexe ($z \mapsto \bar{z}/1 + z$, $z \mapsto \mathrm{e}^z$,...) ;
- aux applications d'une variable matricielle ($\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$, $A \in \mathrm{GL}_n(\mathbb{K}) \mapsto A^{-1}$), aux applications linéaires ou multilinéaires...

17.1 Limites

17.1.1 Convergence

Soit $f : X \subset E \to F$ et a un point adhérent à X.

Définition

On dit que f tend vers $\ell \in F$ en a si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x \in X, \|x - a\|_E \leqslant \alpha \Rightarrow \|f(x) - \ell\|_F \leqslant \varepsilon$$

On note alors $f \underset{a}{\to} \ell$ ou $f(x) \xrightarrow[x \to a]{} \ell$

Exemple Pour f constante égale à C, on obtient $C \underset{a}{\to} C$.

Exemple Pour $f = \mathrm{Id}$, on obtient $x \xrightarrow[x \to a]{} a$.

Exemple Pour $f = \| \,.\, \|$, on obtient $\|x\| \xrightarrow[x \to a]{} \|a\|$.

Théorème

Si $f \underset{a}{\to} \ell$ et $f \underset{a}{\to} \ell'$ alors $\ell = \ell'$.

dém. :
Soit $\varepsilon > 0$. Il existe $\alpha, \alpha' > 0$ tels que

$$\forall x \in X, \|x - a\|_E \leqslant \alpha \Rightarrow \|f(x) - \ell\|_F \leqslant \varepsilon$$

et

$$\forall x \in X, \|x - a\|_E \leqslant \alpha' \Rightarrow \|f(x) - \ell'\|_F \leqslant \varepsilon$$

Pour $\alpha'' = \min(\alpha, \alpha') > 0$ et $x \in B(a, \alpha'') \cap X$ (qui est non vide car a est adhérent à X), on a $\|f(x) - \ell\| \leqslant \varepsilon$ et $\|f(x) - \ell'\| \leqslant \varepsilon$. On en déduit

$$\|\ell - \ell'\| \leqslant \|\ell - f(x)\| + \|f(x) - \ell'\| \leqslant 2\varepsilon$$

Or ceci vaut pour tout $\varepsilon > 0$ donc $\|\ell - \ell'\| = 0$ i.e. $\ell = \ell'$.
□

Définition

On dit que f converge en a s'il existe $\ell \in F$ tel que $f \underset{a}{\to} \ell$.
Cet élément ℓ est alors unique, on l'appelle limite de f en a et on note

$$\ell = \lim_a f \text{ ou } \ell = \lim_{x \to a} f(x)$$

17.1.2 Théorèmes de convergences

a désigne un élément adhérent à X.

17.1.2.1 Caractérisation séquentielle

Théorème

Soit $f : X \subset E \to F$ et $\ell \in F$.
On a équivalence entre :
(i) $f \underset{a}{\to} \ell$;
(ii) $\forall (x_n) \in X^{\mathbb{N}}, x_n \to a \Rightarrow f(x_n) \to \ell$.

dém. :
(i) $\Rightarrow$ (ii) Supposons $f \underset{a}{\to} \ell$.
Soit $(x_n) \in X^{\mathbb{N}}$ telle que $x_n \to a$.
Soit $\varepsilon > 0$. Il existe $\alpha > 0$ tel que

$$\forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow \|f(x) - \ell\| \leqslant \varepsilon$$

Puisque $x_n \to a$ et $\alpha > 0$, il existe $N \in \mathbb{N}$ tel que

$$\forall n \in \mathbb{N}, n \geqslant N \Rightarrow \|x_n - a\| \leqslant \alpha$$

et donc

$$n \geqslant N \Rightarrow \|f(x_n) - \ell\| \leqslant \varepsilon$$

(ii) $\Rightarrow$ (i) Par contraposée.
Supposons $f \underset{a}{\not\to} \ell$. Il existe $\varepsilon > 0$ tel que

$$\forall \alpha > 0, \exists x \in X, \|x - a\| \leqslant \alpha \text{ et } \|f(x) - \ell\| > \varepsilon$$

Soit $n \in \mathbb{N}$, pour $\alpha = \dfrac{1}{n+1} > 0$, il existe $x_n \in X$ tel que $\|x_n - a\| \leqslant \dfrac{1}{n+1}$ et $\|f(x_n) - \ell\| > \varepsilon$.
En faisant varier n, ceci détermine une suite $(x_n) \in X^{\mathbb{N}}$ telle que $x_n \to a$ et $f(x_n) \not\to \ell$.
□

Corollaire

Si f tend vers ℓ en a alors ℓ est adhérent à $f(X)$.
Ce dernier résultat est une extension du théorème de passage à la limite des inégalités larges.

17.1.2.2 Opérations

Théorème

Soit $f, g : X \subset E \to F$ et $\lambda, \mu \in \mathbb{K}$
Si $f \underset{a}{\to} \ell$ et $g \underset{a}{\to} \ell'$ alors $\lambda f + \mu g \underset{a}{\to} \lambda\ell + \mu\ell'$.
Si de plus F est une algèbre normée, $fg \underset{a}{\to} \ell\ell'$

dém. :
Soit $(x_n) \in X^{\mathbb{N}}$ de limite a.
On a $f(x_n) \to \ell$ et $g(x_n) \to \ell'$.
Par opérations sur les suites vectorielles convergentes, $(\lambda f + \mu g)(x_n) \to \lambda\ell + \mu\ell'$.
Or ceci vaut pour toute suite $(x_n) \in X^{\mathbb{N}}$ convergeant vers a donc, par la caractérisation séquentielle des limites, $\lambda f + \mu g \underset{a}{\to} \lambda\ell + \mu\ell'$.
□

Théorème

Soit $\alpha : X \subset E \to \mathbb{K}$, $f : X \subset E \to F$..
Si $\alpha \underset{a}{\to} \lambda$ et $f \underset{a}{\to} \ell$ alors $\alpha.f \underset{a}{\to} \lambda.\ell$.

dém. :
Par la caractérisation séquentielle des limites et opérations sur les suites vectorielles convergentes.
□

Théorème

Soit $f : X \subset E \to F$ et $g : Y \subset F \to G$ telles que $f(X) \subset Y$.
Si $f \underset{a}{\to} b$ et si $g \underset{b}{\to} \ell$ alors $g \circ f \underset{a}{\to} \ell$.

dém. :
Par la caractérisation séquentielle des limites.
Notons que b est adhérent à Y car $b = \lim\limits_{a} f$ est adhérent à $f(X)$ et $f(X) \subset Y$.
□

Corollaire

Si $f \underset{a}{\to} \ell$ alors $\|f\| \underset{a}{\to} \|\ell\|$.

17.1.2.3 Comparaison

Théorème

Soit $f : X \subset E \to F$, $g : X \subset E \to \mathbb{R}$ et a adhérent à X.
Si $\|f(x) - \ell\| \leqslant g(x)$ et $g \underset{a}{\to} 0$ alors $f \underset{a}{\to} \ell$.

dém. :
Par la caractérisation séquentielle des limites et comparaison de suites réelles.
□

17.1.3 Convergence à valeurs dans espace de dimension finie

Soit F un $\mathbb{K}$-espace vectoriel de dimension finie et $e = (e_1, \dots, e_p)$ une base de F.
Considérons $f : X \subset E \to F$.

Pour tout $x \in X$, on peut écrire $f(x) = f_1(x).e_1 + \cdots + f_p(x).e_p = \sum_{j=1}^{p} f_j(x).e_j$ avec $f_j(x) \in \mathbb{K}$.

Définition

Les applications scalaires $f_1, \dots, f_p$ sont appelées fonctions coordonnées (ou composantes) de f relatives à la base $(e_1, \dots, e_p)$.

Théorème

Soit a adhérent à X. On a équivalence entre :
(i) la fonction vectorielle f converge en a ;
(ii) les fonctions numériques $f_1, \dots, f_p$ convergent en a.
De plus, si tel est le cas

$$\lim_a f = \left(\lim_a f_1\right).e_1 + \cdots + \left(\lim_a f_p\right).e_p = \sum_{j=1}^{p} \left(\lim_a f_j\right) e_j$$

dém. :
Par la caractérisation séquentielle des limites.
□

17.1.4 Convergence à valeurs dans un espace normé produit

Soit $F_1, \dots, F_p$ des espaces vectoriels normés respectivement par $N_1, \dots, N_p$ et $F = F_1 \times \ldots \times F_p$ l'espace vectoriel normé produit. Pour $x = (x_1, \dots, x_p) \in F$,

$$\|x\| = \max_{1 \leqslant j \leqslant n} N_j(x_j)$$

Considérons $f : X \subset E \to F$.
Pour tout $x \in X$ $f(x) = (f_1(x), \dots, f_p(x))$ avec $f_j(x) \in F_j$.

Définition

Les applications $f_1, \dots, f_p$ sont appelées applications coordonnées de f.

Théorème

Soit $a \in E$ adhérent X. On a équivalence entre :
(i) f converge en a ;
(ii) $f_1, \dots, f_p$ convergent en a.
De plus, si tel est le cas,

$$\lim_a f = \left(\lim_a f_1, \dots, \lim_a f_p\right)$$

dém. :
Par la caractérisation séquentielle des limites.
□

17.1.5 Convergence et restriction

Soit $f : X \subset E \to F$ et a adhérent à X.

Définition

Soit $X' \subset X$ tel que a soit adhérent à X'.
On appelle limite de f en a selon X' l'éventuelle limite de la restriction $f|_{X'}$ en a. On la note

$$\lim_{x \to a, x \in X'} f(x)$$

Exemple Si a est adhérent à $X^\star = X \setminus \{a\}$, on note

$$\lim_{x \to a, x \neq a} f(x) \underset{\text{déf}}{=} \lim_{x \to a, x \in X^\star} f(x)$$

Exemple Si $X \subset \mathbb{R}$ et a adhérent à $X^+ = X \cap]a, +\infty[$, on note

$$\lim_{x \to a^+} f(x) \text{ ou } \lim_{x \to a, x > a} f(x) \underset{\text{déf}}{=} \lim_{x \to a, x \in X^+} f(x)$$

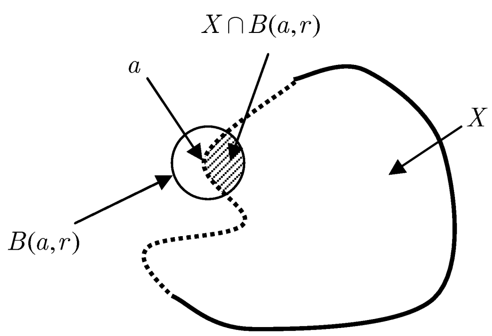
Proposition

Si a est adhérent à $X' \subset X$ et si f converge en a alors la restriction $f|_{X'}$ converge en a vers la même limite.

dém. :
Qui peut le plus, peut le moins.
□

Proposition

Soit $r > 0$ et $X' = B(a, r) \cap X$.
Si la restriction $f|_{X'}$ converge en a alors f converge en a vers la même limite



dém. :
Supposons $f_{\restriction X'}$ converge vers ℓ en a.
Soit $\varepsilon > 0$. Il existe $\alpha > 0$ tel que

$$\forall x \in X', \|x - a\|_E \leqslant \alpha \Rightarrow \|f(x) - \ell\|_F \leqslant \varepsilon$$

Pour $\alpha' = \min(\alpha, r) > 0$, on a

$$\forall x \in X, \|x - a\|_E < \alpha' \Rightarrow \|x - a\|_E \leqslant \alpha \text{ et } x \in X'$$

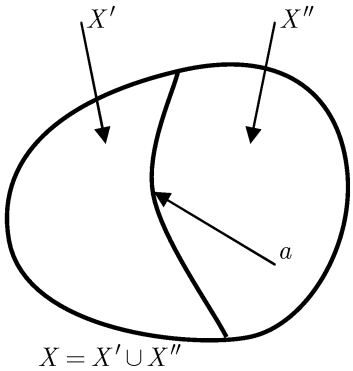
donc

$$\forall x \in X, \|x-a\|_E < \alpha' \Rightarrow \|f(x)-\ell\|_F \leqslant \varepsilon$$

□

Proposition

On suppose $X = X' \cup X''$ avec a adhérent à X' et X''.
Si les restrictions $f|_{X'}$ et $f|_{X''}$ convergent en a vers la même limite alors f converge en a vers cette limite.



dém. :
Notons ℓ la limite commune.
Soit $\varepsilon > 0$. Il existe $\alpha', \alpha'' > 0$ tels que

$$\forall x \in X', \|x-a\|_E \leqslant \alpha' \Rightarrow \|f(x)-\ell\|_F \leqslant \varepsilon \text{ et } \forall x \in X'', \|x-a\|_E \leqslant \alpha'' \Rightarrow \|f(x)-\ell\|_F \leqslant \varepsilon$$

Pour $\alpha = \min(\alpha', \alpha'') > 0$, on a

$$\forall x \in X = X' \cup X'',\ \|x-a\|_E \leqslant \alpha \Rightarrow \|f(x)-\ell\|_F \leqslant \varepsilon$$

□

Remarque Cet outil permet l'étude de limite de fonction définie par une alternative.

17.1.6 Extension « à l'infini »

Définition

Soit $f : X \subset \mathbb{R} \to F$ avec X partie non majorée.
On dit que f tend vers $\ell \in F$ en $+\infty$ si

$$\forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in X, x \geqslant A \Rightarrow \|f(x)-\ell\| \leqslant \varepsilon$$

On note alors $f(x) \xrightarrow[x\to+\infty]{} \ell$.
De façon analogue, pour $X \subset \mathbb{R}$ non minorée, on définit $f \xrightarrow[-\infty]{} \ell$

Définition

Soit $f : X \subset E \to F$ avec X non bornée.
On dit que $f(x)$ tend vers $\ell \in F$ quand $\|x\| \to +\infty$ si

$$\forall \varepsilon > 0, \exists A \in \mathbb{R}, \forall x \in X, \|x\| \geqslant A \Rightarrow \|f(x) - \ell\| \leqslant \varepsilon$$

On note alors $f(x) \xrightarrow[\|x\| \to +\infty]{} \ell$.

Définition

Soit $f : X \subset E \to \mathbb{R}$ et $a \in E$ adhérent à X.
On dit que f tend vers $+\infty$ en a si

$$\forall M \in \mathbb{R}, \exists \alpha > 0, \forall x \in X, \|x - a\| \leqslant \alpha \Rightarrow f(x) \geqslant M$$

On note alors $f(x) \xrightarrow[x \to a]{} +\infty$.
De façon analogue, on définit aussi $f(x) \xrightarrow[x \to a]{} -\infty$, $f(x) \xrightarrow[\|x\| \to +\infty]{} +\infty$, etc.

17.1.7 Exemples

Exemple Dans $\mathbb{R}^2$, étude de $\lim\limits_{(x,y) \to (0,0)} \sqrt{x^2 + xy + y^2}$.
Soit $f : (x, y) \mapsto \sqrt{x^2 + xy + y^2}$ définie sur $X = \mathbb{R}^2$ car

$$x^2 + xy + y^2 \geqslant (x + 1/2)^2 + 3y^2/4$$

$(0, 0)$ est adhérent à $\mathbb{R}^2$.
Quand $(x, y) \to (0, 0)$.
On a $x \to 0$ et $y \to 0$ (car $|x| \leqslant \|(x, y)\|_\infty \to 0$)
Par opérations algébriques $x^2 + xy + y^2 \to 0$.
Par composition $\sqrt{x^2 + xy + y^2} \to 0$.

Exemple Dans $\mathbb{R}^2$, étude de $\lim\limits_{(x,y) \to (0,0)} \dfrac{xy}{\sqrt{x^2 + y^2}}$.
Soit $f : (x, y) \mapsto \dfrac{xy}{\sqrt{x^2 + y^2}}$ définie sur $X = \mathbb{R}^2 \setminus \{(0, 0)\}$.
$(0, 0)$ est adhérent à X
Quand $(x, y) \to (0, 0)$ (avec $(x, y) \in X$)
On pose $x = r \cos\theta$, $y = r \sin\theta$ avec $r = \sqrt{x^2 + y^2} \to 0$ et θ incontrôlable.
Par composition, on a alors

$$f(x, y) = r \cos\theta \sin\theta \to 0$$

Attention : Etudier $\lim\limits_{(x,y) \to (0,0)}$ ne correspond pas à étudier $\lim\limits_{x \to 0} \lim\limits_{y \to 0}$ ou $\lim\limits_{y \to 0} \lim\limits_{x \to 0}$

Exemple Dans $\mathbb{R}^2$, étude de $\lim\limits_{(x,y) \to (0,0)} \dfrac{x^2 - y^2}{x^2 + y^2}$.

Soit $f:(x,y)\mapsto \dfrac{x^2-y^2}{x^2+y^2}$ définie sur $X=\mathbb{R}^2\setminus\{(0,0)\}$.
$(0,0)$ est adhérent à X et

$$\lim_{x\to 0}\lim_{y\to 0} f(x,y)=-1 \text{ et } \lim_{y\to 0}\lim_{x\to 0} f(x,y)=1\ldots$$

Pour $x=r\cos\theta$, $y=r\sin\theta$ avec $r=\sqrt{x^2+y^2}\to 0$, on a

$$f(x,y)=\cos^2\theta-\sin^2\theta$$

qui ne semble pas converger...
Puisque $f(1/n,0)\to 1$ et $f(0,1/n)\to -1$, la fonction f diverge en $(0,0)$.

Exemple Dans $\mathbb{R}^3$, étude de $\displaystyle\lim_{(x,y,z)\to(0,0,0)}\frac{xyz}{x^2+y^2+z^2}$.
Soit $f:(x,y,z)\mapsto \dfrac{xyz}{x^2+y^2+z^2}$ définie sur $X=\mathbb{R}^3\setminus\{(0,0,0)\}$.
Quand $(x,y,z)\to(0,0,0)$ (avec $(x,y,z)\in X$)
On pose $x=r\cos\varphi\sin\theta$, $y=r\sin\varphi\sin\theta$, $z=r\cos\theta$ avec $r=\sqrt{x^2+y^2+z^2}\to 0$ et φ,θ incontrôlables.

$$\frac{xyz}{x^2+y^2+z^2}=r\cos\varphi\sin\varphi\cos^2\theta\sin\theta\to 0$$

Exemple Dans $\mathbb{C}$, étude de $\displaystyle\lim_{z\to 0}\frac{z^2}{|z|}$.
Soit $f:z\mapsto \dfrac{z^2}{|z|}$ définie sur $X=\mathbb{C}^\star$.
0 est adhérent à $\mathbb{C}^\star$.
Quand $z\to 0$ (avec $z\in\mathbb{C}^\star$)
On peut écrire $z=r\mathrm{e}^{i\theta}$ avec $r=|z|\to 0$.
On a alors

$$f(z)=r\mathrm{e}^{2i\theta}\to 0$$

Exemple Dans $\mathbb{C}$, étude de $\displaystyle\lim_{|z|\to+\infty}\frac{1}{z+1}$.
$f:z\mapsto 1/(z+1)$ est définie sur $X=\mathbb{C}\setminus\{-1\}$.
X n'est pas bornée.
Quand $|z|\to+\infty$ (avec $z\in X$).
On a $|z+1|\geqslant|z|-1$ donc $\left|\dfrac{1}{z+1}\right|\leqslant\dfrac{1}{|z|-1}\to 0$ (pour $|z|>1$).
Ainsi $\displaystyle\lim_{|z|\to+\infty}\frac{1}{z+1}=0$

17.2 Continuité

17.2.1 Continuité en un point

Remarque Si $f : X \subset E \to F$ admet une limite en $a \in X$, celle-ci ne peut qu'être égale à $f(a)$.

Définition

On dit que $f : X \subset E \to F$ est continue en $a \in X$ si $f(x) \xrightarrow[x \to a]{} f(a)$.

Théorème

On a équivalence entre :
(i) $f : X \subset E \to F$ est continue en $a \in X$;
(ii) $\forall (x_n) \in X^{\mathbb{N}}, (x_n \to a \Rightarrow f(x_n) \to f(a))$

dém. :
En vertu de la caractérisation séquentielle des limites.
□

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x, y) = \frac{xy}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \text{ et } f(0, 0) = 0$$

La fonction f n'est pas continue en $(0, 0)$.
En effet

$$f(1/n, 1/n) = 1/2 \not\to f(0, 0)$$

17.2.2 Continuité sur le domaine de définition

Définition

On dit que $f : X \subset E \to F$ est continue si f est continue en chaque point $a \in X$.
On note $\mathcal{C}(X, F)$ l'ensemble des fonctions continues de X vers F.

Exemple Les fonctions constantes sont continues.

Exemple La fonction Id_E est continue.

Exemple La fonction $x \mapsto \|x\|$ est continue.

Exemple La fonction $z \mapsto \dfrac{1}{z}$ est continue sur $\mathbb{C}^\star$.
En effet, pour $a \in \mathbb{C}^\star$,

$$\left|\frac{1}{z} - \frac{1}{a}\right| = \frac{|z - a|}{|z|\,|a|} \xrightarrow[z \to a]{} 0$$

Exemple Etudions la continuité de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

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

Soit $(x_0, y_0) \in \mathbb{R}^2$.
Cas $x_0 \neq y_0$.
Sur une boule centrée en (x_0, y_0),

$$f(x,y) = \frac{\sin y - \sin x}{y - x} \xrightarrow[(x,y)\to(x_0,y_0)]{} \frac{\sin y_0 - \sin x_0}{y_0 - x_0} = f(x_0, y_0)$$

Cas $x_0 = y_0$.
Quand $(x, y) \to (x_0, x_0)$ avec $x \neq y$

$$f(x,y) = \frac{\sin y - \sin x}{y - x} = \frac{2 \sin \frac{y-x}{2} \cos \frac{x+y}{2}}{y - x} \to \cos x_0 = f(x_0, x_0)$$

En effet

$$\frac{2 \sin \frac{t}{2}}{t} \xrightarrow[t \to 0]{} 1 \text{ et } y - x \to 0$$

Quand $(x, y) \to (x_0, x_0)$ avec $x = y$

$$f(x,y) = \cos x \to \cos(x_0) = f(x_0, x_0)$$

17.2.3 Applications lipschitziennes

Définition

Une application $f : X \subset E \to F$ est dite lipschitzienne s'il existe $k \in \mathbb{R}^+$ tel que

$$\forall x, y \in X, \|f(y) - f(x)\|_F \leqslant k \|y - x\|_E$$

Exemple L'application $x \mapsto \|x\|$ est lipschitzienne de E vers $\mathbb{R}$.
En effet

$$|\|x\| - \|y\|| \leqslant \|x - y\|$$

Exemple On appelle distance de $x \in E$ à une partie A non vide de E le réel

$$d(x, A) \underset{\text{déf}}{=} \inf \{d(x, A)/a \in A\}$$

L'application $x \in E \mapsto d(x, A)$ est lipschitzienne.
Soit $x, y \in E$.
Pour tout $a \in E$,

$$d(x, A) \leqslant \|x - a\| \leqslant \|x - y\| + \|y - a\|$$

donc
$$d(x, A) - \|x - y\| \leqslant \|y - a\|$$
puis par passage à la borne inférieure
$$d(x, A) - \|x - y\| \leqslant d(y, A)$$
Ainsi
$$d(x, A) - d(y, A) \leqslant \|x - y\|$$
Par un raisonnement symétrique on a aussi $d(y, A) - d(x, A) \leqslant \|y - x\|$ et donc
$$|d(y, A) - d(x, A)| \leqslant \|y - x\|$$
Ainsi l'application $x \mapsto d(x, A)$ est lipschitzienne.

Théorème

Les applications lipschitziennes sont continues.

dém. :
Soit $f : X \subset E \to F$ une fonction lipschitzienne.
Il existe $k \in \mathbb{R}^+$ tel que
$$\forall x, y \in X, \|f(y) - f(x)\|_F \leqslant k \|y - x\|_E$$
Soit $a \in X$.
Quand $x \to a$, $\|f(x) - f(a)\|_F \leqslant k \|x - a\|_E \to 0$ donc $f(x) \to f(a)$.
Ainsi f est continue en chaque $a \in X$.
□

Exemple Soit E un $\mathbb{K}$-espace vectoriel de dimension finie et $e = (e_1, \ldots, e_p)$ une base de E.
Les formes linéaires coordonnées dans la base e sont lipschitziennes.
En effet, notons $\varphi_1, \ldots, \varphi_p$ les formes linéaires coordonnées dans la base e.
Pour $x = x_1.e_1 + \cdots + x_p.e_p \in E$, on a $\varphi_j(x) = x_j$.
Etudions l'application $\varphi_j : E \to \mathbb{K}$.
En choisissant $\|\,.\,\| = \|\,.\,\|_{\infty,e}$, on a pour tout $j \in \{1, \ldots, p\}$ et tout $x, y \in E$,
$$|\varphi_j(y) - \varphi_j(x)| = |y_j - x_j| \leqslant \|y - x\|$$
Ainsi les formes linéaires coordonnées dans une base sont lipschitziennes et donc continues.

Remarque En particulier, les applications suivantes sont continues
$$(x_1, \ldots, x_p) \mapsto x_j,\ z \mapsto \text{Re}(z),\ z \mapsto \text{Im}(z) \text{ et } A \mapsto a_{i,j}$$

Exemple Soit (E_1, N_1),..., (E_p, N_p) des espaces normés et $(E, \|\,.\,\|)$ l'espace normé produit.
Les applications coordonnées $p_j : x = (x_1, \ldots, x_p) \in E \mapsto x_j$ sont lipschitziennes.
En effet, pour tout $x, y \in E$,
$$N_j(p_j(x) - p_j(y)) = N_j(x_j - y_j) \leqslant \|x - y\|$$
Les projections coordonnées p_j sont lipschitziennes donc continues.

Remarque En particulier, les applications suivantes sont continues

$$\begin{array}{rcl} E\times F & \to & E \\ (x,y) & \mapsto & x \end{array} \text{ et } \begin{array}{rcl} E\times F & \to & F \\ (x,y) & \mapsto & y \end{array}$$

17.2.4 Opérations sur les fonctions continues

Théorème

Soit $f,g : X\subset E\to F$ et $\lambda,\mu\in\mathbb{K}$.
Si f et g sont continues alors $\lambda f+\mu g$ est continue.
Si de plus F est une algèbre normée, fg est aussi continue.

dém. :
Par opérations sur les limites en tout point $a\in X$.
□

Corollaire

$\mathcal{C}(X,F)$ est un sous-espace vectoriel (voire une sous-algèbre) de $\mathcal{F}(X,F)$.

Définition

On appelle fonction monôme sur $\mathbb{K}^p$ toute application de la forme

$$x=(x_1,\dots,x_p)\mapsto x_1^{\alpha_1}\times\dots\times x_p^{\alpha_p}$$

On appelle fonction polynôme sur $\mathbb{K}^p$ toute combinaison linéaire de fonctions monômes.

Exemple Les fonctions polynômes sur $\mathbb{K}^p$ sont continues.

Exemple L'application $\det : \mathcal{M}_p(\mathbb{K})\to\mathbb{K}$ est continue.
En effet

$$\det A=\sum_{\sigma\in\mathcal{S}_n}\varepsilon(\sigma)\prod_{i=1}^n a_{\sigma(i),i}$$

et donc l'application $\det$ se comprend comme une somme de produits de fonctions continues.
On dit que le déterminant est une fonction polynôme en les coefficients de la matrice.

Théorème

Soit $\alpha : X\subset E\to\mathbb{K}$ et $f : X\subset E\to F$.
Si α et f sont continues alors $\alpha.f$ est continue.

dém. :
Par opérations sur les limites en tout point $a\in X$.
□

Théorème

Soit $f : X\subset E\to F$ et $g : Y\subset F\to G$ telle que $f(X)\subset Y$.
Si f et g sont continues alors $g\circ f$ est continue.

dém. :
Par opérations sur les limites en tout point $a\in X$.
□

Définition

On appelle fonctions rationnelles sur $\mathbb{K}^p$ toute fonction qui est le rapport de deux fonctions polynômes sur $\mathbb{K}^p$.

Exemple Les fonction rationnelles sur $\mathbb{K}^p$ sont continues sur leur domaine de définition.

Exemple La fonction

$$f : (x, y) \mapsto \frac{\sin(x + y^2)}{2 + \ln(1 + x^2 + y^2)}$$

est continue sur $\mathbb{R}^2$
Par opérations sur les fonctions continues !

Attention : Ne pas argumenter f est continue car « continue en x et continue en y » .
Cette dernière notion correspond à de la continuité partielle, elle est nécessaire mais pas suffisante.

Exemple Soit $f(x, y) = \dfrac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$.
Pour tout $y \in \mathbb{R}$.
Si $y \neq 0$ alors $x \mapsto f(x, y) = \dfrac{xy}{x^2 + y^2}$ est continue.
Si $y = 0$ alors $x \mapsto f(x, y) = 0$ est continue.
Par symétrie, on a aussi $y \mapsto f(x, y)$ est continue pour tout $x \in \mathbb{R}$.
Ainsi la fonction f est « continue en x et en y » .
Cependant, la fonction f n'est pas continue puisque $f(1/n, 1/n) = 1/2 \not\to f(0, 0)$.

Théorème

Si F est de dimension finie alors $f : X \subset E \to F$ est continue si, et seulement si, ses fonctions coordonnées dans une base de F le sont.

Exemple L'application $M \mapsto \text{com}(M)$ est continue de $\mathcal{M}_p(\mathbb{K})$ vers lui-même.
En effet, ses applications coordonnées dans la base canonique sont des polynômes en les coefficients de M.

Exemple L'application $M \mapsto M^{-1}$ est continue sur $\text{GL}_p(\mathbb{K})$.
En effet, on sait

$$M^{-1} = \frac{1}{\det M} {}^t\text{com} M$$

et donc les coefficients de M^{-1} sont des fonctions rationnelles en les coefficients de M.

Théorème

Si F est un espace normé produit alors $f : X \subset E \to F$ est continue si, et seulement si, ses fonctions coordonnées le sont.

Exemple L'application $A \in \mathcal{M}_n(\mathbb{K}) \mapsto (\det A, \text{com} A) \in \mathbb{K} \times \mathcal{M}_n(\mathbb{K})$ est continue.

17.3 Continuité et linéarité

17.3.1 Continuité des applications linéaires

Définition

On note $\mathcal{L}_c(E,F)$ l'ensemble formé des applications linéaires continues de E vers F.

Théorème

$\mathcal{L}_c(E,F)$ est un $\mathbb{K}$-espace vectoriel

dém. :
$\mathcal{L}_c(E,F) = \mathcal{L}(E,F) \cap \mathcal{C}(E,F)$ est un sous-espace vectoriel de $\mathcal{F}(E,F)$.
□

Théorème

Soit une application linéaire $u : E \to F$. On a équivalence entre :
(i) u est continue ;
(ii) u est continue en 0_E ;
(iii) $\exists k \geqslant 0, \forall x \in E, \|u(x)\|_F \leqslant k\|x\|_E$ [lipschitzianité en 0] ;
(iv) u est lipschitzienne.

dém. :
(i) $\Rightarrow$ (ii) : ok
(ii) $\Rightarrow$ (iii) : Supposons u continue en 0.
Pour $\varepsilon = 1$, il existe $\alpha > 0$ tel que

$$\forall x \in E, \|x\|_E \leqslant \alpha \Rightarrow \|u(x)\|_F \leqslant 1$$

Posons $k = 1/\alpha \in \mathbb{R}^+$ et montrons que

$$\forall x \in E,\ \|u(x)\|_F \leqslant k\|x\|_E$$

Pour $x = 0$: ok
Pour $x \neq 0$, posons $x' = \dfrac{\alpha}{\|x\|_E}x$. On a $\|x'\|_E \leqslant \alpha$ donc $\|u(x')\|_F \leqslant 1$.
Or $\|u(x')\|_F = \dfrac{\alpha}{\|x\|_E}\|u(x)\|_F$ donc puis $\|u(x)\|_F \leqslant \dfrac{1}{\alpha}\|x\|_E$.
(iii) $\Rightarrow$ (iv) : Supposons qu'il existe $k \geqslant 0$ tel que $\|u(x)\| \leqslant k\|x\|$ pour tout $x \in E$.
Pour $x, y \in E$,

$$\|u(x) - u(y)\|_F = \|u(x-y)\|_F \leqslant k\|x-y\|_E$$

donc u est lipschitzienne.
(iv) $\Rightarrow$ (i) : ok
□

Exemple Soit $E = \mathcal{C}([0,1],\mathbb{K})$ et $u : E \to \mathbb{K}$ définie par $u(f) = f(1) - f(0)$.
u est une forme linéaire sur E.
Etudions sa continuité pour $\|\,.\,\|_E = \|\,.\,\|_\infty$ et $\|\,.\,\|_E = \|\,.\,\|_1$.
Cas $\|\,.\,\|_E = \|\,.\,\|_\infty$.
Pour tout $f \in E$, $|u(f)| = |f(1)| + |f(0)| \leqslant 2\|f\|_\infty$ donc u est continue.
Cas $\|\,.\,\|_E = \|\,.\,\|_1$.
Pour $f_n : t \mapsto t^n$,

$$|u(f_n)| = 1 \text{ et } \|f_n\|_1 = \frac{1}{n+1} \to 0$$

Par suite, u n'est pas continue (car discontinue en 0_E)

Exemple Soit $E = \mathcal{C}^\infty([a,b], \mathbb{K})$ normé par $\|\,.\,\|_\infty$.
Considérons l'application $I : E \to E$ déterminée par

$$I(f) \text{ est la primitive de } f \text{ s'annulant en } a$$

Etudions la continuité de l'endomorphisme I de E.
Pour tout $f \in E$, on a

$$I(f)(x) = \int_a^x f(t)\,\mathrm{d}t$$

donc

$$|\varphi(f)(x)| \leqslant \int_a^x |f(t)|\,\mathrm{d}t \leqslant (b-a)\,\|f\|_\infty$$

Ainsi

$$\|I(f)\|_\infty \leqslant (b-a)\,\|f\|_\infty$$

et l'application I est continue.
Considérons inversement l'application D de dérivation.
D est un endomorphisme de E .
Pour $f_n : t \mapsto t^n$, on a

$$\|f_n\|_\infty = 1 \text{ et } \|D(f_n)\|_\infty = n \xrightarrow[n\to+\infty]{} 0$$

L'endomorphisme de dérivation n'est pas continue.

17.3.2 Linéarité en dimension finie

Théorème

Si E est de dimension finie, toute application linéaire de E vers F est continue.

dém. :
Cas $\dim E = 0$: ok.
Cas $\dim E = n \in \mathbb{N}^\star$: on introduit $e = (e_1, \ldots, e_n)$ base de E et on considère $\|\,.\,\|_E = \|\,.\,\|_{\infty,e}$.
Soit $u \in \mathcal{L}(E,F)$. Pour $x = x_1.e_1 + \cdots + x_n.e_n$,

$$u(x) = x_1.u(e_1) + \cdots + x_n.u(e_n)$$

et donc

$$\|u(x)\|_F \leqslant |x_1|\,\|u(e_1)\|_F + \cdots + |x_n|\,\|u(e_n)\|_F \leqslant k\,\|x\|$$

avec

$$k = \|u(e_1)\|_F + \cdots + \|u(e_n)\|_F \in \mathbb{R}^+$$

□

Corollaire

Si E est de dimension finie $\mathcal{L}_c(E,F) = \mathcal{L}(E,F)$.

Exemple L'application Tr : $\mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est continue, l'application de transposition de $\mathcal{M}_{n,p}(\mathbb{K})$ vers $\mathcal{M}_{p,n}(\mathbb{K})$,...

Exemple Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^\star$.
Montrons que l'application $\det_{\mathcal{L}(E)} : \mathcal{L}(E) \to \mathbb{K}$ est continue.
Notons que celle-ci n'est pas linéaire !
Cependant, on sait que $\det_{\mathcal{M}_n(\mathbb{K})} : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est continue. Soit e une base de E, l'application de représentation matricielle

$$M_e : \mathcal{L}(E) \to \mathcal{M}_n(\mathbb{K})$$

est linéaire au départ de $\mathcal{L}(E)$ qui est $\mathbb{K}$-espace vectoriel de dimension finie, c'est donc une application continue. On en déduit que

$$\det_{\mathcal{L}(E)} = \det_{\mathcal{M}_n(\mathbb{K})} \circ M_e$$

est continue par composition d'applications continues.

17.3.3 Continuité des applications multilinéaires

Théorème

Soit $B : E \times F \to G$ une application bilinéaire. On a équivalence entre :
(i) B est continue ;
(ii) B est continue en $(0_E, 0_F)$;
(iii) $\exists k \in \mathbb{R}^+, \forall (x,y) \in E \times F, \|B(x,y)\|_G \leqslant k \|x\|_E \|y\|_F$.

dém. :
(i) $\Rightarrow$ (ii) : ok
(ii) $\Rightarrow$ (iii) : Supposons B continue en $(0_E, 0_F)$.
Pour $\varepsilon = 1$, il existe $\alpha > 0$ vérifiant

$$\forall (x,y) \in E \times F,\ \|(x,y)\|_{E\times F} \leqslant \alpha \Rightarrow \|B(x,y)\|_G \leqslant 1$$

Soit $k = 1/\alpha^2 \in \mathbb{R}^+$. Montrons

$$\forall (x,y) \in E \times F,\ \|B(x,y)\| \leqslant k \|x\| \|y\|$$

Si $x = 0_E$ ou $y = 0_F$: ok
Sinon, on pose $x' = \dfrac{\alpha}{\|x\|} x$ et $y' = \dfrac{\alpha}{\|y\|} y$. On a $\|(x', y')\| = \alpha$ donc $\|B(x', y')\| \leqslant 1$.
Or $\|B(x', y')\| = \dfrac{\alpha^2}{\|x\| \|y\|} \|B(x,y)\|$ donc $\|B(x,y)\| \leqslant \dfrac{1}{\alpha^2} \|x\| \|y\|$.
(iii) $\Rightarrow$ (i) Supposons qu'il existe $k \in \mathbb{R}^+$ tel que $\|B(x,y)\| \leqslant k \|x\| \|y\|$ pour tout $x \in E$ et $y \in F$.
Soit $(x_0, y_0) \in E \times F$.

$$\|B(x,y) - B(x_0,y_0)\| = \|B(x,y) - B(x_0,y)\| + \|B(x_0,y) - B(x_0,y_0)\|$$

donc

$$\|B(x,y) - B(x_0,y_0)\| = \|B(x - x_0, y)\| + \|B(x_0, y - y_0)\| \leqslant k \|x - x_0\| \|y\| + k \|x_0\| \|y - y_0\|$$

Quand $(x,y) \to (x_0,y_0)$, $B(x,y) \to B(x_0,y_0)$ et donc B est continue en (x_0,y_0).
□

Corollaire

Si E et F sont de dimensions finies alors toute application bilinéaire au départ de $E \times F$ est continue.

dém. :
Cas $E = \{0_E\}$ ou $F = \{0_F\}$: ok
Cas $E \neq \{0_E\}$ et $F \neq \{0_F\}$: on introduit $e = (e_1, \dots, e_n)$ une base de E, $f = (f_1, \dots, f_p)$ une base de F et on considère $\|\,.\,\|_E = \|\,.\,\|_{\infty,e}$ et $\|\,.\,\|_F = \|\,.\,\|_{\infty,f}$.
Pour $x = \sum_{i=1}^{n} x_i e_i \in E$ et $y = \sum_{j=1}^{p} y_j f_j \in F$ on a

$$b(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{p} x_i y_j b(e_i, f_j)$$

donc

$$\|b(x, y)\| \leqslant k \|x\| \|y\|$$

avec

$$k = \sum_{i=1}^{n} \sum_{j=1}^{p} \|b(e_i, f_j)\|$$

□

Théorème

Soit $m : E = E_1 \times \cdots \times E_p \to F$ une application multilinéaire. On a équivalence entre :
(i) m est continue ;
(ii) $\exists k \in \mathbb{R}^+, \forall x = (x_1, \dots, x_p) \in E, \|m(x)\|_F \leqslant k \|x_1\|_{E_1} \cdots \|x_p\|_{E_p}$.

dém. :
Même principe qu'au dessus.
□

Corollaire

Les applications multilinéaires au départ d'un produit d'espaces dimensions finies sont continues.

dém. :
Semblable à l'étude relative à la bilinéarité.
□

Exemple Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \in \mathbb{N}^\star$ muni d'une base e.
L'application $\det_e : E^n \to \mathbb{K}$ est continue car multilinéaire au départ d'un espace de dimension finie.

17.4 Connexité par arcs

X désigne une partie de E.

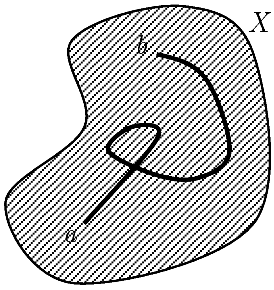
17.4.1 Chemin

Définition

On appelle chemin inscrit dans $X \subset E$ toute application $\gamma : [0, 1] \to E$ continue vérifiant

$$\forall t \in [0, 1], \gamma(t) \in X$$

Les éléments $a = \gamma(0)$ et $b = \gamma(1)$ sont appelés extrémités du chemin.

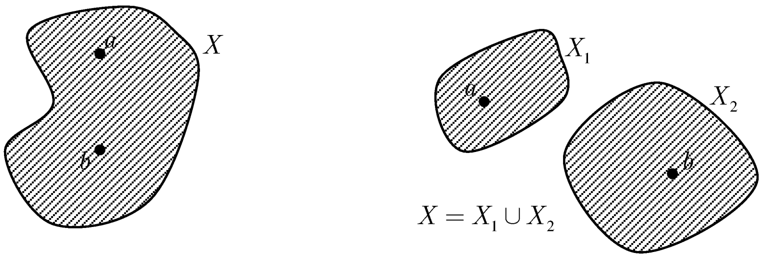


Définition

On dit qu'un élément $a \in X$ peut être relié dans X à un élément $b \in X$ s'il existe un chemin $\gamma : [0,1] \to E$ inscrit dans X vérifiant

$$\gamma(0) = a \text{ et } \gamma(1) = b$$

Exemple



Proposition

Soit $a, b, c \in X$.
a) a peut être relié à lui-même dans X ;
b) si a peut être relié à b dans X, b peut être relié à a dans X ;
c) si a peut être relié à b dans X et si b peut être relié à c dans X alors a peut être relié à c dans X.

dém. :
a) Il suffit de considérer un chemin constant égal à a.
b) Si γ est un chemin inscrit dans X joignant a à b alors $\tilde{\gamma}$ défini par $\tilde{\gamma}(t) = \gamma(1-t)$ détermine un chemin inscrit dans X joignant b à a.
c) Si γ_1 est chemin inscrit dans X joignant a à b et γ_2 joignant b à c alors γ donné par

$$\gamma(t) = \begin{cases} \gamma_1(2t) & \text{si } t \in [0, 1/2] \\ \gamma_2(2t-1) & \text{si } t \in [1/2, 1] \end{cases}$$

détermine un chemin inscrit X joignant a à c.
□

Remarque La relation binaire $\mathcal{R}$ définie sur X par

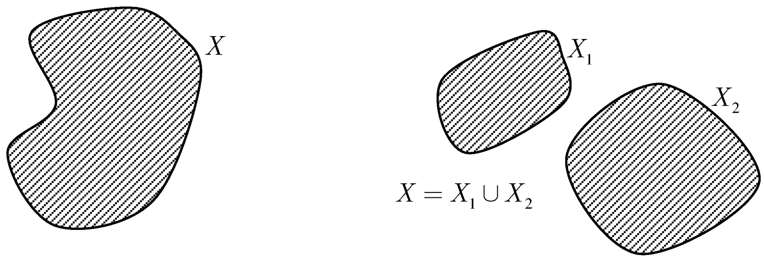
$$a\mathcal{R}b \Leftrightarrow \text{ il existe un chemin inscrit dans } X \text{ joignant } a \text{ à } b$$

définit une relation d'équivalence sur X.
Celle-ci met en relation les éléments qui peuvent être joints et ses classes d'équivalence regroupent ensemble les éléments qui peuvent être joints.

Définition

Les classes d'équivalences de la relation $\mathcal{R}$ sont appelées les composantes connexes par arcs de la partie X.

Exemple



X est la seule composante connexe par arcs

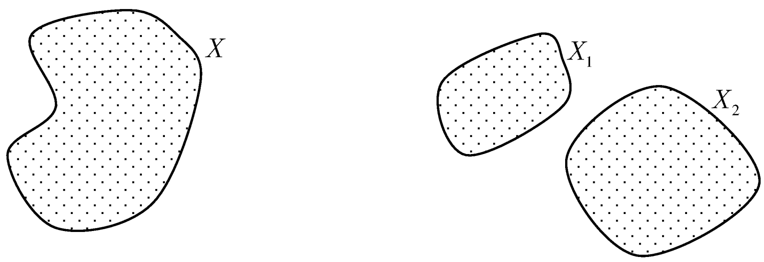
X_1 et X_2 sont les deux composantes connexes par arcs

17.4.2 Parties connexes par arcs

Définition

Une partie X de E est dite connexe par arcs si elle ne possède qu'une seule composante connexe par arcs. Cela signifie encore que pour tout $a, b \in X$, il existe un chemin inscrit dans X d'extrémités a et b.

Exemple



X connexe par arcs

$X_1 \cup X_2$ non connexe par arcs

Exemple Dans $E = \mathbb{R}$, les intervalles sont connexes par arcs.
En revanche, $\mathbb{R}^\star$ n'est pas une partie connexe par arcs.

Proposition

Les parties convexes sont connexes par arcs.

dém. :
Soit X une partie convexe.
Pour tout $a, b \in X$, $[a, b] = \{(1-\lambda)a + \lambda b / \lambda \in [0,1]\} \subset X$.
Considérons alors $\gamma : t \in [0,1] \mapsto \gamma(t) = (1-t)a + tb$.
γ est continue, $\gamma(0) = a$, $\gamma(1) = b$ et $\gamma([0,1]) \subset A$.
□

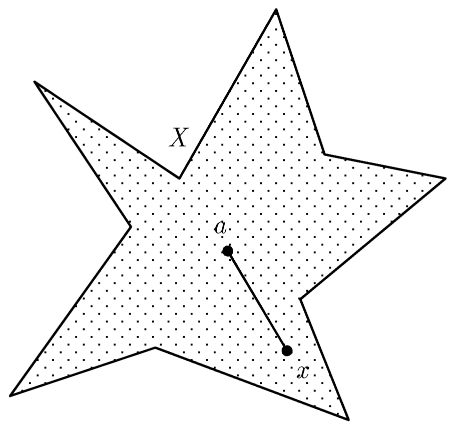
Exemple Les boules, les sous-espaces vectoriels et les sous-espaces affines sont des parties connexes par arcs car convexes.

Définition

Une partie X de E est dite étoilée s'il existe $a \in X$ vérifiant

$$\forall x \in X, [a, x] \subset X$$

Exemple

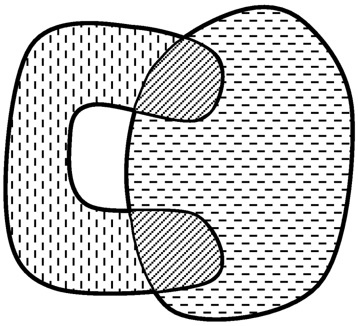


Proposition

Les parties étoilée sont connexes par arcs.

dém. :
Car tout élément de X appartient à la composante connexes par arcs possédant a.
□

Remarque - la réunion de deux connexes par arcs non disjoints est évidemment connexe par arcs ;
- l'intersection de deux connexes par arcs ne l'est pas nécessairement. ;
- le produit cartésien de deux connexes pas arcs est connexe par arcs.



17.4.3 Image continue d'un connexe par arcs

Théorème

L'image directe d'un connexe par arcs par une application continue est connexe par arcs.

dém. :
Soit $f : X \subset E \to F$ continue avec X connexe par arcs.
Pour $a', b' \in f(X)$, il existe $a, b \in X$ tels que $a' = f(a)$ et $b' = f(b)$.
Puisque X est connexe par arcs, il existe $\gamma : [0,1] \to E$ continue telle que $\gamma(0) = a$, $\gamma(1) = b$ et $\gamma([0,1]) \subset X$.
Considérons alors $\gamma' = f \circ \gamma : [0,1] \to F$. γ' est continue, $\gamma'(0) = a'$, $\gamma'(1) = b'$ et $\gamma'([0,1]) = f(\gamma([0,1])) \subset f(X)$.
□

Exemple Le cercle $U = \{z \in \mathbb{C} / |z| = 1\}$ est connexe par arcs.
En effet, c'est l'image du connexe $\mathbb{R}$ par l'application continue $t \mapsto \mathrm{e}^{it}$.

Exemple $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe par arcs.
En effet $\det \mathrm{GL}_n(\mathbb{R}) = \mathbb{R}^\star$ et $\mathbb{R}^\star$ n'est pas connexe par arcs.

17.4.4 Généralisation du théorème des valeurs intermédiaires

Théorème

Les parties connexes par arcs de $\mathbb{R}$ sont ses intervalles.

dém. :
Autrement dit, les parties convexes de $\mathbb{R}$ sont exactement les intervalles de $\mathbb{R}$.
Soit X un intervalle de $\mathbb{R}$, X est convexe donc connexe par arcs.
Inversement, soit X une partie connexe par arcs de $\mathbb{R}$.
Si $X = \emptyset$ alors X est intervalle.

Sinon, pour tout $a \leqslant b \in X$, il existe $\gamma : [0,1] \to \mathbb{R}$ continue telle que $\gamma(0) = a$, $\gamma(1) = b$ et $\gamma([0,1]) \subset X$. Or, par application du théorème des valeurs intermédiaires, la fonction γ prend toutes les comprises entre a et b. Ainsi $[a,b] \subset \gamma([0,1]) \subset X$ et donc

$$\forall a \leqslant b \in X, [a,b] \subset X$$

Posons alors $\alpha = \inf X \in \mathbb{R} \cup \{-\infty\}$ et $\beta = \sup X \in \mathbb{R} \cup \{+\infty\}$.
Pour tout $x \in]\alpha,\beta[$, x n'est ni minorant, ni majorant de X et donc il existe $a, b \in X$ tel que $a < x < b$ et donc $x \in [a,b] \subset X$. Ainsi $]\alpha,\beta[\subset X$ et donc $X =]\alpha,\beta[$, $]\alpha,\beta]$, $[\alpha,\beta[$ ou $[\alpha,\beta]$.
Finalement, X est un intervalle de $\mathbb{R}$.
□

Théorème

Si X est une partie connexe par arcs de E et $f : X \to \mathbb{R}$ une application continue alors $f(X)$ est un intervalle de $\mathbb{R}$.
En conséquence, f prend toute valeur intermédiaire entre deux valeurs déjà prises.

dém. :
$f(X)$ est l'image d'un connexe par arcs par une application continue, c'est donc une partie connexe par arcs de $\mathbb{R}$. Or ces dernières sont des intervalles.
□

Exemple Soit $f : \mathbb{R} \to \mathbb{R}$ continue injective.
Montrons que f est strictement monotone.
Considérons $X = \left\{(x,y) \in \mathbb{R}^2 / x < y\right\}$. X est une partie convexe de $\mathbb{R}^2$ donc connexe par arcs.
La fonction $v : X \to \mathbb{R}$ définie par $v(x,y) = f(y) - f(x)$ est continue et ne s'annule pas en vertu de l'injectivité de f. L'image par v de X est donc un intervalle de $\mathbb{R}$ qui ne contient pas 0. Par suite $v(X) \subset \mathbb{R}^{+\star}$ ou $v(X) \subset \mathbb{R}^{-\star}$ et dans les deux cas f est strictement monotone.

Chapitre 18

Compacité

18.1 Valeur d'adhérence

18.1.1 Suite extraite

Définition

On appelle suite extraite (ou sous-suite) d'une suite $u = (u_n)_{n\in\mathbb{N}}$ d'éléments E toute suite $v = (v_k)_{k\in\mathbb{N}}$ pour laquelle il existe une fonction $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante vérifiant

$$\forall k \in \mathbb{N}, v_k = u_{\varphi(k)}$$

Remarque En posant $n_k = \varphi(k)$, une suite extraire peut se comprendre comme une sélection de termes qui se succèdent

$$(u_{n_k})_{k\in\mathbb{N}} \text{ avec } n_k < n_{k+1}$$

Exemple $(u_{2k})_{k\in\mathbb{N}}$ et $(u_{2k+1})_{k\in\mathbb{N}}$ sont deux suites extraites de $(u_n)_{n\in\mathbb{N}}$.

Proposition

Si w est une suite extraire d'une suite v elle-même extraite d'une suite u alors w est extraite de u.

dém. :
On suppose $(v_k) = (u_{\varphi(k)})$ et $(w_\ell) = (v_{\psi(\ell)})$ avec $\varphi, \psi : \mathbb{N} \to \mathbb{N}$ strictement croissantes.
On a alors $(w_\ell) = (u_{\theta(\ell)})$ avec $\theta = \varphi \circ \psi : \mathbb{N} \to \mathbb{N}$ strictement croissante.
□

Théorème

Si (u_n) converge vers ℓ alors toute suite extraite de (u_n) converge aussi vers ℓ.

dém. :
Soit $(v_k) = (u_{\varphi(k)})$ une suite extraite de (u_n) avec $u_n \to \ell$.
Soit $\varepsilon > 0$. Il existe $N \in \mathbb{N}$ tel que pour tout $n \geqslant N$, $\|u_n - \ell\| \leqslant \varepsilon$.

On montre par une récurrence facile que

$$\forall k \in \mathbb{N}, \varphi(k) \geqslant k$$

Pour $k \geqslant N$, $\varphi(k) \geqslant k \geqslant N$ donc

$$\|v_k - \ell\| = \left\|u_{\varphi(k)} - \ell\right\| \leqslant \varepsilon$$

Ainsi $v_k \to \ell$.
□

18.1.2 Valeur d'adhérence d'une suite

Définition

On appelle valeur d'adhérence d'une suite $u = (u_n)$ d'éléments de E toute limite d'une suite convergente extraite de u. On note Adh(u) l'ensemble des valeurs d'adhérence de la suite u.

Exemple Si $u_n \to \ell$ alors Adh$(u) = \{\ell\}$.

Remarque Une suite possédant au moins deux valeurs d'adhérence (ou n'en possédant aucune) diverge.

Exemple Déterminons les valeurs d'adhérence de $u_n = (-1)^n + \dfrac{1}{n+1}$.
On a $u_{2n} \to 1$ et $u_{2n+1} \to -1$ donc Adh$(u) = \{1, -1\}$.

Exemple Déterminons les valeurs d'adhérence de $u = (u_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ telle que $\|u_n\| \to +\infty$.
Aucune suite extraite de u ne converge car aucune suite extraite de u n'est bornée.
On en déduit Adh$(u) = \emptyset$.

Remarque Les valeurs d'adhérence d'une suite sont les valeurs au voisinage desquelles s'accumule une infinité de termes de la suite.

Théorème

Toute suite bornée d'éléments de $\mathbb{K}$ admet au moins une valeur d'adhérence.

18.2 Partie compacte

18.2.1 Définition

Définition

Une partie K de E est dite compacte si toute suite d'éléments de K possède au moins une valeur d'adhérence dans K i.e.

$$\forall (u_n) \in K^{\mathbb{N}}, \exists \varphi : \mathbb{N} \to \mathbb{N} \text{ strictement croissante}, u_{\varphi(n)} \to \ell \in K$$

On dit encore que K est un compact de E.

Remarque Dans une partie compacte K, on ne peut répartir les éléments d'une suite sans qu'il y ait accumulation au voisinage d'un point de K.

Exemple Sur $E = \mathbb{R}$, les segments $[a, b]$ sont des parties compactes.
En effet, une suite d'éléments de $[a, b]$ est bornée donc admet une suite extraite convergente dont la limite sera élément de $[a, b]$.

Exemple Sur $E = \mathbb{C}$, $\overline{D(0, R)} = \{z \in \mathbb{C}/ |z| \leqslant R\}$ est une partie compacte.
En effet, une suite d'éléments de $\overline{D(0, R)}$ est bornée donc admet une suite extraite convergente dont la limite sera élément de $\overline{D(0, R)}$.

Exemple Sur $E = \mathbb{R}$, $[a, +\infty[$ n'est pas compact.
En effet la suite définie par $u_n = a + n$ n'a pas de valeur d'adhérence.

Exemple Sur $E = \mathbb{R}$, $]a, b]$ n'est pas compact.
En effet, la suite définie $u_n = a + (b - a)/(n + 1)$ n'a qu'une valeur d'adhérence et celle-ci n'est pas élément de $]a, b]$.

18.2.2 Topologie des parties compactes

Théorème

Toute partie compacte est fermée et bornée.

dém. :
Soit K une partie compacte.
Montrons que K est fermée.
Soit $(x_n)_{n\in\mathbb{N}}$ une suite convergente d'éléments de K et posons ℓ sa limite.
Puisque K est compact, $(x_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence dans K, or puisque ℓ est la seule valeur d'adhérence de la suite convergente $(x_n)_{n\in\mathbb{N}}$, on peut conclure que $\ell \in K$. En vertu de la caractérisation séquentielle des parties fermées, on obtient la partie K fermée.
Montrons que K est bornée.
Par l'absurde, supposons K non bornée. Pour tout $n \in \mathbb{N}$, il existe $x_n \in K$ tel que $\|x_n\| > n$. En faisant varier n, cela détermine une suite $(x_n) \in K^{\mathbb{N}}$ telle que $\|x_n\| \to +\infty$. Or cette suite n'a pas de valeur d'adhérence. C'est absurde.
□

Théorème

Toute partie fermée d'une partie compacte est elle-même compacte.

dém. :
Soit F une partie fermée d'un compact K.
Soit (x_n) une suite d'éléments de F. La suite (x_n) apparaît aussi comme une suite d'éléments du compact K, elle admet donc une valeur d'adhérence $\ell \in K$ c'est-à-dire qu'il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $x_{\varphi(n)} \to \ell$. La suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ est une suite convergente d'éléments du fermé F

donc $\ell \in F$.
Finalement, $(x_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence dans F.
□

18.2.3 Opérations sur les parties compactes

Proposition

Une intersection de deux parties compactes est un compact.

dém. :
Car détermine une partie fermée à l'intérieur d'un compact.
□

Proposition

Une réunion de deux parties compactes est un compact

dém. :
Soit K_1 et K_2 deux parties compactes de E et $u = (u_n)_{n\in\mathbb{N}}$ une suite d'éléments de $K_1 \cup K_2$.
Cette suite contient une infinité d'éléments de K_1 (ou de K_2) et possède donc une valeur d'adhérence dans K_1 (ou dans K_2).
□

Théorème

Si K_1 et K_2 sont deux parties compactes d'espaces normés E_1 et E_2 alors $K_1 \times K_2$ est une partie compacte de l'espace normé produit $E_1 \times E_2$.

dém. :
Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de $K_1 \times K_2$.
Pour tout $n \in \mathbb{N}$, on peut écrire $u_n = (x_n, y_n)$ avec $x_n \in K_1$ et $y_n \in K_2$.
La suite (x_n) est une suite d'éléments du compact K_1 donc elle admet une valeur d'adhérence x dans K_1. Ainsi, il existe une extractrice φ telle que $x_{\varphi(n)} \to x$ avec $x \in K_1$.
La suite extraite $(y_{\varphi(n)})$ est une suite d'éléments du compact K_2 donc elle admet une valeur d'adhérence y dans K_2. Ainsi, il existe une extractrice ψ telle que $y_{\varphi(\psi(n))} \to y$ avec $y \in K_2$.
Or, par extraction d'une suite convergente, on a encore $x_{\varphi(\psi(n))} \to x$ et donc $u_{\varphi(\psi(n))} = (x_{\varphi(\psi(n))}, y_{\varphi(\psi(n))}) \to (x, y)$ avec $(x, y) \in K_1, K_2$. Finalement, toute suite d'éléments de $K_1 \times K_2$ admet une valeur d'adhérence dans $K_1 \times K_2$.
□

Corollaire

Si $K_1, \ldots, K_p$ sont des parties compactes d'espaces vectoriels normés $E_1, \ldots, E_p$ alors $K = K_1 \times \cdots \times K_p$ est une partie compacte de l'espace vectoriel normé produit $E = E_1 \times \cdots \times E_p$.

dém. :
Par récurrence via

$$K_1 \times \ldots \times K_p \times K_{p+1} = (K_1 \times \ldots \times K_p) \times K_{p+1}$$

□

18.2.4 Compacité en dimension finie

Théorème

En dimension finie, les parties compactes sont exactement les parties fermées et bornées.

dém. :
Les parties compactes sont assurément de cette forme. Etudions la réciproque.

Soit K une partie fermée bornée d'un espace vectoriel normé E de dimension finie $p \in \mathbb{N}$.
Si $p = 0$ alors $E = \{0_E\}$ et $K = \emptyset$ ou $K = \{0_E\}$. Dans les deux cas K est une partie compacte.
Sinon, on peut introduire une base $e = (e_1, \ldots, e_p)$ de E et considérer la norme $\| \cdot \|_{\infty,e}$.
Soit $u = (u(n))_{n\in\mathbb{N}}$ une suite d'éléments de K.
Notons $u_1, \ldots, u_p$ les suites coordonnées de u.
Considérons $v \in (\mathbb{K}^p)^{\mathbb{N}}$ définie par $v(n) = (u_1(n), \ldots, u_p(n))$
Puisque la partie K est bornée, il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall x \in K, \|x\| \leqslant M$$

En particulier

$$\forall n \in \mathbb{N}, \|u(n)\| \leqslant M$$

et donc

$$\forall 1 \leqslant j \leqslant p, \forall n \in \mathbb{N}, |u_j(n)| \leqslant M$$

La suite v est donc une suite d'éléments du compact $[-M, M]^p$ (si $\mathbb{K} = \mathbb{R}$) ou du compact $\overline{D(0, M)}^p$ (si $\mathbb{K} = \mathbb{C}$). La suite v admet donc une valeur d'adhérence et il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $(v(\varphi(n)))_{n\in\mathbb{N}}$ converge. Les suites coordonnées $(u_i(\varphi(n)))_{n\in\mathbb{N}}$ convergent et finalement $(u(\varphi(n)))_{n\in\mathbb{N}}$ converge.
De plus, $(u(\varphi(n)))_{n\in\mathbb{N}} \in K^{\mathbb{N}}$ et K est fermé donc $(u(\varphi(n)))_{n\in\mathbb{N}}$ converge dans K.
□

Exemple En dimension finie, les boules fermées sont compactes.

Exemple $\mathcal{O}_n(\mathbb{R})$ est une partie compacte de $\mathcal{M}_n(\mathbb{R})$.
En effet $\mathcal{O}_n(\mathbb{R})$ est une partie fermée car

$$\mathcal{O}_n(\mathbb{R}) = f^{-1}(\{I_n\}) \text{ avec } f : A \in \mathcal{M}_n(\mathbb{R}) \mapsto {}^tAA \text{ continue}$$

et $\mathcal{O}_n(\mathbb{R})$ est une partie bornée car

$$\forall A \in \mathcal{O}_n(\mathbb{R}), \forall 1 \leqslant i, j \leqslant n, |a_{i,j}| \leqslant 1$$

Corollaire

> En dimension finie, toute suite bornée admet une valeur d'adhérence.

dém. :
Car une telle suite évolue dans une boule fermée qui est compacte.
□

18.2.5 Applications

18.2.5.1 Convergence d'une suite d'éléments d'un compact

Théorème

> Une suite d'éléments d'une partie compacte converge si, et seulement si, elle admet une unique valeur d'adhérence.

dém. :
($\Rightarrow$) On a déjà vu que l'ensemble des valeurs d'adhérence d'une suite convergente est un singleton.

($\Leftarrow$) Soit $u = (u_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un compact K possédant une unique valeur d'adhérence ℓ. Par l'absurde, supposons que la suite u ne converge pas vers ℓ. Il existe $\varepsilon > 0$ vérifiant

$$\forall N \in \mathbb{N}, \exists n \in \mathbb{N}, n \geqslant N \text{ et } \|u_n - \ell\| > \varepsilon$$

Il existe donc une infinité de termes de la suite u en dehors de $B_f(\ell, \varepsilon)$. On peut ainsi définir une suite extraite $(u_{\varphi(n)})_{n\in\mathbb{N}}$ vérifiant

$$\forall n \in \mathbb{N}, \left\|u_{\varphi(n)} - \ell\right\| > \varepsilon$$

Or celle-ci est une suite d'éléments du compact K et admet donc une valeur d'adhérence $m \in K$. Cette valeur d'adhérence vérifie

$$\|m - \ell\| \geqslant \varepsilon$$

C'est absurde, car la suite u ne possède qu'une seule valeur d'adhérence.

□

Corollaire

> En dimension finie, toute suite bornée admettant une unique valeur d'adhérence converge vers celle-ci.

dém. :

Soit $u = (u_n)_{n\in\mathbb{N}}$ une telle suite. Il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall n \in \mathbb{N}, \|u_n\| \leqslant M$$

La suite u apparaît alors comme étant une suite du compact $B_f(0_E, M)$ et comme elle n'admet qu'une valeur d'adhérence, elle converge vers celle-ci.

□

18.2.5.2 Fermeture des sous-espaces vectoriels de dimension finie

Théorème

> Tout sous-espace vectoriel de dimension finie d'un espace normé est une partie fermée.

dém. :

Soit F sous-espace vectoriel de dimension finie d'un espace normé E.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite convergente d'éléments de F de limite u_∞.

La suite $(u_n)_{n\in\mathbb{N}}$ converge, elle est donc bornée et il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall n \in \mathbb{N}, \|u_n\| \leqslant M$$

La suite $(u_n)_{n\in\mathbb{N}}$ est alors une suite du compact $K = B_f(0_E, M) \cap F$, elle admet une valeur d'adhérence dans K qui ne peut qu'être u_∞. En particulier, $u_\infty \in F$.

Le sous-espace vectoriel F est donc fermé puisqu'il contient les limites de ses suites convergentes.

□

18.2.5.3 Distance à un fermé en dimension finie

Exemple Soit F une partie fermée non vide d'un $\mathbb{K}$-espace vectoriel de dimension finie et x un vecteur de E.

Montrons qu'il existe $y \in F$ tel que $d(x, F) = \|y - x\|$.

Par définition

$$d(x, F) = \inf_{y\in F} \|y - x\|$$

Pour tout $n \in \mathbb{N}$, il existe $y_n \in F$ tel que

$$d(x, F) \leqslant \|y_n - x\| < d(x, F) + \frac{1}{n+1}$$

En faisant varier n, cela définit une suite $(y_n) \in F^{\mathbb{N}}$ telle que $\|y_n - x\| \to d(x, F)$.
Puisque $\|y_n\| \leqslant \|x\| + \|y_n - x\|$, la suite (y_n) est bornée. Il existe donc une suite extraite $(y_{\varphi(n)})$ convergente de limite y.
Puisque $(y_{\varphi(n)})$ est une suite d'éléments du fermé F, on a $y \in F$.
Puisque $y_{\varphi(n)} \to y$ et $\left\|y_{\varphi(n)} - x\right\| \to d(x, F)$ on a aussi $\|y - x\| = d(x, F)$.

18.3 Continuité et compacité

18.3.1 Image continue d'un compact

Théorème

L'image d'une partie compacte par une application continue est une partie compacte

dém. :
Soit $f : K \subset E \to F$ continue avec K partie compacte.
Soit $(y_n) \in f(K)^{\mathbb{N}}$, il existe $(x_n) \in K^{\mathbb{N}}$ telle que $y_n = f(x_n)$.
La suite (x_n) admet une valeur d'adhérence dans K et par continuité son image par f est valeur d'adhérence de (y_n) dans $f(K)$.
□

Exemple Si A et B sont des parties compactes de E alors $A + B$ est un compact de E.
En effet, $A + B$ est l'image du compact $A \times B$ par l'application continue $(x, y) \mapsto x + y$.

Corollaire

Soit $f : K \subset E \to F$.
Si K est une partie compacte et si f est continue alors f est bornée.

dém. :
Une fonction continue sur un compact à une image compacte donc bornée.
□

18.3.2 Théorème des bornes atteintes

Théorème

Toute fonction réelle définie et continue sur un compact non vide admet un minimum et un maximum : on dit qu'elle est bornée et qu'elle atteint ses bornes.

dém. :
Soit $f : K \subset E \to \mathbb{R}$ continue avec K partie compacte non vide de E.
$f(K)$ est un compact non vide de $\mathbb{R}$ donc $m = \inf f(K)$ et $M = \sup f(K)$ existent.
Pour tout $n \in \mathbb{N}$, $M - 1/(n+1) < M$ donc il existe $x_n \in K$ tel que

$$M - \frac{1}{n+1} < f(x_n) \leqslant M$$

En faisant varier n, cela détermine une suite $(x_n) \in K^{\mathbb{N}}$ telle que $f(x_n) \to M$.
Puisque la partie K est compacte, il existe φ extractrice telle que $x_{\varphi(n)} \to a \in K$.
Par continuité de f en a, on a $f(x_{\varphi(n)}) \to f(a)$ et par extraction $f(x_{\varphi(n)}) \to M$ donc $M = f(a)$.
□

Exemple Soit K un compact non vide et $x \in E$.
On pose

$$d(x, K) = \inf_{y \in K} \|y - x\|$$

Montrons qu'il existe $y_0 \in K$ tel que $d(x, K) = \|y_0 - x\|$.
La fonction $y \mapsto \|y - x\|$ est continue sur le compact K, elle y admet donc un minimum et par conséquent, il existe $y_0 \in K$ tel que

$$\inf_{y \in K} \|y - x\| = \min_{y \in K} \|y - x\| = \|y_0 - x\|$$

18.3.3 Uniforme continuité

Définition

Une application $f : X \subset E \to F$ est dite uniformément continue si

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x, y \in X, \|y - x\|_E \leqslant \alpha \Rightarrow \|f(y) - f(x)\|_F \leqslant \varepsilon$$

Remarque $f : X \subset E \to F$ continue signifie

$$\forall x \in X, \forall \varepsilon > 0, \exists \alpha > 0, \forall y \in X, \|y - x\| \leqslant \alpha \Rightarrow \|f(y) - f(x)\| \leqslant \varepsilon$$

Pour l'uniforme continuité, on exige que le paramètre α soit indépendant de x.

Proposition

Toute fonction uniformément continue est continue.

dém. :
Qui peut le plus, peut le moins.
□

Proposition

Toute fonction lipschitzienne est uniformément continue.

dém. :
Supposons $f : X \subset E \to F$ lipschitzienne. Il existe $k \in \mathbb{R}^+$ tel que

$$\forall x, y \in X, \|f(y) - f(x)\| \leqslant k \|y - x\|$$

Sans perte de généralité, on peut suppose $k > 0$.
Soit $\varepsilon > 0$. Pour $\alpha = \varepsilon / k > 0$, on a

$$\forall x, y \in X, \|y - x\| \leqslant \alpha \Rightarrow \|f(y) - f(x)\| \leqslant \varepsilon$$

□

18.3.4 Théorème de Heine

Théorème

Soit $f : K \subset E \to F$.
Si K est une partie compacte et si f est continue alors f est uniformément continue.

dém. :
Par l'absurde, supposons que f non uniformément continue.
Il existe $\varepsilon > 0$ tel que

$$\forall \alpha > 0, \exists x, y \in X, \|y - x\| \leqslant \alpha \text{ et } \|f(y) - f(x)\| > \varepsilon$$

Soit $n \in \mathbb{N}$. Pour $\alpha = \dfrac{1}{n+1} > 0$, il existe $x_n, y_n \in K$ vérifiant

$$\|y_n - x_n\| \leqslant \frac{1}{n+1} \text{ et } \|f(y_n) - f(x_n)\| > \varepsilon$$

En faisant varier n, cela détermine deux suites (x_n) et (y_n) d'éléments de K telles que $\|y_n - x_n\| \to 0$ et $\|f(y_n) - f(x_n)\| > \varepsilon$. Puisque la suite (x_n) évolue dans le compact K, il existe une extractrice φ telle que $x_{\varphi(n)} \to x$ avec $x \in K$. Puisque $\left\|y_{\varphi(n)} - x_{\varphi(n)}\right\| \to 0$, on a aussi $y_{\varphi(n)} \to x$. Or f est continue donc $f(x_n) \to f(x)$ et $f(y_n) \to f(x)$. En passant à la limite la relation $\|f(y_n) - f(x_n)\| > \varepsilon$, on obtient alors une absurdité.
□

Corollaire

Toute fonction continue de $[a, b]$ vers F est uniformément continue.

dém. :
Car $[a, b]$ est une partie compacte.
□

18.3.5 Musculation

Exemple Soit $f : [0, +\infty[\to \mathbb{R}$ continue. On suppose que $f \xrightarrow[+\infty]{} \ell$, montrons que f est uniformément continue.Soit $\varepsilon > 0$. Il existe $A \in \mathbb{R}^+$ tel que

$$\forall x \geqslant A, |f(x) - \ell| \leqslant \varepsilon/2$$

et alors

$$\forall x, y \in [A, +\infty[\,, |f(y) - f(x)| \leqslant \varepsilon \text{ (*)}$$

De plus, f est continue sur $[0, A]$ donc uniformément continue et il existe $\alpha > 0$ tel que

$$\forall x, y \in [0, A]\,, |y - x| \leqslant \alpha \Rightarrow |f(y) - f(x)| \leqslant \varepsilon\text{(**)}$$

Soit $x, y \in \mathbb{R}^+$ avec $|y - x| \leqslant \alpha$. On peut supposer $x \leqslant y$.
Si $x, y \in [0, A]$, on a $|f(y) - f(x)| \leqslant \varepsilon$ en vertu de (**)
Si $x, y \in [A, +\infty[$, on a à nouveau $|f(y) - f(x)| \leqslant \varepsilon$ cette fois-ci en vertu de (*).
Si $x \in [0, A]$ et $y \in [A, +\infty[$, on a nécessairement $|x - A| \leqslant \alpha$. (*) et (**) donnent alors

$$|f(x) - f(y)| \leqslant |f(x) - f(A)| + |f(A) - f(y)| \leqslant 2\varepsilon$$

Quitte à adapter le ε de départ, on obtient ce que l'on veut.

Chapitre 19

Dérivation et intégration d'une fonction vectorielle

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$ et $E, F, G, \ldots$ désignent des $\mathbb{K}$-espaces vectoriels de dimensions finies.
I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide.
On étudie ici des fonctions d'une variable réelle à valeurs dans un espace de dimension finie

$$t \mapsto z(t) \in \mathbb{C}, t \mapsto (x(t), y(t), \ldots) \in \mathbb{R}^n, t \mapsto \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}), \ldots$$

19.1 Dérivation

19.1.1 Vecteur dérivé

Définition

On dit que $f : I \to E$ est dérivable en $a \in I$ si le taux d'accroissement

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

converge quand $h \to 0$ (avec $h \neq 0$)
Sa limite est alors appelée vecteur dérivé de f en a, on la note $f'(a)$.

Théorème

Soit $f : I \to E$ et a élément de I. On a équivalence entre :
(i) $f : I \to E$ est dérivable en a ;
(ii) il existe $\ell \in E$ tel que

$$f(t) = f(a) + (t-a).\ell + (t-a)\varepsilon(t) \text{ avec } \varepsilon(t) \xrightarrow[t \to a]{} 0_E$$

De plus, on a alors $\ell = f'(a)$.
L'égalité asymptotique écrite dans (ii) s'appelle un développement limité à l'ordre 1 de f en a.

dém. :

(i) $\Rightarrow$ (ii) Si f est dérivable en a on peut écrire, pour $t \neq a$

$$\frac{1}{t-a}\left(f(t)-f(a)\right) = f'(a) + \varepsilon(t)$$

Avec $\varepsilon(t) \xrightarrow[t \to a]{} 0$. On alors

$$f(t) - f(a) = (t-a).f'(a) + (t-a)\varepsilon(t)$$

et cette relation vaut aussi pour $t = a$ en posant $\varepsilon(a) = 0_E$. On obtient donc

$$f(t) - f(a) \underset{t \to a}{=} (t-a).f'(a) + o(t-a)$$

(ii) $\Rightarrow$ (i) Si $f(t) \underset{t \to a}{=} f(a) + (t-a).\ell + (t-a)\varepsilon(t)$ avec $\varepsilon(t) \xrightarrow[t \to a]{} 0_E$ alors

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = \frac{1}{h}\left(h.\ell + h\varepsilon(a+h)\right) \xrightarrow[h \to 0]{} \ell$$

□

Remarque On écrit alors

$$f(t) \underset{t \to a}{=} f(a) + (t-a).\ell + o\left((t-a)\right)$$

en introduisant le concept de fonction négligeable comme cela a été fait pour les fonctions réelles ou complexes.

Corollaire

Si f est dérivable en a alors f est aussi continue en a.

Remarque Si $t \mapsto f(t)$ est le paramétrage d'un mobile alors $f'(a)$ est le vecteur vitesse du mobile à l'instant $t = a$.

19.1.2 Dérivabilité à droite et à gauche

Définition

Soit $f : I \to E$ et $a \in I$ qui n'est pas extrémité droite de I. On dit que f est dérivable à droite en a si le taux d'accroissement

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

converge quand $h \to 0^+$. Sa limite est appelée vecteur dérivé à droite de f en a. On le note $f'_d(a)$.
De façon analogue, on définit $f'_g(a)$ vecteur dérivé à gauche de f en a.

Proposition

Soit $f : I \to E$ et a élément intérieur à I. On a équivalence entre :
(i) f est dérivable en a ;
(ii) f est dérivable à droite et à gauche en a avec $f'_d(a) = f'_g(a)$.

19.1.3 Fonction dérivable

Définition

Une fonction $f : I \to E$ est dite dérivable si elle l'est en tout point de I.
On peut alors introduire l'application

$$f' : \begin{cases} I \to E \\ t \in I \mapsto f'(t) \end{cases}$$

appelée fonction dérivée de f.

Proposition

Les fonctions dérivables de I vers E sont continues.

dém. :
Si $f : I \to E$ est dérivable alors f est continue en tout $a \in I$.
□

Théorème

Soit $f : I \to E$ de fonctions coordonnées $f_1, \ldots, f_p$ dans une base $e = (e_1, \ldots, e_p)$ de E.
On a équivalence entre :
(i) f est dérivable ;
(ii) $f_1, \ldots, f_p$ sont dérivables.
De plus, si tel est le cas

$$\forall t \in I, f'(t) = \sum_{j=1}^{p} f_j'(t).e_j$$

dém. :
On a

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = \sum_{j=1}^{p} \frac{1}{h}\left(f_j(a+h) - f_j(a)\right).e_j$$

La convergence de la fonction vectorielle en premier membre équivaut à la convergence des fonctions coordonnées mises en exergue dans le second membre.
□

Exemple $z : I \to \mathbb{C}$ est dérivable si, et seulement si, Re(z) et Im(z) le sont. On a alors

$$z'(t) = (\text{Re}z)'(t) + i(\text{Im}z)'(t)$$

Exemple $x : I \to \mathbb{R}^p$ définie par $x(t) = (x_1(t), \ldots, x_p(t))$ est dérivable si, et seulement si, $x_1, \ldots, x_p$ le sont. On a alors

$$x'(t) = (x_1'(t), \ldots, x_p'(t))$$

Exemple $A : I \to \mathcal{M}_{n,p}(\mathbb{K})$ est dérivable si, et seulement si, les fonctions coefficients $t \mapsto a_{i,j}(t)$ le sont. On a alors

$$A'(t) = \begin{pmatrix} a_{1,1}'(t) & \cdots & a_{1,p}'(t) \\ \vdots & & \vdots \\ a_{n,1}'(t) & \cdots & a_{n,p}'(t) \end{pmatrix}$$

19.1.4 Opérations sur les fonctions dérivables

Théorème

Soit $f, g : I \to E$ et $\lambda \in \mathbb{K}$.
Si f et g sont dérivables alors λf et $f + g$ le sont aussi avec

$$(\lambda f)' = \lambda f', (f+g)' = f' + g'$$

dém. :
Par opérations sur les limites ou par les fonctions coordonnées dans une base de E.
□

Corollaire

L'ensemble $\mathcal{D}(I, E)$ des fonctions de I vers E dérivables est un sous-espace vectoriel de $\mathcal{F}(I, E)$ et l'application $f \mapsto f'$ y est linéaire.

Théorème

Soit $\varphi : J \to I$ et $f : I \to E$.
Si f et φ sont dérivables alors $f \circ \varphi$ l'est aussi

$$(f \circ \varphi)' = \varphi'.f' \circ \varphi$$

dém. :
Immédiat par les fonctions coordonnées dans une base de E.
□

Théorème

Soit $f : I \to E$ et $L \in \mathcal{L}(E, F)$.
Si f est dérivable alors $L(f) : t \mapsto L(f(t))$ est dérivable et

$$[L(f)]' = L(f')$$

dém. :
Soit $a \in I$. Pour $h \neq 0$

$$\frac{1}{h}\left(L(f(a+h)) - L(f(a))\right) = L\left(\frac{1}{h}\left(f(a+h) - f(a)\right)\right) \xrightarrow[h \to 0]{} L(f'(a))$$

car L est continue puisque linéaire au départ d'un espace vectoriel de dimension finie.
□

Attention : Ici écrire la formule $(L(f))' = f' \times L'(f)$ n'a pas de sens car L' n'en a pas.

Exemple Si $A : I \to \mathcal{M}_n(\mathbb{K})$ est dérivable alors $t \mapsto \text{tr}(A(t))$ est dérivable et

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\text{tr}(A(t))\right) = \text{tr}(A'(t))$$

Théorème

Soit $f : I \to E$, $g : I \to F$ et $B : E \times F \to G$ bilinéaire.
Si f et g sont dérivables alors $B(f,g) : t \mapsto B(f(t), g(t))$ est dérivable et

$$B(f,g)' = B(f',g) + B(f,g')$$

dém. :
Soit $a \in I$. Pour $h \neq 0$, on peut écrire

$$\begin{aligned}&\frac{1}{h}\left(B\left(f(a+h), g(a+h)\right) - B\left(f(a), g(a)\right)\right)\\ &= B\left(\frac{1}{h}\left(f(a+h) - f(a)\right), g(a+h)\right) + B\left(f(a), \frac{1}{h}\left(g(a+h) - g(a)\right)\right)\end{aligned}$$

Par continuité de l'application bilinéaire B,

$$\frac{1}{h}\left(B\left(f(a+h), g(a+h)\right) - B\left(f(a), g(a)\right)\right) \xrightarrow[h\to 0]{} B\left(f'(a), g(a)\right) + B\left(f(a), g'(a)\right)$$

□

Corollaire

Si $\alpha : I \to \mathbb{K}$ et $f : I \to E$ sont dérivables alors $\alpha.f$ aussi et

$$(\alpha.f)' = \alpha'.f + \alpha.f'$$

dém. :
L'application produit extérieur . : $\mathbb{K} \times E \to E$ est bilinéaire.
□

Corollaire

On suppose que E est une algèbre.
Si $f, g : I \to E$ sont dérivables alors fg l'est aussi

$$(fg)' = f'g + fg'$$

En particulier, $\mathcal{D}(I,E)$ est une sous-algèbre de $\mathcal{F}(I,E)$.

dém. :
L'application produit $E \times E \to E$ est bilinéaire.
□

Corollaire

On suppose E euclidien de produit scalaire $(. \mid .)$.
Si f, g sont dérivables alors $(f \mid g) : t \mapsto (f(t) \mid g(t))$ est dérivable et

$$(f \mid g)' = (f' \mid g) + (f \mid g')$$

dém. :
$(. \mid .)$ est une application bilinéaire.
□

Théorème

Soit $f_1 : I \to E_1, \dots, f_p : I \to E_p$ et $m : E_1 \times E_2 \times \dots \times E_p \to F$ multilinéaire.
Si $f_1, \dots, f_p$ sont dérivables alors $m(f_1, \dots, f_p) : t \mapsto m(f_1(t), \dots, f_p(t))$ est dérivable et

$$m(f_1, \dots, f_p)' = \sum_{j=1}^{p} m(f_1, \dots, f_j', \dots, f_p)$$

Exemple Si $u, v, w : I \to \mathbb{R}$ sont dérivables alors uvw aussi et

$$(uvw)' = u'vw + uv'w + uvw'$$

Plus généralement, on a pour $f_1, \dots, f_p : I \to \mathbb{R}$ dérivables, la relation

$$(f_1 \dots f_p)' = \sum_{i=1}^{p} f_1 \dots (f_i)' \dots f_p$$

Exemple Soit $A : t \mapsto A(t)$ une fonction dérivable de I vers $\mathcal{M}_n(\mathbb{K})$.
La fonction $t \mapsto \det A(t)$ est dérivable car

$$\det A(t) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}(t)$$

Exprimons la dérivée de $t \mapsto \det A(t)$.
Notons $C_1(t), \dots, C_n(t)$ les colonnes de $A(t)$ et $E = (E_1, \dots, E_n)$ la base canonique de $\mathcal{M}_{n,1}(\mathbb{K})$.
Les fonctions $C_1, \dots, C_n$ sont dérivables et puisque

$$\det(A(t)) = \det_E (C_1(t), \dots, C_n(t))$$

avec $\det_E$ application multilinéaire, on a

$$\frac{\mathrm{d}}{\mathrm{d}t} (\det A(t)) = \sum_{i=1}^{n} \det_E (C_1(t), \dots, C_i'(t), \dots, C_n(t))$$

Exemple $\dfrac{\mathrm{d}}{\mathrm{d}t} \begin{vmatrix} a(t) & b(t) \\ c(t) & d(t) \end{vmatrix} = \begin{vmatrix} a'(t) & b(t) \\ c'(t) & d(t) \end{vmatrix} + \begin{vmatrix} a(t) & b'(t) \\ c(t) & d'(t) \end{vmatrix}$

Remarque On pourrait aussi raisonner par ligne plutôt que par colonne.

19.1.5 Dérivées d'ordres supérieurs

Définition

Soit $f : I \to E$.
On pose $f^{(0)} = f$ appelée dérivée d'ordre 0 de f.
Pour $n \in \mathbb{N}$, si $f^{(n)}$ existe et est dérivable, on pose $f^{(n+1)} = \left(f^{(n)}\right)'$ appelée dérivée d'ordre $n+1$ de f.
On dit que $f : I \to E$ est n fois dérivable si $f^{(n)}$ existe.

Théorème

Soit $f : I \to E$ de fonctions coordonnées $f_1, \dots, f_p$ dans une base $e = (e_1, \dots, e_p)$ de E.
On a équivalence entre :
(i) f est n fois dérivable ;
(ii) $f_1, \dots, f_p$ sont n fois dérivables.
De plus, si tel est le cas :

$$\forall t \in I, f^{(n)}(t) = f_1^{(n)}(t).e_1 + \cdots + f_p^{(n)}(t).e_p$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
□

Théorème

Soit $f, g : I \to E$ et $\lambda \in \mathbb{K}$.
Si f et g sont n fois dérivables alors λf et $f + g$ le sont aussi et

$$(\lambda f)^{(n)} = \lambda f^{(n)} \text{ et } (f+g)^{(n)} = f^{(n)} + g^{(n)}$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
□

Corollaire

L'ensemble $\mathcal{D}^n(I, E)$ des fonctions n fois dérivables de I vers E est un sous-espace vectoriel de $\mathcal{F}(I, E)$.

Théorème

Soit $f : I \to E$ et $L \in \mathcal{L}(E, F)$.
Si f est n fois dérivable alors $L(f)$ aussi et

$$(L(f))^{(n)} = L(f^{(n)})$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
□

Théorème

Soit $B : E \times F \to G$ une application bilinéaire.
Si $f : I \to E$ et $g : I \to F$ sont n fois dérivables alors $B(f, g)$ l'est aussi et

$$B(f,g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} B\left(f^{(n-k)}, g^{(k)}\right)$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
Pour $n = 0$: ok.
Supposons la propriété vraie au rang $n \geqslant 0$.
Soit $f : I \to E$ et $g : I \to F$ des fonctions $n + 1$ fois dérivables.
Par hypothèse de récurrence $B(f, g)$ est n fois dérivable et

$$B(f,g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} B\left(f^{(n-k)}, g^{(k)}\right)$$

Or pour tout $k \in \{0, \ldots, n\}$, $f^{(n-k)}$ et $g^{(k)}$ sont dérivables donc $B\left(f^{(n-k)}, g^{(k)}\right)$ aussi.
Par suite, $B(f, g)$ est $n + 1$ fois dérivable et

$$B(f,g)^{(n+1)} = \sum_{k=0}^{n} \binom{n}{k} \left[B\left(f^{(n+1-k)}, g^{(k)}\right) + B\left(f^{(n-k)}, g^{(k+1)}\right)\right]$$

En séparant les deux sommes et par décalage d'indice

$$B(f,g)^{(n+1)} = \sum_{k=0}^{n} \binom{n}{k} B\left(f^{(n+1-k)}, g^{(k)}\right) + \sum_{k=1}^{n+1} \binom{n}{k-1} B\left(f^{(n+1-k)}, g^{(k)}\right)$$

En adjoignant des termes nuls à chaque somme

$$B(f,g)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n}{k} B\left(f^{(n+1-k)}, g^{(k)}\right) + \sum_{k=0}^{n+1} \binom{n}{k-1} B\left(f^{(n+1-k)}, g^{(k)}\right)$$

En réunissant les deux sommes et par la formule du triangle de Pascal

$$B(f,g)^{(n+1)} = \sum_{k=0}^{n+1} \binom{n+1}{k} B\left(f^{(n+1-k)}, g^{(k)}\right)$$

Récurrence établie.
□

Corollaire

Si $\alpha : I \to \mathbb{K}$ et $f : I \to E$ sont n fois dérivables alors $\alpha.f$ aussi.

Corollaire

On suppose que E est une algèbre.
Si $f, g : I \to E$ sont n fois dérivables alors fg aussi.
En particulier, $\mathcal{D}^n(I, E)$ est une sous-algèbre de $\mathcal{C}(I, E)$

Corollaire

Soit E un espace euclidien
Si $f, g : I \to E$ sont n fois dérivables alors $(f \mid g)$ aussi.

Exemple Soit $f : I \to E$ une fonction $n+1$ fois dérivable.

$$(t.f(t))^{(n+1)} = t.f^{(n+1)}(t) + (n+1)f^{(n)}(t)$$

19.1.6 Classe d'une fonction

Définition

Une fonction $f : I \to E$ est dite de classe $\mathcal{C}^n$ si f est n fois dérivable et si $f^{(n)}$ est continue.
Une fonction $f : I \to E$ est dite de classe $\mathcal{C}^\infty$ si elle est de classe $\mathcal{C}^n$ pour tout $n \in \mathbb{N}$.

Les théorèmes présentés ci-dessus se transposent aux fonctions de classe $\mathcal{C}^n$ avec $n \in \mathbb{N} \cup \{\infty\}$. On en déduit :

Proposition

Pour $n \in \mathbb{N} \cup \{\infty\}$, $f : I \to E$ est de classe $\mathcal{C}^n$ si, et seulement si, ses fonctions coordonnées dans une base de E le sont.

Théorème

Pour $n \in \mathbb{N} \cup \{\infty\}$, l'ensemble $\mathcal{C}^n(I, E)$ des fonctions de classe $\mathcal{C}^n$ de I vers E est un sous-espace vectoriel (voire une sous-algèbre) de $\mathcal{F}(I, E)$.

19.2 Intégration sur un segment

19.2.1 Fonctions continues par morceaux

Soit $e = (e_1, \ldots, e_p)$ une base de l'espace E.

Définition

Une fonction $f : I \to E$ est dite continue par morceaux si ses fonctions coordonnées dans la base e le sont.

Proposition

La notion ne dépend pas du choix de la base e de E.

dém. :
Si $\tilde{e} = (\tilde{e}_1, \ldots, \tilde{e}_p)$ désigne une autre base de E et si P est la matrice de passage de e à $\tilde{e}$, la formule de changement de base

$$X = P\tilde{X} \text{ et } \tilde{X} = P^{-1}X$$

montre que les fonctions coordonnées $\tilde{f}_1, \ldots, \tilde{f}_p$ de f dans $\tilde{e}$ sont combinaisons linéaires des fonctions coordonnées de f dans e. Si ces dernières sont continues par morceaux, ces premières aussi.
□

Théorème

L'ensemble $\mathcal{C}^0_{pm}(I, E)$ des fonctions continues par morceaux de I dans E est un sous-espace vectoriel de l'espace $\mathcal{F}(I, E)$.

dém. :
Par opérations sur les fonctions coordonnées.
□

19.2.2 Intégration entre deux bornes

Soit $e = (e_1, \dots, e_p)$ une base de l'espace E.

Définition

Soit $f : I \to E$ une fonction continue par morceaux de fonctions coordonnées $f_1, \dots, f_p$ dans la base e.
Pour tout $a, b \in I$, on appelle intégrale de f de a à b le vecteur

$$\int_a^b f(t)\,\mathrm{d}t \underset{\text{déf}}{=} \sum_{j=1}^{p} \int_a^b f_j(t)\,\mathrm{d}t.e_j$$

Cette intégrale peut aussi être notée $\displaystyle\int_a^b f$ ou $\displaystyle\int_{[a,b]} f$ lorsque $a \leqslant b$.

Proposition

La valeur de l'intégrale ici définie ne dépend pas du choix de la base e de E.

dém. :
Considérons $\tilde{e} = (\tilde{e}_1, \dots, \tilde{e}_p)$ une autre base de E et introduisons $P = (p_{i,j})$ la matrice de passage de e à $\tilde{e}$. On a

$$\tilde{e}_j = \sum_{i=1}^{p} p_{i,j} e_i \text{ et } f(t) = \sum_{j=1}^{p} f_j(t).e_j = \sum_{j=1}^{p} \tilde{f}_j(t).\tilde{e}_j$$

et donc

$$\sum_{j=1}^{p} \int_a^b \tilde{f}_j(t)\,\mathrm{d}t.\tilde{e}_j = \sum_{j=1}^{p} \sum_{i=1}^{p} p_{i,j} \int_a^b \tilde{f}_j(t)\,\mathrm{d}t.e_i = \sum_{i=1}^{p} \int_a^b \sum_{i=1}^{p} p_{i,j} \tilde{f}_j(t)\,\mathrm{d}t.e_i = \sum_{i=1}^{p} \int_a^b f_i(t)\,\mathrm{d}t.e_i$$

□

19.2.3 Opérations

Théorème

Soit $f, g : I \to E$ continues par morceaux, $\lambda, \mu \in \mathbb{K}$ et $a, b \in I$

$$\int_a^b \lambda f + \mu g = \lambda \int_a^b f + \mu \int_a^b g$$

dém. :
Via les fonctions coordonnées dans une base de E.
□

Théorème

Soit $f : I \to E$ continue par morceaux.

$$\forall a, b, c \in I, \int_a^b f = \int_a^c f + \int_c^b f$$

dém. :
Via les fonctions coordonnées dans une base de E.
□

19.2.4 Sommes de Riemann

Théorème

Si $f : [a, b] \to E$ est continue par morceaux alors

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) \xrightarrow[n \to +\infty]{} \int_a^b f(t)\,\mathrm{d}t$$

dém. :
Via les fonctions coordonnées dans une base de E.
□

Remarque On a aussi

$$\frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\frac{b-a}{n}\right) \xrightarrow[n \to +\infty]{} \int_a^b f(t)\,\mathrm{d}t$$

Corollaire

En particulier, pour $f : [0, 1] \to E$ continue par morceaux

$$\frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \text{ et } \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right) \text{ tendent vers } \int_0^1 f(t)\,\mathrm{d}t$$

19.2.5 Inégalité triangulaire

Théorème

Soit $f : [a, b] \to E$ continue par morceaux et $\| \,.\, \|$ une norme sur E.

$$\left\| \int_{[a,b]} f \right\| \leqslant \int_{[a,b]} \|f\|$$

dém. :
D'une part

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k\frac{b-a}{n}\right) \xrightarrow[n \to +\infty]{} \int_a^b f(t)\,\mathrm{d}t$$

et d'autre part
$$\frac{b-a}{n}\sum_{k=0}^{n-1}\left\|f\left(a+k\frac{b-a}{n}\right)\right\| \xrightarrow[n\to+\infty]{} \int_a^b \|f(t)\|\,\mathrm{d}t$$
Or par inégalité triangulaire
$$\left\|\frac{b-a}{n}\sum_{k=0}^{n-1}f\left(a+k\frac{b-a}{n}\right)\right\| \leqslant \frac{b-a}{n}\sum_{k=0}^{n-1}\left\|f\left(a+k\frac{b-a}{n}\right)\right\|$$
On conclut par passage à la limite.
□

19.3 Intégrales et primitives

19.3.1 Primitive

Définition

On appelle primitive de $f : I \to E$, s'il en existe, toute fonction $F : I \to E$ dérivable vérifiant $F' = f$.

Remarque Les primitives de f peuvent se calculer à partir des fonctions coordonnées de f.

Théorème

Si $f : I \to E$ admet des primitives, celles-ci se déduisent les unes des autres par addition d'une constante vectorielle.

dém. :
Si F est primitive de f alors $F + C$ aussi car $(F + C)' = F' = f$.
Si F et G sont deux primitives de f alors $(F - G)' = 0$ et donc $F - G$ est constante (car ses fonctions coordonnées le sont).
□

19.3.2 Intégrale fonction de sa borne supérieure

Théorème

Soit $f : I \to E$ et $a \in I$. Si f est continue alors f possède une unique primitive s'annulant en a, c'est la fonction
$$F : x \mapsto \int_a^x f(t)\,\mathrm{d}t$$

dém. :
La fonction $F : x \mapsto \displaystyle\int_a^x f(t)\,\mathrm{d}t$ est définie de I vers E et s'annule en a.
Soit $x \in I$. Montrons
$$\frac{1}{h}\left(F(x+h) - F(x)\right) \xrightarrow[h\to 0]{} f(x)$$
Soit $h > 0$.
$$\left\|\frac{1}{h}\left(F(x+h) - F(x)\right) - f(x)\right\| = \frac{1}{h}\left\|\int_x^{x+h} f(t) - f(x)\,\mathrm{d}t\right\| \leqslant \frac{1}{h}\int_x^{x+h}\|f(t) - f(x)\|\,\mathrm{d}t$$

Puisque f est continue en x, pour $\varepsilon > 0$, il existe $\alpha > 0$ tel que

$$\forall t \in I, |t - x| \leqslant \alpha \Rightarrow \|f(t) - f(x)\| \leqslant \varepsilon$$

et alors

$$0 < h \leqslant \alpha \Rightarrow t \in [x, x+h], \|f(t) - f(x)\| \leqslant \varepsilon$$

et donc

$$0 < h \leqslant \alpha \Rightarrow \left\|\frac{1}{h}\left(F(x+h) - F(x)\right) - f(x)\right\| \leqslant \varepsilon$$

Ainsi

$$\frac{1}{h}\left(F(x+h) - F(x)\right) \xrightarrow[h \to 0^+]{} f(x)$$

De même on montre

$$\frac{1}{h}\left(F(x+h) - F(x)\right) \xrightarrow[h \to 0^-]{} f(x)$$

□

Remarque On retient la formule

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x)$$

Corollaire

Si $f : I \to E$ est continue de primitive F alors

$$\forall a, b \in I, \int_a^b f(t)\,\mathrm{d}t = [F(t)]_a^b$$

dém. :
Pour tout $x \in I$, on a

$$\int_a^x f(t)\,\mathrm{d}t = F(x) - F(a)$$

car $x \mapsto \displaystyle\int_a^x f(t)\,\mathrm{d}t$ et F sont primitives de f. En particularisant en $x = b$, on obtient la relation voulue.
□

19.3.3 Changement de variable et intégration par parties

Théorème

Soit $\varphi : I \to J$ de classe $\mathcal{C}^1$ et $f : J \to E$ continue.

$$\forall a, b \in I,\ \int_a^b \varphi'(t).f(\varphi(t))\,\mathrm{d}t = \int_{\varphi(a)}^{\varphi(b)} f(s)\,\mathrm{d}s$$

La manipulation consistant à transformer une intégrale en l'autre est appelée changement de variable définie par la relation $s = \varphi(t)$.

dém. :
Soit F une primitive de la fonction continue f.

$$\int_{\varphi(a)}^{\varphi(b)} f(s)\,\mathrm{d}s = [F(s)]_{\varphi(a)}^{\varphi(b)}$$

On vérifie par les fonctions coordonnées que $F \circ \varphi$ est primitive de la fonction continue $\varphi'.f \circ \varphi$ et donc

$$\int_a^b f(\varphi(t))\varphi'(t)\,\mathrm{d}t = [F(\varphi(t))]_a^b$$

□

Théorème

Soit $B : E \times F \to G$ bilinéaire, $u : I \to E$ et $v : I \to F$ de classe $\mathcal{C}^1$.

$$\forall a, b \in I,\ \int_a^b B(u', v) = [B(u, v)]_a^b - \int_a^b B(u, v')$$

dém. :
Puisque la dérivée de $B(u, v)$ est $B(u', v) + B(u, v')$

$$\int_a^b B(u', v) + \int_a^b B(u, v') = \int_a^b (B(u, v))' = [B(u, v)]_a^b$$

□

19.3.4 Inégalité des accroissements finis

Théorème

Soit $f : I \to E$ de classe $\mathcal{C}^1$. S'il existe $M \in \mathbb{R}^+$ vérifiant

$$\forall t \in I, \|f'(t)\| \leqslant M$$

alors

$$\forall a, b \in I,\ \|f(b) - f(a)\| \leqslant M\,|b - a|$$

En d'autres termes, la fonction f est lipschitzienne.

dém. :
Puisque f est de classe $\mathcal{C}^1$, on peut écrire

$$\forall x \in I, f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t$$

Cas $a \leqslant b$

$$\|f(b) - f(a)\| = \left\| \int_{[a,b]} f'(t)\,\mathrm{d}t \right\| \leqslant \int_{[a,b]} \|f'(t)\|\ \mathrm{d}t$$

et donc

$$\|f(b) - f(a)\| \leqslant \int_{[a,b]} M\,\mathrm{d}t = M(b - a)$$

Cas $a > b$: analogue.
□

19.3.5 Formules de Taylor

19.3.5.1 Formule de Taylor avec reste intégral

Théorème

Soit $f : I \to E$ et $a \in I$. Si f est de classe $\mathcal{C}^{n+1}$

$$\forall x \in I, f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t$$

dém. :
Par récurrence en exploitant l'intégration par parties

$$\int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t = \left[-\frac{(x-t)^{n+1}}{(n+1)!} f^{(n+1)}(t)\right]_a^x + \int_a^x \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,\mathrm{d}t$$

□

Remarque Cette formule constitue une généralisation de l'identité

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t$$

Remarque Par le changement de variable affine $t = a + (x-a)u$, on peut réécrire le reste intégrale

$$\int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t = (x-a)^{n+1} \int_0^1 \frac{(1-u)^n}{n!} f^{(n+1)}(a + (x-a)u)\,\mathrm{d}u$$

Cette nouvelle écriture permet de mieux appréhender l'ordre de grandeur du reste.

19.3.5.2 Inégalité de Taylor-Lagrange

Théorème

Soit $f : I \to E$ et $a \in I$. Si f est de classe $\mathcal{C}^{n+1}$ et si $f^{(n+1)}$ bornée alors

$$\forall x \in I, \left\| f(x) - \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) \right\| \leqslant \frac{|x-a|^{n+1}}{(n+1)!} \sup_{t \in I} \left\| f^{(n+1)}(t) \right\|$$

dém. :
On a

$$\left\| \int_0^1 \frac{(1-u)^n}{n!} f^{(n+1)}(a + (x-a)u)\,\mathrm{d}u \right\| \leqslant \frac{1}{n!} \int_0^1 (1-u)^n \left\| f^{(n+1)} \right\|_\infty \mathrm{d}u = \frac{\left\| f^{(n+1)} \right\|_\infty}{(n+1)!}$$

□

Remarque Ce résultat constitue une généralisation de l'inégalité des accroissements finis.

19.3.5.3 Formule de Taylor-Young

Théorème

Soit $f : I \to E$ et $a \in I$. Si f est de classe $\mathcal{C}^n$

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} . f^{(k)}(a) + (x-a)^n \, \varepsilon(x) \text{ avec } \varepsilon(x) \xrightarrow[x \to a]{} 0_E$$

Cette relation est appelée développement limité de f à l'ordre n en a.

dém. :
Puisque que f est classe $\mathcal{C}^n$

$$\forall x \in I, f(x) = \sum_{k=0}^{n-1} \frac{(x-a)^k}{k!} f^{(k)}(a) + \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) \, \mathrm{d}t$$

Puisque $f^{(n)}$ est continue en a, on peut écrire

$$f^{(n)}(t) = f^{(n)}(a) + \varphi(t) \text{ avec } \varphi \underset{a}{\to} 0$$

et alors

$$\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t) \, \mathrm{d}t = \frac{(x-a)^n}{n!} f^{(n)}(a) + \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} \varphi(t) \, \mathrm{d}t$$

Soit $\varepsilon > 0$. Il existe $\alpha > 0$ tel que

$$|t-a| \leqslant \alpha \Rightarrow \|\varphi(t)\| \leqslant \varepsilon$$

et alors pour $|x-a| \leqslant \alpha$,

$$\left\| \int_a^x \frac{(x-t)^{n-1}}{(n-1)!} \varphi(t) \, \mathrm{d}t \right\| \leqslant \varepsilon \frac{|x-a|^n}{n!}$$

On peut alors écrire

$$\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} \varphi(t) \, \mathrm{d}t = (x-a)^n \varepsilon(x) \text{ avec } \varepsilon(x) \xrightarrow[x \to a]{} 0_E$$

□

Remarque En introduisant le concept de fonction négligeable, on peut aussi écrire

$$f(x) \underset{x \to a}{=} \sum_{k=0}^{n} \frac{(x-a)^k}{k!} . f^{(k)}(a) + o\left((x-a)^n\right)$$

Remarque La formule de Taylor-Young est locale : elle ne donne qu'une information sur le comportement asymptotique de f au voisinage de a. La formule de Taylor avec reste intégrale est quant à elle globale, elle fournit une information sur le comportement de la fonction sur l'intervalle I en entier. Il en est de même pour l'inégalité de Taylor-Lagrange.

19.4 Arcs paramétrés

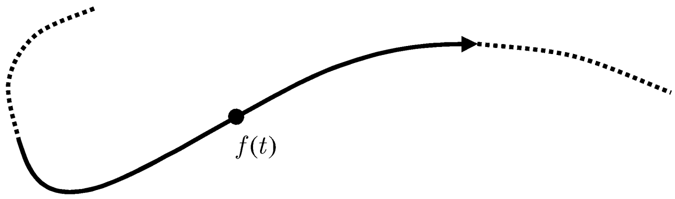
19.4.1 Définition

Définition

On appelle arc paramétré de classe $\mathcal{C}^k$ (avec $k \in \mathbb{N}^\star \cup \{\infty\}$) de E tout couple (I, f) constitué d'un intervalle I de $\mathbb{R}$ et d'une fonction vectorielle $f : I \to E$ de classe $\mathcal{C}^k$.
On s'intéresse alors à l'ensemble de point

$$\Gamma = \{f(t)/t \in I\}$$

appelé support de l'arc (I, f) (et l'on parle aussi de courbe paramétrée).
On dit aussi que la fonction f définit un paramétrage de la courbe Γ.



Remarque La valeur $f(t)$ permet de désigner un point de la courbe Γ, on dit que c'est le point de paramètre t.

Exemple Soit $a \in E$ et $u \neq 0_E$
L'application $t \mapsto a + t.u$ définit un paramétrage de la droite affine $a + \text{Vect}(u)$.

Exemple Considérons $E = \mathbb{C}$.
La fonction $f : t \in [0, 2\pi] \mapsto \mathrm{e}^{it}$ définit un paramétrage de $\mathbb{U} = \{z \in \mathbb{C} / |z| = 1\}$.

Remarque Il est fréquent de confondre l'arc paramétré et le support qu'il définit. C'est cependant maladroit car un arc paramétré détermine aussi une dynamique de parcours sur ce support.

19.4.2 Paramétrage dans le plan géométrique.

En munissant le plan géométrique d'un repère orthonormé $\mathcal{R} = (O; \vec{i}, \vec{j})$, on peut identifier le plan et $\mathbb{R}^2$.
Un arc paramétré donné par $f : I \to \mathbb{R}^2$ détermine alors à une courbe du plan.

Définition

Soit $x, y : I \to \mathbb{R}$ au moins de classe $\mathcal{C}^1$.
On appelle arc du plan défini par le système

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \text{ avec } t \in I$$

l'arc paramétré déterminé par l'application

$$f : t \mapsto (x(t), y(t))$$

Exemple Soit $A(x_0, y_0)$ un point et $\vec{u}(a, b)$ un vecteur non nul

$$\begin{cases} x = x_0 + t.a \\ y = y_0 + t.b \end{cases} \text{ avec } t \in \mathbb{R}$$

définit un paramétrage de la droite passant par A et dirigée par $\vec{u}$.

Exemple Soit $\Omega(a, b)$ un point et $R > 0$

$$\begin{cases} x = x_0 + R\cos(t) \\ y = y_0 + R\sin(t) \end{cases} \text{ avec } t \in [0, 2\pi]$$

définit un paramétrage du cercle de centre Ω et de rayon R.

19.4.3 Tangente en un point

Soit (I, f) un arc paramétré de classe au moins $\mathcal{C}^1$ et $t_0 \in I$.
On suppose qu'au voisinage de t_0,

$$f(t) = f(t_0) \Rightarrow t = t_0$$

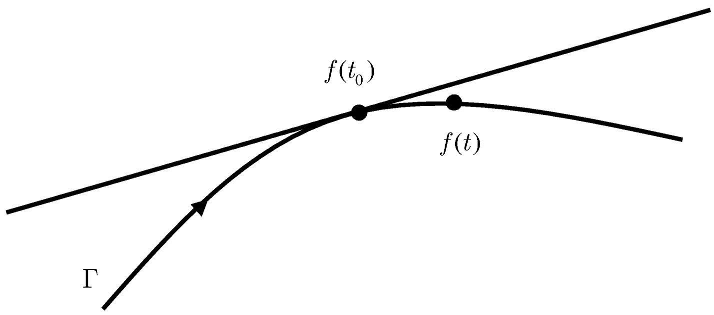
ce qui signifie que la courbe ne se recoupe pas infiniment sur elle-même en t_0...

Définition

On dit que l'arc (I, f) admet une demi-tangente à droite en t_0 si le vecteur unitaire

$$\frac{f(t) - f(t_0)}{\|f(t) - f(t_0)\|}$$

admet une limite en t_0. On dit alors que la droite issue du point $f(t_0)$ dirigée par ce vecteur est la demi-tangente à droite en t_0.
Mutatis mutandis, on définit la demi-tangente à gauche en t_0.
Enfin, si les deux droites demi-tangentes sont confondues, on dit que l'arc (I, f) admet une tangente en t_0 qui est cette droite commune.



Remarque Pour qu'il y ait tangente en t_0, il faut et il suffit que les vecteurs unitaires

$$\lim_{t \to t_0^+} \frac{f(t) - f(t_0)}{\|f(t) - f(t_0)\|} \text{ et } \lim_{t \to t_0^-} \frac{f(t) - f(t_0)}{\|f(t) - f(t_0)\|}$$

existent et soient égaux ou opposés.

19.4.4 Tangente en un point régulier

Soit (I, f) un arc paramétré de classe au moins $\mathcal{C}^1$ et $t_0 \in I$.

Définition

On dit que le paramètre t_0 est régulier lorsque $f'(t_0) \neq 0_E$.
On dit que l'arc est régulier lorsque tous ses paramètres le sont.

Théorème

Si t_0 est un paramètre régulier alors l'arc admet une tangente en $f(t_0)$ et celle-ci est dirigée par $f'(t_0)$.

dém. :
On peut écrire

$$f(t) - f(t_0) = (t - t_0).f'(t_0) + (t - t_0)\,\varepsilon(t) \text{ avec } \varepsilon(t) \xrightarrow[t \to t_0]{} 0_E$$

et donc pour $t \neq t_0$

$$\frac{f(t) - f(t_0)}{\|f(t) - f(t_0)\|} = \frac{t - t_0}{|t - t_0|} \frac{f'(t_0) + \varepsilon(t)}{\|f'(t_0) + \varepsilon(t)\|}$$

donc
$$\lim_{t\to t_0^+} \frac{f(t)-f(t_0)}{\|f(t)-f(t_0)\|} = \frac{f'(t_0)}{\|f'(t_0)\|} \text{ et } \lim_{t\to t_0^-} \frac{f(t)-f(t_0)}{\|f(t)-f(t_0)\|} = -\frac{f'(t_0)}{\|f'(t_0)\|}$$

□

Remarque Si $f'(t_0) = 0_E$ et si $f''(t_0) \neq 0_E$, on peut encore montrer l'existence d'une tangente en $f(t_0)$, cette fois-ci dirigée par $f''(t_0)$ car

$$f(t) - f(t_0) = \frac{1}{2}(t-t_0)^2 f''(t_0) + (t-t_0)^2\varepsilon(t) \text{ avec } \varepsilon(t) \xrightarrow[t\to t_0]{} 0_E$$

Exemple Considérons un arc du plan donné par

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \text{avec } t \in I$$

Si t est un paramètre régulier de cet arc, la tangente en le point de paramètre t_0 passe par le point de coordonnées $(x(t_0), y(t_0))$ et est dirigée par le vecteur de coordonnées $(x'(t_0), y'(t_0))$. Cette tangente a pour équation

$$\begin{vmatrix} x - x(t_0) & x'(t_0) \\ y - y(t_0) & y'(t_0) \end{vmatrix} = 0$$

c'est-à-dire

$$y'(t_0)\,(x - x(t_0)) - x'(t_0)\,(y - y(t_0)) = 0$$

La droite perpendiculaire à la tangente au point de coordonnées $(x(t_0), y(t_0))$ est appelée droite normale à l'arc. Elle a pour équation

$$\left(\begin{array}{c|c} x - x(t_0) & x'(t_0) \\ y - y(t_0) & y'(t_0) \end{array} \right) = 0$$

c'est à dire

$$x'(t_0)\,(x - x(t_0)) + y'(t_0)\,(y - y(t_0)) = 0$$

19.4.5 Vocabulaire cinématique

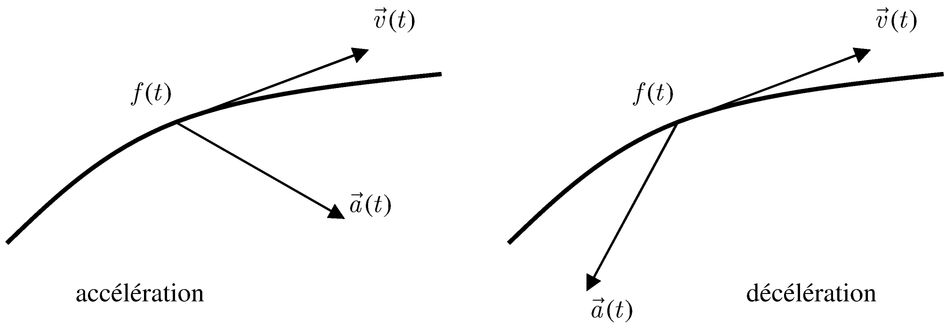
Soit $f : I \to E$ au moins de classe $\mathcal{C}^2$ définissant un arc paramétré.

Définition

En cinématique, les vecteurs $\vec{v}(t) = f'(t)$ et $\vec{a}(t) = f''(t)$ sont appelés vecteurs vitesse et accélération à l'instant t.

Remarque Le vecteur vitesse dirige la tangente (lorsqu'il n'est pas nul) et le vecteur accélération oriente la concavité de la courbe. Selon que l'angle géométrique entre $\vec{v}(t)$ et $\vec{a}(t)$ est aigu ou obtus, il y a accélération ou décélération lors du parcours de la courbe. En effet

$$\frac{\mathrm{d}}{\mathrm{d}t}(v^2) = \frac{\mathrm{d}}{\mathrm{d}t}\,(\vec{v} \mid \vec{v}) = 2\,(\vec{a} \mid \vec{v})$$



19.4.6 Exemples d'arcs plans

Exemple Considérons l'arc déterminé par

$$\begin{cases} x = a\cos t \\ y = b\sin t \end{cases} \text{avec } a > b > 0$$

Posons $x(t) = a\cos t$ et $y(t) = b\sin t$. Les fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$ sont de classe $\mathcal{C}^\infty$ définies sur $\mathbb{R}$. La fonction de paramétrage $f : \mathbb{R} \to \mathbb{R}^2$ définie par $f(t) = (x(t), y(t))$ est de classe $\mathcal{C}^\infty$.

$$\begin{cases} x(t + 2\pi) = x(t) \\ y(t + 2\pi) = y(t) \end{cases}$$

donc $f(t + 2\pi)$ et $f(t)$ sont confondus. Etude sur $[-\pi, \pi]$.

$$\begin{cases} x(-t) = x(t) \\ y(-t) = -y(t) \end{cases}$$

donc $f(-t)$ se déduit de $f(t)$ par une symétrie d'axe (Ox). Etude sur $[0, \pi]$

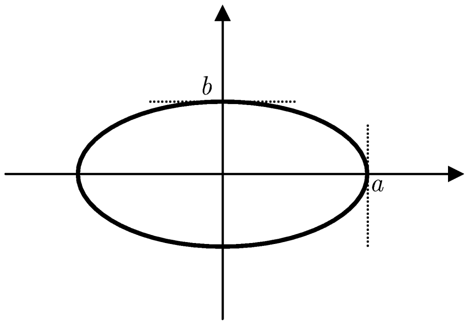
$$\begin{cases} x(\pi - t) = -x(t) \\ y(\pi - t) = y(t) \end{cases}$$

donc $f(\pi - t)$ se déduit de $f(t)$ par une symétrie d'axe (Oy). Etude sur $[0, \pi/2]$

$$\begin{cases} x'(t) = -a\sin t \\ y'(t) = b\cos t \end{cases}$$

t	0		$\pi/2$
$x'(t)$	0	$-$	
$x(t)$	a	$\searrow$	0
$y(t)$	0	$\nearrow$	b
$y'(t)$		$+$	0

En $t = 0$, il y a une tangente verticale.
En $t = \pi$, il y a une tangente horizontale.



Exemple Considérons l'arc déterminé par

$$\begin{cases} x = 3t^2 \\ y = 2t^3 \end{cases}$$

Posons $x(t) = 3t^2$ et $y(t) = 2t^3$. Les fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$ sont de classe $\mathcal{C}^\infty$ définies sur $\mathbb{R}$. La fonction de paramétrage $f : \mathbb{R} \to \mathbb{R}^2$ définie par $f(t) = (x(t), y(t))$ est de classe $\mathcal{C}^\infty$.

$$\begin{cases} x(-t) = x(t) \\ y(-t) = -y(t) \end{cases}$$

donc $f(-t)$ se déduit de $f(t)$ par une symétrie d'axe (Ox).
On peut limiter l'étude à $[0, +\infty[$.

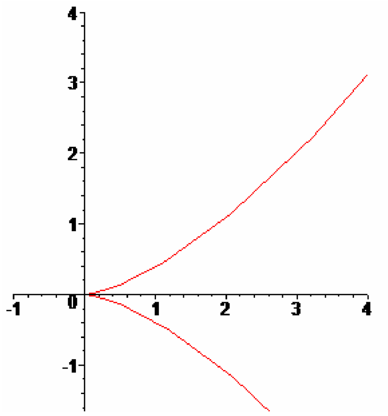
$$\begin{cases} x'(t) = 6t \\ y'(t) = 6t^2 \end{cases}$$

t	0		$+\infty$
$x'(t)$	0	$+$	
$x(t)$	0	$\nearrow$	$+\infty$
$y(t)$	0	$\nearrow$	$+\infty$
$y'(t)$	0	$+$	

Etude en $t = 0$: Le paramètre n'est pas régulier, cependant

$$\frac{f(t) - f(0)}{\|f(t) - f(0)\|} \to (1, 0)$$

Il y a donc une tangente horizontale en ce point.



Déterminons une équation de la tangente en tout point de paramètre $t \neq 0$.
Le point a pour coordonnées $(3t^2, 2t^3)$ et la tangente est dirigée par $(6t, 6t^2)$. Elle a donc pour équation

$$-6t^2\left(x-3t^2\right)+6t\left(y-2t^3\right)=0$$

soit encore

$$tx-y=t^3$$

Il est remarquable que cette équation est aussi valable en $t=0$.

Exemple Etudions l'arc paramétré déterminé par

$$\begin{cases} x = t - \sin t \\ y = 1 - \cos t \end{cases}$$

Posons $x(t) = t - \sin t$ et $y(t) = 1 - \cos t$.
Les fonctions $t \mapsto x(t)$ et $t \mapsto y(t)$ sont de classe $\mathcal{C}^\infty$ définies sur $\mathbb{R}$.
La fonction de paramétrage $f : \mathbb{R} \to \mathbb{R}^2$ définie par $f(t) = (x(t), y(t))$ est de classe $\mathcal{C}^\infty$.

$$\begin{cases} x(t+2\pi) = x(t) + 2\pi \\ y(t+2\pi) = y(t) \end{cases}$$

donc $f(t+2\pi)$ se déduit de $f(t)$ par une translation de vecteur $2\pi\vec{i}$.
On peut limiter l'étude à $[-\pi, \pi]$.

$$\begin{cases} x(-t) = -x(t) \\ y(-t) = y(t) \end{cases}$$

donc $f(-t)$ se déduit de $f(t)$ par une symétrie d'axe (Oy).
On peut limiter l'étude à $[0, \pi]$.

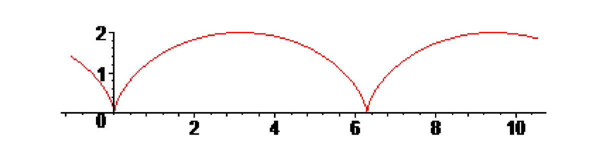
$$\begin{cases} x'(t) = 1 - \cos t \\ y'(t) = \sin t \end{cases}$$

t	0		π
$x'(t)$	0	$+$	2
$x(t)$	0	$\nearrow$	π
$y(t)$	0	$\nearrow$	2
$y'(t)$	0	$+$	0

Etude en $t = \pi$: Le paramètre est régulier, la tangente y est dirigée par $\vec{i}$.
Etude en $t = 0$: Le paramètre n'est pas régulier, cependant

$$\frac{f(t) - f(0)}{\|f(t) - f(0)\|} \to (0, 1)$$

La tangente y est verticale



Déterminons une équation de la tangente en tout point de paramètre $t \neq 0 \quad [2\pi]$.
Le point a pour coordonnées $(t - \sin t, 1 - \cos t)$ et la tangente est dirigée par $(1 - \cos t, \sin t)$. Elle a donc pour équation

$$-\sin(t)\,(x - (t - \sin(t)) + (1 - \cos(t))\,(y - (1 - \cos(t))) = 0$$

soit encore

$$-\sin(t)x + (1 - \cos(t))y = 2 - 2\cos(t) - t\sin(t)$$

19.4.7 Application : vecteurs tangents à une partie d'un espace normé de dimension finie

Soit a un élément d'une partie X d'un espace vectoriel réel de l'espace E.

Définition

On dit qu'un vecteur v de E est tangent à X en a, s'il existe $\varepsilon > 0$ et une fonction f définie sur $]-\varepsilon, \varepsilon[$ à valeurs dans X vérifiant

$$f(0) = a \text{ et } f'(0) = v$$

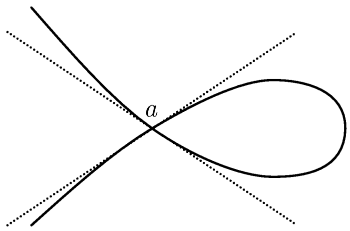
Lorsque le vecteur v est non nul, on dit que la droite

$$a + \mathrm{Vect} v$$

est tangente à X en a.

Exemple Si X correspond à un cercle, les vecteurs tangents correspondent aux vecteurs orthogonaux au vecteur rayon.

Exemple Si X correspond à une courbe se recoupant en a, il peut y avoir deux tangentes distinctes en ce point.



Exemple Si X correspond à une surface de l'espace, la définition qui précède permet aussi de parler de droite tangente à une surface.

Chapitre 20

Suites et séries de fonctions vectorielles

Soit E et F des espaces de dimensions finies. Ces espaces E et F peuvent être normés et le choix des normes n'a pas d'incidence sur la suite.

20.1 Modes de convergence

20.1.1 Suite de fonctions

Soit (u_n) suite de fonctions de $X \subset E$ vers F.

Définition

On dit que (u_n) converge simplement vers $u : X \to F$ si

$$\forall x \in X, u_n(x) \to u(x)$$

Définition

On dit que (u_n) converge uniformément vers $u : X \to F$ si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geqslant N, \forall x \in X, \|u_n(x) - u(x)\|_F \leqslant \varepsilon$$

Théorème

S'il y a convergence uniforme, il y aussi convergence simple et ce vers la même limite.

Remarque Sur $\mathcal{B}(X, F)$ espace des fonctions bornées de X vers F, on peut introduire la norme $\| \,.\, \|_\infty$ définie par

$$\|f\|_\infty = \sup_{x \in X} \|f(x)\|_F$$

On peut alors énoncer de nouveau la convergence uniforme

$$u_n \xrightarrow{CVU} u_n \Leftrightarrow \begin{cases} \exists N \in \mathbb{N}, \forall n \geqslant N, u_n - u \text{ bornée} \\ \|u_n - u\|_\infty \to 0 \end{cases}$$

20.1.2 Séries de fonctions

Soit $\sum u_n$ une série de fonctions de $X \subset E$ vers F i.e. une suite de fonctions (S_n) avec

$$S_n = \sum_{k=0}^{n} u_k$$

On définit la convergence simple et uniforme de la série de fonction $\sum u_n$ à partir de la suite (S_n) de ses sommes partielles.

Théorème

$\sum u_n$ converge simplement si, et seulement si,

$$\forall x \in X, \sum u_n(x) \text{ converge}$$

La somme de la série de fonctions est alors donnée par

$$S(x) = \sum_{n=0}^{+\infty} u_n$$

et son reste de rang n par

$$R_n(x) = \sum_{k=n+1}^{+\infty} u_k(x) = S(x) - S_n(x)$$

Théorème

$\sum u_n$ converge uniformément si, et seulement si,

$$\sum u_n \text{ converge simplement et } R_n \xrightarrow{CVU} \tilde{0}$$

Définition

On dit que $\sum u_n$ converge normalement si
1) chaque u_n est bornée ;
2) la série numérique $\sum \|u_n\|_\infty$ converge

Théorème

La convergence normale entraîne la convergence uniforme.

dém. :
Si $\sum u_n$ converge normalement alors pour tout $x \in X$, la série vectorielle $\sum u_n(x)$ converge absolument car

$$\|u_n(x)\| \leqslant \|u_n\|_\infty$$

et donc $\sum u_n$ converge simplement. De plus

$$\left\|\sum_{k=n+1}^{+\infty} u_k(x)\right\| \leqslant \sum_{k=n+1}^{+\infty} \|u_k(x)\| \leqslant \sum_{k=n+1}^{+\infty} \|u_k\|_\infty \to 0$$

et il y a donc convergence uniforme.
□

Remarque Les théorèmes qui suivront prolongeant ceux pour les fonctions numériques se démontrent de la même manière en substituant $\|\|$ à $||$.

20.2 Limite et continuité

20.2.1 Continuité par convergence uniforme

Soit (u_n) une suite de fonctions de $X \subset E$ vers F.

Théorème

Si
1) chaque u_n est continue ;
2) la suite (u_n) converge uniformément vers $u : X \to F$;
alors la fonction u est continue.

Corollaire

Si
1) chaque u_n est continue ;
2) la série $\sum u_n$ converge uniformément sur X ;
alors la fonction somme $\sum_{n=0}^{+\infty} u_n$ est continue.

Exemple Etude sur $\mathbb{R}^2$ de $S : (x,y) \mapsto \sum_{n=1}^{+\infty} \frac{1}{(n+x^2)(n+y^2)}$.
Définition :
Pour $n \geqslant 1$, on introduit

$$u_n : (x,y) \mapsto \frac{1}{(n+x^2)(n+y^2)}$$

Pour tout $(x,y) \in \mathbb{R}^2$,

$$u_n(x,y) \sim \frac{1}{n^2}$$

Or $\sum 1/n^2$ converge et $1/n^2 \geqslant 0$ donc la série $\sum u_n(x,y)$ converge.
On en déduit que la fonction $S = \sum_{n=0}^{+\infty} u_n$ est définie sur $\mathbb{R}^2$.
Continuité :
Les fonctions u_n sont continues.
Pour tout $(x,y) \in \mathbb{R}^2$,

$$|u_n(x,y)| \leqslant \frac{1}{n^2}$$

Or $\sum 1/n^2$ converge donc $\sum u_n$ converge normalement sur $\mathbb{R}^2$.
On en déduit que S est continue sur $\mathbb{R}^2$.

20.2.2 Continuité par convergence uniforme locale

Si l'on parvient à justifier la convergence uniforme sur des parties suffisamment générales pour déterminer des voisinages de tout $a \in X$, on peut affirmer à nouveau la continuité de l'objet limite.

Exemple Etude de $S(x,y) = \sum_{n=1}^{+\infty} \frac{\cos(ny)}{1+n^2x}$ sur $D = \{(x,y) \in \mathbb{R}^2 / x > 0\}$.
Pour $n \geqslant 1$, on introduit

$$u_n(x,y) = \frac{\cos(ny)}{1+n^2x}$$

Définition :
Pour tout $(x,y) \in D$, on a

$$|u_n(x,y)| \leqslant \frac{1}{1+n^2x} \sim \frac{1}{n^2x}$$

La série $\sum u_n(x,y)$ converge absolument et donc $\sum u_n$ converge simplement sur D.
Continuité :
Chaque fonction u_n est continue sur D.
Pour $a > 0$, considérons $D_a = \{(x,y) \in \mathbb{R}^2 / x \geqslant a\}$.
Pour $(x,y) \in D_a$, on a

$$|u_n(x,y)| \leqslant \frac{1}{n^2a}$$

Or $\sum 1/n^2a$ converge donc $\sum u_n$ converge normalement sur D_a.
La fonction S est donc continue sur D_a et puisque ceci vaut pour tout $a > 0$, elle est continue sur D.

Exemple Etude de $L : z \mapsto \sum_{n=1}^{+\infty} \frac{1}{n} z^n$ sur $D = \{z \in \mathbb{C} / |z| < 1\}$.
Définition :
Pour $n \geqslant 1$, on introduit

$$u_n : z \mapsto \frac{1}{n} z^n$$

Pour tout $z \in D$, on a $|u_n(z)| = o(z^n)$. Or $\sum z^n$ converge absolument donc $\sum u_n(z)$ converge absolument. Ainsi $\sum u_n$ converge simplement sur D.
Continuité :
Soit $r \in [0,1[$. Pour $|z| \leqslant r$, on a

$$|u_n(z)| \leqslant \frac{1}{n} r^n \leqslant r^n$$

Or $\sum r^n$ converge donc $\sum u_n$ converge normalement sur $\overline{D(0,r)}$.
La fonction L est définie et continue sur tous les domaines $\overline{D(0,r)}$ pour $r \in [0,1[$ donc elle est continue sur D.

20.2.3 Théorème de la double limite

Soit (u_n) une suite de fonctions de $X \subset E$ vers F et $a \in \bar{X}$.

Théorème

Si
1) (u_n) converge uniformément sur X vers une fonction u ;
2) pour tout $n \in \mathbb{N}$, $u_n \underset{a}{\to} \ell_n$;
Alors la suite (ℓ_n) converge et en notant ℓ sa limite

$$u(x) \xrightarrow[x \to a]{} \ell$$

Ainsi

$$\lim_{x \to a} \left(\lim_{n \to +\infty} u_n(x) \right) = \lim_{n \to +\infty} \left(\lim_{x \to a} u_n(x) \right)$$

Corollaire

Si
1) $\sum u_n$ converge uniformément sur X ;
2) pour tout $n \in \mathbb{N}$, $u_n \underset{a}{\to} \ell_n$;
Alors la série $\sum \ell_n$ converge et

$$\sum_{n=0}^{+\infty} u_n(x) \xrightarrow[x \to a]{} \sum_{n=0}^{+\infty} \ell_n$$

Exemple Non convergence uniforme de la série définissant $L : z \mapsto \sum_{n=1}^{+\infty} \frac{1}{n} z^n$ sur $D(0,1)$.

On a $1 \in \overline{D(0,1)}$ et

$$u_n(z) = \frac{1}{n} z^n \xrightarrow[z \to 1]{} \frac{1}{n}$$

Or la série $\sum \frac{1}{n}$ diverge donc la série de fonction $\sum u_n$ ne converge par uniformément sur $D(0,1)$.

20.3 Intégration et dérivation

Désormais la variable est supposée réelle

20.3.1 Intégration sur $[a, b]$

Théorème

Soit (u_n) une suite de fonctions de $[a, b]$ vers F.
Si
1) chaque u_n est continue ;
2) (u_n) converge uniformément vers $u : [a, b] \to F$
alors la fonction u est continue et

$$\int_a^b u_n(t)\,\mathrm{d}t \xrightarrow[n\to+\infty]{} \int_a^b u(t)\,\mathrm{d}t$$

Autrement dit

$$\int_a^b \lim_{n\to+\infty} u_n = \lim_{n\to+\infty} \int_a^b u_n$$

Corollaire

Si
1) chaque u_n est continue ;
2) $\sum u_n$ converge uniformément sur $[a, b]$
alors la fonction $\sum_{n=0}^{+\infty} u_n$ est continue et

$$\sum_{n=0}^{+\infty} \int_a^b u_n = \int_a^b \sum_{n=0}^{+\infty} u_n$$

20.3.2 Dérivation

I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide

Théorème

Soit (u_n) une suite de fonctions de I vers F.
Si
1) chaque u_n est de classe $\mathcal{C}^1$;
2) (u_n) converge simplement vers $u : I \to F$;
3) (u'_n) converge uniformément sur tout segment ;
alors u est de classe $\mathcal{C}^1$ et

$$\left(\lim_{n\to+\infty} u_n\right)' = \lim_{n\to+\infty} u'_n$$

Corollaire

Si
1) chaque u_n est de classe $\mathcal{C}^1$;
2) $\sum u_n$ converge simplement sur I ;
3) $\sum u'_n$ converge uniformément sur tout segment de I ;
alors la fonction $\sum_{n=0}^{+\infty} u_n$ est de classe $\mathcal{C}^1$ sur I et

$$\left(\sum_{n=0}^{+\infty} u_n\right)' = \sum_{n=0}^{+\infty} u'_n$$

Remarque On peut aussi énoncer un résultat pour les fonctions de classe $\mathcal{C}^n$.

20.4 Exponentielles

20.4.1 Exponentielle complexe

Théorème

Pour tout $z \in \mathbb{C}$, la série $\sum_{n\geqslant 0} \frac{1}{n!} z^n$ est absolument convergente.

dém. :
Pour $z = 0$: ok.
Pour $z \neq 0$, on introduit $u_n = z^n/n! \neq 0$.
On a

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{|z|}{n+1} \to 0 < 1$$

donc, par la règle de d'Alembert, $\sum_{n\geqslant 0} \frac{1}{n!} z^n$ est absolument convergente.

□

Définition

On pose

$$\exp(z) \underset{\text{déf}}{=} \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

Remarque Cette définition prolonge l'exponentielle réelle car on a déjà vue

$$\forall x \in \mathbb{R}, \mathrm{e}^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$$

Exemple $\exp(0) = 1$ car $0^0 = 1$.

Proposition

$\forall z \in \mathbb{C}, \overline{\exp(z)} = \exp \bar{z}$.

dém. :
Par conjugaison de séries convergentes.
□

Théorème

$$\forall z, z' \in \mathbb{C},\ \exp(z)\exp(z') = \exp(z + z')$$

dém. :

$$\exp(z)\exp(z') = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \sum_{n=0}^{+\infty} \frac{z'^n}{n!}$$

Par produit de Cauchy de séries absolument convergentes

$$\exp(z)\exp(z') = \sum_{n=0}^{+\infty} w_n$$

avec

$$w_n = \sum_{k=0}^{n} \frac{z^k}{k!} \frac{z'^{n-k}}{(n-k)!} = \frac{1}{n!}(z + z')^n$$

en vertu de la formule du binôme de Newton.
Ainsi

$$\exp(z)\exp(z') = \exp(z + z')$$

□

Corollaire

$$\forall \theta \in \mathbb{R}, \exp(i\theta) \in \mathbb{U}$$

dém. :
Pour $\theta \in \mathbb{R}$, on a $|\exp(i\theta)|^2 = \exp(i\theta)\exp(-i\theta) = 1$ donc $\exp(i\theta) \in \mathbb{U}$.
□

Remarque A partir de cette définition de l'exponentielle complexe, on définit les fonctions $\cos$ et $\sin$ par :

$$\cos\theta = \frac{\mathrm{e}^{i\theta} + \mathrm{e}^{-i\theta}}{2} = \mathrm{Re}(\mathrm{e}^{i\theta}) \text{ et } \sin\theta = \frac{\mathrm{e}^{i\theta} - \mathrm{e}^{-i\theta}}{2i} = \mathrm{Im}(\mathrm{e}^{i\theta})$$

On peut alors retrouver les propriétés usuelles de ses fonctions.
Par exemple :
? $|\exp(i\theta)|^2 = 1$ donne $\cos^2\theta + \sin^2\theta = 1$;
- $\exp(-i\theta) = \overline{\exp(i\theta)}$ donne $\cos(-\theta) = \cos(\theta)$ et $\sin(-\theta) = -\sin(\theta)$;
- $\exp(i(a + b)) = \exp(ia)\exp(ib)$ donne

$$\cos(a + b) = \cos a \cos b - \sin a \sin b \text{ et } \sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)\ldots$$

On peut aussi définir précisément le nombre π comme étant le double de la première annulation strictement positive de la fonction cosinus et achever la construction de la trigonométrie...

20.4.2 Exponentielle d'une matrice

Théorème

Pour toute matrice $A \in \mathcal{M}_p(\mathbb{K})$, la série $\sum \frac{1}{n!}A^n$ est absolument convergente.

dém. :
Introduisons $\|\,.\,\|_2$ sur $\mathcal{M}_p(\mathbb{K})$ définie par

$$\|A\|_2 = \sum_{i,j=1}^{n} |a_{i,j}|^2$$

Vérifions que celle-ci est sous multiplicative i.e.

$$\forall A, B \in \mathcal{M}_n(\mathbb{K}), \|AB\|_2 \leqslant \|A\|_2 \|B\|_2$$

On a

$$(AB)_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}$$

et par l'inégalité de Cauchy-Schwarz,

$$\|\,AB\|_2^2 = \sum_{i,j=1}^{n} \left|\sum_{k=1}^{n} a_{i,k}b_{k,j}\right|^2 \leqslant \sum_{i,j=1}^{n} \left(\sum_{k=1}^{n} |a_{i,k}|^2 \sum_{k=1}^{n} |b_{j,k}|^2\right) = \|A\|_2^2 \|B\|_2^2$$

On a alors

$$\left\|\frac{1}{n!}A^n\right\|_2 = \frac{1}{n!}\|A^n\|_2 \leqslant \frac{1}{n!}\|A\|_2^n$$

Or $\sum x^n/n!$ converge pour tout $x \in \mathbb{R}$, donc par comparaison de série à termes positifs, la série $\sum \frac{1}{n!}A^n$ converge absolument.
□

Définition

On appelle exponentielle de la matrice $A \in \mathcal{M}_p(\mathbb{K})$ la somme

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{1}{n!}A^n$$

Exemple $\exp(O_p) = I_p$.

Théorème

Soit $A, B \in \mathcal{M}_p(\mathbb{K})$.
Si $AB = BA$ alors

$$\exp(A + B) = \exp(A)\exp(B)$$

dém. :
C'est la même que pour $\exp(z+z') = \exp(z)\exp(z')$ en admettant que le théorème relatif aux produits de Cauchy de séries absolument convergentes et encore vrai sur $\mathcal{M}_p(\mathbb{K})$. L'hypothèse de commutation est nécessaire à l'usage de la formule du binôme.
□

Corollaire

$\forall A \in \mathcal{M}_p(\mathbb{K})$, $\exp(A)$ est inversible et $\exp(A)^{-1} = \exp(-A)$.

Théorème

L'application $A \mapsto \exp(A)$ est continue.

dém. :
On introduit les fonctions données par $u_n(A) = A^n/n!$ définies pour $A \in \mathcal{M}_p(\mathbb{K})$.
Les fonctions u_n sont toutes continues.
Soit $R \in \mathbb{R}^+$. Pour $\|A\| \leqslant R$, on a

$$\|u_n(A)\|_2 \leqslant \frac{1}{n!}\|A\|_2^n \leqslant \frac{1}{n!}R^n$$

Or $\sum R^n/n!$ converge et donc $\sum u_n$ converge normalement sur $B_f(O_p, R)$.
On en déduit que la fonction $A \mapsto \exp(A)$ est continue sur $B_f(O_p, R)$ et puisque ceci vaut pour tout $R \in \mathbb{R}^+$, la fonction $A \mapsto \exp(A)$ est continue sur $\mathcal{M}_p(\mathbb{K})$.
□

20.4.3 Calcul d'exponentielle de matrices

Pour $A \in \mathcal{M}_n(\mathbb{K})$, calculons

$$\exp(A) = \sum_{k=0}^{+\infty} \frac{1}{k!}A^k$$

20.4.3.1 Cas A est diagonale

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, A^k = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \text{ et } \sum_{k=0}^{N} \frac{1}{k!}A^k = \begin{pmatrix} \displaystyle\sum_{k=0}^{N} \frac{1}{k!}\lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \displaystyle\sum_{k=0}^{N} \frac{1}{k!}\lambda_n^k \end{pmatrix}$$

Ainsi

$$\exp(A) = \begin{pmatrix} \mathrm{e}^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{\lambda_n} \end{pmatrix}$$

20.4.3.2 Cas A diagonalisable

$A = PDP^{-1}$ avec D diagonale. $A^k = PD^kP^{-1}$ et $\displaystyle\sum_{k=0}^{N} \frac{1}{k!}A^k = P\left(\sum_{k=0}^{N} \frac{1}{k!}D^k\right)P^{-1}$. Ainsi

$$\exp(A) = P\exp(D)P^{-1}$$

Exemple Calcul de $\exp(A)$ pour $A = \begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix}$.
$\mathrm{Sp}(A) = \{1, 2\}$.
La matrice A est diagonalisable.
Il existe P tel que $A = PDP^{-1}$ avec $D = \mathrm{diag}(1, 2)$ et alors $\exp(A) = PD'P$ avec $D' = \mathrm{diag}(\mathrm{e}, \mathrm{e}^2)$.
Soit T polynôme tel que $T(1) = \mathrm{e}$ et $T(2) = \mathrm{e}^2$.

$$T(X) = \mathrm{e}(\mathrm{e} - 1)(X - 1) + \mathrm{e} \text{ convient}$$

On a $T(D) = D'$ et par similitude $T(A) = \exp(A)$. Ainsi

$$\exp(A) = \mathrm{e}\,(\mathrm{e} - 1)\,A + \mathrm{e}(2 - \mathrm{e})I_2$$

20.4.3.3 Cas A nilpotente

Supposons $A^p = O_n$.
Pour $N \geqslant p$,

$$\sum_{k=0}^{N} \frac{1}{k!} A^k = \sum_{k=0}^{p} \frac{1}{k!} A^k$$

Ainsi

$$\exp(A) = \sum_{k=0}^{p-1} \frac{1}{k!} A^k$$

20.4.3.4 Cas général

On improvise, par exemple en exploitant un polynôme annulateur...

Exemple Calcul de $\exp(A)$ avec $A = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 1 & 1 \\ -2 & -1 & -1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R})$
On a $\chi_A = (X - 1)^3$ et donc la matrice A est trigonalisable.
Par Cayley-Hamilton, on a $(A - I_3)^3 = O_3$. Posons $N = A - I_3$.
On a $A = I_3 + N$ avec I_3 et N commutant donc

$$\exp(A) = \exp(I_3)\exp(N) = \mathrm{e}\left(I_3 + N + \frac{1}{2}N^2\right)$$

Ainsi

$$\exp(A) = \mathrm{e}\begin{pmatrix} 7/2 & 1 & 5/2 \\ 1 & 1 & 1 \\ -5/2 & -1 & -3/2 \end{pmatrix}$$

20.4.4 Exponentielle d'un endomorphisme

Soit $a \in \mathcal{L}(E)$. Etudions la série $\sum \frac{1}{n!} a^n$.
On introduit e une base de E et on peut définir une norme sur $\mathcal{L}(E)$ en posant

$$\|a\| = \|\mathrm{Mat}_e(a)\|_2$$

Celle-ci vérifie $\|a \circ b\| \leqslant \|a\| \, \|b\|$ et l'on peut dès lors adapter l'étude matricielle aux endomorphismes.

Définition

On appelle exponentielle de a l'endomorphisme

$$\exp(a) = \sum_{n=0}^{+\infty} \frac{1}{n!} a^n$$

Exemple $\exp(\tilde{0}) = \mathrm{Id}_E$.

Exemple Si $A = \mathrm{Mat}_e(a)$ alors

$$\mathrm{Mat}_e\left(\exp(a)\right) = \exp(A)$$

Théorème

Si $a, b \in \mathcal{L}(E)$ vérifie $a \circ b = b \circ a$ alors

$$\exp(a) \circ \exp(b) = \exp(a + b)$$

Corollaire

$\forall a \in \mathcal{L}(E)$, $\exp(a)$ est inversible et $\exp(a)^{-1} = \exp(-a)$.

Théorème

L'application $a \mapsto \exp(a)$ est continue.

20.4.5 Dérivation de l'application $t \mapsto \exp(t.a)$

Fixons $a \in \mathcal{L}(E)$ et considérons la fonction

$$e_a : t \mapsto e_a(t) = \exp(t.a) \in \mathcal{L}(E)$$

avec

$$\exp(t.a) = \sum_{n=0}^{+\infty} \frac{t^n}{n!}.a^n$$

Théorème

L'application $e_a : t \mapsto \exp(t.a)$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et

$$e_a'(t) = a \circ e_a(t) = e_a(t) \circ a$$

dém. :

Introduisons les fonctions $u_n : \mathbb{R} \to E$ définies par

$$u_n(t) = \frac{t^n}{n!}.a^n$$

La série $\sum u_n$ converge simplement et sa somme est la fonction e_a.
Chaque u_n est de classe $\mathcal{C}^1$ et

$$u_n(t) = \frac{t^{n-1}}{(n-1)!}.a^n \text{ si } n \geqslant 1 \text{ et } u_n(t) = 0 \text{ si } n = 0$$

Soit $M \geqslant 0$ et $|t| \leqslant M$.

$$\|u_n(t)\| \leqslant \frac{M^{n-1}}{(n-1)!}\|a^n\| \leqslant \frac{M^{n-1}}{(n-1)!}\|a\|^n = \frac{(M\|a\|)^{n-1}}{(n-1)!}\|a\|$$

Or on sait que pour tout $x \in \mathbb{R}$, $\sum \frac{x^n}{n!}$ converge donc $\sum_{n \geqslant 1} \frac{(M\|a\|)^{n-1}}{(n-1)!}$ converge.

Par comparaison de séries à terme positifs, on obtient la convergence normale de $\sum u_n$ sur $[-M, M]$.
Finalement, par convergence uniforme sur tout segment de $\mathbb{R}$, on peut affirmer que e_a est une fonction de classe $\mathcal{C}^1$ et

$$e_a'(t) = \sum_{n=1}^{+\infty} \frac{t^{n-1}}{(n-1)!}.a^n = \sum_{n=0}^{+\infty} \frac{t^n}{n!}.a^{n+1} = a \circ \exp(t.a) = \exp(t.a) \circ a$$

Enfin, par récurrence, on obtient que $t \to \exp(ta)$ est de classe $\mathcal{C}^\infty$.
□

Corollaire

On a aussi
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\exp(tA)\right) = A\exp(tA) = \exp(tA)A$$

dém. :
En adaptant la démonstration précédente ou en raisonnant via endomorphisme canoniquement associé.
□

Chapitre 21

Intégrales dépendant d'un paramètre

21.1 Passage à la limite sous l'intégrale

21.1.1 Théorème de convergence dominée

On étudie

$$\lim_{n\to+\infty}\int_I f_n(t)\,\mathrm{d}t$$

Rappel :
Cas $I=[a,b]$
Si les fonctions f_n sont continues et si $f_n \xrightarrow[{[a,b]}]{CVU} f$ alors

$$\int_a^b f_n \xrightarrow[n\to+\infty]{} \int_a^b f$$

Cet outil ne suffit pas à résoudre tous les cas possibles.

Théorème

Soit (f_n) une suite de fonctions de I vers $\mathbb{K}$
Si
1) les fonctions (f_n) sont continues par morceaux sur I ;
2) la suite de fonctions (f_n) converge simplement vers une fonction f continue par morceaux ;
3) il existe $\varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall n\in\mathbb{N},\ |f_n|\leqslant\varphi \text{ [hypothèse de domination]}$$

alors les fonctions f_n et f sont intégrables sur I et

$$\int_I f_n \to \int_I f$$

Exemple Etudions

$$\lim_{n\to+\infty}\int_{-\infty}^{+\infty}\frac{1+2\sin(t/n)}{1+t^2}\,\mathrm{d}t$$

Posons $f_n : \mathbb{R} \to \mathbb{R}$ définie par

$$f_n(t) = \frac{1 + 2\sin(t/n)}{1 + t^2}$$

On a $f_n \xrightarrow{CVS} f$ avec $f(t) = \dfrac{1}{1+t^2}$.
Les fonctions f_n et la fonction f sont continues par morceaux.
De plus

$$|f_n(t)| \leqslant \frac{3}{1+t^2} = \varphi(t)$$

avec φ intégrable sur $\mathbb{R}$.
Par convergence dominée, les fonctions f_n et la fonction f sont intégrables et

$$\lim_{n\to+\infty} \int_{-\infty}^{+\infty} \frac{1 + 2\sin(t/n)}{1+t^2}\,\mathrm{d}t = \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{1+t^2} = \pi$$

Exemple Etudions

$$\lim_{n\to+\infty} \int_0^{\pi/2} \sin^n(t)\,\mathrm{d}t$$

Posons $f_n : [0, \pi/2] \to \mathbb{R}$ définie par $f_n(t) = \sin^n(t)$ sur $[0, \pi/2]$.
$f_n \xrightarrow{CVS} f$ avec

$$f(t) = \begin{cases} 1 & \text{si } t = \pi/2 \\ 0 & \text{sinon} \end{cases}$$

Les fonctions f_n et la fonction f sont continues par morceaux.
Pour tout $n \in \mathbb{N}$,

$$|f_n| \leqslant 1 = \varphi$$

φ est intégrable sur $[0, \pi/2]$ car définie et continue sur un segment.
Par convergence dominée

$$\int_0^{\pi/2} f_n \to \int_0^{\pi/2} f$$

et donc

$$\lim_{n\to+\infty} \int_0^{\pi/2} \sin^n(t)\,\mathrm{d}t = 0$$

Remarque Ici la suite de fonctions (f_n) ne converge pas uniformément vers f mais on est parvenu à permuter limite et intégrale.

Exemple Etudions

$$\lim_{n\to+\infty} \int_0^{+\infty} \mathrm{e}^{-t^n}\,\mathrm{d}t$$

Posons $f_n : \mathbb{R}^+ \to \mathbb{R}$ définie par $f_n(t) = \mathrm{e}^{-t^n}$.
Pour $t \in [0, 1[$, $f_n(t) \xrightarrow[n\to+\infty]{} 1$.
Pour $t = 1$, $f_n(t) \xrightarrow[n\to+\infty]{} 1/\mathrm{e}$.

Pour $t > 1$, $f_n(t) \xrightarrow[n\to+\infty]{} 0$.
Ainsi $f_n \xrightarrow{CVS} f$ avec

$$f(t) = \begin{cases} 1 & \text{si } t \in [0,1[\\ 1/\mathrm{e} & \text{si } t = 1 \\ 0 & \text{si } t > 1 \end{cases}$$

Les fonctions f_n et la fonction f sont continues par morceaux.
Pour tout $n \in \mathbb{N}^\star$, $|f_n| \leqslant \varphi$ avec

$$\varphi(t) = \begin{cases} 1 & \text{si } t \in [0,1] \\ \mathrm{e}^{-t} & \text{si } t > 1 \end{cases}$$

Par convergence dominée, les fonctions f_n et la fonction f sont intégrables et

$$\lim_{n\to+\infty} \int_0^{+\infty} \mathrm{e}^{-t^n}\,\mathrm{d}t = \int_0^{+\infty} f(t)\,\mathrm{d}t = \int_0^1 1\,\mathrm{d}t = 1$$

Exemple Etudions

$$\lim_{n\to+\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n \ln t\,\mathrm{d}t$$

Problème : $\displaystyle\int_0^n$ et non $\displaystyle\int_I$.
Solution : $\displaystyle\int_0^n f(t)\,\mathrm{d}t = \int_0^{+\infty} \tilde{f}(t)\,\mathrm{d}t$ avec

$$\tilde{f}(t) = \begin{cases} f(t) & \text{si } t \leqslant n \\ 0 & \text{sinon} \end{cases}$$

Ici, introduisons $f_n :]0,+\infty[\to \mathbb{R}$ définie par

$$f_n(t) = \begin{cases} \left(1 - \dfrac{t}{n}\right)^n \ln t & \text{si } t \in]0,n[\\ 0 & \text{sinon} \end{cases}$$

Soit $t \in]0,+\infty[$.
Quand $n \to +\infty$, pour n assez grand $t < n$ et

$$f_n(t) = \left(1 - \frac{t}{n}\right)^n \ln t \xrightarrow[n\to+\infty]{} \mathrm{e}^{-t} \ln t$$

Ainsi $f_n \xrightarrow{CVS} f$ avec $f : t \mapsto \mathrm{e}^{-t} \ln t$
Les fonctions f_n et la fonction f sont continues par morceaux.
Sachant $\ln(1+u) \leqslant u$ on a pour $t \in]0,n[$

$$|f_n(t)| = \exp\left(n \ln(1 - t/n)\right) |\ln t| \leqslant \exp(-t)\,|\ln t| = \varphi(t)$$

La fonction φ est continue par morceaux sur $]0,+\infty[$ et intégrable car

$$\sqrt{t}\varphi(t) \xrightarrow[t\to0^+]{} 0 \text{ et } t^2\varphi(t) \xrightarrow[t\to+\infty]{} 0$$

Par convergence dominée, les fonctions f_n et la fonction f sont intégrables et

$$\lim_{n\to+\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n \ln t\,\mathrm{d}t = \int_0^{+\infty} \mathrm{e}^{-t} \ln t\,\mathrm{d}t$$

Remarque En calculant $\displaystyle\int_0^n \left(1-\frac{t}{n}\right)^n \ln t \,\mathrm{d}t$, on parvient à montrer alors

$$\int_0^{+\infty} \mathrm{e}^{-t} \ln t \,\mathrm{d}t = -\gamma$$

21.1.2 Autres techniques pour étudier une limite

Convergence uniforme sur un segment $[a,b]$ et convergence dominée ne suffisent pas toujours pour déterminer

$$\lim_{n\to+\infty} \int_I f_n$$

On peut aussi :
- procéder par comparaison ;
- réexprimer l'intégrale (par changement de variable, intégration par parties, astuce,...) ;
- raisonner par les ε.

Exemple Montrons que pour tout $f \in \mathcal{C}^1([a,b],\mathbb{K})$,

$$\int_a^b f(t)\mathrm{e}^{int} \,\mathrm{d}t \to 0$$

Par intégration par parties,

$$\int_a^b f(t)\mathrm{e}^{int} \,\mathrm{d}t = \left[\frac{1}{in}\mathrm{e}^{int} f(t)\right]_a^b - \frac{1}{in}\int_a^b f'(t)\mathrm{e}^{int} \,\mathrm{d}t$$

Par suite

$$\left|\int_a^b f(t)\mathrm{e}^{int} \,\mathrm{d}t\right| \leqslant \frac{1}{n}\left(|f(a)| + |f(b)| + \int_a^b |f'(t)| \,\mathrm{d}t\right) \to 0$$

Ainsi

$$\int_a^b f(t)\mathrm{e}^{int} \,\mathrm{d}t \to 0$$

21.1.3 Intégration terme à terme

On étudie si

$$\int_I \sum_{n=0}^{+\infty} f_n(t) \,\mathrm{d}t = \sum_{n=0}^{+\infty} \int_I f_n(t) \,\mathrm{d}t$$

Rappel : Cas $I = [a,b]$
Si les fonctions f_n sont continues et si la série $\sum f_n$ converge uniformément sur $[a,b]$ alors

$$\int_a^b \sum_{n=0}^{+\infty} f_n = \sum_{n=0}^{+\infty} \int_a^b f_n$$

Théorème

Soit $\sum f_n$ une série de fonctions de I vers $\mathbb{K}$.
Si
1) les fonctions f_n sont continues par morceaux et intégrables sur I ;
2) la série de fonctions $\sum f_n$ converge simplement vers une fonction $\sum_{n=0}^{+\infty} f_n$ continue par morceaux ;
3) la série numérique $\sum \int_I |f_n|$ converge
Alors la fonction $\sum_{n=0}^{+\infty} f_n$ est intégrable sur I et

$$\int_I \sum_{n=0}^{+\infty} f_n = \sum_{n=0}^{+\infty} \int_I f_n$$

Exemple Montrons

$$\int_0^1 \frac{\ln t}{t-1}\,\mathrm{d}t = \sum_{n=1}^{+\infty} \frac{1}{n^2}$$

On a

$$\frac{1}{1-t} = -\sum_{n=0}^{+\infty} t^n \text{ sur } [0,1[$$

donc

$$\frac{\ln t}{t-1} = \sum_{n=0}^{+\infty} (-\ln t)t^n \text{ sur }]0,1[$$

On a alors

$$\int_0^1 \frac{\ln t}{t-1}\,\mathrm{d}t = \int_{]0,1[} \sum_{n=0}^{+\infty} f_n(t)\,\mathrm{d}t$$

avec $f_n :]0,1[\to \mathbb{R}$ définie par $f_n(t) = (-\ln t)t^n$.
Par les calculs qui précédent, la série de fonctions $\sum f_n$ converge simplement et sa somme $t \mapsto \frac{\ln t}{t-1}$ est continue par morceaux.
Chaque fonction f_n est continue par morceaux et intégrable sur $]0,1[$ car

$$\sqrt{t} f_n(t) \xrightarrow[t\to 0]{} 0 \text{ et } f_n(t) \xrightarrow[t\to 1]{} 0$$

Enfin, par intégration par parties

$$\int_0^1 |f_n(t)|\,\mathrm{d}t = \int_0^1 (-\ln(t))t^n\,\mathrm{d}t = \frac{1}{(n+1)^2}$$

La série numérique $\sum \int_0^1 |f_n|$ converge donc par théorème d'intégration terme à terme

$$\int_0^1 \frac{\ln t}{t-1}\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{1}{(n+1)^2} = \sum_{n=1}^{+\infty} \frac{1}{n^2}$$

21.1.4 Autre technique d'intégration terme à terme

Pour résoudre des situations « plus délicates » , on peut aussi intégrer terme à terme en revenant aux sommes partielles. Notons

$$S_n = \sum_{k=0}^{n} f_k \text{ et } S = \sum_{n=0}^{+\infty} f_n$$

Par convergence dominée ou comparaison, supposons avoir montré

$$\int_I S_n(t)\,\mathrm{d}t \xrightarrow[n\to+\infty]{} \int_I S(t)\,\mathrm{d}t$$

En remarquant

$$\int_I S_n = \int_I \sum_{k=0}^{n} f_k = \sum_{k=0}^{n} \int_I f_k$$

on affirme

$$\sum_{k=0}^{n} \int_I f_k \to \int_I \sum_{n=0}^{+\infty} f_n$$

et donc

$$\sum_{n=0}^{+\infty} \int_I f_n = \int_I \sum_{n=0}^{+\infty} f_n$$

Exemple Montrons

$$\int_0^1 \frac{\mathrm{d}t}{1+t^2} = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1}$$

On peut écrire

$$\frac{1}{1+t^2} = \frac{1}{1-q} = \sum_{n=0}^{+\infty} (-1)^n t^{2n} \text{ sur } [0,1[$$

Par suite

$$\int_0^1 \frac{\mathrm{d}t}{1+t^2} = \int_{[0,1[} \sum_{n=0}^{+\infty} f_n$$

avec $f_n(t) = (-1)^n t^{2n}$ définie sur $[0,1[$
Ici $\sum \int_{[0,1[} |f_n| = \sum \frac{1}{2n+1}$ diverge et on ne peut pas appliquer le théorème d'intégration terme à terme. Transitons alors par les sommes partielles
On pose

$$S_n(t) = \sum_{k=0}^{n} (-1)^k t^{2k}$$

On a $S_n \xrightarrow{CVS} S$ avec $S(t) = \dfrac{1}{1+t^2}$.
Les fonctions S_n et S sont continues par morceaux.

$$|S_n(t)| = \frac{\left|1-(-1)^{n+1}t^{2n+2}\right|}{1+t^2} \leqslant \frac{2}{1+t^2} = \varphi(t)$$

avec φ intégrable.

Par convergence dominée

$$\int_0^1 S_n(t)\,\mathrm{d}t \to \int_0^1 S(t)\,\mathrm{d}t$$

Or

$$\int_0^1 S_n(t)\,\mathrm{d}t = \int_0^1 \sum_{k=0}^{n} (-1)^k t^{2k}\,\mathrm{d}t = \sum_{k=0}^{n} \int_0^1 (-1)^k t^{2k}\,\mathrm{d}t = \sum_{k=0}^{n} \frac{(-1)^k}{2k+1}$$

donc

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = \int_0^1 \frac{\mathrm{d}t}{1+t^2}$$

avec en substance la convergence de la série introduite.

21.2 Continuité d'une intégrale à paramètre

On étudie dans cette partie les fonctions de la forme

$$g : x \in X \mapsto \int_I f(x,t)\,\mathrm{d}t$$

Dans un premier temps X désigne un intervalle de $\mathbb{R}$.

21.2.1 Continuité par domination

Théorème

Si $f : X \times I \to \mathbb{K}$ vérifie
1) $\forall x \in X$, $t \mapsto f(x,t)$ est continue par morceaux sur I ;
2) $\forall t \in I$, $x \mapsto f(x,t)$ est continue sur X ;
3) $\exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in X \times I,\ |f(x,t)| \leqslant \varphi(t) \text{ [hypothèse de domination]}$$

Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et continue sur X.

dém. :
Pour tout $x \in X$, la fonction $t \mapsto f(x,t)$ est intégrable sur I et donc $g(x)$ est bien définie.
Etudions la continuité en $a \in X$ via la caractérisation séquentielle des limites.
Soit (x_n) une suite d'éléments de X convergeant vers a.
$g(x_n) = \displaystyle\int_I f(x_n,t)\,\mathrm{d}t = \int_I u_n(t)\,\mathrm{d}t$ avec $u_n(t) = f(x_n,t)$.
Pour tout $t \in I$, $u_n(t) = f(x_n,t) \xrightarrow[n\to+\infty]{} f(a,t) = u_\infty(t)$,
Ainsi (u_n) converge simplement vers la fonction $u_\infty : t \mapsto u(a,t)$.
Chaque u_n et u_∞ sont continues par morceaux.
Pour tout $n \in \mathbb{N}$, $|u_n(t)| \leqslant \varphi(t)$ avec φ intégrable.
Par convergence dominée $\displaystyle\int_I u_n(t)\,\mathrm{d}t \xrightarrow[n\to+\infty]{} \int_I u_\infty(t)\,\mathrm{d}t$ i.e. $g(x_n) \to g(a)$.
Par caractérisation séquentielle de la continuité g est continue en a.
□

Exemple Définition et continuité de $g(x) = \displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-xt}}{1+t^2}\,\mathrm{d}t$ avec $x \in \mathbb{R}^+$.
Considérons $f : (x,t) \mapsto \dfrac{\mathrm{e}^{-xt}}{1+t^2}$ définie sur $\mathbb{R}^+ \times [0,+\infty[$.
$\forall x \in \mathbb{R}$, $t \mapsto f(x,t)$ est continue par morceaux sur $[0,+\infty[$.
$\forall t \in [0,+\infty[$, $x \mapsto f(x,t)$ est continue sur $\mathbb{R}^+$.
$\forall (x,t) \in \mathbb{R}^+ \times [0,+\infty[$, $|f(x,t)| \leqslant \dfrac{1}{1+t^2} = \varphi(t)$
avec $\varphi : [0,+\infty[\to \mathbb{R}^+$ continue par morceaux et intégrable sur $[0,+\infty[$ car $\varphi(t) \underset{t\to+\infty}{\sim} \dfrac{1}{t^2}$.
Par domination, la fonction g est définie et continue sur $\mathbb{R}^+$.

Exemple Définition et continuité de $g(x) = \displaystyle\int_0^{\pi} \cos(x\sin\theta)\,\mathrm{d}\theta$ avec $x \in \mathbb{R}$.
Considérons $f : (x,\theta) \mapsto \cos(x\sin\theta)$ définie sur $\mathbb{R} \times [0,\pi]$.
$\forall x \in \mathbb{R}, \theta \mapsto \cos(x\sin\theta)$ est continue par morceaux sur $[0,\pi]$.
$\forall \theta \in [0,\pi]$, $x \mapsto \cos(x\sin\theta)$ est continue sur $\mathbb{R}$.
$\forall (x,\theta) \in \mathbb{R} \times [0,\pi]$, $|f(x,\theta)| \leqslant 1 = \varphi(\theta)$.
La fonction constante φ est évidemment intégrable sur $[0,\pi]$.
Par domination, g est définie et continue sur $\mathbb{R}$.

Remarque Les hypothèses 1) et 2) du théorème sont évidemment réunies lorsque f est continue sur $X \times I$. En pratique, elles sont faciles à obtenir, c'est surtout l'hypothèse 3 qui importe.

21.2.2 Continuité par domination sur tout segment

Pour obtenir la continuité de g, il n'est pas toujours possible de vérifier l'hypothèse de domination directement sur l'intégralité de l'intervalle X.

Théorème

Si $f : X \times I \to \mathbb{K}$ vérifie
1) $\forall x \in X$, $t \mapsto f(x,t)$ est continue par morceaux sur I ;
2) $\forall t \in I$, $x \mapsto f(x,t)$ est continue sur X ;
3) $\forall\, [a,b] \subset X$, $\exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in [a,b] \times I,\ |f(x,t)| \leqslant \varphi(t) \text{ [hypothèse de domination locale]}$$

Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et continue sur X.

dém. :
g est définie et continue sur chaque $[a,b] \subset X$ donc définie et continue sur X.
□

Exemple Définition et continuité de $g(x) = \displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-xt}}{1+t}\,\mathrm{d}t$ avec $x > 0$.
On introduit

$$f(x,t) = \frac{\mathrm{e}^{-xt}}{1+t} \text{ définie sur } \mathbb{R}^{+\star} \times [0,+\infty[$$

Définition
Soit $x > 0$. La fonction $t \mapsto f(x,t)$ est continue par morceaux sur $[0,+\infty[$ et

$$f(x,t) \underset{t\to+\infty}{\sim} \frac{\mathrm{e}^{-xt}}{t}$$

donc

$$t^2 f(x,t) \xrightarrow[t\to+\infty]{} 0$$

$t \mapsto f(x,t)$ est donc intégrable sur $[0,+\infty[$ et par conséquent $g(x)$ est bien définie pour tout $x > 0$
Continuité
La fonction f est continue sur $\mathbb{R}^{+\star} \times [0,+\infty[$.
Soit $[a,b] \subset]0,+\infty[$. Pour $x \in [a,b]$

$$|f(x,t)| \leqslant \frac{\mathrm{e}^{-at}}{1+t} = \varphi(t)$$

Par l'étude au dessus, la fonction $\varphi : [0,+\infty[\to \mathbb{R}$ est continue par morceaux et intégrable.
Par domination sur tout segment, on en déduit que g est continue sur $\mathbb{R}^{+\star}$.

Exemple Définition et continuité de $g(x) = \displaystyle\int_0^{+\infty} \frac{\ln(1+xt)}{1+t^2}\,\mathrm{d}t$ avec $x \geqslant 0$.
On introduit

$$f(x,t) = \frac{\ln(1+xt)}{1+t^2} \text{ définie sur } \mathbb{R}^+ \times [0,+\infty[$$

Définition
Soit $x \in \mathbb{R}^+$. La fonction $t \mapsto f(x,t)$ est continue par morceaux sur $[0,+\infty[$ et

$$f(x,t) \underset{t\to+\infty}{\sim} \frac{\ln t}{t^2}$$

donc

$$t^{3/2} f(x,t) \xrightarrow[t\to+\infty]{} 0$$

$t \mapsto f(x,t)$ est donc intégrable sur $[0,+\infty[$ et par conséquent $g(x)$ est bien définie pour tout $x \geqslant 0$
Continuité
La fonction f est continue sur $\mathbb{R}^+ \times [0,+\infty[$.
Soit $[a,b] \subset [0,+\infty[$. Pour $x \in [a,b]$

$$|f(x,t)| \leqslant \frac{\ln(1+bt)}{1+t^2} = \varphi(t)$$

Par l'étude au dessus, la fonction $\varphi : [0,+\infty[\to \mathbb{R}$ est continue par morceaux et intégrable.
Par domination sur tout segment, on en déduit que g est continue sur $\mathbb{R}^+$.

21.2.3 Limite

Soit a une extrémité de l'intervalle X.
On désire étudier la limite de $g(x)$ quand $x \to a$.

Théorème

Si $f : X \times I \to \mathbb{K}$ vérifie :
1) $\forall x \in X, f(x,.)$ est continue par morceaux sur I ;
2) $\forall t \in I$, $f(x,t) \xrightarrow[x \to a]{} \ell(t)$ avec ℓ continue par morceaux ;
3) $\exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in X \times I,\ |f(x,t)| \leqslant \varphi(t) \text{ [hypothèse de domination]}$$

alors

$$g(x) = \int_I f(x,t)\,\mathrm{d}t \xrightarrow[x \to a]{} \int_I \ell(t)\,\mathrm{d}t$$

dém. :
Soit (x_n) une suite d'éléments de X convergeant vers a.

$$\int_I f(x_n,t)\,\mathrm{d}t = \int_I u_n(t)\,\mathrm{d}t$$

avec $u_n(t) = f(x_n,t)$.
Pour tout $t \in I$, $u_n(t) = f(x_n,t) \xrightarrow[n \to +\infty]{} \ell(t)$,
Ainsi (u_n) converge simplement vers la fonction ℓ.
Chaque u_n et ℓ sont continues par morceaux.
Pour tout $n \in \mathbb{N}$, $|u_n(t)| \leqslant \varphi(t)$ avec φ intégrable.
Par convergence dominée $\displaystyle\int_I u_n(t)\,\mathrm{d}t \xrightarrow[n \to +\infty]{} \int_I \ell(t)\,\mathrm{d}t$ i.e.

$$\int_I f(x_n,t)\,\mathrm{d}t \xrightarrow[x \to a]{} \int_I \ell(t)\,\mathrm{d}t$$

Par caractérisation séquentielle des limites,

$$\int_I f(x,t)\,\mathrm{d}t \xrightarrow[x \to a]{} \int_I \ell(t)\,\mathrm{d}t$$

□

Remarque L'hypothèse de domination peut être avantageusement remplacée par une hypothèse de domination exprimée sur un intervalle inclus dans X dont a est extrémité, mais pas par une hypothèse de domination sur tout segment.

Exemple Limite quand $x \to +\infty$ de $g(x) = \displaystyle\int_0^{+\infty} \ln(t)\mathrm{e}^{-xt}\,\mathrm{d}t$.
Posons $f(x,t) = \ln(t)\mathrm{e}^{-xt}$ définie sur $\mathbb{R}^{+\star} \times [0,+\infty[$.

$$f(x,t) \xrightarrow[x \to +\infty]{} 0$$

Pour $x \geqslant 1$,

$$|f(x,t)| \leqslant \ln(t)\mathrm{e}^{-t} = \varphi(t)$$

avec φ intégrable sur $]0,+\infty[$.
Par domination, on obtient

$$\lim_{x\to+\infty} g(x) = \int_0^{+\infty} 0\,\mathrm{d}t = 0$$

Remarque Il est souvent tout aussi efficace de raisonner par comparaison de limites lorsque cela est possible.

Exemple Limite quand $x \to +\infty$ de $g(x) = \displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-xt}}{1+t}\,\mathrm{d}t$.
On a

$$0 \leqslant \int_0^{+\infty} \frac{\mathrm{e}^{-xt}}{1+t}\,\mathrm{d}t \leqslant \int_0^{+\infty} \mathrm{e}^{-xt}\,\mathrm{d}t = \frac{1}{x} \xrightarrow[x\to+\infty]{} 0$$

donc par encadrement g tend vers 0 en $+\infty$.
Etudions $\lim\limits_{x\to0^+} g(x)$.
On a

$$g(x) \geqslant \int_0^{1/x} \frac{\mathrm{e}^{-1}}{1+t}\,\mathrm{d}t = \frac{1}{\mathrm{e}}\ln\left(1+\frac{1}{x}\right) \xrightarrow[x\to0^+]{} +\infty$$

donc par comparaison g tend vers $+\infty$ en 0.

21.2.4 Extension aux fonctions d'une variable vectorielle

Ici X désigne une partie d'un espace normé de dimension finie ($X \subset \mathbb{R}, \mathbb{C}, \mathbb{R}^n, \ldots$)

Théorème

Si $f : X \times I \to \mathbb{K}$ vérifie
1) $\forall x \in X$, $t \mapsto f(x,t)$ est continue par morceaux sur I ;
2) $\forall t \in I$, $x \mapsto f(x,t)$ est continue sur X ;
3) $\exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in X \times I,\ |f(x,t)| \leqslant \varphi(t) \text{ [hypothèse de domination]}$$

Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et continue sur X.

dém. :
Il suffit de reprendre à l'identique la démonstration précédente du résultat analogue vu quand X est un intervalle.
□

Remarque Il n'est pas toujours possible d'obtenir l'hypothèse de domination sur X entier. Cependant, il peut suffire de l'obtenir sur des domaines suffisamment généraux si ceux-ci incluent des voisinages de tout $a \in X$.

Exemple Définition et continuité de $g(z) = \displaystyle\int_0^1 \frac{\ln t}{t+z}\,\mathrm{d}t$ avec $z \in \mathbb{C}$ vérifiant $\mathrm{Re}(z) > 0$.
On introduit

$$f(z,t) = \frac{\ln t}{z+t} \text{ définie sur } \Omega \times]0,1] \text{ avec } \Omega = \{z \in \mathbb{C}/\mathrm{Re}(z) > 0\}$$

Définition
Soit $z \in \Omega$. La fonction $t \mapsto f(z,t)$ est continue par morceaux sur $]0,1]$ et

$$f(z,t) \underset{t\to 0^+}{\sim} \frac{\ln t}{z}$$

donc

$$\sqrt{t}f(z,t) \xrightarrow[t\to 0^+]{} 0$$

$t \mapsto f(z,t)$ est donc intégrable sur $]0,1]$ et par conséquent $g(z)$ est bien définie pour tout $z \in \mathbb{C}$.
Continuité
La fonction f est continue sur $\Omega \times]0,1]$

$$|f(z,t)| = \frac{|\ln t|}{|z+t|} \leqslant \frac{|\ln t|}{t + \mathrm{Re}(z)}$$

Soit $a > 0$ et $\Omega_a = \{z \in \mathbb{C}/\mathrm{Re}(z) \geqslant a\}$.
Pour $z \in \Omega_a$ et $t \in]0,1]$,

$$|f(z,t)| \leqslant \frac{\ln t}{t+a} = \varphi_a(t)$$

Par l'étude au dessus, la fonction $\varphi_a : [0,+\infty[\to \mathbb{R}$ est continue par morceaux et intégrable.
On en déduit que la fonction g est continue sur Ω_a pour tout $a > 0$, elle est donc continue sur Ω.

21.3 Dérivation d'une intégrale à paramètre

On étudie dans cette partie les fonctions de la forme

$$g : x \in X \mapsto \int_I f(x,t)\,\mathrm{d}t$$

avec X un intervalle d'intérieur non vide de $\mathbb{R}$.

21.3.1 Formule de Leibniz

Définition

Soit $f : (x,t) \mapsto f(x,t)$ définie sur $X \times I$.
On dit que f admet une dérivée partielle $\dfrac{\partial f}{\partial x}$ si

$$\forall t \in I, \text{ la fonction } x \mapsto f(x,t) \text{ est dérivable}$$

On pose alors

$$\frac{\partial f}{\partial x}(x,t) = \frac{\mathrm{d}}{\mathrm{d}x}\left(f(x,t)\right)$$

Théorème

Soit $f : X \times I \to \mathbb{K}$. On suppose que f admet une dérivée partielle

$$\frac{\partial f}{\partial x}$$

Si
1) $\forall x \in I, t \mapsto f(x,t)$ est continue par morceaux et intégrable sur I ;
2) $\forall x \in I, t \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue par morceaux sur I ;
3) $\forall t \in I, x \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue sur X ;
4) $\exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in X \times I,\ \left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant \varphi(t)$$

Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et de classe $\mathcal{C}^1$ sur X avec

$$g'(x) = \int_I \frac{\partial f}{\partial x}(x,t)\,\mathrm{d}t$$

dém. :
Etudions la dérivabilité en $a \in X$

$$\frac{g(x)-g(a)}{x-a} \xrightarrow[x \to a]{} ?$$

Pour $x \neq a$

$$\frac{g(x)-g(a)}{x-a} = \int_I u(x,t)\,\mathrm{d}t$$

avec

$$u(x,t) = \frac{f(x,t)-f(a,t)}{x-a}$$

Soit $t \in I$.

$$u(x,t) = \frac{h(x)-h(a)}{x-a}$$

en introduisant la fonction $h : x \mapsto f(x,t)$.
Par hypothèse, la fonction h est dérivable et donc

$$u(x,t) \xrightarrow[x \to a]{} h'(a) = \frac{\partial f}{\partial x}(a,t) = \ell(t)$$

La fonction ℓ est continue par morceaux sur I.
Soit $t \in I$.
L'application $h : x \mapsto f(x,t)$ est dérivable et sa dérivée vérifie

$$|h'(x)| = \left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant \varphi(t)$$

Par l'inégalité des accroissements finis, $h : x \mapsto f(x,t)$ est $\varphi(t)$-lipschitzienne.
Par suite

$$|u(x,t)| = \frac{|h(x)-h(a)|}{|x-a|} \leqslant \varphi(t)$$

avec φ continue par morceaux et intégrable.
Par domination
$$\int_I u(x,t)\,\mathrm{d}t \xrightarrow[x\to a]{} \int_I \ell(t)\,\mathrm{d}t$$
i.e.
$$\frac{g(x)-g(a)}{x-a} \to \int_I \frac{\partial f}{\partial x}(a,t)\,\mathrm{d}t$$
Finalement g est dérivable en a et
$$g'(a) = \int_I \frac{\partial f}{\partial x}(a,t)\,\mathrm{d}t$$
Enfin g' est continue par application du théorème de continuité par domination.
□

Remarque Le résultat énoncé est encore vrai si l'on remplace l'hypothèse « $t \mapsto f(x,t)$ est intégrable sur I » par celle de « convergence de $\int_I f(x,t)\,\mathrm{d}t$ » .

Exemple Calcul de $g(x) = \int_0^{+\infty} \mathrm{e}^{-t^2}\cos(xt)\mathrm{d}t$ avec $x \in \mathbb{R}$.
Posons $u(x,t) = \mathrm{e}^{-t^2}\cos(xt)$.
La fonction u est définie sur $\mathbb{R}\times[0,+\infty[$ et admet une dérivée partielle
$$\frac{\partial u}{\partial x}(x,t) = -t\mathrm{e}^{-t^2}\sin(xt)$$
$\forall x \in \mathbb{R}$, $t \mapsto u(x,t)$ est continue par morceaux et intégrable sur $[0,+\infty[$ car négligeable devant $1/t^2$ en $+\infty$.
$\forall x \in \mathbb{R}$, $t \mapsto \frac{\partial u}{\partial x}(x,t)$ est continue par morceaux sur $[0,+\infty[$.
$\forall t \in [0,+\infty[$, $x \mapsto \frac{\partial u}{\partial x}(x,t)$ est continue sur $\mathbb{R}$.
Enfin
$$\forall (x,t) \in \mathbb{R}\times[0,+\infty[\,, \left|\frac{\partial u}{\partial x}(x,t)\right| \leqslant t\mathrm{e}^{-t^2} = \varphi(t)$$
avec $\varphi : [0,+\infty[\to \mathbb{R}$ continue par morceaux et intégrable sur $[0,+\infty[$.
Par domination, la fonction g est de classe $\mathcal{C}^1$ et
$$g'(x) = \int_0^{+\infty} -t\mathrm{e}^{-t^2}\sin(xt)\mathrm{d}t$$
Procédons à une intégration par parties avec les fonctions de classe $\mathcal{C}^1$
$$u(t) = \frac{1}{2}\mathrm{e}^{-t^2} \text{ et } v(t) = \sin(xt)$$
Puisque le produit uv converge en 0 et $+\infty$, l'intégration par parties impropre est possible et
$$g'(x) = \left[\frac{1}{2}\mathrm{e}^{-t^2}\sin(xt)\right]_0^{+\infty} - \frac{1}{2}\int_0^{+\infty} x\mathrm{e}^{-t^2}\cos(xt)\,\mathrm{d}t$$

Ainsi, on obtient

$$g'(x) = -\frac{1}{2}xg(x)$$

g est solution d'une équation différentielle linéaire d'ordre 1 et $g(0) = \sqrt{\pi}/2$ on conclut

$$\varphi(x) = \frac{\sqrt{\pi}}{2}\mathrm{e}^{-\frac{1}{4}x^2}$$

21.3.2 Dérivation par domination sur tout segment

Théorème

Soit $f : X \times I \to \mathbb{K}$. On suppose que f admet une dérivée partielle

$$\frac{\partial f}{\partial x}$$

Si
1) $\forall x \in I, t \mapsto f(x,t)$ est continue par morceaux et intégrable sur I ;
2) $\forall x \in I, t \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue par morceaux sur I ;
3) $\forall t \in I, x \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue sur X ;
4) $\forall [a,b] \subset X, \exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in [a,b] \times I, \left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant \varphi(t)$$

Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et de classe $\mathcal{C}^1$ sur X avec

$$g'(x) = \int_I \frac{\partial f}{\partial x}(x,t)\,\mathrm{d}t$$

dém. :
La fonction g est de classe $\mathcal{C}^1$ sur tout segment $[a,b] \subset X$ donc de classe $\mathcal{C}^1$ sur l'intervalle X.
□

Exemple Calcul de $g(x) = \displaystyle\int_0^1 \frac{t^x - 1}{\ln t}\,\mathrm{d}t$ avec $x \in \left]-1,+\infty\right[$. Considérons $f : (x,t) \mapsto \dfrac{t^x-1}{\ln t}$ définie sur $\left]-1,+\infty\right[\times \left]0,1\right[$.
Soit $x > -1$. La fonction $t \mapsto f(x,t)$ est continue par morceaux sur $\left]0,1\right[$.
Quand $t \to 1^-$.
$t = 1 - h$ avec $h \to 0^+$.

$$f(x,t) = \frac{(1+h)^x - 1}{\ln(1+h)} \to x$$

et donc f est intégrable sur $[1/2, 1[$.
Quand $t \to 0^+$.
On a

$$t^x \xrightarrow[t \to 0^+]{} \begin{cases} 0 & \text{si } x > 0 \\ 1 & \text{si } x = 0 \\ +\infty & \text{si } x \in \left]-1,0\right[\end{cases}$$

Si $x \geqslant 0$, on obtient $f(x,t) \to 0$ ce qui permet un prolongement par continuité.
Si $x < 0$, on a $f(x,t) = o\left(t^x\right) = o\left(1/t^{-x}\right)$ avec $-x < 1$.
Dans les deux cas, $t \mapsto f(x,t)$ est intégrable sur $]0, 1/2]$.
Finalement $t \mapsto f(x,t)$ est intégrable sur $]0,1[$ et donc g est définie sur $]-1,+\infty[$.
La fonction $x \mapsto f(x,t) = \dfrac{t^x - 1}{\ln t}$ est dérivable donc f admet une dérivée partielle $\dfrac{\partial f}{\partial x}$ et

$$\frac{\partial f}{\partial x}(x,t) = t^x$$

$\forall x \in]-1,+\infty[$, $t \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue par morceaux sur $]0,1[$
$\forall t \in]0,1[$, $x \mapsto \dfrac{\partial f}{\partial x}(x,t)$ est continue sur $]-1,+\infty[$.
Soit $[a,b] \subset]-1,+\infty[$. Pour $x \in [a,b]$,

$$\left|\frac{\partial f}{\partial x}(x,t)\right| \leqslant t^a = \varphi(t)$$

avec $\varphi :]0,1[\to \mathbb{R}^+$ continue par morceaux et intégrable sur $]0,1[$.
Par domination sur tout segment, g est de classe $\mathcal{C}^1$ et

$$g(x) = \int_0^1 t^x \,\mathrm{d}t = \frac{1}{x+1}$$

On en déduit

$$g(x) = g(0) + \int_0^x \frac{\mathrm{d}t}{1+t} = \ln(1+x)$$

21.3.3 Dérivées d'ordres supérieurs

Définition

On dit que $f : (x,t) \mapsto f(x,t)$ admet une dérivée partielle $\dfrac{\partial^j f}{\partial x^j}$ si pour chaque valeur de t, la fonction $x \mapsto f(x,t)$ est j fois dérivable et on pose alors

$$\frac{\partial^j f}{\partial x^j}(x,t) = \frac{\mathrm{d}^j}{\mathrm{d}x^j}\left(f(x,t)\right)$$

Théorème

Soit $f : X \times I \to \mathbb{K}$. On suppose que f admet des dérivées partielles

$$\frac{\partial f}{\partial x}, \ldots, \frac{\partial^n f}{\partial x^n}$$

Si

1) $\forall j \in \{0, \ldots, n-1\}$, $\forall x \in X$, $t \mapsto \dfrac{\partial^j f}{\partial x^j}(x,t)$ est continue par morceaux et intégrable sur I.

et si

2) $\forall x \in X$, $t \mapsto \dfrac{\partial^n f}{\partial x^n}(x,t)$ est continue par morceaux

3) $\forall t \in I$, $x \mapsto \dfrac{\partial^n f}{\partial x^n}(x,t)$ est continue ;

4) $\forall\, [a,b] \subset I$, $\exists \varphi : I \to \mathbb{R}^+$ continue par morceaux et intégrable vérifiant

$$\forall (x,t) \in [a,b] \times I,\ \left|\frac{\partial^n f}{\partial x^n}(x,t)\right| \leqslant \varphi(t)$$

Alors la fonction $g : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et de classe $\mathcal{C}^n$ sur X et pour tout $j \in \{1, \ldots, n\}$

$$g^{(j)}(x) = \int_I \frac{\partial^j f}{\partial x^j}(x,t)\,\mathrm{d}t$$

dém. :

Par récurrence sur $n \geqslant 1$.

Cas $n = 1$: résolu ci-dessus

Supposons le théorème vrai au rang $n \geqslant 1$.

Soit f vérifiant les hypothèses données au rang $n+1$.

Pour $[a,b] \subset X$, il existe $\varphi_{a,b} : I \to \mathbb{R}^+$ continue par morceaux intégrable vérifiant

$$\forall (x,t) \in X \times I,\ \left|\frac{\partial^{n+1} f}{\partial x^{n+1}}(x,t)\right| \leqslant \varphi(t)$$

Par calcul intégral

$$\frac{\partial^n f}{\partial x^n}(x,t) = \frac{\partial^n f}{\partial x^n}(a,t) + \int_a^x \frac{\partial^{n+1} f}{\partial x^{n+1}}(y,t)\,\mathrm{d}y$$

et donc

$$\left|\frac{\partial^n f}{\partial x^n}(x,t)\right| \leqslant \left|\frac{\partial^n f}{\partial x^n}(a,t)\right| + (b-a)\varphi(t) = \psi(t)$$

La fonction ψ étant intégrable, on peut employer l'hypothèse de récurrence et affirmer que g est de classe $\mathcal{C}^n$ avec

$$\forall 1 \leqslant j \leqslant n, g^{(j)}(x) = \int_I \frac{\partial^j f}{\partial x^j}(x,t)\,\mathrm{d}t$$

En particulier

$$g^{(n)}(x) = \int_I \frac{\partial^n f}{\partial x^n}(x,t)\,\mathrm{d}t$$

et les hypothèses vérifiées par f au rang $n+1$ assurent que $g^{(n)}$ est de classe $\mathcal{C}^1$ avec

$$(g^{(n)})'(x) = \int_I \frac{\partial^{n+1} f}{\partial x^{n+1}}(x,t)\,\mathrm{d}t$$

Ce qui permet de conclure.
Récurrence établie.
□

Exemple Montrons que $g(x) = \displaystyle\int_0^{+\infty} \frac{\mathrm{e}^{-xt}}{1+t^2}\,\mathrm{d}t$ définit une solution sur $\mathbb{R}^{+\star}$ de l'équation différentielle

$$y'' + y = \frac{1}{x}$$

Considérons $f : (x,t) \mapsto \dfrac{\mathrm{e}^{-xt}}{1+t^2}$ définie sur $]0,+\infty[\times[0,+\infty[$
Pour $t \in [0,+\infty[$, la fonction $x \mapsto f(x,t)$ est deux fois dérivable sur $]0,+\infty[$ donc les dérivées partielles $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial^2 f}{\partial x^2}$ existent et

$$\frac{\partial f}{\partial x}(x,t) = -t\frac{\mathrm{e}^{-xt}}{1+t^2} \text{ et } \frac{\partial^2 f}{\partial x^2}(x,t) = t^2\frac{\mathrm{e}^{-xt}}{1+t^2}$$

Pour tout $x \in]0,+\infty[$, $t \mapsto f(x,t)$ et $t \mapsto \dfrac{\partial f}{\partial x}(x,t)$ sont continues par morceaux sur $[0,+\infty[$ et intégrables sur $[0,+\infty[$ car

$$t^2 f(x,t) \xrightarrow[t\to+\infty]{} 0 \text{ et } t^2\frac{\partial f}{\partial x}(x,t) \xrightarrow[t\to+\infty]{} 0$$

De plus
$\forall x \in]0,+\infty[$, $t \mapsto \dfrac{\partial^2 f}{\partial x^2}(x,t)$ est continue par morceaux.
$\forall t \in [0,+\infty[$, $x \mapsto \dfrac{\partial^2 f}{\partial x^2}(x,t)$ est continue.
Enfin, pour $[a,b] \subset]0,+\infty[$. On a

$$\forall (x,t) \in [a,b]\times[0,+\infty[\,, \left|\frac{\partial^2 f}{\partial x^2}(x,t)\right| \leqslant \mathrm{e}^{-at}$$

avec $\varphi : t \mapsto \mathrm{e}^{-at}$ continue par morceaux et intégrable sur $[0,+\infty[$.
Par domination sur tout segment, la fonction g est de classe $\mathcal{C}^2$ sur $\mathbb{R}^{+\star}$ et

$$g''(x) + g(x) = \int_0^{+\infty} t^2\frac{\mathrm{e}^{-xt}}{1+t^2}\,\mathrm{d}t + \int_0^{+\infty} \frac{\mathrm{e}^{-xt}}{1+t^2}\,\mathrm{d}t = \int_0^{+\infty} \mathrm{e}^{-xt}\,\mathrm{d}t = \frac{1}{x}$$

21.4 Applications

Les résultats qui suivent ne sont pas explicitement au programme : on ne peut les utiliser qu'en les redémontrant.

21.4.1 Transformée de Laplace

Soit $f : [0, +\infty[\to \mathbb{C}$ continue bornée

Définition

On appelle transformée de Laplace de f l'application $L(f)$ définie par

$$\forall x > 0, L(f)(x) = \int_0^{+\infty} f(t)\mathrm{e}^{-xt}\,\mathrm{d}t$$

Exemple Pour $f(t) = 1$, on obtient $L(f)(x) = 1/x$.

Exemple Pour $f(t) = \sin(\omega t)$, on obtient

$$L(f)(x) = \mathrm{Im}\left(\int_0^{+\infty} \mathrm{e}^{(-x+i\omega)t}\,\mathrm{d}t\right) = \frac{\omega}{x^2+\omega^2}$$

Théorème

L'application L est linéaire de $L^\infty([0,+\infty[\,,\mathbb{C})$ vers $\mathcal{C}(]0,+\infty[\,,\mathbb{C})$.

dém. :
Soit $f : [0, +\infty[\to \mathbb{C}$ continue et bornée.
Posons $u(x,t) = f(t)\mathrm{e}^{-xt}$ définie sur $\mathbb{R}^{+\star} \times [0, +\infty[$.
Pour chaque $x > 0$, la fonction $t \mapsto u(x,t)$ est continue par morceaux sur $[0, +\infty[$.
Pour chaque $t \in [0, +\infty[$, la fonction $x \mapsto u(x,t)$ est continue sur $]0, +\infty[$.
Pour $[a, b] \subset]0, +\infty[$, on a

$$\forall (x,t) \in [a,b] \times [0,+\infty[\,, |u(x,t)| \leqslant \|f\|_\infty \mathrm{e}^{-at} = \varphi(t)$$

avec $\varphi : \mathbb{R}^+ \to \mathbb{R}$ continue par morceaux et intégrable.
Par domination sur tout segment, l'application

$$L(f) : x \mapsto \int_0^{+\infty} u(x,t)\,\mathrm{d}t$$

est définie et continue sur $]0, +\infty[$
Ainsi, l'application L est bien définie de l'espace $L^\infty([0,+\infty[\,,\mathbb{C})$ vers $\mathcal{C}(]0,+\infty[\,,\mathbb{C})$.
Sa linéarité est évidente par linéarité du calcul intégral.
□

Remarque On peut aussi montrer que cette application L est injective.

Théorème

Si $f : [0, +\infty[\to \mathbb{C}$ est de classe $\mathcal{C}^1$ et si les fonctions f et f' sont bornées alors

$$\forall x > 0, L(f')(x) = xL(f)(x) - f(0)$$

dém. :
Soit $x > 0$. On a

$$L(f')(x) = \int_0^{+\infty} f'(t)\mathrm{e}^{-xt}\,\mathrm{d}t$$

Procédons à une intégration par parties avec $u'(t) = f'(t)$ et $v(t) = \mathrm{e}^{-xt}$.
Les fonctions u et v sont de classe $\mathcal{C}^1$ et le produit uv admet des limites en 0 et $+\infty$ donc

$$L(f')(x) = \left[f(t)\mathrm{e}^{-xt}\right]_0^{+\infty} - \int_0^{+\infty} (-x)f(t)\mathrm{e}^{-xt}\,\mathrm{d}t$$

Ainsi

$$L(f')(x) = xL(f)(x) - f(0)$$

□

Proposition

$$\lim_{x\to+\infty} xL(f)(x) = f(0)$$

dém. :
Par le changement de variable $u = xt$, on obtient

$$xL(f)(x) = \int_0^{+\infty} f(s/x)\mathrm{e}^{-s}\,\mathrm{d}s$$

Posons $u(x,s) = f(s/x)\mathrm{e}^{-s}$.
$\forall x > 0, s \mapsto u(x,s)$ est continue par morceaux sur $[0,+\infty[$
$\forall s \in [0,+\infty[, u(x,s) \xrightarrow[x\to+\infty]{} f(0)\mathrm{e}^{-s} = \ell(s)$ avec ℓ continue par morceaux
Enfin

$$\forall (x,s) \in \left]0,+\infty\right[\times \left[0,+\infty\right[, |u(x,s)| \leqslant \|f\|_\infty \mathrm{e}^{-s} = \varphi(s)$$

avec $\varphi : \mathbb{R}^+ \to \mathbb{R}$ continue par morceaux et intégrable.
Par domination

$$xL(f)(x) \xrightarrow[x\to+\infty]{} \int_0^{+\infty} \ell(s)\,\mathrm{d}s = f(0)$$

□

Proposition

Si f admet une limite en $+\infty$ alors

$$\lim_{x\to 0^+} xL(f)(x) = \lim_{t\to+\infty} f(t)$$

dém. :
Ce sont les mêmes calculs avec cette fois-ci

$$\ell(s) = L\mathrm{e}^{-s} \text{ où } L = \lim_{t\to+\infty} f(t)$$

□

21.4.2 Transformée de Fourier

Soit $f : \mathbb{R} \to \mathbb{C}$ continue intégrable.

Définition

On appelle transformée de Fourier de f l'application $\hat{f} : \mathbb{R} \to \mathbb{C}$ définie par

$$\forall x \in \mathbb{R}, \hat{f}(x) = \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-ixt}\,\mathrm{d}t$$

Théorème

L'application $f \mapsto \hat{f}$ est une application linéaire de l'espace $L^1(\mathbb{R}, \mathbb{C})$ vers $L^\infty(\mathbb{R}, \mathbb{C})$.

dém. :
Soit $f : \mathbb{R} \to \mathbb{C}$ continue intégrable.
Posons $u(x,t) = f(t)\mathrm{e}^{-ixt}$ définie sur $\mathbb{R} \times \left]-\infty, +\infty\right[$.
Pour chaque $x \in \mathbb{R}$, la fonction $t \mapsto u(x,t)$ est continue par morceaux sur $\left]-\infty, +\infty\right[$.
Pour chaque $t \in \left]-\infty, +\infty\right[$, la fonction $x \mapsto u(x,t)$ est continue sur $\mathbb{R}$.
On a

$$\forall (x,t) \in \mathbb{R} \times \left]-\infty, +\infty\right[, |u(x,t)| \leqslant |f(t)| = \varphi(t)$$

avec $\varphi : \mathbb{R} \to \mathbb{R}$ continue par morceaux et intégrable.
Par domination, la fonction

$$\hat{f} : x \mapsto \int_{-\infty}^{+\infty} u(x,t)\,\mathrm{d}t$$

est définie et continue sur $\mathbb{R}$.
De plus, elle est bornée car

$$\forall x \in \mathbb{R}, \left|\hat{f}(x)\right| \leqslant \int_{-\infty}^{+\infty} |f(t)|\,\mathrm{d}t$$

Enfin l'application $f \mapsto \hat{f}$ est évidemment linéaire par linéarité de l'intégrale.
□

Remarque On peut aussi montrer que cette application linéaire est continue car

$$\left\|\hat{f}\right\|_\infty \leqslant \|f\|_1$$

On peut encore établir, mais c'est difficile, que cette application est injective.

Théorème

Si pour $n \in \mathbb{N}$, l'application $t \mapsto t^n f(t)$ est intégrable alors $\hat{f}$ est de classe $\mathcal{C}^n$ et

$$\forall k \in \{1, \ldots, n\}, \left(\hat{f}\right)^{(k)}(x) = \int_{-\infty}^{+\infty} (-it)^k f(t)\mathrm{e}^{-ixt}$$

dém. :
Posons $u(x,t) = f(t)\mathrm{e}^{-ixt}$ définie sur $\mathbb{R} \times \left]-\infty, +\infty\right[$.

u admet des dérivée partielles $\dfrac{\partial^k u}{\partial x^k}$ à tout ordre $k \in \{0, \ldots, n\}$ avec

$$\frac{\partial^k u}{\partial x^k}(x,t) = (-it)^k f(t)\mathrm{e}^{-ixt}$$

Pour $k \in \{0, \ldots, n-1\}$
$\forall x \in \mathbb{R}, t \mapsto \dfrac{\partial^k u}{\partial x^k}(x,t)$ continue par morceaux sur $]-\infty, +\infty[$ et intégrable car

$$\left|\frac{\partial^k u}{\partial x^k}(x,t)\right| = |f(t)| + |t|^n\,|f(t)|$$

puisque

$$\forall t \in \mathbb{R}, \left|t^k\right| \leqslant 1 + |t|^n$$

Pour $k = n$
$\forall x \in \mathbb{R}, t \mapsto \dfrac{\partial^n u}{\partial x^n}(x,t)$ est continue par morceaux sur $]-\infty, +\infty[$,
$\forall t \in\,]-\infty, +\infty[\,, x \mapsto \dfrac{\partial^n u}{\partial x^n}(x,t)$ est continue sur $\mathbb{R}$ et
Pour tout $[a,b] \subset \mathbb{R}$, on a

$$\forall (x,t) \in [a,b] \times]-\infty, +\infty[\,, \left|\frac{\partial^n u}{\partial x^n}(x,t)\right| \leqslant |t|^n\,|f(t)| = \varphi(t)$$

avec $\varphi : \mathbb{R} \to \mathbb{R}$ continue par morceaux et intégrable.
Par domination sur tout segment, la fonction $\hat{f}$ est de classe $\mathcal{C}^n$ et

$$\forall k \in \{1, \ldots, n\}\,, \left(\hat{f}\right)^{(k)}(x) = \int_{-\infty}^{+\infty} (-it)^k f(t)\mathrm{e}^{-ixt}$$

□

Exemple Calcul de la transformée de Fourier de $f(t) = \mathrm{e}^{-t^2/2}$.
Puisque $t \mapsto tf(t)$ est intégrable, on a

$$\hat{f}'(x) = -i\int_{-\infty}^{+\infty} \mathrm{t}\mathrm{e}^{-t^2/2}\mathrm{e}^{-ixt}$$

Par intégration par parties

$$\hat{f}'(x) = -x\hat{f}(x)$$

$\hat{f}$ est donc solution sur $\mathbb{R}$ de l'équation différentielle

$$y' + xy = 0$$

C'est une équation différentielle linéaire d'ordre 1 homogène de solution générale $y(x) = \lambda\mathrm{e}^{-x^2/2}$.
Sachant que $\hat{f}(0) = \sqrt{\pi}$ (intégrale de Gauss) on obtient $\lambda = \sqrt{\pi}$ puis

$$\forall x \in \mathbb{R}, \hat{f}(x) = \sqrt{\pi}\mathrm{e}^{-x^2/2}$$

21.4.3 La fonction Γ d'Euler

21.4.3.1 Définition

Lemme

Soit $x \in \mathbb{R}$.
L'intégrale $\displaystyle\int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$ converge si, et seulement si, $x > 0$.

dém. :
La fonction $g : t \mapsto t^{x-1}\mathrm{e}^{-t}$ est définie et continue par morceaux sur $]0, +\infty[$.
Cette fonction est positive donc

$$\int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t \text{ converge si, et seulement si, } g \text{ est intégrable sur }]0, +\infty[$$

Quand $t \to +\infty$, $t^2 g(t) = t^2 t^{x-1}\mathrm{e}^{-t} \to 0$ donc

$$g \text{ est intégrable sur } [1, +\infty[\text{ pour tout } x \in \mathbb{R}$$

Quand $t \to 0^+$, $g(t) \sim t^{x-1} = 1/t^{1-x}$ donc

$$g \text{ est intégrable sur }]0, 1] \text{ si, et seulement si, } 1 - x < 1 \text{ i.e. } x > 0$$

□

Définition

Pour tout $x > 0$, on pose

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$$

Exemple $\Gamma(1) = \displaystyle\int_0^{+\infty} \mathrm{e}^{-t}\,\mathrm{d}t = 1$.

Proposition

$\forall x > 0, \Gamma(x+1) = x\Gamma(x)$.

dém. :
On a $\Gamma(x+1) = \displaystyle\int_0^{+\infty} t^x\mathrm{e}^{-t}\,\mathrm{d}t$
On procède à une intégration par parties avec

$$u(t) = t^x \text{ et } v(t) = -\mathrm{e}^{-t}$$

Les fonctions u et v sont de classe $\mathcal{C}^1$ et uv converge en 0^+ et $+\infty$.
Par intégration par parties impropre

$$\Gamma(x+1) = \left[-t^x\mathrm{e}^{-t}\right]_0^{+\infty} + \int_0^{+\infty} xt^x\mathrm{e}^{-t}\,\mathrm{d}t$$

Ainsi $\Gamma(x+1) = x\Gamma(x)$
□

Exemple Par récurrence
$$\forall n \in \mathbb{N}^\star, \Gamma(n) = (n-1)!$$

21.4.3.2 Continuité

Théorème

La fonction Γ est définie et continue sur $\mathbb{R}^{+\star}$.

dém. :
Considérons $g(x,t) = t^{x-1}\mathrm{e}^{-t}$ définie sur $\mathbb{R}^{+\star} \times \left]0,+\infty\right[$.
$\forall x > 0$, $t \mapsto g(x,t)$ est continue par morceaux sur $\left]0,+\infty\right[$.
$\forall t \in \left]0,+\infty\right[$, $x \mapsto g(x,t)$ est continue sur $\mathbb{R}^{+\star}$.
Soit $[a,b] \subset \mathbb{R}^{+\star}$.
Pour tout $x \in [a,b]$, si $t \geqslant 1$, $t^{x-1} \leqslant t^{b-1}$ et si $t \leqslant 1$, $t^{x-1} \leqslant t^{a-1}$. Dans les deux cas
$$t^{x-1} \leqslant t^{a-1} + t^{b-1}$$
Par suite
$$|g(x,t)| \leqslant (t^{a-1} + t^{b-1})\mathrm{e}^{-t} = \varphi_{a,b}(t)$$
avec $\varphi_{a,b}$ intégrable sur $\left]0,+\infty\right[$ car somme de deux fonctions intégrables.
La fonction Γ est continue sur $[a,b]$ et puisque ceci vaut pour tout $[a,b] \subset \mathbb{R}^{+\star}$, Γ est continue sur $\mathbb{R}^{+\star}$.
□

21.4.3.3 Dérivabilité

Lemme

$\forall x > 0, \forall n \in \mathbb{N}^\star$, $t \mapsto (\ln t)^n t^{x-1}\mathrm{e}^{-t}$ est intégrable sur $\left]0,+\infty\right[$.

dém. :
La fonction $h : t \mapsto (\ln t)^n t^{x-1}\mathrm{e}^{-t}$ est continue par morceaux sur $\left]0,+\infty\right[$.
Quand $t \to +\infty$, $t^2 h(t) = t^2(\ln t)^n t^{x-1}\mathrm{e}^{-t} \to 0$.
Quand $t \to 0^+$, pour $\rho \in \left]0,x\right[$, $t^{1-\rho}h(t) \sim (\ln t)^n t^{x-\rho} \to 0$ avec $1-\rho < 1$
□

Théorème

La fonction Γ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^{+\star}$ et
$$\forall n \in \mathbb{N}, \Gamma^{(n)}(x) = \int_0^{+\infty} (\ln t)^n t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$$

dém. :
$g(x,t) = t^{x-1}\mathrm{e}^{-t} = \mathrm{e}^{(x-1)\ln t}\mathrm{e}^{-t}$.
La fonction $x \mapsto g(x,t)$ est de classe $\mathcal{C}^\infty$ donc, la fonction g admet une dérivée partielle $\dfrac{\partial^n g}{\partial x^n}$ pour tout $n \in \mathbb{N}^\star$ et
$$\frac{\partial^n g}{\partial x^n}(x,t) = (\ln t)^n t^{x-1}\mathrm{e}^{-t}$$
La fonction $\dfrac{\partial^n g}{\partial x^n}$ est continue sur $\mathbb{R}^{+\star} \times \left]0,+\infty\right[$.
Pour tout $[a,b] \subset \left]0,+\infty\right[$ et tout $(x,t) \in [a,b] \times \left]0,+\infty\right[$,
$$\left|\frac{\partial^n g}{\partial x^n}(x,t)\right| \leqslant (\ln t)^n(t^{a-1} + t^{b-1})\mathrm{e}^{-t} = \varphi_{n,a,b}(t)$$

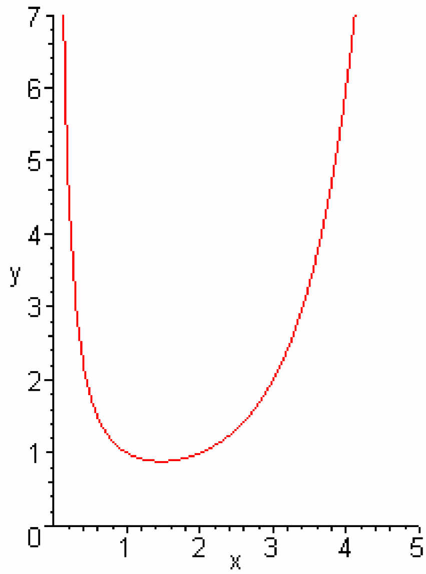
avec $\varphi_{n,a,b}$ intégrable sur $]0,+\infty[$.
Par domination Γ est de classe $\mathcal{C}^\infty$ sur $[a,b]$ et puisque ceci vaut pour $[a,b] \subset \mathbb{R}^{+\star}$, Γ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^{+\star}$.
□

21.4.3.4 Allure

Le signe de $\Gamma'(x) = \displaystyle\int_0^{+\infty} \ln(t) t^{x-1} \mathrm{e}^{-t}\,\mathrm{d}t$ est incertain.
En revanche $\Gamma''(x) = \displaystyle\int_0^{+\infty} (\ln t)^2 t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t > 0$ en tant qu'intégrale d'une fonction positive, continue qui n'est pas la fonction nulle.
On en déduit que Γ' est strictement croissante.
$\Gamma(1) = 1 = \Gamma(2)$ donc par théorème de Rolle il existe $\alpha \in]0,1[$ tel que $\Gamma'(\alpha) = 0$.
Sur $]0,\alpha[$, $\Gamma'(x) < 0$ et Γ est strictement décroissante.
Sir $]\alpha,+\infty[$, $\Gamma'(x) > 0$ et Γ est strictement croissante.
Numériquement $\alpha = 1,46$ à 10^{-2} près et $\Gamma(\alpha) = 0,89$ à 10^{-2} près.
Quand $x \to 0^+$
$\Gamma(x+1) = x\Gamma(x)$ donc $\Gamma(x) = \dfrac{\Gamma(x+1)}{x} \sim \dfrac{1}{x}$ car $\Gamma(x+1) \to \Gamma(1) = 1$.
Quand $x \to +\infty$
Γ est croissante donc la limite de Γ en $+\infty$ existe dans $\mathbb{R} \cup \{+\infty\}$.
Puisque $\Gamma(n+1) = n! \to +\infty$ on peut conclure $\Gamma(x) \to +\infty$.
De plus

$$\frac{\Gamma(x)}{x} = \frac{x-1}{x}\Gamma(x-1) \to +\infty$$

donc Γ présente une branche parabolique verticale.



Chapitre 22

Séries entières

On souhaite étudier les fonctions de la forme

$$x \mapsto \sum_{n=0}^{+\infty} a_n x^n$$

Ce sont des sommes de séries de fonctions, on étudiera le problème de convergence, on observera leur régularité et on verra qu'un grand nombre de fonctions usuelles peuvent s'écrire ainsi.

22.1 Convergence des séries entières

22.1.1 Série entière

Définition

On appelle série entière définie par la suite de coefficients $(a_n) \in \mathbb{C}^{\mathbb{N}}$, la série des fonctions

$$u_n : z \in \mathbb{C} \mapsto a_n z^n$$

Par abus, cette série de fonctions $\sum u_n$ est notée $\sum a_n z^n$.
L'ensemble $\mathcal{D}$ des $z \in \mathbb{C}$ pour lesquels la série numérique $\sum a_n z^n$ converge est appelé domaine de convergence de la série entière et la fonction $S : \mathcal{D} \to \mathbb{C}$ définie par

$$S(z) = \sum_{n=0}^{+\infty} a_n z^n$$

est appelée somme de cette série entière.

Exemple La série entière $\sum a_n z^n$ converge en $z = 0$ et $\sum_{n=0}^{+\infty} a_n 0^n = a_0$.
En effet $0^0 = 1$ et $0^n = 0$ pour $n \in \mathbb{N}^\star$.

Exemple La série entière $\sum z^n$ converge pour tout $z \in \mathbb{C}$ tel que $|z| < 1$ et on a

$$\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$$

Exemple La série entière $\sum \frac{1}{n!} z^n$ converge pour tout $z \in \mathbb{C}$ et par définition

$$\sum_{n=0}^{+\infty} \frac{1}{n!} z^n = \mathrm{e}^z$$

Exemple Si à partir d'un certain rang $a_n = 0$ alors la série entière $\sum a_n z^n$ converge sur $\mathbb{C}$ et sa somme est une fonction polynôme.

Déterminons la forme du domaine de convergence d'une série entière $\sum a_n z^n$.

22.1.2 Rayon de convergence

Lemme

Soit $z_0 \in \mathbb{C}$ tel que la suite $(a_n z_0^n)_{n \in \mathbb{N}}$ soit bornée.
Pour tout $z \in \mathbb{C}$ tel que $|z| < |z_0|$, la série numérique $\sum a_n z^n$ est absolument convergente.

dém. :
Il existe $M \in \mathbb{R}^+$ tel que $|a_n z_0^n| \leqslant M$ pour tout $n \in \mathbb{N}$.
Pour $|z| < |z_0|$, on peut écrire

$$|a_n z^n| = |a z_0^n \times (z/z_0)^n| \leqslant M \left|\frac{z}{z_0}\right|^n$$

Or $|z/z_0| < 1$ donc $\sum |z/z_0|^n$ est absolument convergente et par comparaison $\sum a_n z^n$ l'est aussi.
□

Définition

On appelle rayon de convergence de la série entière $\sum a_n z^n$, le nombre

$$R \underset{\text{déf}}{=} \sup \{r \geqslant 0/(a_n r^n) \text{ est borne}\} \in \mathbb{R}^+ \cup \{+\infty\}$$

Exemple Rayon de convergence de $\sum z^n$.
$\{r \geqslant 0/(r^n) \text{ est borne}\} = [0,1]$ donc $R = 1$.

Exemple Rayon de convergence de $\sum \frac{1}{n!} z^n$.
$\{r \geqslant 0/(r^n/n!) \text{ est borne}\} = \mathbb{R}^+$ donc $R = +\infty$.

Exemple Rayon de convergence $\sum n! z^n$.
$\{r \geqslant 0/(a_n r^n) \text{ est borne}\} = \{0\}$ donc $R = 0$.

22.1.3 Convergence simple

Théorème

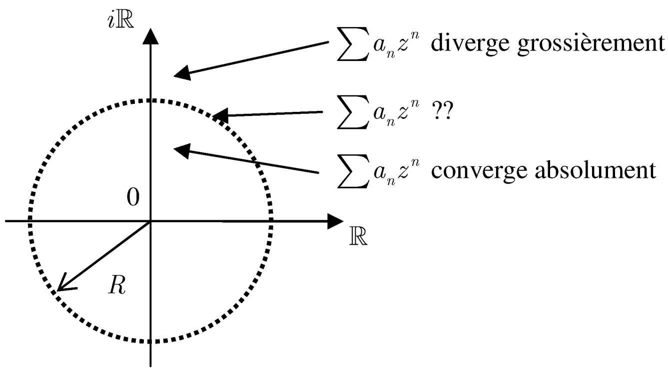
Soit $\sum a_n z^n$ une série entière de rayon de convergence R et $z \in \mathbb{C}$.
Si $|z| < R$ alors la série $\sum a_n z^n$ est absolument convergente.
Si $|z| > R$ alors la série $\sum a_n z^n$ diverge grossièrement (plus précisément la suite $(a_n z^n)$ n'est même pas bornée).

dém. :
Notons $A = \{r \geqslant 0/(a_n r^n) \text{ est borne}\}$ et $R = \sup A$.
Si $|z| < R$ alors $|z|$ ne majore pas A et donc il existe $r > 0$ tel que $|z| < r$ et tel que la suite $(a_n r^n)$ soit bornée. En vertu du lemme d'Abel, la série $\sum a_n z^n$ est absolument convergente.
Si $|z| > R$ alors $|z| \notin A$ et donc $(a_n z^n)$ n'est pas bornée.

□

Corollaire

Soit $\mathcal{D}$ le domaine de convergence d'une série entière de rayon de convergence R.
Si $R = 0$ alors $\mathcal{D} = \{0\}$.
Si $R = +\infty$ alors $\mathcal{D} = \mathbb{C}$.
Si $R \in]0, +\infty[$ alors $D(0, R) \subset \mathcal{D} \subset \overline{D(0, R)}$ en notant $D(0, R) = \{z \in \mathbb{C}/ |z| < R\}$
Sur le cercle de centre 0 et de rayon R, les natures de $\sum a_n z^n$ peuvent être diverses.



Définition

Le disque
$$D(0, R) = \{z \in \mathbb{C}/ |z| < R\}$$
est appelé disque ouvert de convergence de la série entière.

Remarque Sur ce disque, la série entière converge assurément. Elle peut aussi converger en certains points du cercle limite.

22.1.4 Convergence normale

Théorème

> Une série entière de rayon de convergence $R > 0$ converge normalement, et donc uniformément, sur tout disque fermé de centre 0 et de rayon $r < R$.

dém. :
Soit $\sum a_n z^n$ une série entière de rayon de convergence $R > 0$.
Cette série entière est par définition la série des fonctions $u_n : z \mapsto a_n z^n$
Soit $D = \overline{D(0,r)} = \{z \in \mathbb{C} / |z| \leqslant r\}$ avec $r < R$.
Pour tout $z \in D$, $|u_n(z)| \leqslant |a_n| r^n$.
Or il y a convergence absolue de la série $\sum a_n r^n$ donc $\sum u_n$ converge normalement sur D.
□

Corollaire

> La somme d'une série entière de rayon de convergence $R > 0$ est continue sur son disque ouvert de convergence.

dém. :
Par convergence uniforme sur tout compact d'une série de fonctions continues.
□

Exemple La fonction $z \mapsto \mathrm{e}^z$ est continue sur $\mathbb{C}$.

Attention : Il peut ne pas y avoir convergence normale de la série entière sur le disque ouvert de convergence.

Exemple Considérons la série entière $\sum z^n$.
Son rayon de convergence est $R = 1$.
Cependant $\sup\limits_{|z|<1} |z^n| = 1$ et il n'y a donc pas convergence normale sur $D(0,1) = \{z \in \mathbb{C} / |z| < 1\}$.

22.1.5 Calcul du rayon de convergence

*Idée :*On sait
$|z| < R \Rightarrow \sum a_n z^n$ converge
$|z| > R \Rightarrow \sum a_n z^n$ diverge.
Par contraposition :
Si $\sum a_n z^n$ converge alors $|z| \leqslant R$.
Si $\sum a_n z^n$ diverge alors $R \leqslant |z|$.

22.1.5.1 Exploitation de la règle de d'Alembert

Rappel :
Soit $\sum u_n$ une série numérique à termes non nuls à partir d'un certain rang.
On suppose

$$\left|\frac{u_{n+1}}{u_n}\right| \to \ell \in \mathbb{R}^+ \cup \{+\infty\}$$

Si $\ell < 1$ alors $\sum u_n$ est absolument convergente.
Si $\ell > 1$ alors $\sum u_n$ est grossièrement divergente.
En exploitant ce critère, on peut étudier la convergence de $\sum a_n z^n$ et préciser le rayon de convergence R.

Exemple Rayon de convergence de

$$\sum (-1)^{\frac{n(n+1)}{2}}(n-1)2^n z^n$$

Soit $z \in \mathbb{C}$.
Posons $u_n(z) = (-1)^{\frac{n(n+1)}{2}}(n-1)2^n z^n$.
Pour $z \neq 0$ et $n \geqslant 2$, on a $u_n \neq 0$.

$$\left|\frac{u_{n+1}(z)}{u_n(z)}\right| = \frac{n}{n-1}\frac{2^{n+1}}{2^n}\left|\frac{z^{n+1}}{z^n}\right| \to 2\,|z|$$

Si $|z| < 1/2$ alors $\sum u_n(z)$ est absolument convergente.
Si $|z| > 1/2$ alors $\sum u_n(z)$ diverge grossièrement.
On en déduit $R = 1/2$.

Exemple Rayon de convergence de

$$\sum \frac{1}{(2n)!}z^n$$

Posons $u_n(z) = \dfrac{1}{(2n)!}z^n$ pour $z \in \mathbb{C}^\star$.

$$\left|\frac{u_{n+1}(z)}{u_n(z)}\right| = \frac{1}{(2n+2)(2n+1)}\,|z| \to 0$$

Pour tout $z \in \mathbb{C}^\star$, $\sum u_n(z)$ est absolument convergente (et aussi pour $z = 0$) donc $R = +\infty$.

Exemple Rayon de convergence de

$$\sum \frac{n-1}{n^2+1}z^n$$

$u_n(z) = \dfrac{n+1}{n^2+1}z^n$ avec $z \neq 0$.
$\left|\dfrac{u_{n+1}(z)}{u_n(z)}\right| \sim \dfrac{1/(n+1)}{1/n}\,|z| \to |z|$.
On en déduit $R = 1$.

Remarque Plus généralement, soit $F \in \mathbb{C}(X)\setminus\{0\}$, le rayon de convergence de $\sum F(n)z^n$ vaut 1 car pour $z \neq 0$

$$\frac{\left|F(n+1)z^{n+1}\right|}{|F(n)z^n|} = \left|\frac{F(n+1)}{F(n)}\right|\,|z| \to |z|$$

en effet

$$F(n) = \frac{a_p n^p + \cdots}{b_q n^q + \cdots} \sim \frac{a_p n^p}{b_q n^q} = \lambda n^{p-q}$$

donc

$$\frac{F(n+1)}{F(n)} \sim \frac{\lambda (n+1)^{p-q}}{\lambda n^{p-q}} \to 1$$

22.1.5.2 Cas des séries lacunaires

Remarque La série de fonctions $\sum a_n z^{2n}$ peut se comprendre comme une série entière. En effet

$$\sum a_n z^{2n} = \sum b_n z^n$$

avec

$$b_{2p} = a_p \text{ et } b_{2p+1} = 0$$

Le rayon de convergence d'une telle série peut souvent se déterminer par la démarche précédente.

Exemple Rayon de convergence de

$$\sum \frac{(-1)^n}{n+1} z^{2n+1}$$

Soit $z \neq 0$.
$\sum \frac{(-1)^n}{n+1} z^{2n+1} = \sum u_n(z)$ avec $u_n(z) = \frac{(-1)^n}{n+1} z^{2n+1} \neq 0$.

$$\left| \frac{u_{n+1}(z)}{u_n(z)} \right| = \frac{n+1}{n+2} |z|^2 \to |z|^2$$

Si $|z| < 1$ alors $\sum \frac{(-1)^n}{n+1} z^{2n+1}$ est absolument convergente.
Si $|z| > 1$ alors $\sum \frac{(-1)^n}{n+1} z^{2n+1}$ est grossièrement divergente.
On en déduit $R = 1$

Exemple Rayon de convergence de

$$\sum \binom{2n}{n} z^{3n}$$

Posons $u_n(z) = \binom{2n}{n} z^{3n} = \frac{(2n)!}{(n!)^2} z^{3n}$ pour $z \in \mathbb{C}^\star$.

$$\left| \frac{u_{n+1}(z)}{u_n(z)} \right| = \frac{(2n+2)(2n+1)}{(n+1)^2} \left| \frac{z^{3(n+1)}}{z^{3n}} \right| = 2\frac{2n+1}{n+1} |z|^3 \to 4 |z|^3$$

On en déduit $R = \sqrt[3]{1/4}$.

Remarque La démarche exploitant le critère de d'Alembert possède deux inconvénients majeurs :
- elle ne possède pas de réciproque ;
- il se peut que le rapport $|u_{n+1}/u_n|$ n'ait pas de limite...

Pour déterminer un rayon de convergence, on procède alors généralement par double inégalité comme on le verra par exemple pour la série entière $\sum \sin(n) z^n$

22.1.5.3 Par comparaison

Soit R_a et R_b les rayons de convergence de deux séries entières $\sum a_n z^n$ et $\sum b_n z^n$.

Théorème

Si $a_n = O(b_n)$ alors $R_a \geqslant R_b$.

dém. :
Soit $z \in \mathbb{C}$ tel que $|z| < R_b$. La série $\sum b_n z^n$ est absolument convergente et par comparaison $\sum a_n z^n$ l'est aussi. Puisque $\sum a_n z^n$ converge pour tout $|z| < R_b$, on a nécessairement $R_b \leqslant R_a$.
□

Corollaire

1) Si $|a_n| \leqslant |b_n|$ alors $R_a \geqslant R_b$
2) Si $a_n = o(b_n)$ alors $R_a \geqslant R_b$.
3) Si $a_n \sim b_n$ alors $R_a = R_b$.

Exemple Les séries entières $\sum a_n z^n \sum |a_n| z^n$ ont même rayon de convergence.

Exemple Rayon R de convergence de $\sum \sin(n) z^n$.
On a $|a_n| \leqslant 1$, or $\sum z^n$ est de rayon de convergence 1, donc $R \geqslant 1$.
De plus (a_n) ne tend pas vers 0 donc $\sum a_n z^n$ diverge pour $z = 1$ et donc $R \leqslant 1$.
On peut conclure $R = 1$.

Remarque Plus généralement, si (a_n) est bornée et ne tend pas vers 0 alors $\sum a_n z^n$ a un rayon de convergence égal à 1.

22.1.5.4 Rayon de $\sum n a_n z^n$

Théorème

Les séries entières $\sum a_n z^n$ et $\sum n a_n z^n$ ont même rayon de convergence.

dém. :
Notons R et R' les deux rayons de convergence de ces séries entières.
Puisque $a_n = o(n a_n)$, on a déjà $R \geqslant R'$.
Inversement, soit $z \in \mathbb{C}$ tel que $|z| < R$. Introduisons ρ tel que $|z| < \rho < R$, on a

$$n a_n z^n = n \left(z/\rho\right)^n a_n \rho^n = o(a_n \rho^n)$$

Or il y a convergence absolue de $\sum a_n \rho^n$, donc $\sum n a_n z^n$ converge absolument.
Ainsi $R' \geqslant R$ puis il y a égalité.
□

Exemple Montrons que pour tout $\alpha \in \mathbb{R}$, $\sum a_n z^n$ et $\sum n^\alpha a_n z^n$ ont même rayon de convergence.
Par récurrence, on obtient aisément l'égalité des rayons de convergence de $\sum a_n z^n$ et $\sum n^k a_n z^n$ pour $k \in \mathbb{Z}$.
En considérant $k = \lfloor \alpha \rfloor$, on a $n^k |a_n| \leqslant n^\alpha |a_n| \leqslant n^{k+1} |a_n|$ ce qui permet de conclure.

22.1.6 Somme et produit de séries entières

22.1.6.1 Somme

Définition

On appelle somme des séries entières $\sum a_n z^n$ et $\sum b_n z^n$ la série entière $\sum (a_n + b_n) z^n$.

Théorème

Si R_a et R_b sont les rayons de convergence des séries entières $\sum a_n z^n$ et $\sum b_n z^n$ alors le rayon de convergence R de la série entière somme $\sum (a_n + b_n) z^n$ vérifie

$$R \geqslant \min(R_a, R_b)$$

De plus, pour $|z| < \min(R_a, R_b)$,

$$\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n$$

dém. :
On remarque $a_n z^n + b_n z^n = (a_n + b_n) z^n$.
Soit $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$.
Les séries numériques $\sum a_n z^n$ et $\sum b_n z^n$ convergent absolument donc par somme la série numérique $\sum (a_n + b_n) z^n$ converge aussi et de plus

$$\sum_{n=0}^{+\infty} (a_n + b_n) z^n = \sum_{n=0}^{+\infty} a_n z^n + \sum_{n=0}^{+\infty} b_n z^n$$

Puisque $\sum (a_n + b_n) z^n$ converge pour tout $|z| < \min(R_a, R_b)$, on a

$$\min(R_a, R_b) \leqslant R$$

□

Remarque Il est possible que $R > \min(R_a, R_b)$, par exemple quand $b_n = -a_n$.

Proposition

Si $R_a \neq R_b$ alors $R = \min(R_a, R_b)$.

dém. :
Quitte à échanger supposons $R_a < R_b$.
On sait déjà que $R \geqslant R_a$.
Pour $R_a < |z| < R_b$, $\sum a_n z^n$ diverge alors que $\sum b_n z^n$ converge donc $\sum (a_n + b_n) z^n$ diverge.
On en déduit $R = R_a = \min(R_a, R_b)$.
□

Exemple Soit $\sum a_n z^n$ une série entière de rayon de convergence R.
Considérons $\sum a_{2p} z^{2p}$ et $\sum a_{2p+1} z^{2p+1}$ de rayons de convergence R' et R''.
Montrons

$$R = \min(R', R'')$$

Remarquons
$\sum a_{2p} z^{2p} = \sum b_n z^n$ avec $b_{2p} = a_{2p}$ et $b_{2p+1} = 0$
$\sum a_{2p+1} z^{2p+1} = \sum c_n z^n$ avec $c_{2p} = 0$ et $c_{2p+1} = a_{2p+1}$
D'une part $a_n = b_n + c_n$ pour tout $n \in \mathbb{N}$ donc $\sum a_n z^n$ est la somme des séries entières $\sum a_{2p} z^{2p}$ et $\sum a_{2p+1} z^{2p+1}$ puis $R \geqslant \min(R', R'')$.
D'autre part, $|b_n|, |c_n| \leqslant |a_n|$ donc $R', R'' \geqslant R$ puis $\min(R', R'') \geqslant R$.
Finalement $R = \min(R', R'')$.

22.1.6.2 Produit

Définition

On appelle produit des séries entières $\sum a_n z^n$ et $\sum b_n z^n$ la série entière $\sum c_n z^n$ avec $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.

Théorème

Si R_a et R_b sont les rayons de convergence des séries entières $\sum a_n z^n$ et $\sum b_n z^n$ alors le rayon de convergence R de la série entière produit $\sum c_n z^n$ vérifie

$$R \geqslant \min(R_a, R_b)$$

De plus, pour $|z| < \min(R_a, R_b)$, on a

$$\sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n\right)\left(\sum_{n=0}^{+\infty} b_n z^n\right)$$

dém. :
On remarque

$$c_n z^n = \sum_{k=0}^{n} (a_k z^k)(b_{n-k} z^{n-k})$$

Ainsi la série numérique $\sum c_n z^n$ est le produit de Cauchy des séries numériques $\sum a_n z^n$ et $\sum b_n z^n$.
Pour $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$, $\sum a_n z^n$ et $\sum b_n z^n$ sont absolument convergentes donc par produit de Cauchy $\sum c_n z^n$ est absolument convergente et de plus

$$\sum_{n=0}^{+\infty} c_n z^n = \left(\sum_{n=0}^{+\infty} a_n z^n\right)\left(\sum_{n=0}^{+\infty} b_n z^n\right)$$

Puisque $\sum c_n z^n$ converge pour tout $|z| < \min(R_a, R_b)$, on a $\min(R_a, R_b) \leqslant R$.
□

Exemple Soit $\sum a_n z^n$ une série entière de rayon de convergence $R \geqslant 1$.

Etudions la série entière $\sum S_n z^n$ avec $S_n = \sum_{k=0}^{n} a_k$.

Pour tout $n \in \mathbb{N}$, $S_n = \sum_{k=0}^{n} a_k \times 1$ donc $\sum S_n z^n$ est le produit des séries entières $\sum a_n z^n$ et $\sum z^n$.

Par suite $\sum S_n z^n$ est de rayon de convergence $\geqslant \min(R, 1) = 1$ et pour tout $z \in \mathbb{C}$ tel que $|z| < 1$,

$$\sum_{n=0}^{+\infty} S_n z^n = \frac{1}{1-z} \sum_{n=0}^{+\infty} a_n z^n$$

22.2 Série entière d'une variable réelle

Désormais, nous étudions $z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ pour $z \in \mathbb{R}$, on préfère alors noter la variable x (ou t).

22.2.1 Particularisation

Soit $\sum a_n x^n$ une série entière de rayon de convergence $R > 0$.
Pour tout $x \in \left]-R, R\right[$, $\sum a_n x^n$ converge absolument.
Pour tout $|x| > R$: $\sum a_n x^n$ diverge grossièrement.
Pour $x = R$ ou $x = -R$: ça dépend.

Définition

L'intervalle $\left]-R, R\right[$ est appelé intervalle ouvert de convergence de la série $\sum a_n x^n$.

Définition

L'ensemble I des x pour lesquels la série numérique converge vérifie

$$\left]-R, R\right[\subset I \subset [-R, R]$$

on l'appelle intervalle de convergence de la série entière étudiée.

Théorème

La série entière $\sum a_n x^n$ converge normalement sur tout segment inclus dans $\left]-R, R\right[$.

dém. :
Car $\sum a_n z^n$ converge normalement sur tout disque fermé inclus dans le disque ouvert $D(0,R)$.
□

Corollaire

> La fonction $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est continue sur $]-R,R[$.

Exemple Etudions

$$S : x \mapsto \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{2n+1} x^n$$

S est une série entière de rayon de convergence $R = 1$.
S est donc assurément définie et continue sur $]-1,1[$.
Etude en $x = -1$
$\sum \frac{(-1)^{n-1}}{2n+1}(-1)^n = \sum \frac{-1}{2n+1}$ diverge.
S n'est pas définie en -1.
Etude en $x = 1$
$\sum \frac{(-1)^{n-1}}{2n+1} 1^n = \sum \frac{(-1)^{n-1}}{2n+1}$ est une série alternée convergente en vertu du critère spécial.
S est définie en 1.
Continuité en 1
Considérons $u_n : [0,1] \to \mathbb{R}$ définie par $u_n(x) = \frac{(-1)^{n-1}}{2n+1} x^n$ avec $n \geqslant 1$.
Les fonctions u_n sont continues.
$\sum u_n(x)$ converge par le critère spécial.

$$|R_n(x)| \leqslant |u_{n+1}(x)| \leqslant \frac{1}{2n+1} x^{n+1} \leqslant \frac{1}{2n+1} \to 0$$

Il y a convergence uniforme sur $[0,1]$ donc S est continue sur $[0,1]$.

22.2.2 Intégration

Définition

> On appelle série entière primitive de $\sum a_n x^n$ la série entière $\sum \frac{a_n}{n+1} x^{n+1}$.

Proposition

> $\sum a_n x^n$ et $\sum \frac{a_n}{n+1} x^{n+1}$ ont même rayon de convergence.

dém. :
Le rayon de convergence de $\sum \frac{a_n}{n+1} x^{n+1}$ est le même que celui de

$$\sum (n+1) \times \frac{a_n}{n+1} x^{n+1} = \sum a_n x^{n+1}$$

qui est aussi celui de $\sum a_n x^n$.
□

Théorème

Si $\sum a_n x^n$ est une série entière de rayon de convergence $R > 0$ alors

$$x \mapsto \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

est sur $]-R, R[$ la primitive s'annulant en 0 de

$$x \mapsto \sum_{n=0}^{+\infty} a_n x^n$$

dém. :

Sur $]-R, R[$, la primitive s'annulant en 0 de la fonction continue $x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est

$$x \mapsto \int_0^x \sum_{n=0}^{+\infty} a_n t^n \,\mathrm{d}t$$

Pour tout $x \in]-R, R[$, la série entière converge uniformément sur le segment d'extrémités 0 et x. On peut donc intégrer terme à terme et affirmer

$$\int_0^x \sum_{n=0}^{+\infty} a_n t^n \,\mathrm{d}t = \sum_{n=0}^{+\infty} \int_0^x a_n t^n \,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

□

Exemple On sait que pour $x \in]-1, 1[$

$$\sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}$$

Par intégration de série entière, on obtient

$$\forall x \in]-1, 1[\,, \sum_{n=0}^{+\infty} \frac{x^{n+1}}{n+1} = \int_0^x \frac{\mathrm{d}t}{1-t} = -\ln(1-x)$$

On peut retenir la formule

$$\forall x \in]-1, 1[\,, -\ln(1-x) = \sum_{n=1}^{+\infty} \frac{x^n}{n}$$

22.2.3 Dérivation

Définition

On appelle série entière dérivée d'une série entière $\sum a_n x^n$ la série entière

$$\sum_{n \geqslant 1} n a_n x^{n-1} = \sum (n+1) a_{n+1} x^n$$

Proposition

$\sum a_n x^n$ et $\sum_{n\geqslant 1} n a_n x^{n-1}$ ont même rayon de convergence.

dém. :
$\sum a_n x^n$ a le rayon de convergence de $\sum n a_n x^n$ qui est aussi celui de $\sum_{n\geqslant 1} n a_n x^{n-1}$

□

Proposition

Si $\sum a_n x^n$ est une série entière de rayon de convergence $R > 0$ alors sa somme $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$
est de classe $\mathcal{C}^1$ sur $]-R, R[$ et

$$\forall x \in \left]-R, R\right[, S'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n$$

dém. :
Introduisons $u_n : x \mapsto a_n x^n$.
Les fonctions u_n sont de classe $\mathcal{C}^1$, $\sum u_n$ converge simplement sur $]-R, R[$ et $\sum u_n'$ converge normalement sur tout segment inclus dans $]-R, R[$ car la série entière dérivée a pour rayon de convergence R.

□

Théorème

Si $\sum a_n x^n$ est une série entière de rayon de convergence $R > 0$ alors sa somme $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est de classe $\mathcal{C}^\infty$ sur $]-R, R[$ et ses dérivées successives s'obtiennent en dérivant terme à terme :

$$\forall p \in \mathbb{N}, \forall x \in \left]-R, R\right[, S^{(p)}(x) = \sum_{n=p}^{+\infty} n(n-1)\dots(n-p+1) a_n x^{n-p}$$

ou encore

$$\forall p \in \mathbb{N}, \forall x \in \left]-R, R\right[, S^{(p)}(x) = \sum_{n=0}^{+\infty} (n+p)(n+p-1)\dots(n+1) a_{n+p} x^n$$

Attention : En $\pm R$, on ne peut rien dire à partir de la seule connaissance du rayon de convergence.

22.2.4 Expression des coefficients d'une série entière

Théorème

Si $\sum a_n x^n$ est une série entière de rayon de convergence $R > 0$ et de somme S alors

$$\forall n \in \mathbb{N}, a_n = \frac{S^{(n)}(0)}{n!}$$

dém. :
S est de classe $\mathcal{C}^\infty$ sur $]-R, R[$ et

$$S^{(p)}(x) = \sum_{n=0}^{+\infty} (n+p)(n+p-1)\dots(n+1)a_{n+p}x^n$$

En particularisant en $x = 0$, on obtient $S^{(p)}(0) = p!a_p$.
□

Corollaire

Soit $\sum a_n x^n$ et $\sum b_n x^n$ sont deux séries entières de rayons de convergence $R_a, R_b > 0$.
S'il existe un voisinage de 0 sur lequel

$$\sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} b_n x^n$$

alors

$$\forall n \in \mathbb{N}, a_n = b_n$$

dém. :
Notons $S_a : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ et $S_b : x \mapsto \sum_{n=0}^{+\infty} b_n x^n$.
Par hypothèse, il existe $r > 0$ tel que

$$\forall x \in \left]-r, r\right[, S_a(x) = S_b(x)$$

On a alors

$$\forall p \in \mathbb{N}, \forall x \in \left]-r, r\right[, S_a^{(p)}(x) = S_b^{(p)}(x)$$

donc

$$a_p = \frac{S_a^{(p)}(0)}{p!} = \frac{S_b^{(p)}(0)}{p!} = b_p$$

□

Exemple Soit $\sum a_n x^n$ une série entière de rayon de convergence $R > 0$ et de somme

$$S : x \in]-R, R[\mapsto \sum_{n=0}^{+\infty} a_n x^n$$

Montrons

$$S \text{ est paire si, et seulement si, } \forall p \in \mathbb{N}, a_{2p+1} = 0$$

($\Leftarrow$) Supposons $\forall p \in \mathbb{N}, a_{2p+1} = 0$.

$S(x) = \sum_{n=0}^{+\infty} a_n x^n = \sum_{p=0}^{+\infty} a_{2p} x^{2p}$ donc S est une fonction paire définie sur $]-R, R[$ ou $[-R, R]$.

($\Rightarrow$) Supposons S paire.

Pour tout $x \in]-R, R[$, $S(x) = S(-x)$ donc $\sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} (-1)^n a_n x^n$.

Par identification des coefficients de séries entières de rayons de convergence > 0, on a pour tout $n \in \mathbb{N}$, $a_n = (-1)^n a_n$ et donc

$$\forall p \in \mathbb{N}, a_{2p+1} = 0$$

De même, on montre :

$$S \text{ est impaire si, et seulement si, } \forall p \in \mathbb{N}, a_{2p} = 0$$

22.3 Développements en série entière

I désigne un intervalle de $\mathbb{R}$ qui est voisinage de 0.
Soit $r \in \mathbb{R}^{+\star} \cup \{+\infty\}$ tel que $]-r, r[\subset I$.

22.3.1 Fonctions développables en série entière

Définition

On dit que $f : I \to \mathbb{C}$ est développable en série entière sur $]-r, r[$ s'il existe une série entière $\sum a_n x^n$ telle que

$$\forall x \in]-r, r[, \ \sum a_n x^n \text{ converge et } f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

Remarque Cette série entière est nécessairement de rayon de convergence $R \geqslant r$

Exemple Considérons $f : x \mapsto \dfrac{1}{1-x}$ définie sur $]-\infty, 1[$

f est développable en série entière sur $]-1, 1[$ car on sait

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$$

et donc $f(x)$ apparaît sur $]-1, 1[$ comme égale à la somme d'une série entière convergente.

Exemple Considérons $f : x \mapsto \dfrac{1}{1+x^2}$ définie sur $\mathbb{R}$.

f est développable en série entière sur $]-1,1[$ car

$$\frac{1}{1+x^2} \underset{u=-x^2}{=} \frac{1}{1-u} \underset{|u|<1}{=} \sum_{n=0}^{+\infty} u^n = \sum_{n=0}^{+\infty} (-1)^n x^{2n}$$

et donc $f(x)$ apparaît sur $]-1,1[$ comme égale à la somme d'une série entière convergente.

Exemple $x \mapsto \mathrm{e}^x$ est développable en série entière sur $\mathbb{R}$ avec

$$\mathrm{e}^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$$

Définition

On dit que $f : I \to \mathbb{C}$ est développable en série entière en 0 s'il existe $r > 0$ telle que f est développable en série entière sur $]-r,r[$.

Exemple Les fonctions $x \mapsto \dfrac{1}{1-x}, \dfrac{1}{1+x^2}, \mathrm{e}^x$ sont développables en série entière en 0.

22.3.2 Série de Taylor

Définition

On appelle série de Taylor (en 0) d'une fonction $f : I \to \mathbb{C}$ de classe $\mathcal{C}^\infty$ la série entière

$$\sum \frac{f^{(n)}(0)}{n!} x^n$$

Théorème

Si $f : I \to \mathbb{C}$ est développable en série entière sur $]-r,r[$ avec

$$\forall x \in \,]-r,r[\,, f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

alors f est de classe $\mathcal{C}^\infty$ sur $]-r,r[$ et

$$\forall n \in \mathbb{N}, a_n = \frac{f^{(n)}(0)}{n!}$$

Autrement dit, il n'y a qu'une seule série entière qui puisse correspondre à f, à savoir sa série de Taylor.

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$$

dém. :
Il existe une série entière $\sum a_n x^n$ de rayon de convergence $R \geqslant r$ tel que sur $]-r, r[$

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

Considérons alors la fonction $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$.
La fonction S est définie et de classe $\mathcal{C}^\infty$ sur $]-R, R[$ donc sur $]-r, r[$
Puisque f et S coïncident sur $]-r, r[$, f est de classe $\mathcal{C}^\infty$ sur $]-r, r[$.
De plus, pour tout $n \in \mathbb{N}$,

$$a_n = \frac{S^{(n)}(0)}{n!} = \frac{f^{(n)}(0)}{n!}$$

donc la série entière introduite n'est autre que la série de Taylor de f.
□

Remarque Une fonction qui n'est pas de classe $\mathcal{C}^\infty$ sur $]-r, r[$ ne peut y être développable en série entière.

Remarque Si f est de classe $\mathcal{C}^\infty$, on peut étudier si f est développable en série entière en vérifiant si

$$\sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k \xrightarrow[n \to +\infty]{} f(x)$$

On peut pour cela exploiter l'inégalité de Taylor-Lagrange ou l'égalité de Taylor avec reste intégral.

Exemple Soit $f : [-1, 1] \to \mathbb{C}$ de classe $\mathcal{C}^\infty$ et vérifiant

$$\left\| f^{(n)} \right\|_\infty \leqslant M K^n n!$$

avec $M \in \mathbb{R}^+$ et $K > 0$. Montrons que f est développable en série entière en 0.
Pour tout $x \in [-1, 1]$,

$$\left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k \right| \leqslant \frac{\left\| f^{(n+1)} \right\|_\infty}{(n+1)!} \left| x^{n+1} \right| \leqslant \frac{\left\| f^{(n+1)} \right\|_\infty \left| x^{n+1} \right|}{(n+1)!} \leqslant M K^{n+1} \left| x \right|^{n+1}$$

Pour $|x| < r = \min(1, 1/|K|)$ on a $(K\,|x|)^{n+1} \to 0$ et donc

$$\sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k \to f(x)$$

Ainsi la série $\sum \frac{f^{(n)}(0)}{n!} x^n$ converge et

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$$

La fonction f s'écrit sur $]-r, r[$ comme égale à la somme d'une série entière convergente, elle est donc développable en série entière sur $]-r, r[$.

Attention : Il existe des fonctions de classe $\mathcal{C}^\infty$ qui ne sont pas développables en série entière !

22.3.3 Opérations sur les fonctions développables en série entière

Théorème

Si $f, g : I \to \mathbb{C}$ sont développables en série entière sur $]-r, r[$ alors pour tout $\lambda \in \mathbb{C}$, λf, $f + g$ et fg sont développables en série entière sur $]-r, r[$.

dém. :
Il existe des séries entières $\sum a_n x^n$ et $\sum b_n x^n$ de rayons de convergence $R_a, R_b \geqslant r$ telles que sur $]-r, r[$,

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n \text{ et } g(x) = \sum_{n=0}^{+\infty} b_n x^n$$

Pour tout $x \in]-r, r[$, on a

$$(\lambda f)(x) = \lambda f(x) = \lambda \sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} \lambda a_n x^n$$

La fonction λf est sur $]-r, r[$ somme d'une série entière convergente, elle est donc développable en série entière.
Pour tout $x \in]-r, r[$, on a

$$(f + g)(x) = f(x) + g(x) = \sum_{n=0}^{+\infty} a_n x^n + \sum_{n=0}^{+\infty} b_n x^n = \sum_{n=0}^{+\infty} (a_n + b_n) x^n$$

La fonction $f + g$ est sur $]-r, r[$ somme d'une série entière convergente, elle est donc développable en série entière.
Enfin, par produit de Cauchy de séries absolument convergentes

$$(fg)(x) = f(x)g(x) = \left(\sum_{n=0}^{+\infty} a_n x^n\right)\left(\sum_{n=0}^{+\infty} b_n x^n\right) = \sum_{n=0}^{+\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) x^n$$

La fonction fg est sur $]-r, r[$ somme d'une série entière convergente, elle est donc développable en série entière.
□

Exemple Pour tout $x \in \mathbb{R}$, $\mathrm{e}^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$ et $\mathrm{e}^{-x} = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} x^n$ donc les fonctions ch et sh sont développables en série entière sur $\mathbb{R}$ avec

$$\mathrm{ch} x = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^{2n} \text{ et } \mathrm{sh} x = \sum_{n=0}^{+\infty} \frac{1}{(2n+1)!} x^{2n+1}$$

Théorème

Si $f : I \to \mathbb{C}$ est développable en série entière sur $]-r, r[$ alors $\bar{f}$, $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ l'est aussi.

dém. :
Si $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ sur $]-r, r[$ alors $\overline{f(x)} = \sum_{n=0}^{+\infty} \overline{a_n} x^n$, $\text{Re}(f(x)) = \sum_{n=0}^{+\infty} \text{Re}(a_n) x^n$, $\text{Im}(f(x)) = \sum_{n=0}^{+\infty} \text{Im}(a_n) x^n$.
Les fonction s $\bar{f}$, $\text{Re}(f)$ et $\text{Im}(f)$ sont donc développables en série entière sur $]-r, r[$ car sommes de séries entières convergentes sur cet intervalle.
□

Exemple Pour tout $x \in \mathbb{R}$,

$$\mathrm{e}^{ix} = \sum_{n=0}^{+\infty} \frac{i^n}{n!} x^n$$

donc les fonctions $\cos$ et $\sin$ sont développables en série entière sur $\mathbb{R}$ avec

$$\cos x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} x^{2n} \text{ et } \sin x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$$

Théorème

Si $f : I \to \mathbb{C}$ est développable en série entière sur $]-r, r[$ alors ses dérivées successives le sont aussi.

dém. :
Si $f(x) = \sum_{n=0}^{+\infty} a_n x^n$ sur $]-r, r[$ alors par dérivation de la somme d'une série entière

$$f'(x) = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n$$

et donc f' est développable en série entière sur $]-r, r[$. Il en est de même de $f'', \ldots, f^{(n)}, \ldots$.
□

Exemple On sait

$$\forall x \in]-1, 1[\, , \frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$$

Par dérivation d'un développement en série entière

$$\forall x \in]-1, 1[\, , \frac{1}{(1-x)^2} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right) = \sum_{n=0}^{+\infty} (n+1) x^n$$

Théorème

Si $f : I \to \mathbb{C}$ est développable en série entière sur $]-r, r[$ avec

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n$$

alors les primitives F de f le sont aussi avec

$$F(x) = F(0) + \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$$

dém. :

On sait que $x \mapsto \sum_{n=0}^{+\infty} \frac{a_n}{n+1} x^{n+1}$ est la primitive s'annulant en 0 de $x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ donc F ne diffère de cette fonction sur $]-r, r[$ que de la valeur $F(0)$.

□

Exemple $x \mapsto \ln(1+x)$ est définie sur $]-1, +\infty[$ et

$$\frac{\mathrm{d}}{\mathrm{d}x} (\ln(1+x)) = \frac{1}{1+x}$$

Or

$$\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n \text{ sur }]-1, 1[$$

Par intégration d'un développement en série entière, on a

$$\ln(1+x) = \ln(1) + \sum_{n=0}^{+\infty} \frac{(-1)^n}{n+1} x^{n+1} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} x^n \text{ sur }]-1, 1[$$

Par une étude de série de fonctions, on peut établir la définition et la continuité du second membre en $x = 1$. Cela permet de prolonger l'identité en $x = 1$.

Exemple $x \mapsto \arctan x$ est définie sur $\mathbb{R}$ et

$$\frac{\mathrm{d}}{\mathrm{d}x}(\arctan x) = \frac{1}{1+x^2} = \sum_{n=0}^{+\infty} (-1)^n x^{2n} \text{ sur }]-1, 1[$$

Par intégration d'un développement en série entière, on obtient

$$\arctan x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \text{ sur }]-1, 1[$$

Comme ci-dessus, on peut prolonger cette identité à $x = 1$ et $x = -1$.

22.3.4 Développement du binôme $(1+x)^\alpha$

Théorème

Pour tout $\alpha \in \mathbb{R}$, la fonction $x \mapsto (1+x)^\alpha$ est développable en série entière sur $]-1,1[$ et

$$(1+x)^\alpha = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n$$

dém. :
Posons

$$a_n = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}$$

et étudions la série entière $\sum a_n x^n$
On a

$$a_0 = 1,\, a_1 = \alpha,\, a_2 = \frac{\alpha(\alpha-1)}{2}, \dots, a_{n+1} = \frac{\alpha-n}{n+1} a_n$$

Déterminons le rayon de convergence R de la série entière $\sum a_n x^n$.
Cas $\alpha \in \mathbb{N}$
Pour $n > \alpha$, $a_n = 0$ et donc $R = +\infty$ (polynôme)
Cas $\alpha \notin \mathbb{N}$
Pour tout $n \in \mathbb{N}$, $a_n \neq 0$
Pour $x \in \mathbb{R}^\star$, considérons $u_n = a_n x^n$
et $\left|\frac{u_{n+1}}{u_n}\right| = \frac{|\alpha-n|}{n+1}|x| \to |x|$ donc $R = 1$.
Dans les deux cas, la fonction

$$S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$$

est définie et de classe $\mathcal{C}^\infty$ sur $]-1,1[$ et

$$S'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n$$

donc

$$S'(x) = \sum_{n=0}^{+\infty} (n+1) a_{n+1} x^n = \sum_{n=0}^{+\infty} (\alpha-n) a_n x^n$$

puis

$$S'(x) = \alpha \sum_{n=0}^{+\infty} a_n x^n - x \sum_{n=1}^{+\infty} n a_n x^{n-1} = \alpha S(x) - x S'(x)$$

La fonction S est donc solution sur $]-1,1[$ de l'équation différentielle

$$(1+x)' + \alpha y = 0$$

de solution générale $y(x) = \lambda(1+x)^\alpha$.
Il existe donc $\lambda \in \mathbb{R}$ tel que pour tout $x \in]-1,1[$

$$S(x) = \lambda(1+x)^\alpha$$

Or $\lambda = S(0) = a_0 = 1$ donc
$$S(x) = (1+x)^\alpha$$

□

Exemple Cas $\alpha \in \mathbb{N}$
Si $\alpha = p \in \mathbb{N}$
$$(1+x)^p = \sum_{k=0}^{+\infty} \frac{p(p-1)\dots(p-k+1)}{k!} x^k = \sum_{k=0}^{p} \binom{p}{k} x^k$$
On retrouve la formule du binôme.

Exemple Cas $\alpha \in \mathbb{Z}\backslash\mathbb{N}$.
On écrit $\alpha = -(p+1)$ avec $p \in \mathbb{N}$
$$\frac{1}{(1+x)^{p+1}} = \sum_{n=0}^{+\infty} (-1)^n \frac{(p+1)(p+2)\dots(p+n)}{n!} x^n = \sum_{n=0}^{+\infty} (-1)^n \binom{n+p}{n} x^n$$

Exemple Cas $\alpha = -1/2$.
$$\frac{1}{\sqrt{1+x}} = \sum_{n=0}^{+\infty} (-1)^n \frac{(2n)!}{(2^n n!)^2} x^n$$

22.3.5 Calcul de développements en série entière

22.3.5.1 Cas des fonctions rationnelles

Exemple Soit $a \in \mathbb{C}^\star$. La fonction $x \mapsto \dfrac{1}{x-a}$ est développable en série entière sur $]-r, r[$ avec $r = |a|$.
En effet, pour $|x| < |a|$,
$$\frac{1}{x-a} = -\frac{1}{a}\frac{1}{1-x/a} = \sum_{n=0}^{+\infty} \frac{-1}{a^{n+1}} x^n$$

Exemple Soit $a \in \mathbb{C}^\star$. La fonction $x \mapsto \dfrac{1}{(x-a)^2}$ est développable en série entière sur $]-r, r[$ avec $r = |a|$.
En effet, en dérivant le développement précédent
$$\frac{1}{(x-a)^2} = \sum_{n=0}^{+\infty} \frac{n+1}{a^{n+2}} x^n$$

Remarque Plus généralement, et par dérivations successives, on peut former le développement de $1/(x-a)^p$.

Exemple Formons le développement en série entière en 0 de

$$f : x \mapsto \frac{1}{(x-1)^2(x+2)}$$

La partie entière de f est nulle, 1 est pôle double et -2 est pôle simple. La décomposition en éléments simples de f est alors de la forme

$$f(x) = \frac{1}{(x-1)^2(x+2)} = \frac{a}{x+2} + \frac{b}{x-1} + \frac{c}{(x-1)^2}$$

avec

$$a = \left.\frac{1}{(x-1)^2}\right|_{x=-2} = \frac{1}{9},\ c = \left.\frac{1}{(x+2)}\right|_{x=1} = \frac{1}{3} \text{ et } b = \left.\left(\frac{1}{(x+2)}\right)'\right|_{x=1} = -\frac{1}{9}$$

Sur $]-1,1[$,

$$f(x) = \frac{1}{18}\frac{1}{1+\frac{x}{2}} + \frac{1}{9}\frac{1}{1-x} + \frac{1}{3}\frac{1}{(1-x)^2} = \sum_{n=0}^{+\infty}\left(\frac{(-1)^n}{18.2^n} + \frac{3n+4}{9}\right)x^n$$

Exemple Formons le développement en série entière en 0 de

$$f : x \mapsto \frac{1}{x^2+x+1}$$

Pour $x \in]-1,1[$,

$$f(x) = \frac{1-x}{1-x^3} = \sum_{n=0}^{+\infty}(1-x)x^{3n} = \sum_{n=0}^{+\infty} a_n x^n$$

avec $a_{3n} = 1, a_{3n+1} = -1$ et $a_{3n+2} = 0$.

22.3.5.2 Calcul par dérivation puis intégration

Exemple Formons le développement en série entière en 0 de

$$f : x \mapsto \ln(1+x+x^2)$$

On a

$$f'(x) = \frac{1+2x}{1+x+x^2} = \frac{(1+2x)(1-x)}{1-x^3} = \frac{1+x-2x^2}{1-x^3} = (1+x-2x^2)\sum_{n=0}^{+\infty} x^{3n}$$

pour $|x| < 1$.
Ainsi

$$f'(x) = \sum_{n=0}^{+\infty} x^{3n} + x^{3n+1} - 2x^{3n+2} = \sum_{n=0}^{+\infty} a_n x^n$$

avec $a_{3n} = 1$, $a_{3n+1} = 1$ et $a_{3n+2} = -2$.
Par intégration d'un développement en série entière

$$f(x) = f(0) + \sum_{n=0}^{+\infty}\frac{a_n}{n+1}x^{n+1} = \sum_{n=0}^{+\infty}\frac{a_n}{n+1}x^{n+1}$$

Exemple Formons le développement en série entière en 0 de la fonction arcsin.

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}(-x^2)^n$$

pour $x \in \left]-1,1\right[$ et $\alpha = 1/2$.

$$\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} = \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{2n-1}{2}\right)}{n!} = (-1)^n\frac{(2n)!}{(2^n n!)^2}$$

donc $(\arcsin x)' = \sum_{n=0}^{+\infty} \frac{(2n)!}{(2^n n!)^2} x^{2n}$ puis par intégration d'un développement en série entière

$$\arcsin x = \sum_{n=0}^{+\infty} \frac{(2n)!}{(2^n n!)^2}\frac{x^{2n+1}}{2n+1}$$

On peut aussi former le développement en série entière de la fonction arccos via
$\arccos x = \pi/2 - \arcsin x$.

22.3.5.3 Calcul en exploitant une équation différentielle

Exemple Formons le développement en série entière en 0 de

$$f : x \mapsto \frac{\arcsin x}{\sqrt{1-x^2}}$$

Les fonctions $x \mapsto 1/\sqrt{1-x^2}$ et $x \mapsto \arcsin x$ sont développables en série entière sur $\left]-1,1\right[$ donc f l'est aussi par produit. On pourrait calculer ce développement en procédant à un produit, mais l'expression finale ne serait pas très explicite. On va plutôt calculer ce développement en exploitant une équation différentielle vérifiée par f. La fonction f est dérivable sur $\left]-1,1\right[$ et

$$f'(x) = \frac{1}{1-x^2} + \frac{x \arcsin x}{(1-x^2)^{3/2}}$$

Ainsi, f vérifie l'équation différentielle

$$(1-x^2)y' - xy = 1$$

La fonction f étant impaire, son développement en série entière sur $\left]-1,1\right[$ peut s'écrire

$$f(x) = \sum_{n=0}^{+\infty} a_n x^{2n+1}$$

Par dérivation de série entière sur $\left]-1,1\right[$, on peut écrire

$$f'(x) = \sum_{n=0}^{+\infty} (2n+1)a_n x^{2n}$$

La relation $(1-x^2)f'(x) - xf(x) = 1$ donne alors

$$a_0 + \sum_{n=1}^{+\infty} \left((2n+3)a_{n+1} - (2n+2)a_n\right) x^{2n+2} = 1$$

Par unicité des coefficients d'un développement en série entière

$$a_0 = 1 \text{ et } \forall n \geqslant 1,\, a_{n+1} = \frac{2n+2}{2n+3}a_n$$

Ainsi

$$a_{n+1} = \frac{2n}{2n+1}\frac{2n-2}{2n-1}\cdots\frac{2}{3}a_0 = \frac{(2^n n!)^2}{(2n+1)!}$$

Finalement

$$f(x) = \sum_{n=0}^{+\infty} \frac{(2^n n!)^2}{(2n+1)!}x^{2n+1}$$

22.4 Applications

22.4.1 Régularité d'un prolongement continu

Exemple Soit $f : \mathbb{R}^\star \to \mathbb{R}$ définie par $f(x) = \dfrac{\mathrm{e}^x - 1}{x}$. Prolongeons f en 0.
Quand $x \to 0$, $\mathrm{e}^x = 1 + x + o(x)$ donc

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

On peut prolonger f par continuité en 1 en posant $f(0) = 1$.
Montrer que la fonction f ainsi prolongée est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.
Pour tout $x \in \mathbb{R}$,

$$\mathrm{e}^x - 1 = \sum_{n=1}^{+\infty} \frac{1}{n!}x^n$$

Pour tout $x \in \mathbb{R}^\star$,

$$\frac{\mathrm{e}^x - 1}{x} = \sum_{n=1}^{+\infty} \frac{1}{n!}x^{n-1} = \sum_{n=0}^{+\infty} \frac{1}{(n+1)!}x^n$$

puis pour tout $x \in \mathbb{R}$,

$$f(x) = \sum_{n=0}^{+\infty} \frac{1}{(n+1)!}x^n$$

Ainsi f est développable en série entière sur $\mathbb{R}$ et c'est donc une fonction de classe $\mathcal{C}^\infty$.
De plus

$$\forall n \in \mathbb{N},\, f^{(n)}(0) = \frac{1}{n+1}$$

car par série de Taylor

$$\frac{f^{(n)}(0)}{n!} = \frac{1}{(n+1)!}$$

Exemple De même, on obtient que la fonction sinus cardinal est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$.

Remarque On en déduit que la fonction $x \mapsto \dfrac{\sin x}{\mathrm{e}^x - 1}$ se prolonge en une fonction de classe $\mathcal{C}^\infty$ car

$$\frac{\sin x}{\mathrm{e}^x - 1} = \frac{\sin x}{x} \frac{x}{\mathrm{e}^x - 1}$$

est produit des deux fonctions $x \mapsto \dfrac{x}{\mathrm{e}^x - 1}$ et $x \mapsto \dfrac{\sin x}{x}$ qui se prolongent en des fonctions de classe $\mathcal{C}^\infty$.

22.4.2 Calcul de sommes

Exemple Calcul de $\displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n+1}}{(n+1)!}$.
On a immédiatement $R = +\infty$.
Pour $x \in \mathbb{R}$, par décalage d'indice

$$S(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n+1}}{(n+1)!} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n!} x^{2n-1}$$

donc

$$xS(x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n!} x^{2n} = 1 - \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!} \left(x^2\right)^n$$

Finalement

$$S(x) = \frac{1 - \mathrm{e}^{-x^2}}{x} \text{ pour } x \neq 0 \text{ et } S(0) = 0$$

Exemple Calcul de $\displaystyle\sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^n$.
On a immédiatement $R = +\infty$.
Si $x \geqslant 0$ alors

$$\sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^n = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} \sqrt{x}^{2n} = \operatorname{ch}\sqrt{x}$$

Si $x \leqslant 0$ alors

$$\sum_{n=0}^{+\infty} \frac{1}{(2n)!} x^n = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} |x|^n = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} \sqrt{|x|}^{2n} = \cos\sqrt{|x|}$$

Exemple Calcul de $\displaystyle\sum_{n=2}^{+\infty} \frac{(-1)^n}{n^2-1} x^n$.
On a immédiatement $R = 1$.
Puisque la série converge en $x = 1$ et $x = -1$, l'intervalle de convergence est $[-1, 1]$
Par décomposition en éléments simples

$$\frac{1}{n^2-1} = \frac{1}{2}\left(\frac{1}{n-1} - \frac{1}{n+1}\right)$$

Pour $x \in]-1, 1[$,

$$\sum_{n=2}^{+\infty} \frac{(-1)^n x^n}{n-1} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n} x^{n+1} = x\ln(1+x)$$

Pour $x \in]-1, 1[$ et $x \neq 0$,

$$\sum_{n=2}^{+\infty} \frac{(-1)^n x^n}{n+1} = \frac{1}{x}\sum_{n=3}^{+\infty} \frac{(-1)^{n-1}}{n} x^n = \frac{1}{x}\left(\ln(1+x) - x + \frac{1}{2}x^2\right)$$

Ainsi, pour $x \in]-1, 1[$ et $x \neq 0$,

$$\sum_{n=2}^{+\infty} \frac{(-1)^n}{n^2-1} x^n = \frac{1}{2}\left(x - \frac{1}{x}\right)\ln(1+x) + \frac{1}{2} - \frac{1}{4}x$$

Pour $x = 0$, la somme est nulle (car le coefficient constant est nul)
Etude en $x = \pm 1$
Posons $u_n(x) = \dfrac{(-1)^n}{n^2-1} x^n$. Les fonctions $u_n : [-1, 1] \to \mathbb{R}$ sont continues et $\|u_n\|_\infty = \dfrac{1}{n^2-1}$ est sommable. La série $\sum u_n$ converge normalement sur $[-1, 1]$ et sa somme y est continue.

$$S(1) = \lim_{x \to 1^-} S(x) = \frac{1}{4} \text{ et } S(-1) = \lim_{x \to (-1)^+} S(x) = \frac{3}{4}$$

Exemple Calcul de $\displaystyle\sum_{n=0}^{+\infty} \frac{x^{2n+1}}{2n+1}$.
On a immédiatement $R = 1$.
Pour $x \in]-1, 1[$, on peut écrire

$$S(x) = \sum_{n=1}^{+\infty} \frac{x^n}{n} - \sum_{n=1}^{+\infty} \frac{x^{2n}}{2n}$$

avec convergence des séries écrites. On a alors

$$S(x) = -\ln(1-x) + \frac{1}{2}\ln(1-x^2) = \frac{1}{2}\ln\frac{1+x}{1-x}$$

On aurait aussi pu calculer directement $S'(x)$.

22.4.3 Intégration terme à terme

22.4.3.1 Intégration sur $I = [a,b] \subset]-R,R[$

Une série entière converge normalement sur tout $[a,b]$ inclus dans $]-R,R[$, cela permet d'intégrer terme à terme.

Exemple Montrons $\displaystyle\int_0^\pi \operatorname{sinc}(t)\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!}\frac{\pi^{2n+1}}{2n+1}$
La fonction sinus cardinale est développable en série entière

$$\operatorname{sinc}(t) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!}t^{2n}$$

avec un rayon de convergence $R = +\infty$. Cette série entière converge donc normalement sur tout segment inclus dans $\mathbb{R}$ et donc en particulier sur $[0,\pi]$.
Puisque les fonctions sommées sont continues et que la série de fonctions converge uniformément

$$\int_0^\pi \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!}t^{2n}\,\mathrm{d}t = \sum_{n=0}^{+\infty}\int_0^\pi \frac{(-1)^n}{(2n+1)!}t^{2n}\,\mathrm{d}t$$

ce qui donne la formule proposée.

22.4.3.2 Intégration sur $I = [0,R[$

On peut intégrer terme à terme sous réserver de vérifier la convergence de $\displaystyle\sum\int_I |u_n|$.

Exemple Calcul de $I = \displaystyle\int_0^1 \frac{\ln(1+t)}{t}\,\mathrm{d}t$.
Sur $]0,1[$,

$$f(t) = \frac{\ln(1+t)}{t} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}t^{n-1} = \sum_{n=1}^{+\infty} \frac{(-1)^{n-1}}{n}t^{n-1}$$

(et la relation vaut aussi 1 et peut valoir en 0 par prolongement par continuité)
Posons $u_n :]0,1[\to \mathbb{R}$ définie par

$$u_n(t) = \frac{(-1)^{n-1}}{n}t^{n-1}$$

La série de fonctions $\sum u_n$ converge simplement et sa somme $\displaystyle\sum_{n=1}^{+\infty} u_n = f$ est continue par morceaux.
Chaque u_n est continue par morceaux et intégrable sur $]0,1[$.
Enfin, la série $\displaystyle\sum\int |u_n|$ converge car

$$\int_{]0,1[} |u_n| = \int_0^1 \frac{t^{n-1}}{n}\,\mathrm{d}t = \frac{1}{n^2}$$

Par théorème, f est intégrable sur $]0,1[$ et

$$I = \int_{]0,1[} f = \sum_{n=1}^{+\infty}\int_0^1 u_n(t)\,\mathrm{d}t = \sum_{n=1}^{+\infty}\frac{(-1)^n}{n^2}$$

Sachant $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, on peut achever le calcul de I,

$$I = \sum_{p=0}^{+\infty} \frac{1}{(2p+1)^2} - \sum_{p=1}^{+\infty} \frac{1}{(2p)^2} = \left(\sum_{p=0}^{+\infty} \frac{1}{(2p+1)^2} + \sum_{p=1}^{+\infty} \frac{1}{(2p)^2}\right) - 2\sum_{p=1}^{+\infty} \frac{1}{(2p)^2}$$

et donc

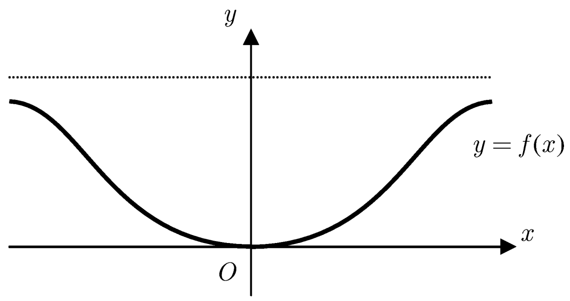
$$I = \sum_{n=1}^{+\infty} \frac{1}{n^2} - \frac{1}{2}\sum_{p=1}^{+\infty} \frac{1}{p^2} = \frac{1}{2}\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{12}$$

22.4.4 Musculation : fonction $\mathcal{C}^\infty$ non développable en série entière.

Soit $f : \mathbb{R}^\star \to \mathbb{R}$ définie par

$$f(x) = \mathrm{e}^{-1/x^2}$$

f est de classe $\mathcal{C}^\infty$ sur $]-\infty, 0[$ et $]0, +\infty[$.
Quand $x \to 0$, $f(x) \to 0$.
On prolonge f par continuité en 0 en posant $f(0) = 0$.



Montrons par récurrence sur $n \in \mathbb{N}$

$$\forall n \in \mathbb{N}, \forall x \neq 0, f^{(n)}(x) = P_n\left(\frac{1}{x}\right)\mathrm{e}^{-1/x^2} \text{ avec } P_n \in \mathbb{R}[X]$$

Cas $n = 0$: $P_0(X) = 1$ convient.
Cas $n = 1$: $P_1(X) = 2X^3$ convient.
Supposons la propriété vérifiée au rang $n \geqslant 0$

$$f^{(n+1)}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(P_n\left(\frac{1}{x}\right)\mathrm{e}^{-1/x^2}\right) = \left(-\frac{1}{x^2}P_n'\left(\frac{1}{x}\right) + \frac{2}{x^3}P_n\left(\frac{1}{x}\right)\right)\mathrm{e}^{-1/x^2}$$

Le polynôme $P_{n+1}(X) = X^2 P_n'(X) + 2X^3 P_n(X)$ convient.
Récurrence établie.
Quand $x \to 0^+$ (ou 0^-) (avec $x \neq 0$)

$$f^{(n)}(x) = P_n\left(\frac{1}{x}\right)\mathrm{e}^{-1/x^2} \underset{X=1/x}{=} P_n(X)\mathrm{e}^{-X^2} \to 0$$

On peut alors conclure que f est de classe $\mathcal{C}^\infty$ avec

$$\forall n \in \mathbb{N}, f^{(n)}(0) = 0$$

Finalement, f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et sa série de Taylor est nulle.
On en déduit que f n'est pas développable en série entière car, si par l'absurde, f l'est sur $]-r, r[$ alors

$$\forall x \in \left]-r, r\right[, f(x) = \sum_{n=0}^{+\infty} a_n x^n = 0 \text{ car } a_n = \frac{f^{(n)}(0)}{n!} = 0$$

C'est absurde, puisque f n'est pas nulle sur un voisinage de 0.

22.4.5 Musculation : fonction absolument monotone

Soit $r \in \mathbb{R}^{+\star} \cup \{+\infty\}$ et $f :]-r, r[\to \mathbb{R}$ de classe $\mathcal{C}^\infty$ telle que $f^{(n)} \geqslant 0$ pour tout $n \in \mathbb{N}$.
Montrer que f est développable en série entière sur $]-r, r[$.
Soit $x \in]-r, r[$. On peut écrire

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k + R_n(x)$$

avec

$$R_n(x) = \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) \,\mathrm{d}t$$

Par le changement de variable $t = xu$, on peut écrire

$$R_n(x) = x^{n+1} \int_0^1 \frac{(1-u)^n}{n!} f^{(n+1)}(xu) \,\mathrm{d}u$$

Choisissons y tel que $|x| < y < r$. Puisque $f^{(n+1)}$ est croissante, on a

$$\forall u \in [0, 1], f^{(n+1)}(xu) \leqslant f^{(n+1)}(yu)$$

et donc

$$|R_n(x)| \leqslant |x|^{n+1} \int_0^1 \frac{(1-u)^n}{n!} f^{(n+1)}(yu) \,\mathrm{d}u \leqslant |x/y|^{n+1} R_n(y)$$

De plus $R_n(y) \leqslant f(y)$ car les termes de la somme partielle de Taylor en y sont tous positifs et donc

$$|R_n(x)| \leqslant |x/y|^{n+1} f(y) \xrightarrow[n \to +\infty]{} 0$$

Finalement, f est aussi égale à la somme de sa série de Taylor sur $]-r, r[$.

Chapitre 23

Equations différentielles linéaires vectorielles

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^\star$
I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide.

23.1 Les équations vectorielles

23.1.1 Equation et systèmes différentiels

Définition

On appelle équation différentielle vectorielle linéaire d'ordre 1, définie sur I et à valeurs dans E, toute équation de la forme

$$(E) : x' = a(t)(x) + b(t)$$

avec $t \mapsto a(t)$ fonction continue de I vers $\mathcal{L}(E)$, $t \mapsto b(t)$ fonction continue de I vers E et d'inconnue $t \mapsto x(t)$ fonction dérivable de I vers E.

Exemple Cas $E = \mathbb{K}$.
Les endomorphismes sur $\mathbb{K}$ correspondent aux applications $x \mapsto ax$ avec $a \in \mathbb{K}$.
Une équation scalaire s'apparente alors à une équation vectorielle à valeurs dans $E = \mathbb{K}$ et inversement.

Remarque En introduisant une base $e = (e_1, \ldots, e_n)$ de E et en posant

$$A(t) = \text{Mat}_e\,(a(t)) \in \mathcal{M}_n(\mathbb{K}),\ B(t) = \text{Mat}_e\,(b(t)) \in \mathcal{M}_{n,1}(\mathbb{K}) \text{ et } X(t) = \text{Mat}_e\,(x(t)) \in \mathcal{M}_{n,1}(\mathbb{K}),$$

l'équation vectorielle

$$x' = a(t)(x) + b(t)$$

équivaut à l'équation matricielle

$$X' = A(t)X + B(t)$$

En notant $a_{i,j}(t)$ les coefficients de la matrice $A(t)$, $b_i(t)$ ceux de la colonne $B(t)$ et $x_i(t)$ ceux de la colonne $X(t)$, l'équation étudiée équivaut encore au système différentiel

$$(\Sigma) : \begin{cases} x_1' = a_{1,1}(t)x_1 + \cdots + a_{1,n}(t)x_n + b_1(t) \\ \vdots \\ x_n' = a_{n,1}(t)x_1 + \cdots + a_{n,n}(t)x_n + b_n(t) \end{cases}$$

En pratique, c'est fréquemment sous la forme d'un système différentiel que sont présentés les équations linéaires vectorielles.

Exemple Le système

$$\begin{cases} x_1' = t.x_1 + 2.x_2 + \mathrm{e}^t \\ x_2' = (1-t).x_1 + t.x_2 \end{cases}$$

définit un système différentiel de taille 2.

Exemple Résoudre l'équation différentielle scalaire

$$(E) : x'' = a(t)x' + b'(t)x + c(t)$$

revient à résoudre le système différentiel

$$(\Sigma) : \begin{cases} x' = y \\ y' = a(t)y + b(t)x + c(t) \end{cases}$$

car x est solution de (E) si, et seulement si, (x, x') est solution de (Σ).

Proposition

Les solutions de l'équation $(E) : x' = a(t)(x) + b(t)$ sont des fonctions de classe $\mathcal{C}^1$.

dém. :
Soit x une solution de (E). La fonction x est dérivable et

$$\forall t \in I, x'(t) = a(t)(x(t)) + b(t)$$

Introduisons l'application $V : \mathcal{L}(E) \times E \to E$ définie par $V(u, x) = u(x)$.
L'application V est bilinéaire donc continue (car $\dim E < +\infty$).
Puisque $x' = V(a, x) + b$, la fonction x' est continue et donc x est de classe $\mathcal{C}^1$.
□

23.1.2 Problème de Cauchy

Soit $a : I \to \mathcal{L}(E)$ et $b : I \to E$ des fonctions continues. On étudie l'équation différentielle

$$(E) : x' = a(t)(x) + b(t)$$

Définition

Soit $(t_0, x_0) \in I \times E$. Un problème de Cauchy associé à l'équation (E) en t_0 consiste à déterminer les solutions de l'équation de l'équation

$$(E) : x' = a(t)(x) + b(t)$$

vérifiant la condition initiale $x(t_0) = x_0$.

Exemple Pour les équations scalaires, on a vu qu'un problème de Cauchy détermine une solution unique.

Proposition

Soit $x : I \to E$ une fonction continue. On a équivalence entre :
(i) x est solution sur I du problème de Cauchy

$$\begin{cases} x' = a(t)(x) + b(t) \\ x(t_0) = x_0 \end{cases}$$

(ii) x vérifie

$$x(t) = x_0 + \int_{t_0}^{t} a(u)(x(u)) + b(u) \,\mathrm{d}u$$

dém. :
(i) $\Rightarrow$ (ii) Supposons (i)
Puisque la fonction x est de classe $\mathcal{C}^1$,

$$x(t) = x(t_0) + \int_{t_0}^{t} x'(u) \,\mathrm{d}u$$

donc

$$x(t) = x_0 + \int_{t_0}^{t} a(u)(x(u)) + b(u) \,\mathrm{d}u$$

(ii) $\Rightarrow$ (i) Supposons (ii)

$$x(t_0) = x_0 + \int_{t_0}^{t_0} \ldots \,\mathrm{d}u = x_0$$

et puisque

$$t \mapsto \int_{t_0}^{t} a(u)(x(u)) + b(u) \,\mathrm{d}u$$

est dérivable, x est dérivable avec

$$x'(t) = a(t)(x(t)) + b(t)$$

□

Théorème

(admis)
Soit $(t_0, x_0) \in I \times E$. Le problème de Cauchy

$$\begin{cases} x' = a(t)(x) + b(t) \\ x(t_0) = x_0 \end{cases}$$

possède une unique solution définie sur I.

23.1.3 Structure de l'ensemble solution

Soit $a : I \to \mathcal{L}(E)$ et $b : I \to E$ des fonctions continues. On étudie l'équation différentielle

$$(E) : x' = a(t)(x) + b(t)$$

23.1.3.1 Équation homogène

Définition

L'équation $(E_0) : x' = a(t)(x)$ est appelée équation homogène associée à l'équation (E).
Ses solutions sont appelées solutions homogènes de l'équation (E).

Théorème

L'ensemble $\mathcal{S}_0$ des solutions sur I de l'équation homogène (E_0) est un sous-espace vectoriel de $\mathcal{C}^1(I, E)$ de dimension $n = \dim E$.

dém. :
Les solutions de l'équation (E_0) sont de classe $\mathcal{C}^1$ donc $\mathcal{S}_0 \subset \mathcal{C}^1(I, E)$.
Considérons la fonction $\Phi : \mathcal{C}^1(I, \mathbb{K}) \to \mathcal{C}(I, \mathbb{K})$ définie par

$$\Phi(x) = x' - a(x)$$

En fait, $\Phi(x)$ désigne la fonction $t \mapsto x'(t) - a(t)\,(x(t))$
La fonction Φ est linéaire et $\mathcal{S}_0 = \ker \Phi$ donc $\mathcal{S}_0$ est un sous-espace vectoriel de $\mathcal{C}^1(I, E)$.
Pour $t_0 \in I$, considérons l'application $E_{t_0} : \mathcal{S}_0 \to E$ définie par

$$E_{t_0} : x \mapsto x(t_0)$$

E_{t_0} est une application linéaire car

$$E_{t_0}(\lambda_1 x_1 + \lambda_2 x_2) = (\lambda_1 x_1 + \lambda_2 x_2)(t_0) = \lambda_1 x_1(t_0) + \lambda_2 x_2(t_0) = \lambda_1 E_{t_0}(x_1) + \lambda_2 E_{t_0}(x_2)$$

Par le théorème de Cauchy linéaire, on peut affirmer que l'application E_{t_0} est bijective.
Par suite E_{t_0} est un isomorphisme et donc $\dim \mathcal{S}_0 = \dim E$.
□

Exemple L'ensemble des solutions d'un système différentiel

$$(\Sigma) : \begin{cases} x' = a(t)x + b(t)y \\ y' = c(t)x + d(t)y \end{cases}$$

est un plan vectoriel.

23.1.3.2 Système fondamental de solutions

Puisque l'espace $\mathcal{S}_0$ est de dimension n, il possède une base à n éléments.

Définition

> On appelle système fondamental de solutions de l'équation homogène (E_0) toute base $(\varphi_1, \ldots, \varphi_n)$ de l'espace $\mathcal{S}_0$.

Remarque Si $(\varphi_1, \ldots, \varphi_n)$ est un système fondamental de solution de (E_0), la solution générale homogène est

$$x(t) = \lambda_1 \varphi_1(t) + \cdots + \lambda_n \varphi_n(t) \text{ avec } \lambda_1, \ldots, \lambda_n \in \mathbb{K}$$

23.1.3.3 Résolution de l'équation complète

Théorème

> L'ensemble $\mathcal{S}$ des solutions sur I de l'équation
>
> $$(E) : x' = a(t)(x) + b(t)$$
>
> est un sous-espace affine de $\mathcal{C}^1(I, E)$ de direction l'espace $\mathcal{S}_0$.
> C'est donc un sous-espace affine de dimension $n = \dim E$.

dém. :
Les solutions sont de classe $\mathcal{C}^1$ donc $\mathcal{S} \subset \mathcal{C}^1(I, E)$.
Par le théorème de Cauchy linéaire, en fixant une condition initiale, on peut assurer l'existence d'au moins une solution $\tilde{x}$ à l'équation étudiée.
Soit $x \in \mathcal{C}^1(I, E)$. En introduisant à nouveau l'application Φ présentée dans le théorème ci-dessus, l'équation (E) s'écrit $\Phi(x) = b$. On a alors

$$x \in \mathcal{S} \Leftrightarrow \Phi(x) = \Phi(\tilde{x})$$

En ramenant au premier membre

$$x \in \mathcal{S} \Leftrightarrow \Phi(x - \tilde{x}) = 0$$

et donc

$$x \in \mathcal{S} \Leftrightarrow x - \tilde{x} \in \mathcal{S}_0$$

Ainsi $\mathcal{S} = \tilde{x} + \mathcal{S}_0$ est un sous-espace affine de direction $\mathcal{S}_0$.
□

Protocole : Pour résoudre (E) :
- on identifie le type l'équation (E) ;
- on résout l'équation homogène (E_0) : $x_0(t) = \ldots$;
- on cherche une solution particulière : $\tilde{x}(t) = \ldots$;
- on exprime la solution générale : $x(t) = \tilde{x}(t) + x_0(t)$.

Proposition

Si $b(t) = b_1(t) + b_2(t)$ avec b_1 et $b_2 : I \to E$ fonctions continues et si x_1 et x_2 sont respectivement solutions des équations

$$(E_1) : x' = a(t)(x) + b_1(t) \text{ et } (E_2) : x' = a(t)(x) + b_2(t)$$

alors $\tilde{x}$ est solution de l'équation

$$(E) : x' = a(t)(x) + b(t)$$

dém. :
$\Phi(x_1) = b_1$ et $\Phi(x_2) = b_2$ donc $\Phi(x_1 + x_2) = b_1 + b_2 = b$.
Pour tout $t \in I$

$$x'(t) = x_1'(t) + x_2'(t) = a(t)\left(x_1(t)\right) + b_1(t) + a(t)\left(x_2(t)\right) + b_2(t)$$

et donc, par linéarité de l'endomorphisme $a(t)$

$$x'(t) = a(t)\left(x_1(t) + x_2(t)\right) + b_1(t) + b_2(t) = a(t)\left(x(t)\right) + b(t)$$

□

23.1.4 Méthode de variation des constantes

On cherche une solution à l'équation complète

$$(E) : x' = a(t)(x) + b(t)$$

Supposons résolue l'équation homogène associée

$$(E_0) : x' = a(t)(x)$$

On connaît alors $(\varphi_1, \ldots, \varphi_n)$ système fondamental de solutions de l'équation homogène.
La solution générale homogène s'écrit

$$x(t) = \lambda_1.\varphi_1(t) + \cdots + \lambda_n.\varphi_n(t)$$

Théorème

On peut trouver par quadrature une solution particulière de l'équation complète

$$(E) : x' = a(t)(x) + b(t)$$

de la forme

$$x(t) = \lambda_1(t).\varphi_1(t) + \cdots + \lambda_n(t).\varphi_n(t)$$

avec $\lambda_1, \ldots, \lambda_n$ fonctions dérivables.

dém. :
Soit $x(t) = \lambda_1(t).\varphi_1(t) + \cdots + \lambda_n(t).\varphi_n(t)$ avec $\lambda_1, \ldots, \lambda_n$ fonctions dérivables.
On a

$$x'(t) = \lambda_1'(t).\varphi_1(t) + \cdots + \lambda_n'(t).\varphi_n(t) + \lambda_1(t).\varphi_1'(t) + \cdots + \lambda_n(t).\varphi_n'(t)$$

et puisque $\varphi_i'(t) = a(t)\,(\varphi_i(t))$, on obtient

$$x'(t) = a(t)\,(x(t)) + b(t) \Leftrightarrow \lambda_1'(t).\varphi_1(t) + \cdots + \lambda_n'(t).\varphi_n(t) = b(t)$$

Soit $e = (e_1, \ldots, e_n)$ une base de l'espace E. Posons

$$X_j(t) = \mathrm{Mat}_e(\varphi_j(t)) \text{ et } B(t) = \mathrm{Mat}_e b(t)$$

L'équation précédente s'écrit

$$\lambda_1(t)X_1(t) + \cdots + \lambda_n(t)X_n(t) = B(t)$$

Considérons encore

$$W(t) = \mathrm{Mat}_e\,(\varphi_1(t), \ldots, \varphi_n(t)) = (X_1(t) \mid \ldots \mid X_n(t)) \in \mathcal{M}_n(\mathbb{K})$$

et $Y(t) = {}^t(\lambda_1'(t), \ldots, \lambda_n'(t))$. L'équation devient le système linéaire

$$W(t)Y(t) = B(t)$$

Or la matrice $W(t)$ est inversible. En effet, pour chaque $t_0 \in I$, l'application $E_{t_0} : \mathcal{S}_0 \to E$ définie par $x \mapsto x(t_0)$ est un isomorphisme. Celui-ci transforme en une base en une base et donc

$$W(t_0) = \mathrm{Mat}_e\,(\varphi_1(t_0), \ldots, \varphi_n(t_0)) \text{ est inversible}$$

On a alors

$$x'(t) = a(t)x(t) + b(t) \Leftrightarrow Y(t) = W(t)^{-1}B(t)$$

Enfin, la fonction $t \mapsto W(t)^{-1}B(t)$ est continue, on peut donc déterminer par quadrature des fonctions $\lambda_1, \ldots, \lambda_n$ telles que la fonction donnée par $x(t) = \lambda_1(t)\varphi_1(t) + \cdots + \lambda_n(t)\varphi_n(t)$ est alors solution particulière de l'équation (E).
□

Remarque Cette méthode explique la méthode de variation de la constante vue pour les équations scalaires d'ordre 2.

23.1.5 Un exemple de résolution

Exemple Résoudre l'équation

$$(\Sigma) : \begin{cases} x_1' = 3x_1 - 2x_2 + \mathrm{e}^t \\ x_2' = x_1 + \mathrm{e}^t \end{cases}$$

C'est un système différentiel de taille 2 de système homogène associé

$$(\Sigma_0) : \begin{cases} x_1' = 3x_1 - 2x_2 \\ x_2' = x_1 \end{cases}$$

On peut observer que

$$X_1(t) = \begin{pmatrix} \mathrm{e}^t \\ \mathrm{e}^t \end{pmatrix}, X_2(t) = \begin{pmatrix} 2\mathrm{e}^{2t} \\ \mathrm{e}^{2t} \end{pmatrix}$$

sont deux solutions indépendantes de Σ_0, elles forment donc un système fondamental de solutions et la solution générale homogène est

$$X(t) = \lambda_1 X_1(t) + \lambda_2 X_2(t)$$

Déterminons une solution particulière à l'équation complète de la forme

$$X(t) = \lambda_1(t) X_1(t) + \lambda_2(t) X_2(t)$$

avec λ_1, λ_2 fonctions dérivables.
On injectant dans (Σ) on obtient

$$\begin{cases} \lambda_1'(t)\mathrm{e}^t + 2\lambda_2'(t)\mathrm{e}^t = \mathrm{e}^t \\ \lambda_1'(t)\mathrm{e}^t + \lambda_2'(t)\mathrm{e}^{2t} = \mathrm{e}^t \end{cases}$$

La résolution donne

$$\begin{cases} \lambda_1'(t) = 1 \\ \lambda_2'(t) = 0 \end{cases}$$

puis la solution particulière

$$X(t) = \begin{pmatrix} t\mathrm{e}^t \\ t\mathrm{e}^t \end{pmatrix}$$

23.2 Equation linéaire d'ordre 1 à coefficient constant

E désigne un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^\star$

23.2.1 Définition

Définition

On appelle équation différentielle linéaire d'ordre 1 à coefficient constant, définie sur I et à valeurs dans E, toute équation différentielle de la forme

$$(E) : x' = a(x) + b(t)$$

avec $a \in \mathcal{L}(E)$, $t \mapsto b(t)$ continue de I vers E et d'inconnue $t \mapsto x(t)$ dérivable de I vers E.

Remarque Via l'introduction d'une base de E, une telle équation différentielle correspond :
- à une équation matricielle

$$X' = AX + B(t) \text{ avec } A \in \mathcal{M}_n(\mathbb{K}) \text{ et } B(t) \in \mathcal{M}_{n,1}(\mathbb{K})$$

- à un système différentiel

$$(\Sigma) : \begin{cases} x_1' = a_{1,1}x_1 + \cdots + a_{1,n}x_n + b_1(t) \\ \vdots \\ x_n' = a_{n,1}x_1 + \cdots + a_{n,n}x_n + b_n(t) \end{cases} \quad \text{avec } a_{i,j} \in \mathbb{K} \text{ et } b_i(t) \in \mathbb{K}$$

Remarque Compte tenu de la méthode de variation des constantes, il suffit de savoir résoudre l'équation homogène (E_0) pour résoudre complètement (E).

23.2.2 Résolution théorique de l'équation homogène

Rappel :

Pour $a \in \mathcal{L}(E)$, $\exp(a) = \sum_{k=0}^{+\infty} \frac{1}{k!} a^k \in \mathcal{L}(E)$. En particulier $\exp(\tilde{0}) = \mathrm{Id}_E$.
Pour $a \in \mathcal{L}(E)$, l'application $t \mapsto \exp(t.a)$ est de classe $\mathcal{C}^\infty$ et

$$\frac{\mathrm{d}}{\mathrm{d}t}(\exp(t.a)) = a \circ \exp(t.a)$$

Théorème

Soit $a \in \mathcal{L}(E)$ et $x_0 \in E$. L'unique solution au problème de Cauchy

$$\begin{cases} x' = a(x) \\ x(0) = x_0 \end{cases}$$

est la fonction

$$x : t \mapsto \exp(t.a)(x_0)$$

dém. :
On sait déjà que le problème de Cauchy possède une solution unique. Vérifions que celle proposée convient.

$$x(t) = \exp(t.a)(x_0)$$

On a déjà $x(0) = \mathrm{Id}_E(x_0) = x_0$. Vérifions que la fonction x est dérivable et calculons $x'(t)$. Introduisons l'application $V : \mathcal{L}(E) \times E \to E$ qui à $(u, x) \in \mathcal{L}(E) \times E$ associe $V(u, x) = u(x)$. Cette application est bilinéaire. Par composition avec les fonctions $t \mapsto \exp(t.a)$ et $t \mapsto x_0$, toutes deux dérivables, on peut affirmer que la fonction $t \mapsto x(t) = V(\exp(t.a), x_0)$ est dérivable avec

$$x'(t) = V(a \circ \exp(t.a), x_0) + V(\exp(t.a), 0_E)$$

et donc

$$x'(t) = a(\exp(t.a)(x_0)) = a(x(t))$$

□

Remarque La solution au problème de Cauchy

$$\begin{cases} x' = a(x) \\ x(t_0) = x_0 \end{cases}$$

est alors $x : t \mapsto \exp((t - t_0).a)(x_0)(x_0)$.

Corollaire

L'espace $\mathcal{S}_0$ des solutions sur $\mathbb{R}$ de l'équation homogène $x' = a(x)$ est

$$\mathcal{S}_0 = \{t \mapsto \exp(t.a)(x_0) / x_0 \in E\}$$

Exemple Soit $e = (e_1, \dots, e_n)$ une base de E.
En posant $\varphi_i : t \mapsto \exp(t.a)(e_i)$ et en écrivant $x_0 = \lambda_1.e_1 + \cdots + \lambda_n.e_n$, la solution générale de E s'exprime

$$x(t) = \lambda_1.\varphi_1(t) + \cdots + \lambda_n.\varphi_n(t)$$

Remarque Matriciellement, la solution de l'équation $X' = AX$ vérifiant $X(0) = X_0$ est

$$X(t) = \exp(t.A)X_0$$

Exemple Si X_0 est vecteur propre de A associée à la valeur propre λ alors

$$\exp(t.A)X_0 = \sum_{n=0}^{+\infty} \frac{1}{n!} t^n A^n X_0 = \sum_{n=0}^{+\infty} \frac{1}{n!} t^n \lambda^n X_0 = \mathrm{e}^{\lambda t} X_0$$

23.2.3 Résolution pratique de l'équation homogène

La résolution de l'équation homogène $x' = a(x)$ (resp. $X' = AX$) se ramène à la détermination de $\exp(t.a)$ (resp. $\exp(t.A)$). Il est alors pertinent d'opérer la réduction de l'endomorphisme a (resp. la matrice A).

Exemple Résoudre

$$(\Sigma) : \begin{cases} x_1' = 3x_1 - 4x_2 \\ x_2' = 2x_1 - 3x_2 \end{cases}$$

C'est un système différentiel de taille 2 linéaire à coefficient constant d'équation matricielle $X' = AX$ avec

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, A = \begin{pmatrix} 3 & -4 \\ 2 & -3 \end{pmatrix}$$

Equation homogène : $X' = AX$.
$\chi_A = X^2 - 1$, $\mathrm{Sp}(A) = \{1, -1\}$, $E_1(A) = \mathrm{Vect}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $E_{-1}(A) = \mathrm{Vect}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
On a

$$A = PDP^{-1} \text{ avec } P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \text{ et } D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

et donc

$$X' = AX \Leftrightarrow X' = PDP^{-1}X \Leftrightarrow P^{-1}X' = DP^{-1}X$$

Posons $Y = P^{-1}X$. On a $Y' = P^{-1}X'$ et donc

$$X' = AX \Leftrightarrow Y' = DY$$

Posons $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

$$Y' = DY \Leftrightarrow \begin{cases} y_1' = y_1 \\ y_2' = -y_2 \end{cases} \Leftrightarrow \begin{cases} y_1(t) = \lambda_1 \mathrm{e}^t \\ y_2(t) = \lambda_2 \mathrm{e}^{-t} \end{cases} \text{ avec } \lambda_1, \lambda_2 \in \mathbb{K}$$

$$X = PY \Leftrightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Leftrightarrow \begin{cases} x_1 = 2y_1 + y_2 \\ x_2 = y_1 + y_2 \end{cases}$$

$$X' = AX \Leftrightarrow X(t) = \begin{pmatrix} 2\lambda_1 \mathrm{e}^t + \lambda_2 \mathrm{e}^{-t} \\ \lambda_1 \mathrm{e}^t + \lambda_2 \mathrm{e}^{-t} \end{pmatrix} = \lambda_1 \begin{pmatrix} 2\mathrm{e}^t \\ \mathrm{e}^t \end{pmatrix} + \lambda_2 \begin{pmatrix} \mathrm{e}^{-t} \\ \mathrm{e}^{-t} \end{pmatrix}$$

$X_1(t) = \begin{pmatrix} 2\mathrm{e}^t \\ \mathrm{e}^t \end{pmatrix}$ et $X_2(t) = \begin{pmatrix} \mathrm{e}^{-t} \\ \mathrm{e}^{-t} \end{pmatrix}$ définissent un système fondamental de solutions.

Exemple Résoudre

$$\begin{cases} x_1' = x_1 - x_2 \\ x_2' = x_1 + x_2 \end{cases}$$

Système différentiel de taille 2 linéaire homogène à coefficients constants.
Equation matricielle : $X' = AX$ avec

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$\chi_A(X) = (X-1)^2 + 1$.
Cas $\mathbb{K} = \mathbb{C}$:
$\mathrm{Sp}(A) = \{1 \pm i\}$, $E_{1+i}(A) = \mathrm{Vect} \begin{pmatrix} 1 \\ -i \end{pmatrix}$ et $E_{1-i}(A) = \mathrm{Vect} \begin{pmatrix} 1 \\ i \end{pmatrix}$.

$$A = PDP^{-1} \text{ avec } P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \text{ et } D = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}$$

et donc

$$X' = AX \Leftrightarrow Y' = DY \text{ avec } Y = P^{-1}X$$

En écrivant $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$,

$$Y' = DY \Leftrightarrow \begin{cases} y_1' = (1+i)y_1 \\ y_2' = (1-i)y_2 \end{cases} \Leftrightarrow \begin{cases} y_1(t) = \lambda_1 \mathrm{e}^{(1+i)t} \\ y_2(t) = \lambda_2 \mathrm{e}^{(1-i)t} \end{cases} \text{ avec } \lambda_1, \lambda_2 \in \mathbb{C}$$

$X = PY = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$X' = AX \Leftrightarrow X(t) = \begin{pmatrix} \lambda_1 \mathrm{e}^{(1+i)t} + \lambda_2 \mathrm{e}^{(1-i)t} \\ -i\lambda_1 \mathrm{e}^{(1+i)t} + i\lambda_2 \mathrm{e}^{(1-i)t} \end{pmatrix} \text{ avec } \lambda_1, \lambda_2 \in \mathbb{C}$$

$X_1(t) = \begin{pmatrix} \mathrm{e}^{(1+i)t} \\ -i\mathrm{e}^{(1+i)t} \end{pmatrix}$ et $X_2(t) = \begin{pmatrix} \mathrm{e}^{(1-i)t} \\ i\mathrm{e}^{(1-i)t} \end{pmatrix} = \overline{X_1(t)}$ définissent un système fondamental de solutions.
Cas $\mathbb{K} = \mathbb{R}$:

$X_1(t) = \begin{pmatrix} \mathrm{e}^{(1+i)t} \\ -i\mathrm{e}^{(1+i)t} \end{pmatrix}$ est solution complexe de l'équation $X' = AX$ or la matrice A est réelle donc

$$\mathrm{Re}(X_1(t)) = \begin{pmatrix} \cos(t)\mathrm{e}^t \\ \sin(t)\mathrm{e}^t \end{pmatrix} \text{ et } \mathrm{Im}(X_1(t)) = \begin{pmatrix} \sin(t)\mathrm{e}^t \\ -\cos(t)\mathrm{e}^t \end{pmatrix}$$

sont des solutions réelles de l'équation $X' = AX$.
Celles-ci sont clairement indépendantes et donc forment un système fondamental de solutions.
Solution générale

$$X(t) = \alpha \begin{pmatrix} \cos(t)\mathrm{e}^t \\ \sin(t)\mathrm{e}^t \end{pmatrix} + \beta \begin{pmatrix} \sin(t)\mathrm{e}^t \\ -\cos(t)\mathrm{e}^t \end{pmatrix} \text{ avec } \alpha, \beta \in \mathbb{R}$$

Remarque On peut aussi procéder efficacement par la transformation de système suivante

$$\begin{cases} x_1' = x_1 - x_2 \\ x_2' = x_1 + x_2 \end{cases} \Leftrightarrow \begin{cases} x_2 = x_1 - x_1' \\ x_1' - x_1'' = x_1 + (x_1 - x_1') \end{cases} \Leftrightarrow \begin{cases} x_2 = x_1 - x_1' \\ x_1'' - 2x_1' + 2x_1 = 0 \end{cases}$$

On sait alors résoudre l'équation définissant x_1 puis exprimer la fonction x_2 associée.

Exemple Résoudre

$$(\Sigma) : \begin{cases} x_1' = 3x_1 + 2x_2 \\ x_2' = -2x_1 - x_2 \end{cases}$$

C'est un système différentielle linéaire d'ordre 1 homogène et à coefficients constants d'équation matricielle $X' = AX$ avec

$$A = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$\chi_A(X) = (X-1)^2$.
$E_1(A) = \mathrm{Vect} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$
Posons $C_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. On a

$$A = PTP^{-1} \text{ avec } P = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ et } T = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

et donc

$$X' = AX \Leftrightarrow Y' = TY \text{ avec } Y = P^{-1}X$$

En posant $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$,

$$Y' = TY \Leftrightarrow \begin{cases} y_1' = y_1 + 2y_2 \\ y_2' = y_2 \end{cases} \Leftrightarrow \begin{cases} y_1(t) = \lambda_1 \mathrm{e}^t + 2\lambda_2 t\mathrm{e}^t \\ y_2(t) = \lambda_2 \mathrm{e}^t \end{cases} \text{ avec } \lambda_1, \lambda_2 \in \mathbb{K}$$

puis

$$X' = AX \Leftrightarrow X(t) = \begin{pmatrix} \lambda_1 e^t + \lambda_2(2t+1)e^t \\ -\lambda_1 e^t + \lambda_2(1-2t)e^t \end{pmatrix}$$

23.2.4 Comportement asymptotique des solutions homogènes

On limite l'étude au cas $n = 2$ et $\mathbb{K} = \mathbb{R}$.
On étudie le système différentiel

$$(\Sigma) : \begin{cases} x_1' = ax_1 + bx_2 \\ x_2' = cx_1 + dy_2 \end{cases}$$

avec $a, b, c, d \in \mathbb{R}$. L'équation matricielle associée est $X' = AX$ avec

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ et } X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

23.2.4.1 Lignes de champ

Définition

On appelle ligne de champ du système (Σ) tout arc de $\mathbb{R}^2$ paramétré par

$$\begin{cases} x = x_1(t) \\ y = x_2(t) \end{cases}$$

avec (x_1, x_2) solution sur $\mathbb{R}$ du système (Σ).

Proposition

En tout point régulier,une ligne de champ est tangente au champ de vecteurs

$$(x, y) \mapsto (ax + by, cx + dy)$$

dém. :
Soit (x_1, x_2) une solution de Σ et $t_0 \in \mathbb{R}$ tel que le point

$$(x_0, y_0) = (x(t_0), y(t_0)) = (x_1(t_0), x_2(t_0))$$

soit régulier.
La tangente en (x_0, y_0) est dirigée par le premier vecteur dérivé qui a pour coordonnées

$$\begin{cases} x'(t_0) = x_1'(t_0) = ax_1(t_0) + bx_2(t_0) = ax_0 + by_0 \\ y'(t_0) = x_2'(t_0) = cx_1(t_0) + dx_2(t_0) = cx_0 + dy_0 \end{cases}$$

C'est le vecteur du champ de vecteur proposé
□

23.2.4.2 Comportement en l'infini

Pour étudier le comportement en $+\infty$ des lignes de champ, on introduit le polynôme caractéristique $\chi_A(X) \in \mathbb{R}[X]$ de discriminant Δ.

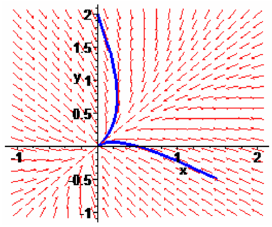
Cas $\Delta > 0$: la matrice A est diagonalisable dans $\mathcal{M}_2(\mathbb{R})$ de valeurs propres $\lambda_1 < \lambda_2$.
Notons V_1, V_2 des vecteurs propres associés aux valeurs propres λ_1, λ_2. Les fonctions définies par $X_1(t) = \mathrm{e}^{\lambda_1 t}V_1$ et $X_2(t) = \mathrm{e}^{\lambda_2 t}V_2$ déterminent un système fondamental de solution de l'équation $X' = AX$.
La solution générale de l'équation est alors de la forme

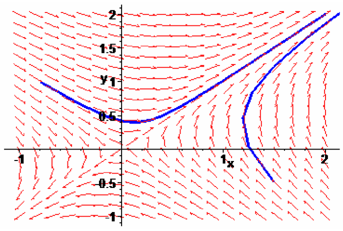
$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \mu_1 \mathrm{e}^{\lambda_1 t}V_1 + \mu_2 \mathrm{e}^{\lambda_2 t}V_2 \text{ avec } \mu_1, \mu_2 \in \mathbb{R}$$

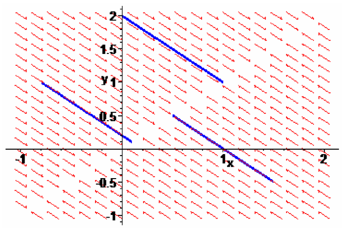
On peut aussi écrire

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \mathrm{e}^{\lambda_2 t}(\mu_1 \mathrm{e}^{(\lambda_1 - \lambda_2)t}V_1 + \mu_2 V_2)$$

Si $\lambda_2 < 0$: les lignes de champ convergent vers 0 en $+\infty$ avec une tangente dirigée par V_2.
Si $0 < \lambda_2$: les lignes de champ divergent vers $+\infty$ en prenant la direction de V_2.
Si $\lambda_2 = 0$: les lignes de champ convergent vers les points d'une droite dirigée par V_2.







Cas $\Delta = 0$: on a une racine réelle double et des comportements proches de ceux présentés ci-dessus.
Cas $\Delta < 0$: la matrice A est diagonalisable dans $\mathcal{M}_2(\mathbb{C})$ avec des valeurs propres $\lambda, \bar{\lambda}$.
Pour V_1 vecteur propre associé la valeur propre λ, la colonne

$$Z(t) = \mathrm{e}^{\lambda t}V_1$$

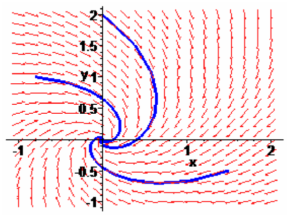
est solution complexe de l'équation $Z' = AZ$ et alors $X_1 = \mathrm{Re}(Z)$ et $X_2 = \mathrm{Im}(Z)$ déterminent un système fondamental de solutions de l'équation $X' = AX$. Puisque

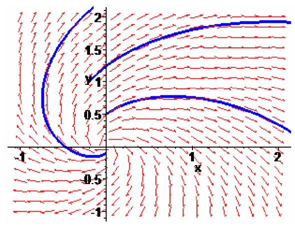
$$\mathrm{e}^{\lambda t} = \mathrm{e}^{\mathrm{Re}(\lambda)t}\left(\cos(\omega t) + i\sin(\omega t)\right) \text{ avec } \omega = \mathrm{Im}(\lambda)$$

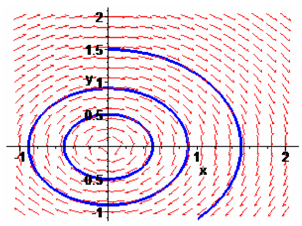
on obtient une écriture générale des solutions de la forme

$$\begin{cases} x_1(t) = (\alpha\cos(\omega t) + \beta\sin(\omega t))\mathrm{e}^{\mathrm{Re}(\lambda)t} \\ x_2(t) = (\gamma\cos(\omega t) + \delta\sin(\omega t))\mathrm{e}^{\mathrm{Re}(\lambda)t} \end{cases}$$

Si $\mathrm{Re}(\lambda) < 0$: les lignes de champ s'enroulent vers $(0, 0)$ en $+\infty$
Si $\mathrm{Re}(\lambda) > 0$: les lignes de champ s'échappent en branche spirale en $+\infty$.
Si $\mathrm{Re}(\lambda) = 0$: les lignes de champ sont refermées sur elles-mêmes.







Remarque Le comportement en $-\infty$ des solutions se déduit de l'étude précédente par renversement temporelle. Celui-ci nous ramène aux études précédentes en ayant passé à l'opposé la matrice et donc aussi ses valeurs propres.

23.3 Equations scalaires d'ordre n

23.3.1 Présentation

Définition

On appelle équation différentielle scalaire linéaire d'ordre n définie sur I toute équation de la forme

$$(E) : x^{(n)} = a_{n-1}(t)x^{(n-1)} + a_{n-2}(t)x^{(n-2)} + \cdots + a_1(t)x' + a_0(t)x + b(t)$$

avec $a_0, \ldots, a_n : I \to \mathbb{K}$ et $b : I \to \mathbb{K}$ continues, et d'inconnues $x : I \to \mathbb{K}$ fonction n fois dérivable.

Proposition

Les solutions d'une telle équation sont de classe $\mathcal{C}^n$.

Lemme

Soit $x : I \to \mathbb{K}$ dérivable. On a équivalence entre :
(i) x est solution de l'équation (E) ;
(ii) x est le premier élément d'un tuple $(x_1, \ldots, x_n)$ solution du système différentiel

$$(\Sigma) : \begin{cases} x_1' = x_2 \\ x_2' = x_3 \\ \vdots \\ x_{n-1}' = x_n \\ x_n' = a_{n-1}(t)x_n + \cdots + a_1(t)x_2 + a_0(t)x_1 + b(t) \end{cases}$$

dém. :
(i) $\Rightarrow$ (ii) Si x est solution sur I de l'équation alors x est n fois dérivable et le tuple $(x, x', \ldots, x^{(n-1)})$ est solution sur I du système.
(ii) $\Rightarrow$ (i) Si x est le premier élément d'un tuple $(x_1, \ldots, x_n)$ solution sur I du système alors les premières équations fournissent $x_2 = x_1' = x'$, $x_3 = x''$,..., $x_n = x^{(n-1)}$ et la dernière fournit la vérification par x de l'équation (E).
□

23.3.2 Problème de Cauchy

Soit $a_0, \ldots, a_n : I \to \mathbb{K}$ et $b : I \to \mathbb{K}$ continues. On étudie l'équation

$$(E) : x^{(n)} = a_{n-1}(t)x^{(n-1)} + a_{n-2}(t)x^{(n-2)} + \cdots + a_1(t)x' + a_0(t)x + b(t)$$

et on considère le système (Σ) associé comme défini dans la section ci-dessus.

Définition

Soit $(t_0, x_0, x_1 \ldots, x_{n-1}) \in I \times \mathbb{K}^n$. Un problème de Cauchy associé à l'équation (E) en t_0 consiste à déterminer les solutions de l'équation (E) vérifiant les conditions initiales

$$\forall 0 \leqslant k \leqslant n-1, x^{(k)}(t_0) = x_k$$

Remarque Ce problème est naturellement associé à un problème de Cauchy relatif au système (Σ) où la condition initiale sur ce système transpose les multiples conditions initiales imposées pour l'équation (E).

Théorème

> Le problème de Cauchy proposé possède une solution unique définie sur I.

dém. :
Car le problème de Cauchy associé au système différentiel admet une solution unique.
□

23.3.3 Structure de l'ensemble des solutions

Soit $a_0, \ldots, a_n : I \to \mathbb{K}$ et $b : I \to \mathbb{K}$ continues. On étudie l'équation

$$(E) : x^{(n)} = a_{n-1}(t)x^{(n-1)} + a_{n-2}(t)x^{(n-2)} + \cdots + a_1(t)x' + a_0(t)x + b(t)$$

23.3.3.1 Équation homogène

Définition

> L'équation (E_0) : $x^{(n)} = a_{n-1}(t)x^{(n-1)} + \cdots + a_1(t)x' + a_0(t)x$ est appelée équation homogène associée à (E) .

Théorème

> L'ensemble $\mathcal{S}_0$ des solutions sur I de l'équation homogène (E_0) est un sous-espace vectoriel de dimension n de l'espace $\mathcal{C}^n(I, \mathbb{K})$.

dém. :
Les solutions de l'équation homogène sont de classe $\mathcal{C}^n$ donc $\mathcal{S}_0 \subset \mathcal{C}^n(I, \mathbb{K})$.
Considérons l'application $\Phi : \mathcal{C}^n(I, \mathbb{K}) \to \mathcal{C}(I, \mathbb{K})$ définie par

$$\Phi(x) = x^{(n)} - \left(a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0x\right)$$

L'application Φ est linéaire et $\mathcal{S}_0 = \ker \Phi$ donc $\mathcal{S}_0$ est un sous-espace vectoriel de $\mathcal{C}^n(I, \mathbb{K})$.
Soit $t_0 \in I$. Considérons $E_{t_0} : \mathcal{S}_0 \to \mathbb{K}^n$ définie par $E_{t_0}(x) = (x(t_0), x'(t_0), \ldots, x^{(n-1)}(t_0))$.
L'application E_{t_0} est linéaire et comme un problème de Cauchy possède une solution unique, elle est bijective. C'est donc un isomorphisme et par conséquent

$$\dim \mathcal{S}_0 = \dim \mathbb{K}^n = n$$

□

23.3.3.2 Équation complète

Théorème

> L'ensemble $\mathcal{S}$ des solutions sur I de l'équation complète (E) est un sous-espace affine de $\mathcal{C}^n(I, \mathbb{K})$ de direction $\mathcal{S}_0$.

dém. :
Les solutions de l'équation complète sont de classe $\mathcal{C}^n$ et donc $\mathcal{S} \subset \mathcal{C}^n(I, \mathbb{K})$.

Par le théorème de Cauchy, on peut assurer l'existence d'une solution $\tilde{x}$.
Considérons à nouveau l'application Φ de la démonstration du théorème précédent.
Pour $x \in \mathcal{C}^2(I, \mathbb{K})$

$$x \in \mathcal{S} \Leftrightarrow \Phi(x) = \Phi(\tilde{x})$$

et donc

$$x \in \mathcal{S} \Leftrightarrow x - \tilde{x} \in \mathcal{S}_0$$

Ainsi l'ensemble $\mathcal{S}$ des solutions sur I est le sous-espace affine $\tilde{x} + \mathcal{S}_0$.
□

Remarque Pour résoudre $(E) : x'' = a(t)x' + b(t)x + c(t)$:
- on reconnaît le type l'équation ;
- on résout l'équation homogène : $x_0(t) = \dots$;
- on détermine une solution particulière : $\tilde{x}(t) = \dots$;
- on exprime la solution générale : $x(t) = \tilde{x}(t) + x_0(t)$.

23.3.4 Musculation : résolution des équations à coefficients constants

On étudie l'ensemble $\mathcal{S}$ des solutions à valeurs complexes de l'équation différentielle linéaire d'ordre n à coefficients constants :

$$x^{(n)} + a_{n-1}x^{(n-1)} + \dots + a_1 x' + a_0 x = 0$$

avec $a_i \in \mathbb{C}$ d'inconnue $x : \mathbb{R} \to \mathbb{C}$ n fois dérivable.

Proposition

Les solutions sur $\mathbb{R}$ de cette équation sont des fonctions de classe $\mathcal{C}^\infty$.

Considérons l'espace $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ et l'endomorphisme de celui-ci $D : x \mapsto x'$.
Pour $P = X^n + a_{n-1}X^{n-1} + \dots + a_0$ on a

$$\mathcal{S} = \ker P(D)$$

Dans $\mathbb{C}[X]$, on peut factoriser

$$P = (X - \lambda_1)^{\alpha_1} \dots (X - \lambda_m)^{\alpha_m}$$

avec $\lambda_k \in \mathbb{C}$ deux à deux distincts et $\alpha_k \in \mathbb{N}^\star$.
Pour $k \neq \ell$, $(X - \lambda_k)^{\alpha_k} \wedge (X - \lambda_\ell)^{\alpha_\ell} = 1$ donc

$$\ker P(D) = \bigoplus_{k=1}^{m} \ker(D - \lambda_k \mathrm{Id})^{\alpha_i}$$

Reste à déterminer : $\ker(D - \lambda\mathrm{Id})^\alpha$ avec $\lambda \in \mathbb{C}$ et $\alpha \in \mathbb{N}^\star$.
Cas $\alpha = 1$:

$$(D - \lambda\mathrm{Id})(x) = 0 \Leftrightarrow x' - \lambda x = 0 \Leftrightarrow \exists C \in \mathbb{C}, \forall t \in \mathbb{R}, x(t) = C\mathrm{e}^{\lambda t}$$

Introduisons $e_\lambda : t \mapsto \mathrm{e}^{\lambda t}$. On a donc

$$\ker(D - \lambda\mathrm{Id}) = \mathrm{Vect}(e_\lambda)$$

Cas général :
Soit $x \in E$ et y la fonction définie de sorte $x = e_\lambda y$ i.e. $y : t \to \mathrm{e}^{-\lambda t}x(t)$. On a

$$(D - \lambda\mathrm{Id})(x) = e_\lambda D(y),\ (D - \lambda\mathrm{Id})^2(x) = e_\lambda D^2(y), \dots,\ (D - \lambda\mathrm{Id})^\alpha(x) = e_\lambda D^\alpha(y)$$

donc

$$x \in \ker(D - \lambda \mathrm{Id})^\alpha \Leftrightarrow y \in \ker D^\alpha$$

Or la solution générale de l'équation $y^{(\alpha)} = 0$ est

$$y(t) = c_0 + c_1 t + \cdots + c_{\alpha-1} t^{\alpha-1}$$

avec $c_0, c_1, \ldots, c_{\alpha-1} \in \mathbb{C}$.
Ainsi

$$\ker(D - \lambda \mathrm{Id})^\alpha = \left\{ t \mapsto (c_0 + c_1 t + \cdots + c_{\alpha-1} t^{\alpha-1}) \mathrm{e}^{\lambda t} / c_0, c_1, \ldots, c_{\alpha-1} \in \mathbb{C} \right\}$$

Exemple Résoudre l'équation

$$y^{(4)} - 2y'' + y = 0$$

C'est une équation différentielle linéaire d'ordre 4 homogène à coefficients constants d'équation caractéristique $r^4 - 2r^2 + 1 = 0$ i.e. $(r-1)^2(r+1)^2 = 0$.
1 et -1 sont racines doubles.
La solution générale est

$$y : t \mapsto (at + b)\mathrm{e}^t + (ct + d)\mathrm{e}^{-t}$$

Remarque On a

$$\dim \ker P(D) = \sum_{k=1}^{m} \dim \ker(D - \lambda_k \mathrm{Id})^{\alpha_k} = \sum_{k=1}^{m} \alpha_k = n$$

et l'on retrouve que l'espace des solutions est de dimension n.

Chapitre 24

Equations différentielles linéaires scalaires

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide.

24.1 Equations linéaires d'ordre 1

24.1.1 Equation différentielle scalaire

Définition

On appelle équation différentielle (scalaire) linéaire d'ordre 1 définie sur I toute équation de la forme

$$(E) : x' = a(t)x + b(t)$$

avec $t \mapsto a(t)$ et $t \mapsto b(t)$ fonctions continues de I vers $\mathbb{K}$ et d'inconnue $t \mapsto x(t)$ fonction dérivable de I vers $\mathbb{K}$.

Remarque L'usage veut qu'on n'exprime pas la variable pour la fonction inconnue. Néanmoins, vérifier que la fonction x est solution sur I consiste à observer

$$\forall t \in I, x'(t) = a(t)x(t) + b(t)$$

Remarque Pour la théorie la fonction inconnue est notée x. En pratique, elle est souvent notée y.

Exemple Pour $a \in \mathbb{C}$, la solution générale de l'équation

$$(E) : y' + ay = 0$$

est

$$y(t) = \lambda \mathrm{e}^{-at} \text{ avec } \lambda \in \mathbb{C}$$

Proposition

Les fonctions solutions de (E) sont de classe $\mathcal{C}^1$ et même de classe $\mathcal{C}^{n+1}$ si a et b sont de classe $\mathcal{C}^n$.

24.1.2 Problème de Cauchy

Soit $a, b : I \to \mathbb{K}$ continue. On étudie l'équation

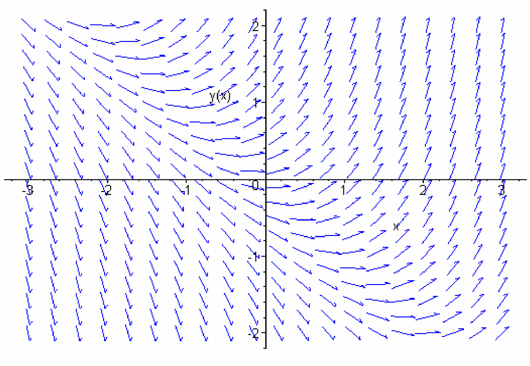
$$(E) : x' = a(t)x + b(t)$$

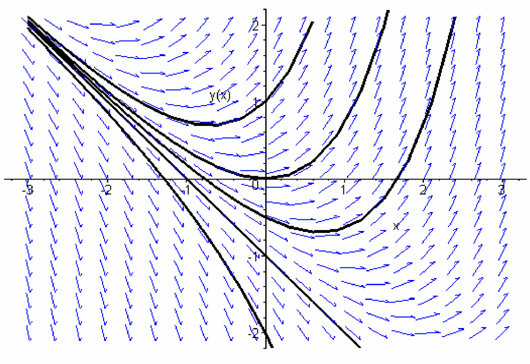
Définition

On appelle courbe intégrale de l'équation différentielle (E) tout graphe dans $\mathbb{R}^2$ d'une solution de celle-ci.

Remarque En chaque point d'une courbe intégrale, la tangente est déterminée par l'expression du second membre de l'équation différentielle. On peut alors figurer un champ de vecteurs dans le plan permettant d'anticiper l'allure des courbes intégrales.

Exemple Champ de vecteurs et quelques courbes intégrales associées à l'équation différentielle $y' = x + y$.





Définition

Soit $(t_0, x_0) \in I \times \mathbb{K}$. Un problème de Cauchy associé à l'équation (E) en t_0 consiste à déterminer les solutions de $(E) : x' = a(t)x + b(t)$ vérifiant la condition initiale

$$x(t_0) = x_0$$

Théorème

Soit $(t_0, x_0) \in I \times \mathbb{K}$. Le problème de Cauchy

$$\begin{cases} x' = a(t)x + b(t) \\ x(t_0) = x_0 \end{cases}$$

possède une unique solution définie sur I.

dém. :
Introduisons A la primitive s'annulant en t_0 de la fonction continue $a : I \to \mathbb{K}$.
Unicité : Si x est solution alors

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(x(t)\mathrm{e}^{-A(t)}\right) = \left(x'(t) - a(t)x(t)\right)\mathrm{e}^{-A(t)} = b(t)\mathrm{e}^{-A(t)}$$

donc $t \mapsto x(t)\mathrm{e}^{-A(t)}$ est de classe $\mathcal{C}^1$ et

$$x(t)\mathrm{e}^{-A(t)} = x(t_0) + \int_{t_0}^{t} \mathrm{e}^{-A(u)}b(u)\,\mathrm{d}u$$

puis

$$x(t) = \mathrm{e}^{A(t)}\left(x_0 + \int_{t_0}^{t} \mathrm{e}^{-A(u)}b(u)\,\mathrm{d}u\right)$$

Existence : La fonction définie par

$$x(t) = \left(x_0 + \int_{t_0}^{t} b(u)\mathrm{e}^{-A(u)}\,\mathrm{d}u\right)\mathrm{e}^{A(t)}$$

est bien solution.
□

Corollaire

Par chaque point de coordonnées $(t_0, x_0) \in I \times \mathbb{K}$ passe une courbe intégrale et une seule. En particulier, les courbes intégrales ne se recoupent pas, elles constituent une partition du domaine $I \times \mathbb{K}$ du plan.

24.1.3 Structure de l'ensemble solution

Soit $a, b : I \to \mathbb{K}$ continues. On étudie l'équation différentielle

$$(E) : x' = a(t)x + b(t)$$

24.1.3.1 Équation homogène

Définition

L'équation est appelée équation homogène associée à l'équation (E).
Ses solutions sont appelées solutions homogènes de l'équation (E).

Théorème

L'ensemble $\mathcal{S}_0$ des solutions sur I de l'équation homogène (E_0) est la droite vectorielle engendrée par

$$t \mapsto \mathrm{e}^{A(t)}$$

où A désigne une primitive de la fonction continue a.

dém. :
Soit x une fonction dérivable. On a

$$x'(t) = a(t)x(t) \Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(x(t)\mathrm{e}^{-A(t)}\right) = 0$$

et donc x est solution de (E_0) sur I si, et seulement si, x est de la former

$$t \mapsto \lambda \mathrm{e}^{A(t)} \text{ avec } \lambda \in \mathbb{K}$$

□

24.1.3.2 Résolution de l'équation complète

Rappel :
On appelle sous-espace affine d'un espace vectoriel E tout ensemble de la forme

$$V = a + F = \{a + x / x \in F\}$$

avec F un sous-espace vectoriel de E
Le sous-espace vectoriel F est unique, on l'appelle direction de V.
Il n'y a pas unicité de l'élément a décrivant le sous-espace affine V, au contraire, pour tout $a \in V$, on peut écrire

$$V = a + F$$

Un sous-espace affine est donc entièrement déterminé par la connaissance de sa direction et de l'un de ses éléments.

Théorème

L'ensemble $\mathcal{S}$ des solutions sur I de l'équation complète

$$(E) : x' = a(t)x + b(t)$$

est une droite affine de $\mathcal{C}^1(I, \mathbb{K})$ de direction l'espace $\mathcal{S}_0$.

dém. :
Les solutions sont de classe $\mathcal{C}^1$ donc $\mathcal{S} \subset \mathcal{C}^1(I, \mathbb{K})$.
Par problème de Cauchy, on peut assurer l'existence d'au moins une solution $\tilde{x}$ à l'équation étudiée.
Soit $x \in \mathcal{C}^1(I, \mathbb{K})$. On a alors

$$x \in \mathcal{S} \Leftrightarrow \forall t \in I, x'(t) - a(t)x(t) = \tilde{x}'(t) - a(t)\tilde{x}(t)$$

En ramenant au premier membre

$$x \in \mathcal{S} \Leftrightarrow \forall t \in I, (x - \tilde{x})'(t) - a(t)\,(x(t) - x(t)) = 0$$

et donc

$$x \in \mathcal{S} \Leftrightarrow x - \tilde{x} \in \mathcal{S}_0$$

Ainsi $\mathcal{S} = \tilde{x} + \mathcal{S}_0$ est un sous-espace affine de direction $\mathcal{S}_0$.
□
Protocole : Pour résoudre $(E) : x' = a(t)x + b(t)$:
- on identifie le type de l'équation (E) en reconnaissant a et b fonctions continues ;
- on résout l'équation homogène (E_0) : $x_0(t) = \dots$;
- on cherche une solution particulière : $\tilde{x}(t) = \dots$;
- on exprime la solution générale : $x(t) = \tilde{x}(t) + x_0(t)$.

Remarque Si $b(t) = b_1(t) + b_2(t)$ avec b_1 et $b_2 : I \to E$ fonctions continues et si x_1 et x_2 sont respectivement solutions des équations

$$(E_1) : x' = a(t)x + b_1(t) \text{ et } (E_2) : x' = a(t)x + b_2(t)$$

alors $\tilde{x} = x_1 + x_2$ est solution de l'équation

$$(E) : x' = a(t)(x) + b(t)$$

24.1.3.3 Méthode de la variation de la constante

Supposons la solution générale homogène de la forme

$$x_0(t) = \lambda\varphi(t) \text{ avec } \lambda \in \mathbb{K}$$

Théorème

> Par quadrature, on peut déterminer une solution particulière de l'équation complète (E) de la forme $x(t) = \lambda(t)\varphi(t)$ avec λ fonction dérivable bien choisie.

dém. :
x est solution de (E) si, et seulement si,

$$\forall t \in I, \lambda'(t)\varphi(t) = b(t)$$

Puisque la fonction φ est continue ne s'annule pas (c'est une fonction composée avec une exponentielle), on peut déterminer λ convenable pour que x soit solution de (E).
□

Exemple Résolvons l'équation

$$(E) : (1 + t^2)y' + 2ty = 1$$

On a

$$(E) \Leftrightarrow y' + \frac{2t}{1 + t^2}y = \frac{1}{1 + t^2}$$

(E) est équivalente à une équation différentielle linéaire d'ordre 1 définie sur $\mathbb{R}$.
Equation homogène :

$$(1 + t^2)y' + 2ty = 0 \Leftrightarrow y' = -\frac{2t}{1 + t^2}y$$

On a
$$\int \frac{-2t}{1+t^2}\,\mathrm{d}t = -\ln(1+t^2)$$
Solution homogène : $y(t) = \dfrac{\lambda}{1+t^2}$ avec $\lambda \in \mathbb{R}$
Solution particulière : $y(t) = \dfrac{\lambda(t)}{1+t^2}$ avec $t \mapsto \lambda(t)$ fonction dérivable.
$$(1+t^2)y'(t) + 2ty(t) = 1 \Leftrightarrow \lambda'(t) = 1$$
$\lambda(t) = t$ convient et $y(t) = \dfrac{t}{1+t^2}$ est solution particulière.
Solution générale
$$y(t) = \frac{\lambda + t}{1+t^2} \text{ avec } \lambda \in \mathbb{R}$$

24.2 Equation linéaire d'ordre 2

24.2.1 Définition

Définition

On appelle équation différentielle linéaire (scalaire) d'ordre 2 définie sur I toute équation de la forme
$$(E) : x'' = a(t)x' + b(t)x + c(t)$$
avec $a, b, c : I \to \mathbb{K}$ continues et d'inconnue $x : I \to \mathbb{K}$ deux fois dérivable.

Exemple Lorsque les fonctions a et b sont constantes, on parle d'équation à coefficients constants.

Exemple $y'' + 2ty' + (1-t^2)y = \mathrm{e}^t$ est une équation linéaire d'ordre 2 définie sur $\mathbb{R}$.

Exemple $(1+t^2)y'' + 2ty' + y = 0$ est équivalente sur $\mathbb{R}$ à une équation linéaire d'ordre 2 car
$$\forall t \in \mathbb{R}, (1+t^2) \neq 0$$

Proposition

Les solutions de (E) sont de classe $\mathcal{C}^2$ et plus généralement de classe $\mathcal{C}^{n+2}$ si a, b, c sont $\mathcal{C}^n$.

24.2.2 Problème de Cauchy.

Soit $a, b, c : I \to \mathbb{K}$ des fonctions continues.

Définition

Soit $(t_0, x_0, x'_0) \in I \times \mathbb{K}^2$. Un problème de Cauchy associé à l'équation (E) en t_0 consiste à déterminer les solutions de l'équation $(E) : x'' = a(t)x' + b(t)x + c(t)$ vérifiant les conditions initiales
$$x(t_0) = x_0 \text{ et } x'(t_0) = x'_0$$

Théorème

Soit $(t_0, x_0, x'_0) \in I \times \mathbb{K}^2$. Le problème de Cauchy

$$\begin{cases} x'' = a(t)x' + b(t)x + c(t) \\ x(t_0) = x_0 \\ x'(t_0) = x'_0 \end{cases}$$

possède une unique solution définie sur I. (admis)

Attention : Il ne faut pas confondre un problème de Cauchy avec un problème de conditions aux bords. Par exemple, les conditions $y(0) = 0$ et $y(2\pi) = 0$ ne déterminent pas une solution unique pour l'équation différentielle $y'' + y = 0$.

Exemple Considérons l'équation

$$(E) : y'' + p(x)y' + q(x)y = 0$$

avec $p, q : I \to \mathbb{R}$ continues.
Montrons que s'il existe $x_0 \in I$ vérifiant $y(x_0) = y'(x_0) = 0$ alors y est la fonction nulle.
En effet, la fonction nulle et la fonction y sont solutions au problème de Cauchy :

$$\begin{cases} y'' + p(x)y' + q(x)y = 0 \\ y(x_0) = y'(x_0) = 0 \end{cases}$$

Or ce problème de Cauchy détermine une solution unique.

24.2.3 Structure de l'ensemble des solutions

Soit $a, b, c : I \to \mathbb{K}$ continues. On étudie l'équation

$$(E) : x'' = a(t)x' + b(t)x + c(t)$$

24.2.3.1 Équation homogène

Définition

L'équation $(E_0) : x'' = a(t)x' + b(t)x$ est appelée équation homogène associée à (E).
Ses solutions sont appelées solutions homogènes de l'équation (E).

Théorème

L'ensemble $\mathcal{S}_0$ des solutions sur I de l'équation homogène (E_0) est un sous-espace vectoriel de $\mathcal{C}^2(I, \mathbb{K})$ de dimension 2.

dém. :
Les solutions de l'équation homogène sont de classe $\mathcal{C}^2$ donc $\mathcal{S}_0 \subset \mathcal{C}^2(I, \mathbb{K})$.
Considérons la fonction $\Phi : \mathcal{C}^2(I, \mathbb{K}) \to \mathcal{C}(I, \mathbb{K})$ définie par

$$\Phi(x) = x'' - (ax' + bx)$$

En fait, la fonction $\Phi(x)$ désigne l'application $t \mapsto x''(t) - (a(t)x'(t) + b(t)x(t))$.
La fonction Φ est linéaire et $\mathcal{S}_0 = \ker \Phi$ donc $\mathcal{S}_0$ est un sous-espace vectoriel de $\mathcal{C}^2(I, \mathbb{K})$.
Soit $t_0 \in I$. Considérons l'application $E_{t_0} : \mathcal{S}_0 \to \mathbb{K}^2$ définie par

$$E_{t_0}(x) = (x(t_0), x'(t_0))$$

L'application E_{t_0} est linéaire, par résolution d'un problème de Cauchy

$$\forall (x_0, x_0') \in \mathbb{K}^2, \exists! x \in \mathcal{S}_0, E_{t_0}(x) = (x_0, x_0')$$

L'application E_{t_0} est donc bijective et c'est par conséquent un isomorphisme. On en déduit

$$\dim \mathcal{S}_0 = \dim \mathbb{K}^2 = 2$$

□

24.2.3.2 Système fondamental de solutions

Définition

On appelle système fondamental de solutions de l'équation homogène $x'' = a(t)x' + b(t)x$ toute base (φ, ψ) de l'espace $\mathcal{S}_0$.

Remarque Si (φ, ψ) est un système fondamental de solutions alors on peut exprimer la solution générale de l'équation (E_0) qui est

$$x(t) = \lambda\varphi(t) + \mu\psi(t) \text{ avec } \lambda, \mu \in \mathbb{K}$$

Exemple Les solutions φ, ψ de l'équation homogène vérifiant les conditions initiales

$$\begin{cases} \varphi(t_0) = 1 \\ \varphi'(t_0) = 0 \end{cases} \text{ et } \begin{cases} \psi(t_0) = 0 \\ \psi'(t_0) = 1 \end{cases}$$

forment un système fondamental de solutions.

24.2.3.3 Wronskien

Définition

On appelle wronskien de deux solutions (φ, ψ) de l'équation homogène (E_0) la fonction

$$t \mapsto w(t) = \begin{vmatrix} \varphi(t) & \psi(t) \\ \varphi'(t) & \psi'(t) \end{vmatrix}$$

Théorème

Le wronskien w de deux solutions de l'équation $(E_0) : x'' = a(t)x' + b(t)x$ est solution de l'équation différentielle d'ordre 1

$$w'(t) = a(t)w(t)$$

dém. :
Par dérivation par ligne du déterminant

$$w'(t) = \begin{vmatrix} \varphi'(t) & \psi'(t) \\ \varphi'(t) & \psi'(t) \end{vmatrix} + \begin{vmatrix} \varphi(t) & \psi(t) \\ \varphi''(t) & \psi''(t) \end{vmatrix} = \begin{vmatrix} \varphi(t) & \psi(t) \\ a(t)\varphi'(t) + b(t)\varphi(t) & a(t)\psi'(t) + b(t)\psi(t) \end{vmatrix}$$

En décomposant la deuxième ligne en combinaison linéaire de deux lignes

$$w'(t) = a(t)\begin{vmatrix} \varphi(t) & \psi(t) \\ \varphi'(t) & \psi'(t) \end{vmatrix} + b(t)\begin{vmatrix} \varphi(t) & \psi(t) \\ \varphi(t) & \psi(t) \end{vmatrix} = a(t)w(t)$$

□

Exemple Le wronskien d'un couple de solutions de l'équation $x'' + q(t)x = 0$ est constant.

Corollaire

Un wronskien qui s'annule est la fonction nulle.

Théorème

Si φ, ψ sont solutions de l'équation homogène alors on a équivalence entre :
(i) (φ, ψ) est un système fondamental de solutions ;
(ii) $\forall t \in I, w(t) \neq 0$;
(iii) $\exists t_0 \in I, w(t_0) \neq 0$.

dém. :
Soit $t_0 \in I$, l'application $E_{t_0} : \mathcal{S}_0 \to \mathbb{K}^2$ définie par $E_{t_0}(x) = (x(t_0), x'(t_0))$ est un isomorphisme d'espaces vectoriels. Par conséquent la famille (φ, ψ) est un système fondamental de solutions de (E_0) si, et seulement si, la famille $(E_{t_0}(\varphi), E_{t_0}(\psi))$ est une base de $\mathbb{K}^2$ i.e. si, et seulement si, $w(t_0) \neq 0$.
□

24.2.3.4 Équation complète

Théorème

L'ensemble $\mathcal{S}$ des solutions sur I de l'équation complète

$$(E) : x'' = a(t)x' + b(t)x + c(t)$$

est un plan affine de $\mathcal{C}^2(I, \mathbb{K})$ de direction $\mathcal{S}_0$.

dém. :
Les solutions sur I de l'équation complète sont de classe $\mathcal{C}^2$ donc $\mathcal{S} \subset \mathcal{C}^2(I, \mathbb{K})$.
Considérons à nouveau l'application $\Phi : \mathcal{C}^2(I, \mathbb{K}) \to \mathcal{C}(I, \mathbb{K})$ définie par

$$\Phi(x) = x'' - (ax' + bx)$$

L'équation (E) se comprend alors comme l'équation $\Phi(x) = c$.
Par résolution d'un problème de Cauchy, on peut assurer l'existence d'une solution particulière $\tilde{x}$.
Pour $x \in \mathcal{C}^2(I, \mathbb{K})$

$$x \in \mathcal{S} \Leftrightarrow \Phi(x) = \Phi(\tilde{x})$$

et donc

$$x \in \mathcal{S} \Leftrightarrow x - \tilde{x} \in \mathcal{S}_0 \Leftrightarrow x \in \tilde{x} + \mathcal{S}_0$$

Ainsi l'ensemble $\mathcal{S}$ des solutions sur I est un sous-espace affine de direction $\mathcal{S}_0$.
□

Remarque Pour résoudre $(E) : x'' = a(t)x' + b(t)x + c(t)$:
- on identifie le type de l'équation (E) en reconnaissant a, b, c fonctions continues ;
- on résout l'équation homogène (E_0) : $x_0(t) = \dots$;
- on détermine une solution particulière : $\tilde{x}(t) = \dots$;
- on exprime la solution générale : $x(t) = \tilde{x}(t) + x_0(t)$.

Remarque On peut aussi énoncer un principe de superposition des solutions.

24.2.4 Cas des équations à coefficients constants

On étudie l'équation
$$(E) : y'' + ay' + by = 0$$
avec $a, b \in \mathbb{K}$ et $c : I \to \mathbb{K}$ continue.

24.2.4.1 Solution homogène

Considérons l'équation homogène associée
$$(E_0) : y'' + ay' + by = 0$$
Soit $\lambda \in \mathbb{K}$. La fonction $t \mapsto \mathrm{e}^{\lambda t}$ est solution de (E_0) si, et seulement si, λ est racine de l'équation
$$r^2 + ar + b = 0$$

Définition

L'équation $r^2 + ar + b = 0$ est appelée équation caractéristique associée à l'équation (E) (ou (E_0)).

Cas $\mathbb{K} = \mathbb{C}$.
Si $\Delta \neq 0$: deux solutions α, β
$\varphi(t) = \mathrm{e}^{\alpha t}$ et $\psi(t) = \mathrm{e}^{\beta t}$ sont solutions de (E_0).
$$w(0) = \begin{vmatrix} 1 & 1 \\ \alpha & \beta \end{vmatrix} = \beta - \alpha \neq 0$$
(φ, ψ) est un système fondamental de solutions de (E_0).
La solution générale est alors et
$$x(t) = \lambda \mathrm{e}^{\alpha t} + \mu \mathrm{e}^{\beta t} \text{ avec } \lambda, \mu \in \mathbb{C}$$
Si $\Delta = 0$: une solution double α
$\varphi(t) = \mathrm{e}^{\alpha t}$ et (après calculs) $\psi(t) = t\mathrm{e}^{\alpha t}$ sont solutions de (E_0).
$$w(0) = \begin{vmatrix} 1 & 0 \\ \alpha & 1 \end{vmatrix} = 1 \neq 0$$
La solution générale est alors
$$x(t) = \lambda \mathrm{e}^{\alpha t} + \mu t \mathrm{e}^{\alpha t} \text{ avec } \lambda, \mu \in \mathbb{C}$$

Cas $\mathbb{K} = \mathbb{R}$.
Si $\Delta > 0$ ou $\Delta = 0$: idem avec $\lambda, \mu \in \mathbb{R}$
Si $\Delta < 0$, 2 solutions conjuguées $\alpha \pm i\omega$ avec $\omega \neq 0$.
La fonction $t \mapsto \mathrm{e}^{(\alpha+i\omega)t}$ est solution complexe de (E_0) donc ses parties réelle et imaginaire $\varphi(t) = \mathrm{e}^{\alpha t}\cos(\omega t)$ et $\psi(t) = \mathrm{e}^{\alpha t}\sin(\omega t)$ sont solutions réelles de (E_0).

$$\omega(0) = \begin{vmatrix} 1 & 0 \\ \alpha & \omega \end{vmatrix} = \omega \neq 0$$

La solution générale est alors

$$x(t) = (\lambda\cos(\omega t) + \mu\sin(\omega t))\mathrm{e}^{\alpha t} \text{ avec } \lambda, \mu \in \mathbb{R}$$

24.2.4.2 Solution particulière

Cas $c(t) = A\mathrm{e}^{\alpha t}$ avec $A \in \mathbb{K}$
On peut trouver une solution particulière de la forme

$$y(t) = \begin{cases} C\mathrm{e}^{\alpha t} & \text{si } \alpha \text{ n'est pas racine de } r^2 + ar + b = 0 \\ Ct\mathrm{e}^{\alpha t} & \text{si } \alpha \text{ est racine simple de } r^2 + ar + b = 0 \\ Ct^2\mathrm{e}^{\alpha t} & \text{si } \alpha \text{ est racine double de } r^2 + ar + b = 0 \end{cases}$$

avec $C \in \mathbb{K}$ bien choisi
Cas $\mathbb{K} = \mathbb{R}$ et $c(t) = B\cos(\omega t)$ ou $B\sin(\omega t)$.
On peut aussi trouver une solution particulière en étudiant l'équation complexe

$$z'' + az' + bz = B\mathrm{e}^{i\omega t}$$

et en considérant la partie réelle ou imaginaire d'une solution particulière.

24.2.5 Méthode de la variation des constantes

Soit $a, b, c : I \to \mathbb{K}$ continues. On cherche une solution particulière de l'équation

$$(E) : x'' = a(t)x' + b(t)x + c(t)$$

Supposons connu un système fondamental (φ, ψ) de solutions de l'équation homogène

$$(E_0) : x'' = a(t)x' + b(t)x$$

La solution générale de l'équation homogène est

$$x(t) = \lambda\varphi(t) + \mu\psi(t)$$

Théorème

Par quadrature, on peut trouver une solution particulière sur I de l'équation

$$(E) : x'' = a(t)x' + b(t)x + c(t)$$

de la forme

$$x(t) = \lambda(t)\varphi(t) + \mu(t)\psi(t)$$

avec $\lambda, \mu : I \to \mathbb{K}$ fonctions dérivables vérifiant :

$$\begin{cases} \lambda'(t)\varphi(t) + \mu'(t)\psi(t) = 0 \\ \lambda'(t)\varphi'(t) + \mu'(t)\psi'(t) = c(t) \end{cases}$$

dém. :
Le système proposé est de Cramer car de déterminant

$$\begin{vmatrix} \varphi(t) & \psi(t) \\ \varphi'(t) & \psi'(t) \end{vmatrix} = w(t) \neq 0$$

On peut donc trouver des fonctions λ et μ dérivables vérifiant

$$\lambda'(t) = \frac{\begin{vmatrix} 0 & \psi(t) \\ c(t) & \psi'(t) \end{vmatrix}}{w(t)} \text{ et } \mu'(t) = \frac{\begin{vmatrix} \varphi(t) & 0 \\ \varphi'(t) & c(t) \end{vmatrix}}{w(t)}$$

Considérons alors la fonction $x = \lambda\varphi + \mu\psi$.
x est dérivable et $x' = (\lambda'\varphi + \mu'\psi) + (\lambda\varphi' + \mu\psi')$
Puisque $\lambda'\varphi + \mu'\psi = 0$, on simplifie $x' = \lambda\varphi' + \mu\psi'$.
x est alors deux fois dérivable et $x'' = \lambda'\varphi' + \mu'\psi' + \lambda\varphi'' + \mu\psi''$.
On vérifie alors $x'' = a(t)x' + b(t)x + c(t)$ puisque φ, ψ sont solutions de l'équation homogène et $\lambda'\varphi' + \mu'\psi' = c$
□

Exemple Résolvons

$$y'' - 2y' + y = \frac{\mathrm{e}^t}{1 + t^2}$$

C'est une équation différentielle linéaire d'ordre 2 à coefficients constants.
Equation caractéristique $r^2 - 2r + 1 = 0$ de racine double 1.
Solution générale homogène $y(t) = (\lambda t + \mu)\mathrm{e}^t$ avec $\lambda, \mu \in \mathbb{R}$.
Solution particulière $y(t) = \lambda(t)t\mathrm{e}^t + \mu(t)\mathrm{e}^t$ avec λ, μ fonction dérivables.
$y'(t) = \lambda'(t)t\mathrm{e}^t + \mu'(t)\mathrm{e}^t + \lambda(t)(t+1)\mathrm{e}^t + \mu(t)\mathrm{e}^t$.
On pose $\lambda'(t)t\mathrm{e}^t + \mu'(t)\mathrm{e}^t = 0$.
$y''(t) = \lambda'(t)(t+1)\mathrm{e}^t + \mu'(t)\mathrm{e}^t + \lambda(t)(t+2)\mathrm{e}^t + \mu(t)\mathrm{e}^t$.
$y''(t) - 2y'(t) + y(t) = \dfrac{\mathrm{e}^t}{1+t^2}$ si, et seulement si, $\lambda'(t)(t+1)\mathrm{e}^t + \mu'(t)\mathrm{e}^t = \dfrac{\mathrm{e}^t}{1+t^2}$.
Résolvons

$$\begin{cases} \lambda'(t)t\mathrm{e}^t + \mu'(t)\mathrm{e}^t = 0 \\ \lambda'(t)(t+1)\mathrm{e}^t + \mu'(t)\mathrm{e}^t = \dfrac{\mathrm{e}^t}{1+t^2} \end{cases}$$

On obtient

$$\begin{cases} \lambda'(t) = \dfrac{1}{1+t^2} \\ \mu'(t) = -\dfrac{t}{1+t^2} \end{cases}$$

$\lambda(t) = \arctan t$ et $\mu(t) = -\dfrac{1}{2}\ln\left(1 + t^2\right)$ conviennent.
Solution particulière

$$y(t) = \frac{1}{2}\left(2t\arctan(t) - \ln(1+t^2)\right)\mathrm{e}^t$$

Solution génératrice

$$y(t) = \frac{1}{2}\left(2t\arctan(t) - \ln(1+t^2)\right)\mathrm{e}^t + \lambda t\mathrm{e}^t + \mu\mathrm{e}^t \text{ avec } \lambda, \mu \in \mathbb{R}$$

Exemple Résolvons $y'' + y = f(t)$ avec $f : \mathbb{R} \to \mathbb{R}$ continue.
C'est une équation différentielle linéaire d'ordre 2 à coefficients constants.
Equation caractéristique $r^2 + 1 = 0$ de racines $\pm i$.
Solution générale homogène $y(t) = \lambda\cos(t) + \mu\sin(t)$ avec $\lambda, \mu \in \mathbb{R}$.
Solution particulière

$$y(t) = \lambda(t)\cos(t) + \mu(t)\sin(t)$$

avec λ et μ fonctions dérivables solutions du système

$$\begin{cases} \lambda'(t)\cos t + \mu'(t)\sin t = 0 \\ -\lambda'(t)\sin t + \mu'(t)\cos t = f(t) \end{cases}$$

Par les formules de Cramer, on obtient

$$\begin{cases} \lambda'(t) = -\sin(t)f(t) \\ \mu'(t) = \cos(t)f(t) \end{cases}$$

Pour

$$\lambda(t) = -\int_0^t \sin(u)f(u)\,\mathrm{d}u \text{ et } \mu(t) = \int_0^t f(u)\cos(u)\,\mathrm{d}u$$

on a

$$y(t) = \lambda(t)\cos(t) + \mu(t)\sin(t) = \int_0^t \sin(t-u)f(u)\,\mathrm{d}u$$

solution particulière.
Solution générale

$$y(t) = \int_0^t \sin(t-u)f(u)\,\mathrm{d}u + \lambda\cos(t) + \mu\sin(t) \text{ avec } \lambda, \mu \in \mathbb{R}$$

24.2.6 Résolution pratique de l'équation homogène

En dehors des équations à coefficients constants, il n'y a pas de méthode systématique (et surtout pas d'équation caractéristique).

24.2.6.1 Recherche de solutions polynomiales

Exemple Résolvons

$$(E) : (t^2 + 2t + 2)y'' - 2(t+1)y' + 2y = 0$$

Pour tout $t \in \mathbb{R}$, $t^2 + 2t + 2 \neq 0$ donc (E) est équivalente à une équation différentielle linéaire d'ordre 2 homogène définie sur $\mathbb{R}$.
Recherchons les fonctions polynomiales solutions.
Soit $y(t) = t^n + \cdots$ une fonction polynomiale.

$$(t^2 + 2t + 2)y'' - 2(t+1)y'(t) + 2y(t) = (n(n-1) - 2n + 2)t^n + \cdots$$

Si y est solution de (E) alors $n^2 - 3n + 2 = 0$ donc $n = 1$ ou 2.
On recherche désormais y de la forme $y(t) = at^2 + bt + c$.

$$(t^2 + 2t + 2)y'' - 2(t+1)y'(t) + 2y(t) = 2(c - b + 2a)$$

$y(t) = at^2 + bt + c$ est solution de $(E) \Leftrightarrow c = b - 2a \Leftrightarrow y(t) = a(t^2 - 2) + b(t + 1)$
Posons $\varphi(t) = t^2 - 2$ et $\psi(t) = t + 1$.
φ et ψ sont solutions de (E), elles sont visiblement indépendantes, elles forment donc un système fondamental de solutions.
Solution générale de (E) :

$$y(t) = \lambda(t^2 - 2) + \mu(t + 1) \text{ avec } \lambda, \mu \in \mathbb{R}$$

24.2.6.2 Recherche de solutions développables en séries entières

Exemple Résolvons sur $]-1, 1[$

$$(E) : (1 - t^2)y'' - 4ty' - 2y = 0$$

Pour tout $t \in]-1, 1[$, $1 - t^2 \neq 0$ donc (E) est équivalente à une équation différentielle linéaire d'ordre 2 homogène définie sur $]-1, 1[$.
Recherchons les fonctions développables en série entière au voisinage de 0.
Analyse :
Soit y la somme de la série entière $\sum a_n t^n$ de rayon de convergence $R > 0$.
Sur $]-R, R[$,

$$y(t) = \sum_{n=0}^{+\infty} a_n t^n, \, y'(t) = \sum_{n=1}^{+\infty} n a_n t^{n-1}$$

et

$$y''(t) = \sum_{n=2}^{+\infty} n(n-1)a_n t^{n-2} = \sum_{n=0}^{+\infty} (n+2)(n+1)a_{n+2} t^n$$

ce qui donne

$$(1 - t^2)y'' - 4ty' - 2y = \sum_{n=0}^{+\infty} (n+2)(n+1)(a_{n+2} - a_n)t^n$$

Par unicité des coefficients d'un développement en série entière

$$\forall t \in]-R, R[\, , \sum_{n=0}^{+\infty} (n+2)(n+1)(a_{n+2} - a_n)t^n = 0 \Leftrightarrow \forall n \in \mathbb{N}, (n+2)(n+1)(a_{n+2} - a_n) = 0$$

Ainsi y est solution de (E) sur $]-R, R[$ si, et seulement si,

$$\forall n \in \mathbb{N}, a_{n+2} - a_n = 0$$

On a alors pour tout $p \in \mathbb{N}$, $a_{2p} = a_0$ et $a_{2p+1} = a_1$ puis, par sommabilité

$$y(t) = \sum_{p=0}^{+\infty} a_{2p} t^{2p} + \sum_{p=0}^{+\infty} a_{2p+1} t^{2p+1} = \sum_{p=0}^{+\infty} a_0 t^{2p} + \sum_{p=0}^{+\infty} a_1 t^{2p+1}$$

ce qui donne

$$y(t) = \frac{a_0}{1 - t^2} + \frac{a_1 t}{1 - t^2} \text{ pour } t \in]-R, R[\text{ avec nécessairement } R \leqslant 1$$

Synthèse :
Soit
$$\varphi(t)=\frac{1}{1-t^2} \text{ et } \psi(t)=\frac{t}{1-t^2}$$
φ est développable en série entière sur $]-1,1[$ et par les calculs qui précèdent est solutions de l'équation différentielle (E) sur $]-1,1[$. Il en est de même pour ψ. Les fonctions φ et ψ sont deux solutions indépendantes, elles forment donc un système fondamental de solutions de (E).
Solution générale :
$$y(t)=\frac{\lambda+\mu t}{1-t^2} \text{ avec } \lambda,\mu\in\mathbb{R}$$

24.2.7 Autres démarches

24.2.7.1 Changement de fonction inconnue

Résoudre une équation différentielle par changement de fonction inconnue consiste à traduire l'équation étudiée en une nouvelle équation en la fonction inconnue proposée, généralement plus simple à résoudre.

Exemple Résolvons sur $\mathbb{R}$ l'équation
$$(E):(1+t^2)y''+4ty'+(1-t^2)y=0$$
en posant $z=(1+t^2)y$.
Soient $y:\mathbb{R}\to\mathbb{R}$ deux fois dérivable et $z:\mathbb{R}\to\mathbb{R}$ définie par $z(t)=(1+t^2)y(t)$.
z est deux fois dérivable
$$z(t)=(1+t^2)y(t)$$
$$z'(t)=(1+t^2)y'(t)+2ty(t)$$
$$z''(t)=(1+t^2)y''(t)+4ty'(t)+2y(t)$$
On remarque
$$(1+t^2)y''(t)+4ty'(t)+(1-t^2)y(t)=z''(t)-z(t)$$
donc
$$y \text{ est solution de } (E) \text{ sur } \mathbb{R} \Leftrightarrow z \text{ est solution sur } \mathbb{R} \text{ de } (E'):z''-z=0$$
(E') est une équation différentielle linéaire d'ordre 2 homogène à coefficients constants.
Solution générale
$$z(t)=\lambda e^t+\mu e^{-t}$$
et
$$y(t)=\frac{\lambda e^t+\mu e^{-t}}{1+t^2} \text{ avec } \lambda,\mu\in\mathbb{R}$$

Remarque Lorsque $\varphi(t)$ détermine une solution ne s'annulant pas de l'équation homogène associée à une équation
$$y''+a(t)y'+b(t)y=c(t)$$
alors le changement de fonction inconnue $y(t)=z(t)\varphi(t)$ permet de résoudre cette équation. En effet, on a alors
$$y''+a(t)y'+b(t)y=c(t)\Leftrightarrow\varphi(t)z''+(2\varphi'(t)+a(t)\varphi(t))z'=c(t)$$
qui apparaît comme une équation d'ordre 1 en la fonction inconnue z'.

Exemple Résolvons sur $]0,+\infty[$ l'équation

$$(E) : t^2y'' + ty' - y = t^2$$

La fonction $t \mapsto t$ est solution de l'équation homogène associée.
Réalisons alors le changement de fonction inconnue $y(t) = tz(t)$.
Pour $y :]0,+\infty[\to \mathbb{R}$ deux fois dérivable, la fonction z est aussi deux fois dérivable et

$$y'(t) = tz'(t) + z(t) \text{ et } y''(t) = tz''(t) + 2z'(t)$$

La fonction y est alors solution de (E) si, et seulement si,

$$\forall t > 0, t^3z''(t) + 3t^2z'(t) = t^2$$

La résolution de cette équation d'ordre 1 en la fonction z' donne

$$z'(t) = \frac{\lambda}{t^3} + \frac{1}{3} \text{ avec } \lambda \in \mathbb{R}$$

En intégrant

$$z(t) = \frac{\lambda'}{t^2} + \frac{t}{3} + \mu \text{ avec } \lambda', \mu \in \mathbb{R}$$

et enfin la solution générale de (E) est

$$y(t) = \frac{\lambda}{t} + \mu t + \frac{t^2}{3} \text{ avec } \lambda, \mu \in \mathbb{R}$$

24.2.7.2 Changement de variable

Résoudre une équation différentielle par changement de variable consiste à traduire l'équation étudiée en une nouvelle équation en la fonction inconnue de la nouvelle variable. Cette nouvelle équation est généralement plus simple à résoudre.

Exemple Résolvons sur $]0,+\infty[$ l'équation

$$(E) : x^2y'' + 3xy' + y = 0$$

On procède au changement de variable $x = \mathrm{e}^t$.
Soit $y :]0,+\infty[\to \mathbb{R}$ deux fois dérivable et $z : \mathbb{R} \to \mathbb{R}$ définie par $z(t) = y(x) = y\left(\mathrm{e}^t\right)$.
La fonction z est deux fois dérivable et

$$y(x) = z(\ln x), y'(x) = \frac{1}{x}z'(\ln x) \text{ et } y''(x) = \frac{1}{x^2}z''(\ln x) - \frac{1}{x^2}z'(\ln x)$$

La fonction y est alors solution de (E) sur $]0,+\infty[$ si, et seulement si,

$$\forall x > 0, z''(\ln x) + 2z'(\ln x) + z(\ln x) = 0$$

ce qui revient à dire que z est solution sur $\mathbb{R}$ de l'équation

$$z''(t) + 2z'(t) + z(t)$$

La solution générale de cette équation est

$$z(t) = (\lambda t + \mu)\mathrm{e}^{-t} \text{ avec } \lambda, \mu \in \mathbb{R}$$

La solution générale de l'équation (E) est donc

$$y(x) = \frac{\lambda \ln x + \mu}{x}$$

Remarque Le changement de variable $x = \mathrm{e}^t$ est adapté à la résolution sur $]0, +\infty[$ des équations de la forme

$$x^2y'' + axy' + by = 0$$

qu'il transforme en équation à coefficients constants

$$z'' + (a-1)z' + bz = 0$$

Pour résoudre sur $]-\infty, 0[$, il suffit de poser $x = -\mathrm{e}^t$.
Ce sont ici les équations différentielles d'Euler.

24.3 L'épineux problème des raccords

24.3.1 Rappel

Théorème

Soit $a \in I$ et $f : I\setminus\{a\} \to \mathbb{R}$ continue sur I et dérivable sur $I\setminus\{a\}$.
Si $f'(t) \xrightarrow[t\to a, t\neq a]{} \ell \in \mathbb{R}$ alors f est dérivable en a et $f'(a) = \ell$.
Si $f'(t) \xrightarrow[t\to a, t\neq a]{} +\infty$ (ou $-\infty$) alors f n'est pas dérivable en a, mais y présente une tangente verticale.

dém. :
Supposons $f'(t) \xrightarrow[t\to a]{} \ell \in \bar{\mathbb{R}}$. On étudie le taux de variation

$$\frac{1}{h}\left(f(a+h) - f(a)\right)$$

Cas a est intérieur à I :
Quand $h \to 0^+$, en appliquant le théorème des accroissements finis entre a et $a+h$, il existe c_h compris entre a et $a+h$ tel que

$$f(a+h) - f(a) = f'(c_h)h$$

et alors

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = f'(c_h) \to \ell$$

car $c_h \to a$ par encadrement. On en déduit $f_d'(a) = \ell$.
L'étude quand $h \to 0^-$ est analogue et fournit $f_g'(a) = \ell$ ce qui permet de conclure.
Cas a est extrémité de I : Une seule des deux études précédentes suffit pour conclure.
□

24.3.2 Résolution de l'équation $a(t)y' + b(t)y = c(t)$

Soit $a, b, c : I \to \mathbb{K}$ continues. On étudie l'équation différentielle

$$(E) : a(t)y' + b(t)y = c(t)$$

Si a ne s'annule pas sur I alors l'équation (E) est équivalente à

$$y' = \alpha(t)y + \beta(t) \text{ avec } \alpha = -b/a \text{ et } \beta = c/a \text{ qu'on sait résoudre}$$

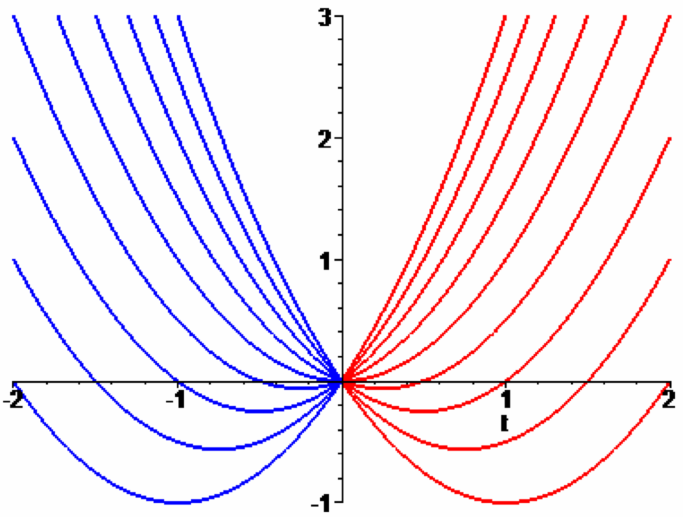
Si a s'annule alors
- on commence par résoudre (E) sur les plus grands intervalles $J \subset I$ sur lesquels a ne s'annule pas ;

- on procède ensuite au raccord des solutions aux points où a s'annule.
Pour raccorder les solutions en un point t_0 où a s'annule :
- on exprime une solution à droite et à gauche de t_0 ;
- on étudie s'il est possible de la prolonger par continuité en t_0 ;
- on étudie si ce prolongement est dérivable en t_0 ;
- on vérifie que l'équation différentielle est alors satisfaite en t_0.

Exemple Résolvons l'équation $(E) : ty' - y = t^2$ sur $\mathbb{R}$.
Sur $I = \mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$: $(E) \Leftrightarrow y' - \dfrac{1}{t}y = t$.
C'est une équation linéaire d'ordre 1.
Solution générale sur I : $y(t) = t^2 + \lambda t$ avec $\lambda \in \mathbb{R}$.
Déterminons les solutions de (E) sur $\mathbb{R}$.
Soit $y : \mathbb{R}^\star \to \mathbb{R}$ une solution de (E) sur $\mathbb{R}^{+\star}$ et $\mathbb{R}^{-\star}$.
Il existe $\lambda, \lambda' \in \mathbb{R}$ tels que

$$\forall t > 0, y(t) = t^2 + \lambda t \text{ et } \forall t < 0, y(t) = t^2 + \lambda' t$$

A quelle(s) condition(s) sur λ et λ' peut-on prolonger y en 0 pour obtenir une solution sur $\mathbb{R}$?
Continuité en 0 :
Quand $t \to 0^+$, $y(t) = t^2 + \lambda t \to 0$.
Quand $t \to 0^-$, $y(t) = t^2 + \lambda' t \to 0$.
Le prolongement en 0 est possible avec $y(0) = 0$ sans conditions sur λ, λ'.
Dérivabilité en 0 :
Quand $t \to 0^+$, $y'(t) = 2t + \lambda \to \lambda$ donc $y'_d(0) = \lambda$.
Quand $t \to 0^-$, $y'(t) = 2t + \lambda' \to \lambda'$ donc $y'_g(0) = \lambda'$.
Le prolongement en 0 est dérivable si, et seulement si, $\lambda = \lambda'$ et alors $y'(0) = \lambda$
Equation différentielle en 0 :
$0y'(0) - y(0) = 0$: ok.
Finalement :
Solution générale sur $\mathbb{R}$: $y(t) = t^2 + \lambda t$ avec $\lambda \in \mathbb{R}$.

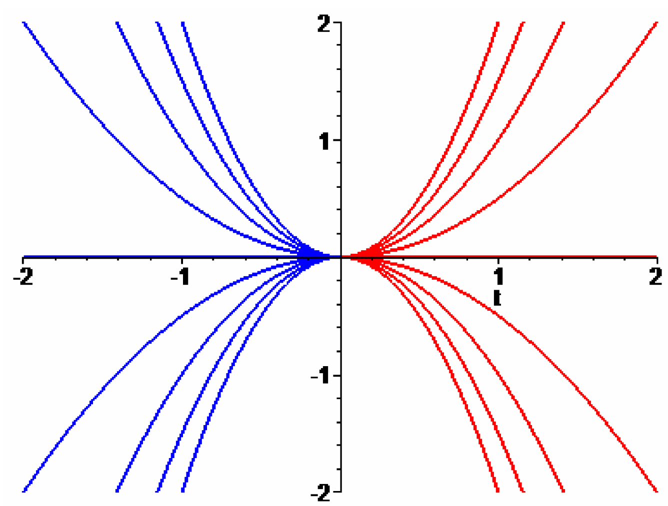


Exemple Résolvons l'équation $(E) : ty' - 2y = 0$ sur $\mathbb{R}$.
Sur $I = \mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$, $(E) \Leftrightarrow y' = \dfrac{2}{t}y$.
C'est une équation différentielle linéaire d'ordre 1.
Solution générale sur I, $y(t) = \lambda t^2$ avec $\lambda \in \mathbb{R}$.
Recherchons les solutions sur $\mathbb{R}$.
Soit $y : \mathbb{R}^\star \to \mathbb{R}$ une fonction solution sur $\mathbb{R}^{+\star}$ et $\mathbb{R}^{-\star}$.
Il existe $\lambda, \lambda' \in \mathbb{R}$ tels que

$$\forall t > 0, y(t) = \lambda t^2 \text{ et } \forall t < 0, y(t) = \lambda' t^2$$

A quelle(s) condition(s) sur λ et λ' peut-on prolonger y en 0 pour obtenir une solution sur $\mathbb{R}$?
Continuité en 0 :
Quand $t \to 0^+$, $y(t) = \lambda t^2 \to 0$.
Quand $t \to 0^-$, $y(t) = \lambda' t^2 \to 0$.
On peut prolonger y par continuité en 0 par $y(0) = 0$ sans conditions sur λ, λ'.
Quand $t \to 0^+$, $y'(t) = 2\lambda t \to 0$ donc $y'_d(0) = 0$.
Quand $t \to 0^-$, $y(t) = 2\lambda' t \to 0$ donc $y'_g(0) = 0$
Le prolongement en 0 est dérivable avec $y'(0) = 0$ sans conditions sur λ, λ'.
Equation différentielle en 0 :
$0y'(0) - 2y(0) = 0$: ok.
Finalement :
Solution générale sur $\mathbb{R}$

$$y(t) = \begin{cases} \lambda t^2 & \text{si } t > 0 \\ 0 & \text{si } t = 0 \\ \lambda' t^2 & \text{si } t < 0 \end{cases} \quad \text{avec } \lambda, \lambda' \in \mathbb{R}$$

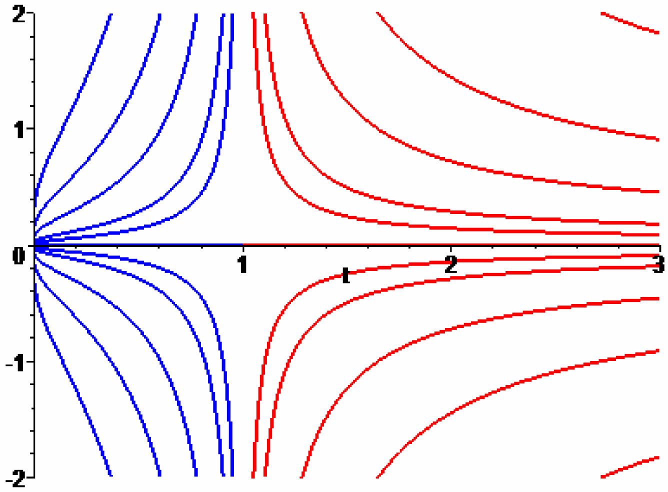


Exemple Résolvons l'équation $(E) : t\ln(t)y' + y = 0$ sur $]0, +\infty[$.
Sur $I =]0, 1[$ ou $]1, +\infty[$, $(E) \Leftrightarrow y' = -\dfrac{1}{t\ln t}y$.

C'est une équation différentielle linéaire d'ordre 1 homogène.
Solution générale sur I, $y(t) = \dfrac{\lambda}{\ln t}$.
Recherchons les solutions sur $]0, +\infty[$.
Soit $y :]0, 1[\cup]1, +\infty[\to \mathbb{R}$ une solution sur $]0, 1[$ et $]1, +\infty[$.
Il existe $\lambda, \lambda' \in \mathbb{R}$ tels que

$$\forall t \in]0, 1[\,, y(t) = \frac{\lambda}{\ln t} \text{ et } \forall t > 1, y(t) = \frac{\lambda'}{\ln t}$$

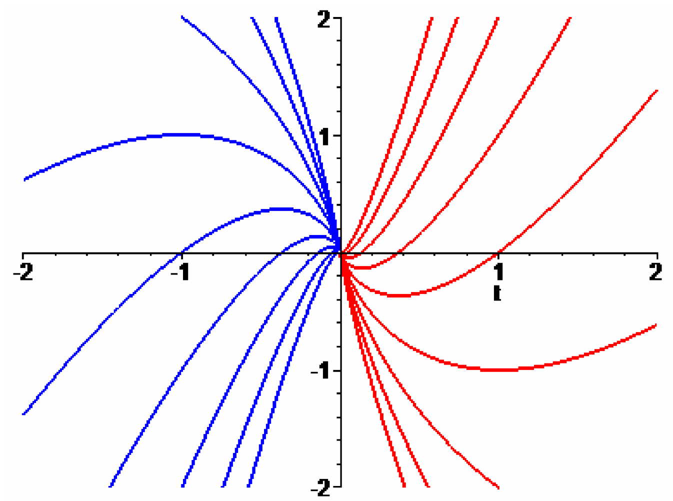
Continuité en 1 :
Quand $t \to 1^+$, $y(t) \to \begin{cases} +\infty & \text{si } \lambda' > 0 \\ 0 & \text{si } \lambda' = 0 \\ -\infty & \text{si } \lambda' < 0 \end{cases}$.
Quand $t \to 1^-$, $y(t) \to \begin{cases} -\infty & \text{si } \lambda > 0 \\ 0 & \text{si } \lambda = 0 \\ +\infty & \text{si } \lambda < 0 \end{cases}$
Le prolongement par continuité en 1 n'est possible que si $\lambda = \lambda' = 0$ et alors $y(t) = 0$ sur $]0, +\infty[$.
Inversement, cette fonction est évidemment solution sur $]0, +\infty[$
Solution générale sur $]0, +\infty[$: $y(t) = 0$.



Exemple Résolvons l'équation $(E) : ty' - y = t$ sur $\mathbb{R}$.
Sur $I = \mathbb{R}^{+\star}$ ou $\mathbb{R}^{-\star}$, $(E) \Leftrightarrow y' = \dfrac{1}{t}y + 1$.
C'est une équation différentielle linéaire d'ordre 1.
Solution générale sur I, $y(t) = t\ln|t| + \lambda t$ avec $\lambda \in \mathbb{R}$.
Recherchons les solutions sur $\mathbb{R}$.
Soit $y : \mathbb{R}^\star \to \mathbb{R}$ une fonction solution sur $\mathbb{R}^{+\star}$ et $\mathbb{R}^{-\star}$.
Il existe $\lambda, \lambda' \in \mathbb{R}$ tels que

$$\forall t > 0, y(t) = t\ln t + \lambda t \text{ et } \forall t < 0, y(t) = t\ln(-t) + \lambda' t$$

A quelle(s) condition(s) sur λ et λ' peut-on prolonger y en 0 pour obtenir une solution sur $\mathbb{R}$?
Continuité en 0 :
Quand $t \to 0^+$, $y(t) = t\ln t + \lambda t \to 0$.
Quand $t \to 0^-$, $y(t) = t\ln|t| + \lambda' t \to 0$.
On peut prolonger y par continuité en 0 par $y(0) = 0$ sans conditions sur λ, λ'.
Quand $t \to 0^+$, $y'(t) = \lambda + 1 + \ln t \to -\infty$.
Le prolongement en 0 n'est pas être dérivable en 0.
Il n'y a pas de solutions sur $\mathbb{R}$ à l'équation (E)



24.3.3 Résolution de l'équation $a(t)y'' + b(t)y' + c(t)y = d(t)$

La problématique est identique, cependant les raccords aux points où a s'annule s'obtiennent en étudiant la dérivabilité jusqu'à l'ordre 2.

Exemple Résolvons l'équation (E) : $(t-1)y'' - ty' + y = 0$ sur $\mathbb{R}$.
Sur $I =]-\infty, 1[$ ou $]1, +\infty[$:

$$(E) \Leftrightarrow y'' - \frac{t}{t-1}y' + \frac{1}{t-1}y = 0$$

C'est une équation linéaire homogène d'ordre 2.
$t \mapsto t$ et $t \mapsto \mathrm{e}^t$ sont solutions linéairement indépendantes donc forment un système fondamental de solutions sur $\mathbb{R}$.
La solution générale sur I est $y(t) = \lambda t + \mu \mathrm{e}^t$.
Notons que l'argument ne vaut pas sur $I = \mathbb{R}$, car on ne sait pas a priori si l'espace des solutions de (E) est de dimension 2.
Déterminons les solutions de (E) sur $\mathbb{R}$:
Soit $y : \mathbb{R}\setminus\{1\} \to \mathbb{R}$ une solution sur $]-\infty, 1[$ et $]1, +\infty[$.
Il existe $\lambda, \lambda', \mu, \mu' \in \mathbb{R}$ tels que

$$\forall t > 1, y(t) = \lambda t + \mu \mathrm{e}^t \text{ et } \forall t < 1, y(t) = \lambda' t + \mu' \mathrm{e}^t$$

Continuité en 1 :
Quand $t \to 1^+$, $y(t) \to \lambda + \mu\mathrm{e}$.
Quand $t \to 1^-$, $y(t) \to \lambda' + \mu'\mathrm{e}$.
On peut prolonger y en 1 si, et seulement si, $\lambda + \mu\mathrm{e} = \lambda' + \mu'\mathrm{e}$ et alors $y(1) = \lambda + \mu\mathrm{e}$.
Dérivabilité en 1 :
Quand $t \to 1^+$, $y'(t) = \lambda + \mu\mathrm{e}^t \to \lambda + \mu\mathrm{e}$ donc $y'_d(1) = \lambda + \mu\mathrm{e}$
Quand $t \to 1^-$, $y'(t) = \lambda' + \mu'\mathrm{e}^t \to \lambda' + \mu'\mathrm{e}$ donc $y'_g(1) = \lambda + \mu\mathrm{e}$
Le prolongement par continuité en 1 est dérivable et $y'(1) = \lambda + \mu\mathrm{e}$.
Dérivabilité à l'ordre 2 en 1 :
Quand $t \to 1^+$, $y''(t) = \mu\mathrm{e}^t \to \mu\mathrm{e}$.
Quand $t \to 1^-$, $y''(t) = \mu'\mathrm{e}^t = \mu'\mathrm{e}$.
Le prolongement est dérivable à l'ordre 2 en 1 si, et seulement si, $\mu = \mu'$ et alors $\lambda = \lambda'$ et $y''(1) = \mu$.
Vérification de l'équation différentielle en 1 :
$0y''(1) - y'(1) + y(1) = 0$: ok
Finalement :
Solution générale de (E) sur $\mathbb{R}$ $y(t) = \lambda t + \mu\mathrm{e}^t$ avec $\lambda, \mu \in \mathbb{R}$.

Remarque Comme pour les équations d'ordre 1 différents comportements sont possibles lors des raccords.
Par exemple, pour l'équation différentielle $t^2y'' + ty' - y = 0$, la solution générale sur $\mathbb{R}^{+\star}$ ou sur $\mathbb{R}^{-\star}$ est $y(t) = \lambda t + \mu/t$ et la solution générale sur $\mathbb{R}$ est $y(t) = \lambda t$.

Chapitre 25

Calcul différentiel

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
E, F, G et H désignent des $\mathbb{R}$-espaces vectoriels de dimensions finies non nulles indifféremment normés.
On pose $n = \dim E$ et $m = \dim F$
Ω et Ω' désignent des ouverts de E et F.
I désigne un intervalle ouvert de $\mathbb{R}$.

25.1 Différentielle d'une fonction

25.1.1 Développement limité à l'ordre 1

Soit $f : \Omega \subset E \to F$ et $a \in \Omega$

Définition

On appelle développement limité à l'ordre 1 de f en a toute écriture :

$$f(a+h) = f(a) + \ell(h) + \|h\|\,\varepsilon(h)$$

avec $\ell \in \mathcal{L}(E, F)$ et $\varepsilon(h) \to 0_F$ quand $h \to 0_E$
On dit alors que ℓ est application linéaire tangente à f en a.

Remarque On écrit souvent $o(h)$ pour $\|h\|\,\varepsilon(h)$.

Exemple Pour $f : (x, y) \in \mathbb{R}^2 \to \mathbb{R}$ un développement limité à l'ordre 1 en $(0, 0)$ est de la forme

$$f(x, y) = f(0, 0) + ax + by + o(x, y) \text{ quand } (x, y) \to (0, 0)$$

Proposition

Il y a unicité de l'application linéaire tangente décrivant un développement limité à l'ordre 1 de f en a.

dém. :
Supposons que $\ell, m \in \mathcal{L}(E, F)$ conviennent.

$$\ell(h) - m(h) = o(h) = \|h\|\,\varepsilon(h) \text{ avec } \varepsilon \xrightarrow[0_E]{} 0_F$$

Pour $v \in E$, considérons $h = \lambda.v$ avec $\lambda \to 0^+$.

$$\ell(\lambda v) - m(\lambda v) = \|\lambda.v\| \, \varepsilon(\lambda.v)$$

donne

$$\ell(v) - m(v) = \|v\| \, \varepsilon(\lambda.v)$$

Quand $\lambda \to 0^+$, on obtient $\ell(v) - m(v) \to 0_F$ et donc $\ell(v) = m(v)$ puis $\ell = m$.
□

25.1.2 Différentiabilité en un point

Soit $f : \Omega \subset E \to F$ et $a \in \Omega$

Définition

On dit que f est différentiable en a si f admet un développement limité à l'ordre 1 en a. L'application linéaire tangente à f en a est aussi appelée différentielle de f en a et on la note $\mathrm{d}f(a)$. Ainsi :

$$f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + o(h) \text{ quand } h \to 0_E$$

avec $\mathrm{d}f(a) \in \mathcal{L}(E, F)$.

Remarque On a ici adopté la notation d'opérateur. Il faut comprendre

$$\mathrm{d}f(a) \cdot h = [\mathrm{d}f(a)] \, (h)$$

Cette quantité se lit différentielle de f en a le long du vecteur h.

Théorème

Si f est différentiable en a alors f est continue en a.

dém. :
Par développement limité à l'ordre 1

$$f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + \|h\| \, \varepsilon(h) \text{ avec } \varepsilon(h) \xrightarrow[h \to 0_E]{} 0_F$$

L'application linéaire $\mathrm{d}f(a)$ étant continue puisqu'au départ d'un espace de dimension finie, on obtient

$$f(a+h) \xrightarrow[h \to 0_E]{} f(a) + 0_F + 0_F = f(a)$$

□

Exemple Si $f : E \to F$ est constante alors

$$\forall a \in E, \mathrm{d}f(a) = \tilde{0}$$

En effet, soit $a \in E$. On peut écrire

$$f(a+h) = f(a)$$

Quand $h \to 0_E$, $f(a+h) = f(a) + \ell(h) + o(h)$ avec $\ell = \tilde{0}$ linéaire
Ainsi,f est différentiable en a et $\mathrm{d}f(a) = \tilde{0}$.

Exemple Si $f : E \to F$ est linéaire alors

$$\forall a \in E, \mathrm{d}f(a) = f$$

En effet, soit $a \in E$. On peut écrire

$$f(a+h) = f(a) + f(h)$$

Quand $h \to 0_E$, $f(a+h) = f(a) + \ell(h) + o(h)$ avec $\ell = f$ linéaire.
Ainsi f est différentiable en a et $\mathrm{d}f(a) = f$.

Exemple Soit $f : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R})$ définie par $f(M) = M^2$ et $A \in \mathcal{M}_n(\mathbb{R})$.
Déterminons $\mathrm{d}f(A)$.

$$f(A+H) = (A+H)^2 = A^2 + AH + HA + H^2$$

Ainsi quand $H \to O_n$,

$$f(A+H) = (A+H)^2 = f(A) + \ell(H) + o(H)$$

avec $\ell(H) = AH + HA$, $\ell \in \mathcal{L}(\mathcal{M}_n(\mathbb{R}))$,
donc f est différentiable en A et $\mathrm{d}f(A) : H \to AH + HA$.

Exemple Soit $f : \mathbb{C}^\star \to \mathbb{C}$ définie par $f(z) = 1/z$ et $a \in \mathbb{C}^\star$.
Déterminons $\mathrm{d}f(a)$.

$$f(a+h) = \frac{1}{a+h} = \frac{1}{a}\frac{1}{1+h/a}$$

Or quand $u \in \mathbb{C} \to 0$, $\dfrac{1}{1+u} = 1 - u + o(u)$ car $\dfrac{1}{1+u} - (1-u) = \dfrac{u^2}{1+u} = O(u^2) = o(u)$.
Par suite
Quand $h \to 0$:

$$f(a+h) = \frac{1}{a}\left(1 - \frac{h}{a} + o(h)\right) = f(a) + \ell(h) + o(h)$$

avec $\ell : h \mapsto -h/a^2$ linéaire.
Ainsi f est différentiable en a et

$$\mathrm{d}f(a) : h \mapsto -\frac{h}{a^2}$$

Proposition

Soit $f : I \subset \mathbb{R} \to F$ et $a \in I$. On a équivalence entre :
(i) f est dérivable en a ;
(ii) f est différentiable en a.
De plus, on a alors

$$\mathrm{d}f(a) : h \mapsto h.f'(a) \text{ et } f'(a) = \mathrm{d}f(a) \cdot 1$$

dém. :
(i) $\Rightarrow$ (ii) Supposons f dérivable en a.
Quand $h \to 0$,

$$\frac{1}{h}\left(f(a+h) - f(a)\right) \to f'(a)$$

donc

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = f'(a) + \varepsilon(h) \text{ avec } \varepsilon(h) \xrightarrow[h\to 0]{} 0$$

puis

$$f(a+h) = f(a) + h.f'(a) + h\varepsilon(h) = f(a) + \ell(h) + o(h)$$

avec $\ell : h \mapsto h.f'(a)$, $\ell \in \mathcal{L}(\mathbb{R}, F)$.
Par suite f est différentiable en a et $\mathrm{d}f(a) : h \mapsto h.f'(a)$.
(ii) $\Rightarrow$ (i) Supposons f différentiable en a.
Quand $h \to 0$, $f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + o(h)$ donc

$$\frac{1}{h}\left(f(a+h) - f(a)\right) = \frac{1}{h}\left(\mathrm{d}f(a)\cdot h + o(h)\right) = \mathrm{d}f(a)\cdot 1 + o(1) \to \mathrm{d}f(a)(1)$$

Ainsi f est dérivable en a et $f'(a) = \mathrm{d}f(a) \cdot 1$.
□

25.1.3 Fonctions différentiables

Définition

Une fonction $f : \Omega \subset E \to F$ est dite différentiable si elle est différentiable en tout point $a \in \Omega$. L'application $\mathrm{d}f : \Omega \to \mathcal{L}(E, F)$ est alors appelée différentielle de f.

Théorème

Les fonctions différentiables sont continues.

Exemple Pour $f : I \subset \mathbb{R} \to F$

f est différentiable si, et seulement si, f est dérivable

Exemple Si $f : E \to F$ est constante alors f est différentiable en tout $a \in E$ et $\mathrm{d}f(a) = \tilde{0}$.
Par suite f est différentiable et $\mathrm{d}f = \tilde{0}$.

Exemple Si $f \in \mathcal{L}(E, F)$ alors f est différentiable en tout $a \in E$ et $\mathrm{d}f(a) = f$.
Par suite f est différentiable et $\mathrm{d}f : a \mapsto f$.
En identifiant constante et fonction égale à la constante, on écrit $\mathrm{d}f = f$.
En particulier
- $(x_1, \ldots, x_p) \in \mathbb{K}^p \mapsto x_j$ est différentiable ;
- $z \in \mathbb{C} \mapsto \mathrm{Re}(z), \mathrm{Im}(z)$ sont différentiables ;
- $A \in \mathcal{M}_{n,p}(\mathbb{K}) \mapsto a_{i,j}$ est différentiable.

25.1.4 Opérations

Théorème

Soit $f, g : \Omega \subset E \to F$ et $\lambda, \mu \in \mathbb{R}$.
Si f et g sont différentiables alors $\lambda f + \mu g$ l'est aussi et

$$\mathrm{d}(\lambda f + \mu g) = \lambda \mathrm{d}f + \mu \mathrm{d}g$$

dém. :
Soit $a \in U$.

$$(\lambda f + \mu g)(a+h) = \lambda f(a+h) + \mu g(a+h)$$

donc

$$(\lambda f + \mu g)(a+h) = \lambda\left(f(a) + \mathrm{d}f(a)\cdot h + \|h\|\,\varepsilon(h)\right) + \mu\left(g(a) + \mathrm{d}g(a)\cdot h + \|h\|\,\tilde{\varepsilon}(h)\right)$$

Par suite

$$(\lambda f + \mu g)(a+h) = (\lambda f + \mu g)(a) + \ell(h) + \|h\|\left(\varepsilon(h) + \tilde{\varepsilon}(h)\right)$$

avec $\ell = \lambda \mathrm{d}f(a) + \mu \mathrm{d}g(a) \in \mathcal{L}(E, F)$
Par suite $\lambda f + \mu g$ est différentiable en a et $\mathrm{d}(\lambda f + \mu g)(a) = \lambda \mathrm{d}f(a) + \mu \mathrm{d}g(a)$.
□

Corollaire

L'ensemble des fonctions différentiables de Ω vers F constitue une sous-espace vectoriel de $\mathcal{F}(\Omega, F)$.

Théorème

Soit $f : \Omega \subset E \to F$, $g : \Omega \subset E \to F$ et $b : F \times G \to H$ bilinéaire.
Si f et g sont différentiables alors $b(f, g)$ l'est aussi et

$$\mathrm{d}\left(b(f,g)\right) = b(\mathrm{d}f, g) + b(f, \mathrm{d}g)$$

dém. :
Soit $a \in U$.

$$b(f,g)(a+h) = b\left(f(a+h), g(a+h)\right)$$

donne

$$b\left(f(a), g(a)\right) = b\left(f(a) + \mathrm{d}f(a)\cdot h + \|h\|\,\varepsilon(h), g(a) + \mathrm{d}g(a)\cdot h + \|h\|\,\tilde{\varepsilon}(h)\right)$$

En développant

$$b(f,g)(a+h) = b(f,g)(a) + b\left(\mathrm{d}f(a)\cdot h, g(a)\right) + b\left(f(a),\, \mathrm{d}g(a)\cdot h\right) + \varphi(h)$$

avec

$$\varphi(h) = b\left(f(a), \|h\|\,\tilde{\varepsilon}(h)\right) + b\left(\mathrm{d}f(a)\cdot h,\, \mathrm{d}g(a)\cdot h\right) + \cdots$$

(où les termes de $\cdots$ sont semblables ou pires…)
Les applications linéaires $\mathrm{d}f(a)$ et $\mathrm{d}g(a)$ sont continues et donc il existe $k_f, k_g \in \mathbb{R}^+$ vérifiant

$$\forall h \in E, \|\mathrm{d}f(a)\cdot h\| \leqslant k_f\,\|h\| \text{ et } \|\mathrm{d}g(a)\cdot h\| \leqslant k_g\,\|h\|$$

De plus, la forme bilinéaire b étant au départ d'un produit d'espace de dimension finie, elle est aussi continue et il existe donc $k \in \mathbb{R}^+$ vérifiant

$$\forall (h, h') \in E \times F, \|b(h, h')\| \leqslant k \|h\| \|h'\|$$

On a alors

$$\|\varphi(h)\| \leqslant k \|f(a)\| \|h\| \|\tilde{\varepsilon}(h)\| + kk_f k_g \|h\|^2 + \cdots = o(h)$$

Ainsi

$$b(f, g)(a + h) = b(f, g)(a) + \ell(h) + o(h)$$

avec $\ell : h \mapsto b(\mathrm{d}f(a) \cdot h, g(a)) + b(f(a), \mathrm{d}g(a) \cdot h)$ linéaire.
Ainsi $b(f, g)$ est différentiable en a et

$$\mathrm{d}(b(f, g))(a) : h \mapsto b(\mathrm{d}f(a) \cdot h, g(a)) + b(f(a), \mathrm{d}g(a) \cdot h)$$

Abusivement, on écrit

$$\mathrm{d}(b(f, g))(a) = b(\mathrm{d}f(a), g(a)) + b(f(a), \mathrm{d}g(a))$$

puis

$$\mathrm{d}b(f, g) = b(\mathrm{d}f, g) + b(f, \mathrm{d}g)$$

□

Corollaire

Si F est une algèbre (par exemple $F = \mathbb{R}$, $\mathbb{C}$, $\mathcal{M}_n(\mathbb{R})$,...) alors pour $f, g : \Omega \to F$ différentiables, fg est différentiable et

$$\mathrm{d}(fg) = (\mathrm{d}f)g + f(\mathrm{d}g)$$

L'ensemble des fonctions différentiables de Ω vers F constitue alors une sous-algèbre de $\mathcal{F}(\Omega, F)$.

dém. :
L'application $b : F \times F \to F$ définie par $b(x, y) = xy$ est bilinéaire.
□

Remarque On peut aussi appliquer ce résultat à un produit scalaire, un produit extérieur,...

Exemple Les fonctions polynomiales sur $\mathbb{R}^n$ sont différentiables.

Exemple La fonction $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est différentiable car $\det$ est somme et produit de fonctions différentiables.

Théorème

Soit $f : \Omega \subset E \to F$.
On a équivalence entre :
(i) f est différentiable ;
(ii) les fonctions coordonnées de f dans une base de F le sont.

dém. :
Soit $e' = (e'_1, \ldots, e'_m)$ une base de F.
(i) $\Rightarrow$ (ii) Si f est différentiable en a alors

$$f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + \|h\| \,\varepsilon(h) \text{ avec } \varepsilon(h) \xrightarrow[h \to 0_E]{} 0_F$$

En notant :
- $f_1, \ldots, f_m$ les fonctions coordonnées de f dans la base e' ;
- $\varepsilon_1, \ldots, \varepsilon_m$ les fonctions coordonnées de ε dans la base e' ;
- $(\mathrm{d}f(a))_1, \ldots, (\mathrm{d}f(a))_m$ les fonctions coordonnées de $\mathrm{d}f(a)$ dans la base e' ;
on obtient en passant aux coordonnées le développement limité précédent

$$\forall 1 \leqslant k \leqslant m, f_k(a+h) = f_k(a) + (\mathrm{d}f(a))_k \cdot h + \|h\| \,\varepsilon_k(h)$$

avec $(\mathrm{d}f(a))_k$ linéaire et $\varepsilon_k(h) \xrightarrow[h \to 0_E]{} 0_F$.
(ii) $\Rightarrow$ (i) C'est un raisonnement analogue en sens inverse.
□

Exemple La fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = (x+y, xy)$ est différentiable.
En effet, ses fonctions coordonnées le sont.

Exemple La fonction $M \mapsto \text{com}(M)$ est différentiable.
En effet, les coefficients de com(M) sont des polynômes en les coefficients de M donc des fonctions différentiables.

25.1.5 Composition

Théorème

Soit $f : \Omega \subset E \to F$ et $g : \Omega' \subset F \to G$ telles que $f(\Omega) \subset \Omega'$.
Si f et g sont différentiables alors $g \circ f$ aussi et

$$\forall a \in \Omega, \mathrm{d}(g \circ f)(a) = [\mathrm{d}g(f(a))] \circ \mathrm{d}f(a)$$

dém. :
Soit $a \in \Omega$. On peut écrire

$$f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + \|h\| \,\varepsilon(h) \text{ avec } \varepsilon(h) \xrightarrow[h \to 0_E]{} 0_F$$

Ainsi

$$f(a+h) = f(a) + h' \text{ avec } h' = \mathrm{d}f(a) \cdot h + \|h\| \,\varepsilon(h)$$

Aussi

$$g(f(a) + h') = g(f(a)) + \mathrm{d}g(f(a)) \cdot h' + \|h'\| \,\varepsilon'(h') \text{ avec } \varepsilon'(h') \xrightarrow[h' \to 0_E]{} 0_F$$

puis

$$(g \circ f)(a+h) = g\left(f(a)\right) + \mathrm{d}g(f(a)) \cdot (\mathrm{d}f(a) \cdot h) + \varphi(h)$$

avec

$$\varphi(h) = \|h\| \, \mathrm{d}g(f(a)) \cdot \varepsilon(h) + \|h'\| \, \tilde{\varepsilon}(h')$$

Par continuité de $\mathrm{d}f(a)$, on a $\|\mathrm{d}f(a) \cdot h\| \leqslant k_f \|h\|$ puis $\|h'\| \leqslant (k_f + |\varepsilon(h)|) \|h\|$ ce qui donne $\varphi(h) = o(h)$.
Ainsi

$$(g \circ f)(a+h) = g(f(a)) + (\mathrm{d}g(f(a)) \circ \mathrm{d}f(a)) \cdot h + o(h)$$

avec $\mathrm{d}g(f(a)) \circ \mathrm{d}f(a) \in \mathcal{L}(E, H)$.
Finalement $g \circ f$ est différentiable en a et

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a)$$

□

Exemple Les fonctions rationnelles sur $\mathbb{R}^p$ sont différentiables.
En effet, l'inverse d'une fonction polynomiale est différentiable par un argument de composition.

Exemple La fonction $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $\Phi(r, \theta) = (r\cos\theta, r\sin\theta)$ est différentiable.
En effet, ses fonctions coordonnées le sont par un argument de composition.

Corollaire

> Soit $f : \Omega \subset E \to \mathbb{R}$ et $\varphi : I \subset \mathbb{R} \to \mathbb{R}$ telles que $f(\Omega) \subset I$.
> Si f est différentiable et φ dérivable $\varphi(f) = \varphi \circ f$ l'est aussi
>
> $$\mathrm{d}\varphi(f) = \varphi'(f).\mathrm{d}f$$

dém. :
$\mathrm{d}(\varphi \circ f)(a) = \mathrm{d}\varphi(f(a)) \circ \mathrm{d}f(a)$ or $\mathrm{d}\varphi(f(a)) : h \mapsto \varphi'(f(a)).h$ donc $\mathrm{d}(\varphi \circ f)(a) = \varphi'(f(a)).\mathrm{d}f(a)$.
□

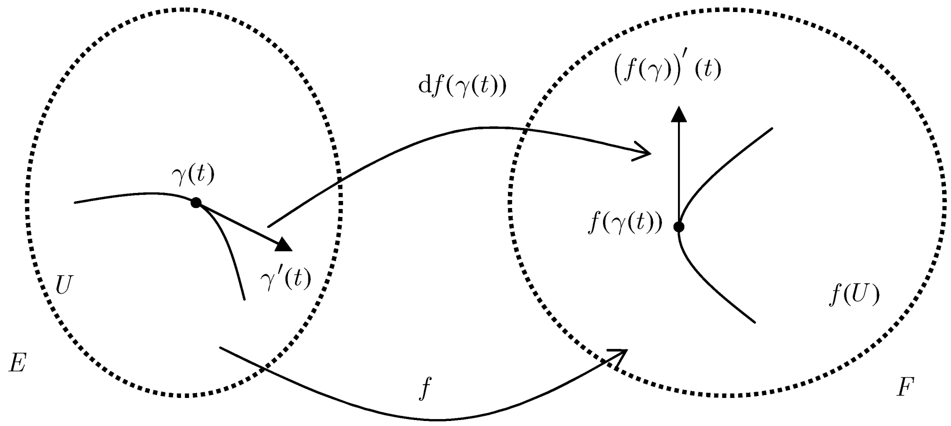
Exemple $\mathrm{d}(f^n) = nf^{n-1}\mathrm{d}f$, $\mathrm{d}\left(\dfrac{1}{f}\right) = -\dfrac{1}{f^2}\mathrm{d}f$, $\mathrm{d}(\ln f) = \dfrac{\mathrm{d}f}{f}$,...

Corollaire

> Soit $\gamma : I \subset \mathbb{R} \to E$ et $f : \Omega \subset E \to F$ telles que $\gamma(I) \subset \Omega$.
> Si γ est dérivable et f différentiable alors $t \mapsto f(\gamma(t))$ est dérivable et
>
> $$(f \circ \gamma)'(t) = \mathrm{d}f(\gamma(t)) \cdot \gamma'(t)$$

dém. :
$(f \circ \gamma)'(t) = \mathrm{d}(f \circ \gamma)(t) \cdot 1 = (\mathrm{d}f(\gamma(t)) \circ \mathrm{d}\gamma(t)) \cdot 1 = \mathrm{d}f(\gamma(t))\gamma'(t)$ car $\gamma'(t) = \mathrm{d}\gamma(t) \cdot 1$.
□

Remarque L'application γ se comprend comme le paramétrage d'un mobile inscrit évoluant dans E. Si l'on comprend f comme une transformation géométrique, $f \circ \gamma$ est un paramétrage de l'arc transformé. La formule de dérivation montre que le vecteur vitesse en un point de l'arc γ est transformé par la différentielle à f en ce point pour former le vecteur vitesse à l'arc transformé.

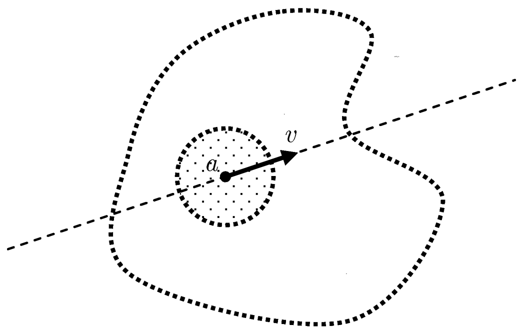


25.2 Dérivées partielles

La différentielle est une application compliquée. Par la notion de dérivée partielle, nous allons accéder simplement à ses valeurs.

25.2.1 Dérivation selon un vecteur

Soit $f : \Omega \subset E \to F$ et $a \in \Omega$. Puisque Ω est ouvert, il existe $\alpha > 0$ tel que $B(a, \alpha) \subset \Omega$.
Pour $v \in E$ fixé, la fonction $t \in \mathbb{R} \mapsto f(a + t.v)$ est définie au voisinage de 0, elle étudie les valeurs prises par f sur la droite affine $a + \mathrm{Vect}v$.



Définition

On dit que f est dérivable en a selon le vecteur v si la fonction $t \mapsto f(a + t.v)$ est dérivable en 0.
On pose alors
$$\mathrm{D}_v f(a) = \lim_{t \to 0} \frac{1}{t} \left(f(a + t.v) - f(a)\right)$$
appelé vecteur dérivé de f en a selon le vecteur v.

Théorème

Si f est différentiable en a alors f est dérivable en a selon tout vecteur $v \in E$ et

$$\mathrm{D}_v f(a) = \mathrm{d}f(a) \cdot v$$

dém. :
Quand $h \to 0_E$,

$$f(a+h) = f(a) + \mathrm{d}f(a) \cdot h + \|h\| \, \varepsilon(h) \text{ avec } \varepsilon(h) \xrightarrow[h \to 0_E]{} 0_F$$

Pour $v \in E$ fixé.
Quand $t \to 0$,

$$f(a+t.v) = f(a) + \mathrm{d}f(a) \cdot (t.v) + \|t.v\| \, \varepsilon(t.v) = f(a) + t.\mathrm{d}f(a) \cdot h + o(t)$$

car $\mathrm{d}f(a)$ est linéaire.
Par suite

$$\frac{1}{t}\left(f(a+t.v) - f(a)\right) \to \mathrm{d}f(a) \cdot v$$

□

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3/y$ pour $y \neq 0$ et $f(x,0) = 0$.
Soit $v = (v_x, v_y) \in \mathbb{R}^2$. Etudions $\mathrm{D}_v f(0,0)$.

$$\frac{1}{t}\left(f((0,0)+t.v) - f(0,0)\right) = \frac{1}{t} f(t.v_x, t.v_y)$$

Si $v_y \neq 0$ alors

$$\frac{1}{t} f(t.v_x, t.v_y) = \frac{t^3 v_x^3}{t^2 v_y} \xrightarrow[t \to 0]{} 0$$

Si $v_y = 0$ alors

$$\frac{1}{t} f(t.v_x, t.v_y) = 0 \xrightarrow[t \to 0]{} 0$$

Ainsi f est dérivable en $(0,0)$ selon tout vecteur v et $\mathrm{D}_v f(0,0) = 0$.
Cependant f n'est pas continue en (0,0) (et a fortiori n'y est pas différentiable) car
$f(1/n, 1/n^3) = 1 \xrightarrow[n \to +\infty]{} 1 \neq f(0,0)$ alors $\left(1/n, 1/n^3\right) \to (0,0)$.

25.2.2 Dérivées partielles

Choisissons arbitrairement une base $e = (e_1, \ldots, e_n)$ de E.
Soit $f : \Omega \subset E \to F$.

Définition

Sous réserve d'existence, on appelle i-ème dérivé partiel de f (dans la base e) en $a \in \Omega$ le vecteur dérivé de f en a selon le vecteur e_i. On note alors

$$\partial_i f(a) = \mathrm{D}_{e_i} f(a) = \lim_{t \to 0} \frac{1}{t}\left(f(a+t.e_i) - f(a)\right)$$

Exemple Calculons les dérivées partielles de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x_1, x_2) = x_1 x_2^2$ relatives à la base canonique.
Notons $c = (c_1, c_2)$ la base canonique de $\mathbb{R}^2$.
Les dérivées partielles de f dans c en (x_1, x_2) sont

$$\partial_1 f(x_1, x_2) = \lim_{t \to 0} \frac{1}{t} \left(f(x_1 + t, x_2) - f(x_1, x_2)\right) = x_2^2$$

$$\partial_2 f(x_1, x_2) = \lim_{t \to 0} \frac{1}{t} \left(f(x_1, x_2 + t) - f(x_1, x_2)\right) = 2x_1 x_2$$

Définition

Sous réserve d'existence, l'application $\partial_i f : \Omega \subset E \to F$ est appelée i-ème dérivée partielle de f (dans la base e).

Théorème

Si $f : \Omega \subset E \to F$ est différentiable alors les dérivées partielles de f dans la base $e = (e_1, \ldots, e_n)$ existent et pour tout $a \in \Omega$ on a

$$\partial_i f(a) = \mathrm{d}f(a) \cdot e_i$$

De plus,

$$\forall h = \sum_{i=1}^{n} h_i . e_i \in E, \mathrm{d}f(a) \cdot h = \mathrm{D}_h f(a) = \sum_{i=1}^{n} h_i . \partial_i f(a)$$

dém. :
Si f est différentiable alors pour tout $a \in U$ et tout $h \in E$, f est dérivable a selon le vecteur h et

$$\mathrm{D}_h f(a) = \mathrm{d}f(a) \cdot h$$

En particulier, pour $h = e_i$,

$$\partial_i f(a) = \mathrm{D}_{e_i} f(a) = \mathrm{d}f(a) \cdot e_i$$

De plus, si $h_1 = h_1.e_1 + \cdots + h_n.e_n$ alors

$$\mathrm{d}f(a) \cdot h = \mathrm{d}f(a) \cdot \left(\sum_{i=1}^{n} h_i . e_i\right) = \sum_{i=1}^{n} h_i . \mathrm{d}f(a) \cdot e_i = \sum_{i=1}^{n} h_i . \partial_i f(a)$$

car $\mathrm{d}f(a)$ est une application linéaire.
□

Corollaire

Le développement limité à l'ordre 1 de f en a s'écrit alors

$$f(a + h) = f(a) + \sum_{i=1}^{n} h_i . \partial_i f(a) + o(h) \text{ quand } h \to 0_E$$

Remarque Sous l'hypothèse « f est différentiable en a » , les dérivées partielles permettent de calculer la différentielle de f... Il reste à savoir calculer les dérivées partielles de f !

25.2.3 Dérivées partielles d'une fonction de n variables réelles

Soit $f : \Omega \subset \mathbb{R}^n \to F$ donnée par

$$f : x = (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$$

On étudie les dérivées partielles de f dans la base canonique $e = (e_1, \ldots, e_n)$ de $\mathbb{R}^n$.

Théorème

Sous réserve d'existence

$$\partial_i f(a) = \frac{\mathrm{d}}{\mathrm{d}x_i} \left(f(a_1, \ldots, x_i, \ldots, a_n) \right)\big|_{x_i = a_i}$$

dém. :
Sous réserve d'existence

$$\partial_i f(a) = \lim_{t \to 0} \left(\frac{1}{t} \left(f(a + te_i) - f(a) \right) \right) = \lim_{t \to 0} \left(\frac{1}{t} \left(f(a_1, \ldots, a_i + t, \ldots, a_n) - f(a_1, \ldots, a_n) \right) \right)$$

Ainsi $\partial_i f(a)$ apparaît comme la dérivée en $x_i = a_i$ de l'application $x_i \mapsto f(a_1, \ldots, x_i, \ldots, a_n)$.
□

Remarque Ainsi et de façon synthétique

$$\partial_i f(x) = \frac{\mathrm{d}}{\mathrm{d}x_i} \left(f(x_1, \ldots, x_i, \ldots, x_n) \right)$$

Définition

Si l'on a convenu de noter $x_1, \ldots, x_n$ les éléments du n-uplet x, il est usuel de noter

$$\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$$

plutôt que $\partial_1 f, \ldots, \partial_n f$ les dérivées partielles de f. Ainsi

$$\frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) = \frac{\mathrm{d}}{\mathrm{d}x_i} \left(f(x_1, \ldots, x_n) \right) = \lim_{t \to 0} \frac{1}{t} \left(f(x + te_i) - f(x) \right)$$

Exemple Calcul des dérivées partielles dans la base canonique de $f : \mathbb{R}^3 \to \mathbb{R}$ définie par $f(x, y, z) = x^2 + z\sin(xy)$.
Les dérivées partielles de f sont

$\dfrac{\partial f}{\partial x}(x, y, z) = \dfrac{\mathrm{d}}{\mathrm{d}x} \left(x^2 + z\sin(xy) \right) = 2x + yz\cos(xy)$,

$\dfrac{\partial f}{\partial y}(x, y, z) = \dfrac{\mathrm{d}}{\mathrm{d}y} \left(x^2 + z\sin(xy) \right) = xz\cos(xy)$ et

$\dfrac{\partial f}{\partial z}(x, y, z) = \dfrac{\mathrm{d}}{\mathrm{d}z} \left(x^2 + z\sin(xy) \right) = \sin(xy)$

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

Calcul des dérivées partielles dans la base canonique de f en $(0,0)$.

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \to 0} \frac{1}{t}\left(f(t,0) - f(0,0)\right) = 1 \text{ et } \frac{\partial f}{\partial y}(0,0) = \lim_{t \to 0} \frac{1}{t}\left(f(0,t) - f(0,0)\right) = -1$$

25.2.4 Dérivées partielles d'une fonction d'une variable vectorielle

Soit $f : \Omega \subset E \to F$ et $e = (e_1, \dots, e_n)$ une base de E.
Pour $x \in \Omega$, convenons de noter $x_1, \dots, x_n \in \mathbb{R}$ les coordonnées de x dans la base e. On a alors

$$f(x) = f(x_1 e_1 + \cdots + x_n e_n)$$

Il est alors usuel d'identifier la fonction f avec la fonction de n variables réelles donnée par

$$f(x_1, \dots, x_n) = f(x)$$

Exemple Soit $f : \mathbb{C} \to \mathbb{C}$. En munissant $\mathbb{C}$ de la base canonique $(1, i)$, on identifie $f : z \mapsto f(z)$ avec la fonction

$$f : (x,y) \mapsto f(x + i.y)$$

Exemple Soit $f : \mathcal{M}_2(\mathbb{R}) \to \mathbb{R}$. En munissant $\mathcal{M}_2(\mathbb{R})$ de sa base canonique, on identifie $f : M \mapsto f(M)$ avec l'application

$$f : (a,b,c,d) \mapsto f\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Théorème

Sous réserve d'existence, les dérivées partielles dans la base $e = (e_1, \dots, e_n)$ de f en a sont alors données par

$$\partial_i f(a) = \frac{\mathrm{d}}{\mathrm{d}x_i}\left(f(a_1, \dots, x_i, \dots, a_n)\right)\Big|_{x_i = a_i}$$

dém. :

$$\partial_i f(a) = \lim_{t \to 0}\left(\frac{1}{t}\left(f(a + te_i) - f(a)\right)\right) = \lim_{t \to 0}\left(\frac{1}{t}\left(f(a_1, \dots, a_i + t, \dots, a_n) - f(a_1, \dots, a_n)\right)\right)$$

Ainsi $\partial_i f(a)$ apparaît comme la dérivée en $x_i = a_i$ de l'application $x_i \mapsto f(a_1, \dots, x_i, \dots, a_n)$.
□

Remarque Ainsi

$$\partial_i f(x) = \frac{\mathrm{d}}{\mathrm{d}x_i}\left(f(x_1, \ldots, x_n)\right)$$

Définition

Si l'on a convenu de noter $x_1, \ldots, x_n$ les coordonnées de la variable x dans la base e, il est usuel de noter $\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_n}$ les dérivées partielles de f. Ainsi

$$\frac{\partial f}{\partial x_i}(x) = \frac{\mathrm{d}}{\mathrm{d}x_i}\left(f(x_1, \ldots, x_n)\right) = \lim_{t \to 0} \frac{1}{t}\left(f(x + te_i) - f(x)\right)$$

Exemple Soit $f : \mathbb{C}^\star \to \mathbb{C}$ définie par $f(z) = 1/z$.
Calculons les dérivées partielles dans la base canonique de f en $z = x + iy$.

$$\frac{\partial f}{\partial x}(z) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{z}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x+iy}\right) = -\frac{1}{(x+iy)^2} = -\frac{1}{z^2}$$

et

$$\frac{\partial f}{\partial y}(z) = \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{1}{x+iy}\right) = -\frac{i}{z^2}$$

Exemple Soit $f : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ définie par $f(M) = M^2$.
Calculons les dérivées partielles dans la base canonique de f en $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

$$\frac{\partial f}{\partial a}(M) = \frac{\mathrm{d}}{\mathrm{d}a}\left(M^2\right) = \frac{\mathrm{d}}{\mathrm{d}a}\begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} = \begin{pmatrix} 2a & b \\ c & 0 \end{pmatrix}, \frac{\partial f}{\partial b}(M) = \begin{pmatrix} c & a+d \\ 0 & c \end{pmatrix}, \ldots$$

25.2.5 Matrice jacobienne

On suppose les espaces E et F munis de bases $e = (e_1, \ldots, e_n)$ et $e' = (e'_1, \ldots, e'_m)$.
Soit $f : \Omega \subset E \to F$ différentiable en $a \in \Omega$.

Définition

On appelle matrice jacobienne de f en a la matrice de l'application linéaire $\mathrm{d}f(a)$ relative aux bases e et e'

$$\mathrm{Jac} f(a) \underset{\text{déf}}{=} \mathrm{Mat}_{e,e'}(\mathrm{d}f(a)) \in \mathcal{M}_{m,n}(\mathbb{R})$$

Théorème

En notant $f_1, \ldots, f_m$ les fonctions coordonnées de f alors

$$\mathrm{Jac} f(x) = \left(\partial_i f_k(x)\right)_{1 \leqslant k \leqslant m, 1 \leqslant i \leqslant n} = \begin{pmatrix} \partial_1 f_1(x) & \cdots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \cdots & \partial_n f_m(x) \end{pmatrix}$$

dém. :
Les colonnes de la matrice $\text{Jac}(f)(x) = \text{Mat}_{e,e'}(\mathrm{d}f(x))$ sont formées par les coordonnées dans e' des images des vecteurs de la base e. Or

$$\mathrm{d}f(x)e_i = \partial_i f(x) = \frac{\mathrm{d}}{\mathrm{d}x_i}\left(f_1(x).e'_1 + \cdots + f_m(x).e'_m\right) = \sum_{k=1}^{m} \partial_i f_k(x).e'_k$$

et l'on remplit la matrice jacobienne comme proposé.
□

Remarque Si l'on convient de noter $x_1, \ldots, x_n$ les coordonnées de la variable x

$$\text{Jac}f(x) = \left(\frac{\partial f_k}{\partial x_i}(x)\right)_{1\leqslant k\leqslant m, 1\leqslant i\leqslant n} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(x) & \cdots & \dfrac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(x) & \cdots & \dfrac{\partial f_m}{\partial x_n}(x) \end{pmatrix}$$

Remarque Pour une fonction $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, l'usage veut que l'on travaille relativement aux bases canoniques pour définir la matrice jacobienne.

Exemple Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par $f(x,y,z) = \left(x^2+y^2+z^2, xyz\right)$.

$$\text{Jac}f(x,y,z) = \begin{pmatrix} 2x & 2y & 2z \\ yz & xz & xy \end{pmatrix}$$

Exemple Soit $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $\Phi(r,\theta) = (r\cos\theta, r\sin\theta)$.

$$\text{Jac}\Phi(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Remarque Cette matrice jacobienne caractérise la différentielle de f en a et donne ainsi accès au développement limité à l'ordre 1 de f en a.

Exemple Pour l'application Φ ci-dessus

$$\Phi(r+r', \theta+\theta') \underset{(r',\theta')\to(0,0)}{=} \Phi(r,\theta) + (\cos(\theta)r' - r\sin(\theta)\theta', \sin(\theta)r' + r\cos(\theta)\theta') + o(r',\theta')$$

et la relation revêt même une certaine élégance en écrivant $\mathrm{d}r, \mathrm{d}\theta$ au lieu de r', θ'...

25.2.6 Opération sur les dérivées partielles

On munit E d'une base $e = (e_1, \dots, e_n)$.

Théorème

Soit $f, g : \Omega \subset E \to F$ et $\lambda, \mu \in \mathbb{R}$.
Si f et g admettent des dérivées partielles alors $\lambda f + \mu g$ aussi et

$$\partial_i(\lambda f + \mu g) = \lambda \partial_i f + \mu \partial_i g$$

dém. :
Soit $x \in E$. On écrit $x = x_1.e_1 + \cdots + x_n.e_n$ et l'on comprend les fonctions f et g comme des fonctions de n variables réelles. La dérivée partielle $\partial_i f$ s'obtient par dérivation d'application partielle

$$\partial_i f(x) = \frac{\mathrm{d}}{\mathrm{d}x_i}\left(f(x_1, \dots, x_n)\right)$$

et alors

$$\partial_i(\lambda.f + \mu g)(x) = \frac{\mathrm{d}}{\mathrm{d}x_i}\left(\lambda.f(x_1, \dots, x_n) + \mu.g(x_1, \dots, x_n)\right)$$

Par dérivation d'une fonction d'une variable réelle

$$\partial_i(\lambda.f + \mu g)(x) = \lambda.\partial_i f(x) + \mu.\partial_i g(x)$$

□

Remarque Dans le cas où f et g sont différentiables, ce résultat se retrouve aussi par

$$\mathrm{d}(\lambda f + \mu g)(a) = \lambda\,\mathrm{d}f(a) + \mu\,\mathrm{d}g(a)$$

Théorème

Soit $f : \Omega \subset E \to F$, $g : \Omega \subset E \to G$ et $b : F \times G \to H$ bilinéaire.
Si f et g admettent des dérivées partielles alors $b(f, g)$ aussi et

$$\partial_i b(f, g) = b\left(\partial_i f, g\right) + b\left(f, \partial_i g\right)$$

dém. :
Comme au-dessus par dérivation des applications partielles.
□

Théorème

Soit $f : \Omega \subset E \to F$.
On a équivalence entre :
(i) f admet des dérivées partielles ;
(ii) les fonctions coordonnées de f admettent des dérivées partielles
De plus, on a alors

$$\left(\partial_i f\right)_k = \partial_i\left(f_k\right)$$

en notant f_k et $\left(\dfrac{\partial f}{\partial x_i}\right)_k$ les fonctions coordonnées de f et $\dfrac{\partial f_k}{\partial x_i}$.

dém. :
Comme au-dessus par dérivation des applications partielles.
□

25.2.7 Dérivées partielles d'une fonction composée de fonctions différentiables

On suppose E et F munis de bases $e = (e_1, \dots, e_n)$ et $e' = (e'_1, \dots, e'_m)$.

Théorème

Soit $f : \Omega \subset E \to F$ et $g : \Omega' \subset F \to G$ telles que $f(\Omega) \subset \Omega'$.
Si f et g sont différentiables alors les dérivées partielles de $g \circ f$ sont données par

$$\partial_i \left(g \circ f\right)(a) = \sum_{k=1}^{m} \partial_i f_k(a).\partial_k g(f(a))$$

dém. :
f et g sont différentiables donc $g \circ f$ l'est aussi et

$$\mathrm{d}(g \circ f)(a) = [\mathrm{d}g(f(a))] \circ \mathrm{d}f(a)$$

Or

$$\partial_i(g \circ f)(a) = \mathrm{d}(g \circ f)(a) \cdot e_i$$

donc

$$\partial_i \left(g \circ f\right)(a) = [(\mathrm{d}g)(f(a))] \cdot \partial_i f(a)$$

avec

$$\partial_i f(a) = \sum_{k=1}^{m} \partial_i f_k(a) \cdot e'_k$$

puis par linéarité

$$\partial_i \left(g \circ f\right)(a) = \sum_{k=1}^{m} \partial_i f_k(a). \left[(\mathrm{d}g)(f(a))\right] \cdot e'_k$$

ce qui donne

$$\partial_i \left(g \circ f\right)(a) = \sum_{k=1}^{m} \partial_i f_k(a).\partial_k g(f(a))$$

□

Remarque Si l'on convient de noter $x_1, \dots, x_n$ les coordonnées d'un vecteur générique $x \in E$ et $y_1, \dots, y_m$ les coordonnées d'un vecteur générique $y \in F$ la formule se réécrit

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{k=1}^{m} \frac{\partial f_k}{\partial x_i}(a) \frac{\partial g}{\partial y_k}(f(a))$$

Exemple Soit $\gamma : I \subset \mathbb{R} \to E$ et $f : \Omega \subset E \to F$ telles que $\gamma(I) \subset \Omega$.
On note $x_1, \dots, x_n$ les coordonnées d'un vecteur générique $x \in E$ et on note encore $x_1, \dots, x_n$ les fonctions coordonnées de γ de sorte que

$$f(x) = f(x_1, \dots, x_n) \text{ et } \gamma(t) = x_1(t)e_1 + \dots + x_n(t)e_n$$

et donc
$$f(\gamma(t)) = f(x_1(t), \dots, x_n(t))$$
Si f est différentiable et γ dérivable alors $t \mapsto f(\gamma(t))$ est dérivable et
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(f(x_1(t), \dots, x_n(t))\right) = x_1'(t)\frac{\partial f}{\partial x_1}(\gamma(t)) + \cdots + x_n'(t)\frac{\partial f}{\partial x_n}(\gamma(t))$$

Exemple Soit $f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) \in \mathbb{R}$ différentiable
Calculons la dérivée de $t \in \mathbb{R} \mapsto f(2t, 1 + t^2)$.
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(f(2t, 1 + t^2)\right) = 2\frac{\partial f}{\partial x}(2t, 1 + t^2) + 2t\frac{\partial f}{\partial y}(2t, 1 + t^2)$$

Attention : Ici, écrire $\dfrac{\partial f}{\partial t}$ n'aurait pas de sens.

Exemple Soit $f : (u, v) \in \mathbb{R}^2 \mapsto f(u, v) \in \mathbb{R}$ différentiable.
Calculons la dérivée de $t \mapsto f(\cos(t), \sin(t))$
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(f(\cos(t), \sin(t))\right) = -\sin t\frac{\partial f}{\partial u}(\cos t, \sin t) + \cos t\frac{\partial f}{\partial v}(\cos t, \sin t)$$

Exemple Soit $f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) \in \mathbb{R}$ et $\Phi : (u, v) \in \mathbb{R}^2 \mapsto (\varphi(u, v), \psi(u, v)) \in \mathbb{R}^2$ différentiables.
Calculons ses dérivées partielles de $g = f \circ \Phi : (u, v) \mapsto f\left(\varphi(u, v), \psi(u, v)\right)$.
$$\frac{\partial g}{\partial u}(u, v) = \frac{\mathrm{d}}{\mathrm{d}u}\left(f(\varphi(u, v), \psi(u, v)\right)$$
$$= \frac{\partial \varphi}{\partial u}(u, v)\frac{\partial f}{\partial x}\left(\varphi(u, v), \psi(u, v)\right) + \frac{\partial \psi}{\partial u}(u, v)\frac{\partial f}{\partial y}\left(\varphi(u, v), \psi(u, v)\right)$$
$$\frac{\partial g}{\partial v}(u, v) = \frac{\mathrm{d}}{\mathrm{d}v}\left(f(\varphi(u, v), \psi(u, v)\right)$$
$$= \frac{\partial \varphi}{\partial v}(u, v)\frac{\partial f}{\partial x}\left(\varphi(u, v), \psi(u, v)\right) + \frac{\partial \psi}{\partial v}(u, v)\frac{\partial f}{\partial y}\left(\varphi(u, v), \psi(u, v)\right)$$

Attention : Ici, écrire $\dfrac{\partial f}{\partial u}$ n'aurait pas de sens.

Exemple Soit $f : (a, b) \in \mathbb{R}^2 \mapsto f(a, b) \in \mathbb{R}$ différentiable.
Calculons les dérivées partielles de $g : (x, y) \in \mathbb{R}^2 \mapsto f(x + y, xy)$.
$\dfrac{\partial g}{\partial x}(x, y) = \dfrac{\partial f}{\partial a}(x + y, xy) + y\dfrac{\partial f}{\partial b}(x + y, xy)$,
$\dfrac{\partial g}{\partial y}(x, y) = \dfrac{\partial f}{\partial a}(x + y, xy) + x\dfrac{\partial f}{\partial b}(x + y, xy)$.

Exemple Soit $f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) \in \mathbb{R}$ différentiable.
Calculons les dérivées partielles de $g : (r, \theta) \in \mathbb{R}^2 \mapsto f(r\cos\theta, r\sin\theta)$.
$\dfrac{\partial g}{\partial r}(r, \theta) = \cos\theta\dfrac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta\dfrac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)$,
$\dfrac{\partial g}{\partial \theta}(r, \theta) = -r\sin\theta\dfrac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta\dfrac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)$.

Remarque Les résultats qui précèdent se retiennent sous la forme de « la règle de la chaîne » :

$$\frac{\partial}{\partial u}\left(f(x_1, \ldots, x_n)\right) = \frac{\partial x_1}{\partial u}\frac{\partial f}{\partial x_1} + \cdots + \frac{\partial x_n}{\partial u}\frac{\partial f}{\partial x_n}$$

25.3 Classe d'une fonction

25.3.1 Fonction de classe $\mathcal{C}^1$

Théorème

Soit $f : \Omega \subset E \to F$. On a équivalence entre :
(i) f est différentiable et $\mathrm{d}f$ est continue ;
(ii) les dérivées partielles de f dans une base de E existent et sont continues.

dém. :
(i) $\Rightarrow$ (ii) Supposons f différentiable et $\mathrm{d}f$ continue.
Les dérivées partielles de f dans une base $e = (e_1, \ldots, e_n)$ existent et sont données par

$$\partial_j f(a) = \mathrm{d}f(a) \cdot e_j$$

Puisque l'application $a \mapsto \mathrm{d}f(a)$ est continue, que l'application constante $a \mapsto e_j$ est continue et que l'application $b : \mathcal{L}(E, F) \times E \to F$ est bilinéaire, on peut affirmer que l'application $a \mapsto \partial_j f(a)$ est continue par opérations sur les fonctions continues.
(ii) $\Rightarrow$ (i) Supposons f de classe $\mathcal{C}^1$ dans la base $e = (e_1, \ldots, e_n)$.
Cas $n = 2$
On identifie la fonction f avec l'application

$$f : (x_1, x_2) \mapsto f(x_1, x_2) = f(x_1 e_1 + x_2 e_2)$$

En raisonnant moyennant les fonctions coordonnées dans une base de F, on peut supposer $F = \mathbb{R}$.
Soit $a = (a_1, a_2) \in \Omega$.
Quand $h = (h_1, h_2) \to (0, 0)$, écrivons

$$f(a + h) = f(a) + \ell(h) + o(h)$$

On a

$$f(a + h) - f(a) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) + f(a_1, a_2 + h_2) - f(a_1, a_2)$$

En appliquant le théorème des accroissements finis aux applications $x_1 \mapsto f(x_1, a_2 + h_2)$ et $x_2 \mapsto f(a_1, x_2)$, il existe, d'une part, c_h compris entre a_1 et $a_1 + h_1$ et, d'autre part, d_h compris entre a_2 et $a_2 + h_2$ vérifiant :

$$f(a + h) - f(a) = h_1\frac{\partial f}{\partial x_1}(c_h, a_2 + h_2) + h_2\frac{\partial f}{\partial x_2}(a_1, d_h)$$

Quand $h \to (0,0)$, $(c_h, a_2 + h_2) \to (a_1, a_2)$ et $(a_1, d_h) \to (a_1, a_2)$ donc par continuité des dérivées partielles de f, on obtient

$$f(a+h) - f(a) = h_1 \frac{\partial f}{\partial x_1}(a_1, a_2) + h_2 \frac{\partial f}{\partial x_2}(a_1, a_2) + o(h)$$

Ainsi

$$f(a+h) = f(a) + \ell(h) + o(h)$$

avec l'application linéaire

$$\ell : (h_1, h_2) \mapsto h_1 \frac{\partial f}{\partial x_1}(a) + h_2 \frac{\partial f}{\partial x_2}(a)$$

On en déduit que f est différentiable en a et

$$\forall h \in E, \mathrm{d}f(a) \cdot h = \frac{\partial f}{\partial x_1}(a) h_1 + \frac{\partial f}{\partial x_2}(a) h_2$$

Considérons les applications $p_1 : (h_1, h_2) \mapsto h_1$ et $p_2 : (h_1, h_2) \mapsto h_2$. On peut écrire

$$\mathrm{d}f(a) = \frac{\partial f}{\partial x_1}(a).p_1 + \frac{\partial f}{\partial x_2}(a).p_2$$

Par opérations sur les fonctions continues, la différentielle $\mathrm{d}f$ apparaît continue.
En effet, les applications $a \mapsto \dfrac{\partial f}{\partial x_1}(a)$, $a \mapsto \dfrac{\partial f}{\partial x_2}(a)$ sont continues, les applications $a \mapsto p_1$ et $a \mapsto p_2$ sont continues car constantes et enfin l'application produit extérieur est bilinéaire.
□

Définition

> On dit qu'une fonction $f : \Omega \subset E \to F$ est de classe $\mathcal{C}^1$ si ses dérivées partielles de f dans une base existent et sont continues.

Remarque La notion ne dépend pas du choix de la base utilisée.

Proposition

> Les fonctions de classe $\mathcal{C}^1$ sont continues.

dém. :
Car différentiables.
□

Exemple Les fonctions constantes sont de classe $\mathcal{C}^1$.
En effet leurs dérivées partielles sont nulles donc continues.

Exemple Les applications linéaires sont de classe $\mathcal{C}^1$.
En effet, pour $f \in \mathcal{L}(E, F)$, les dérivées partielles de f dans $e = (e_1, \ldots, e_n)$ sont les applications données par

$$\partial_i f(a) = \mathrm{d}f(a) \cdot e_i = f(e_i)$$

Ce sont des applications constantes donc continues.
En particulier, les applications $(x_1, \ldots, x_p) \in \mathbb{R}^p \mapsto x_j$, $z \in \mathbb{C} \mapsto \mathrm{Re}(z), \mathrm{Im}(z)$ et $A \in \mathcal{M}_{n,p}(\mathbb{R}) \mapsto a_{i,j}$ sont de classe $\mathcal{C}^1$.

25.3.2 Formule d'intégration

Théorème

Soit $f : \Omega \subset E \to F$ une application de classe $\mathcal{C}^1$.
Si $\gamma : [0,1] \to E$ est un arc de classe $\mathcal{C}^1$ inscrit dans Ω d'extrémités $a = \gamma(0)$ et $b = \gamma(1)$ alors

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(\gamma(t)) \cdot \gamma'(t) \, \mathrm{d}t$$

dém. :
Soit $\varphi : [0,1] \to F$ définie par $\varphi(t) = f(\gamma(t))$.
Par composition la fonction φ est dérivable

$$\varphi'(t) = \mathrm{d}f(\gamma(t)) \cdot \gamma'(t) = \sum_{i=1}^{n} x_i'(t) \frac{\partial f}{\partial x_i}(\gamma(t))$$

La fonction φ est donc de classe $\mathcal{C}^1$ et alors

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) \, \mathrm{d}t$$

Or $\varphi(1) = f(b)$, $\varphi(0) = f(a)$ et $\varphi'(t) = \mathrm{d}f(\gamma(t)) \cdot \gamma'(t)$.
□

Exemple Si $[a,b] \subset \Omega$ alors

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(a + t(b-a)) \cdot (b-a) \, \mathrm{d}t$$

En effet, $\gamma(t) = a + t.(b-a)$ définit un paramétrage de classe $\mathcal{C}^1$ du segment $[a,b]$.

Corollaire

Si Ω est un ouvert connexe par arcs et si $f : \Omega \subset E \to F$ est de classe $\mathcal{C}^1$ alors

$$f \text{ est constante si, et seulement si, } \mathrm{d}f = \tilde{0}$$

dém. :
Le sens direct est déjà connu. Supposons maintenant $\mathrm{d}f = \tilde{0}$.
Cas Ω convexe : Par l'exemple ci-dessus, on obtient

$$\forall a, b \in \Omega, f(b) = f(a)$$

Cas général : C'est plus technique, contentons-nous de quelques idées... Par l'étude précédente, on peut affirmer que f est localement constante i.e.

$$\forall a \in \Omega, \exists \alpha > 0, \forall x \in B(a, \alpha), f(x) = f(a)$$

Pour $a, b \in \Omega$, il existe $\gamma : [0,1] \to E$ chemin inscrit dans Ω d'extrémités $\gamma(0) = a$ et $\gamma(1) = b$. On montre alors

$$\sup \{t \in [0,1] \,/ \forall s \in [0,t], f(\gamma(t)) = f(a)\} = 1$$

ce qui fournit $f(b) = f(a)$.
□

25.3.3 Dérivées partielles successives

Définition

Soit $f : \Omega \subset E \to F$ et e une base de E.
La fonction f est appelée dérivée partielle d'ordre 0 de f.
Pour $k \in \mathbb{N}$, sous réserve d'existence, on appelle dérivées partielles d'ordre $k+1$ de f les dérivées partielles des dérivées partielles d'ordre k de f.

Remarque Si l'on note $x_1, \dots, x_p$ les coordonnées dans la base e de la variable x, on note

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} = \partial_{i_1}(\dots(\partial_{i_k} f)\dots)$$

Exemple Calculons les dérivée partielles d'ordre 1 et 2 de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x\,\mathrm{e}^{xy}$.
Les dérivées partielles d'ordre 1 de f sont

$$\frac{\partial f}{\partial x}(x,y) = (1+xy)\mathrm{e}^{xy} \text{ et } \frac{\partial f}{\partial y}(x,y) = x^2\mathrm{e}^{xy}$$

Les dérivées partielles d'ordre 2 de f sont

$$\frac{\partial^2 f}{\partial x^2}(x,y) = (2y+xy^2)\mathrm{e}^{xy},\ \frac{\partial^2 f}{\partial y \partial x}(x,y) = (2x+x^2y)\mathrm{e}^{xy}$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = (2x+x^2y)\mathrm{e}^{xy},\ \frac{\partial^2 f}{\partial y^2}(x,y) = x^3\mathrm{e}^{xy}$$

25.3.4 Classe d'une fonction

Définition

On dit que $f : \Omega \subset E \to F$ est de classe $\mathcal{C}^k$ si ses dérivées partielles d'ordre k existent et sont continues.
On dit que f est de classe $\mathcal{C}^\infty$ si f est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.

Remarque On peut montrer que cette notion de dépend pas du choix de la base utilisée pour définir les dérivées partielles.

Exemple Les applications de classe $\mathcal{C}^0$ correspondent aux applications continues.

Exemple Les applications constantes sont de classe $\mathcal{C}^\infty$.

Exemple Les application linéaires sont de classe $\mathcal{C}^\infty$.
Leur dérivées partielles sont constantes puisque pour une application linéaire f,

$$\partial_j f(a) = \mathrm{d}f(a) \cdot e_j = f(e_j)$$

Remarque En particulier, les fonctions $(x_1, \ldots, x_p) \mapsto x_j$, $z \mapsto \text{Re}(z), \text{Im}(z)$ et $A \mapsto a_{i,j}$ sont de classe $\mathcal{C}^\infty$.

Proposition

> Si $f : \Omega \subset E \to F$ est de classe $\mathcal{C}^{k+1}$ alors f est de classe $\mathcal{C}^k$.

dém. :
Si f est de classe $\mathcal{C}^{k+1}$ alors les dérivées partielles d'ordre k de f existent et sont de classe $\mathcal{C}^1$ donc continues.
□

25.3.5 Opérations

Soit $k \in \mathbb{N} \cup \{\infty\}$.

Théorème

> Soit $f, g : \Omega \subset E \to F$ et $\lambda, \mu \in \mathbb{R}$.
> Si f et g sont de classe $\mathcal{C}^k$ alors $\lambda f + \mu g$ l'est aussi.

dém. :
Par récurrence pour $k \in \mathbb{N}$.
Pour $k = 0$: ok
Supposons la propriété établie au rang $k \geqslant 0$.
Soit f et g de classe $\mathcal{C}^{k+1}$.
f et g sont de classe $\mathcal{C}^1$ donc f et g sont différentiables. La fonction $\lambda f + \mu g$ l'est alors aussi et

$$\partial_i(\lambda f + \mu g) = \lambda \partial_i f + \mu \partial_i g$$

Puisque $\partial_i f$ et $\partial_i g$ sont de classe $\mathcal{C}^k$, on obtient $\partial_i(\lambda f + \mu g)$ de classe $\mathcal{C}^k$ en vertu de l'hypothèse de récurrence.
Ainsi, les dérivées partielles de $\lambda f + \mu g$ existent et sont de classe $\mathcal{C}^k$.
Or les dérivées partielles de d'ordre k des dérivées partielles de $\lambda f + \mu g$ sont les dérivées partielles d'ordre $k + 1$ de $\lambda f + \mu g$. On peut alors conclure que $\lambda f + \mu g$ est de classe $\mathcal{C}^{k+1}$.
Récurrence établie.
Pour $k = \infty$.
Si f et g sont de classe $\mathcal{C}^\infty$ alors f et g sont de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$ et donc $\lambda f + \mu g$ aussi.
□

Corollaire

> L'ensemble $\mathcal{C}^k(\Omega, F)$ des fonctions de classe $\mathcal{C}^k$ de Ω vers F est un sous-espace vectoriel de $\mathcal{F}(\Omega, F)$.

Théorème

> Soit $f : \Omega \subset E \to F$, $g : \Omega \subset E \to G$ et $b : F \times G \to H$ bilinéaire.
> Si f et g sont de classe $\mathcal{C}^1$ alors $b(f, g)$ l'est aussi.

dém. :
Le protocole démonstratif est similaire au précédent. On y exploite la formule

$$\partial_i \left(b(f, g)\right) = b\left(\partial_i f, g\right) + b\left(f, \partial_i g\right)$$

□

Corollaire

Si F est une algèbre (par ex : $F = \mathbb{R}, \mathbb{C}$ ou $\mathcal{M}_n(\mathbb{K})$) alors $\mathcal{C}^k(\Omega, F)$ est une sous-algèbre de $\mathcal{F}(\Omega, F)$.

Exemple Les fonctions polynomiales sur $\mathbb{R}^p$ sont de classe $\mathcal{C}^\infty$.

Exemple L'application $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est de classe $\mathcal{C}^\infty$ par somme et produit de fonctions $\mathcal{C}^\infty$.

Théorème

Soit $f : \Omega \subset E \to F$. On a équivalence entre :
(i) f est de classe $\mathcal{C}^k$;
(ii) les fonctions coordonnées de f dans une base de F sont de classe $\mathcal{C}^k$.

Exemple L'application $f : (x, y) \mapsto (x^2 + y^2, xy)$ est de classe $\mathcal{C}^\infty$.

Théorème

Soit $f : \Omega \subset E \to F$ et $g : \Omega' \subset F \to G$ telles que $f(\Omega) \subset \Omega'$.
Si f et g sont de classe $\mathcal{C}^k$ alors $g \circ f$ l'est aussi.

dém. :
Via la formule calculant les dérivées partielles d'une fonction composée.
□

Exemple $\Phi(r, \theta) = (r\cos\theta, r\sin\theta)$ définit une fonction $\mathcal{C}^\infty$ de $\mathbb{R}^2$ vers $\mathbb{R}^2$.

25.3.6 Théorème de Schwarz

Théorème

Si $f : \Omega \subset E \to F$ est de classe $\mathcal{C}^2$ alors pour tout $i, j \in \{1, \ldots, n\}$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

Exemple Soit $f : (x, y) \mapsto f(x, y)$ de classe $\mathcal{C}^2$.
Calculons les dérivées partielles d'ordre 2 de $g : (u, v) \mapsto f(u + v, uv)$.
Les dérivées partielles d'ordre 1 de g sont
$\frac{\partial g}{\partial u}(u, v) = \frac{\partial f}{\partial x}(u + v, uv) + v\frac{\partial f}{\partial y}(u + v, uv)$, $\frac{\partial g}{\partial v}(u, v) = \frac{\partial f}{\partial x}(u + v, uv) + u\frac{\partial f}{\partial y}(u + v, uv)$
Les dérivées partielles d'ordre 2 de g sont
$\frac{\partial^2 g}{\partial u^2}(u, v) = \frac{\partial^2 f}{\partial x^2}(u + v, uv) + 2v\frac{\partial^2 f}{\partial x \partial y}(u + v, uv) + v^2\frac{\partial^2 f}{\partial y^2}(u + v, uv)$,
$\frac{\partial^2 g}{\partial u \partial v}(u, v) = \frac{\partial^2 f}{\partial x^2}(u + v, uv) + (u + v)\frac{\partial^2 f}{\partial x \partial y}(u + v, uv) + uv\frac{\partial^2 f}{\partial y^2}(u + v, uv) + \frac{\partial f}{\partial y}(u + v, uv)$,
$\frac{\partial^2 g}{\partial v^2}(u, v) = \frac{\partial^2 f}{\partial x^2}(u + v, uv) + 2u\frac{\partial^2 f}{\partial x \partial y}(u + v, uv) + u^2\frac{\partial^2 f}{\partial y^2}(u + v, uv)$.

Exemple Considérons la fonction

$$f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases}$$

Vérifions que f n'est pas de classe $\mathcal{C}^2$.
Pour $(x,y) \neq (0,0)$

$$\frac{\partial f}{\partial x}(x,y) = \frac{y(4x^2y^2 - x^4 + y^4)}{(x^2+y^2)^2}$$

et

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t \to 0} \left(\frac{f(t,0) - f(0,0)}{t} \right) = 0$$

De plus, en passant en polaires, on vérifie que $\dfrac{\partial f}{\partial x}$ est continue en $(0,0)$.
On mène une étude semblable pour $\dfrac{\partial f}{\partial y}$ avec

$$\frac{\partial f}{\partial y}(x,y) = -\frac{x(4x^2y^2 - x^4 + y^4)}{(x^2+y^2)^2}$$

On en déduit que f est de classe $\mathcal{C}^1$.
Cependant

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{t \to 0} \frac{1}{t} \left(\frac{\partial f}{\partial x}(0,t) - \frac{\partial f}{\partial x}(0,0) \right) = 1 \text{ et } \frac{\partial^2 f}{\partial x \partial y}(0,0) = -1$$

La fonction f n'est donc pas de classe $\mathcal{C}^2$.

25.4 Fonctions numériques

Ici les fonctions étudiées sont supposées à valeurs réelles.

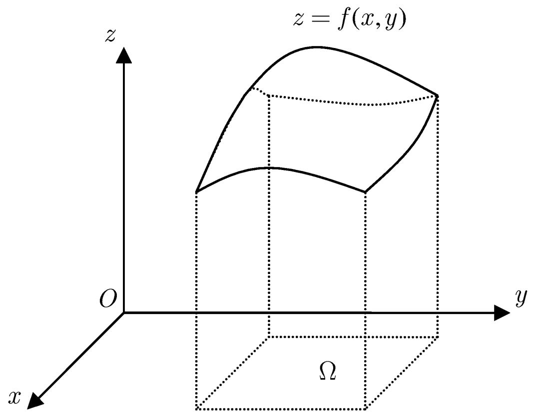
25.4.1 Surface représentant une fonction de deux variables réelles

Soit $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ vue en les deux variables x et y.

Définition

On appelle surface représentative de f l'ensemble formé des $(x, y, z) \in \mathbb{R}^3$ vérifiant l'équation

$$\Sigma_f : z = f(x, y)$$



Définition

Si f est différentiable en (x_0, y_0), le plan d'équation cartésienne

$$z = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + f(x_0, y_0)$$

est appelé plan tangent à Σ_f au point (x_0, y_0, z_0).

Exemple Considérons la surface $z = x^2 + 2y^2$.
Une équation du plan tangent en (x_0, y_0, z_0) est

$$z = 2x_0(x - x_0) + 4y_0(y - y_0) + z_0$$

et puisque $z_0 = x_0^2 + 2y_0^2$, on peut simplifier

$$z = 2x_0 x + 4y_0 y - z_0$$

Rappel :
Soit a un élément d'une partie X d'un espace vectoriel réel E.
On dit qu'un vecteur v de E est tangent à X en a, s'il existe $\varepsilon > 0$ et un arc γ défini sur $]-\varepsilon, \varepsilon[$ inscrit dans a vérifiant

$$\gamma(0) = a \text{ et } \gamma'(0) = v$$

Lorsque le vecteur v est non nul, on dit que la droite

$$a + \text{Vect} v$$

est tangente à X en a.

Théorème

Si f est différentiable en (x_0, y_0) alors les tangentes à Σ_f au point (x_0, y_0, z_0) sont toutes incluses dans le plan tangent à Σ_f en (x_0, y_0, z_0).

dém. :
Soit T une tangente à Σ_f en (x_0, y_0, z_0). Il existe $v \in \mathbb{R}^3$ non nul et un arc $\gamma : t \mapsto (x(t), y(t), z(t))$ défini sur $]-\varepsilon, \varepsilon[$ inscrit dans X vérifiant

$$\gamma(0) = (x_0, y_0, z_0) \text{ et } \gamma'(0) = v = (x'(0), y'(0), z'(0))$$

Puisque $z(t) = f(x(t), y(t))$, on obtient par dérivation en 0

$$z'(0) = x'(0)\frac{\partial f}{\partial x}(x_0, y_0) + y'(0)\frac{\partial f}{\partial y}(x_0, y_0)$$

Les éléments de la droite T sont alors de coordonnées

$$\begin{cases} x = x_0 + \lambda x'(0) \\ y = y_0 + \lambda y'(0) \\ z = z_0 + \lambda z'(0) \end{cases}$$

vérifiant l'équation du plan proposée.
□

Remarque On peut aussi montrer que le plan tangent est exactement la réunion des droites tangentes à Σ_f en (x_0, y_0, z_0).

25.4.2 Gradient

On suppose que E est un espace vectoriel euclidien dont on note $(. \mid .)$ le produit scalaire.

25.4.2.1 Définition

On suppose que E est un espace vectoriel euclidien dont on note $(. \mid .)$ le produit scalaire.
Rappel :
Le théorème de représentation des formes linéaires dans un espace euclidien fournit

$$\forall \varphi \in E^\star, \exists! u \in E, \forall x \in E, \varphi(x) = (u \mid x)$$

Théorème

Si $f : \Omega \subset E \to \mathbb{R}$ est une application différentiable alors pour tout $a \in \Omega$, il existe un unique vecteur de E noté $\nabla f(a)$ vérifiant

$$\forall v \in E, \mathrm{D}_v f(a) = (\nabla f(a) \mid v)$$

Ce vecteur est appelé gradient de f en a, il est déterminé par

$$\nabla f(a) = \partial_1 f(a) e_1 + \cdots + \partial_n f(a) e_n$$

dès que $(e_1, \ldots, e_n)$ désigne une base orthonormée de E.

dém. :
Soit $a \in \Omega$. f est différentiable en a et

$$\forall h \in E, \mathrm{D}_h f(a) = \mathrm{d}f(a) \cdot h$$

Puisque l'application $\mathrm{d}f(a)$ est une forme linéaire sur E, il existe un unique vecteur $\nabla f(a) \in E$ vérifiant

$$\forall h \in E, \mathrm{d}f(a) \cdot h = (\nabla f(a) \mid h)$$

i.e.

$$\forall h \in E, \mathrm{D}_h f(a) = (\nabla f(a) \mid h)$$

De plus, si $(e_1, \ldots, e_n)$ est une base orthonormée

$$\nabla f(a) = \sum_{i=1}^{n} (\nabla f(a) \mid e_i)\, e_i = \sum_{i=1}^{n} \mathrm{D}_i f(a).e_i$$

□

Corollaire

Le développement limité à l'ordre 1 de f en a s'écrit alors

$$f(a+h) = f(a) + (\nabla f(a) \mid h) + o(h) \text{ quand } h \to 0_E$$

Exemple Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = x^2 + 2xy$. f est différentiable.
En munissant $\mathbb{R}^2$ de sa structure euclidienne canonique et en considérant (e_1, e_2) sa base canonique

$$\nabla f(a) = \frac{\partial f}{\partial x}(a) e_1 + \frac{\partial f}{\partial x_2}(a) e_2 = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a)\right)$$

Ainsi

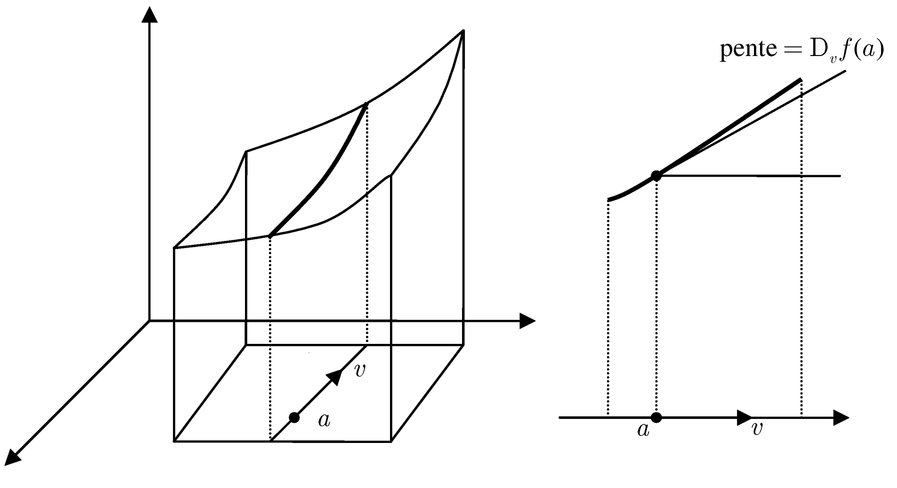
$$\nabla f(x, y) = (2x + 2y, 2x)$$

25.4.2.2 Interprétation

Pour v un vecteur unitaire

$$\mathrm{D}_v f(a) = \lim_{t \to 0} \frac{1}{t} \left(f(a + tv) - f(a)\right)$$

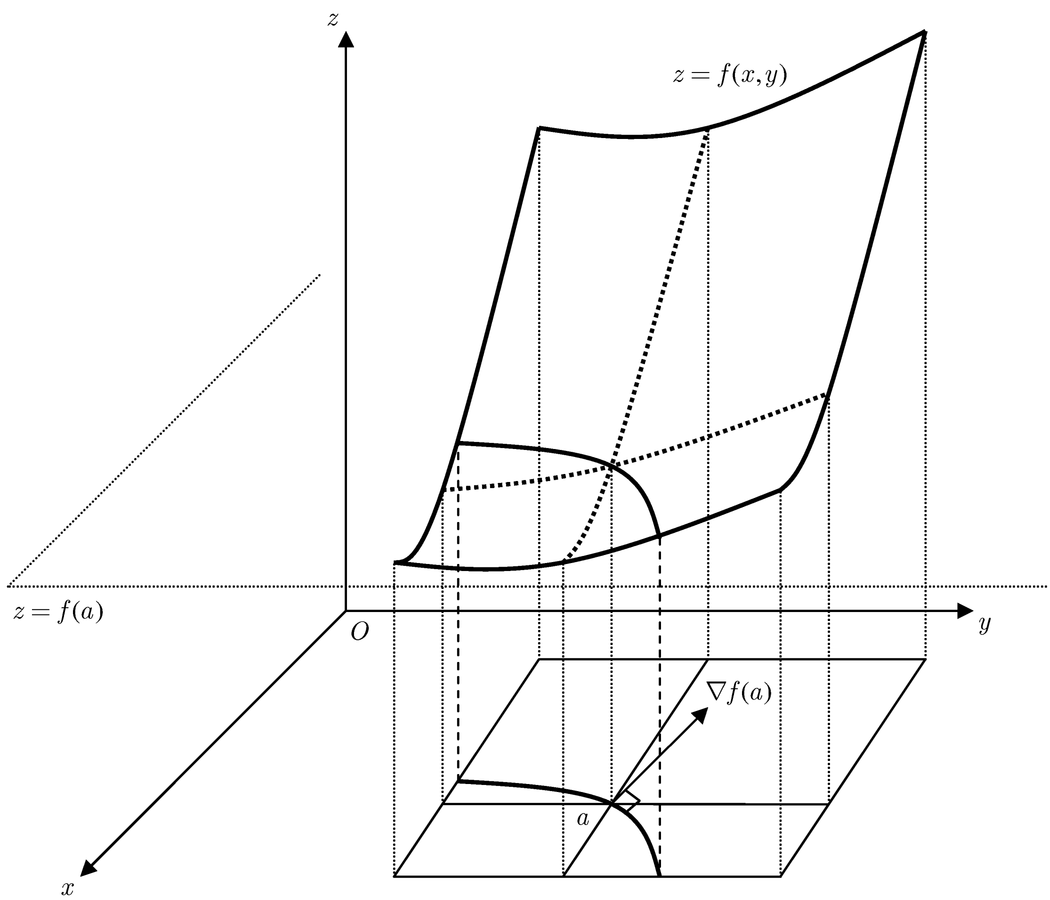
Cette quantité se comprend comme étant la pente de f dans la direction donnée par le vecteur v.



Puisque

$$\mathrm{D}_v f(a) = (\nabla f(a) \mid v) = \|\nabla f(a)\| \, \|v\| \cos\theta \text{ avec } \theta \in [0, \pi]$$

cette pente est maximale quand v a le sens et la direction de $\nabla f(a)$.
Ainsi, lorsqu'il n'est pas nul, le vecteur $\nabla f(a)$ indique la direction de la plus grande pente, son sens donne le sens de progression croissante sur cette pente et $\|\nabla f(a)\|$ donne la valeur de cette pente extrême.



25.4.2.3 Ligne de niveau

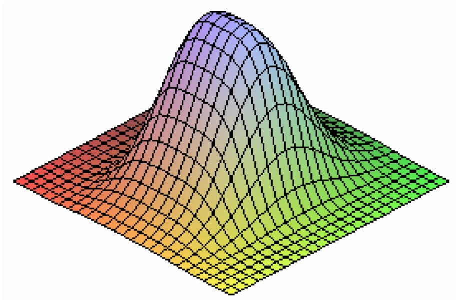
Définition

Soit $\lambda \in \mathbb{R}$ et $f : \Omega \subset E \to \mathbb{R}$. L'ensemble X_λ formé des $x \in \Omega$ vérifiant

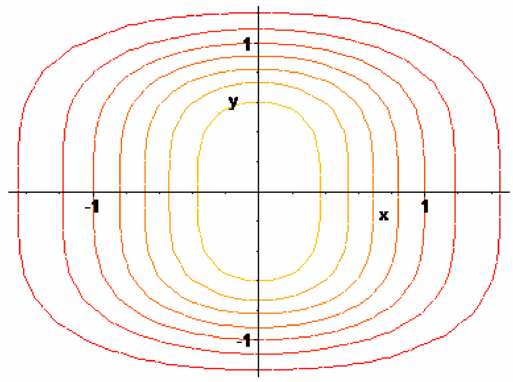
$$f(x) = \lambda$$

est appelé ligne de niveau $\lambda \in \mathbb{R}$ de f.

Exemple Pour $f(x, y) = \mathrm{e}^{-x^2 - y^4}$
On obtient la surface représentative



et les lignes de niveau suivantes



Exemple En électrostatique, le champ électrique est perpendiculaire aux équipotentielles...

Théorème

Les vecteurs tangents au point x d'une ligne de niveau d'une fonction $f : \Omega \subset E \to \mathbb{R}$ différentiable sont orthogonaux au gradient de f en x.

dém. :

On introduit $(e_1, \ldots, e_n)$ une base orthonormée de E. On sait

$$\nabla f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).e_i$$

Soit v un vecteur tangent au point x d'une ligne de niveau X de f. Il existe un arc $\gamma : t \mapsto \gamma(t)$ défini sur $]-\varepsilon, \varepsilon[$ inscrit dans X vérifiant

$$\gamma(0) = x \text{ et } \gamma'(0) = v$$

En notant $x_1(t), \ldots, x_n(t)$ les coordonnées de $\gamma(t)$, on a

$$v = \gamma'(0) = x_1'(0).e_1 + \cdots + x_n'(0).e_n$$

Puisque γ inscrit dans X, la fonction $t \mapsto f(\gamma(t)) = f(x_1(t), \dots, x_n(t))$ est constante. Par dérivation de fonctions composées en 0, on obtient

$$0 = x_1'(0)\frac{\partial f}{\partial x_1}(x) + \cdots + x_n'(0)\frac{\partial f}{\partial x_n}(x)$$

et donc

$$(\nabla f(a) \mid v) = 0$$

□

25.4.3 Recherche d'extremum

25.4.3.1 Point critique

Définition

Soit $f : X \subset E \to \mathbb{R}$.
On dit que f admet un minimum (global) en $a \in A$ si

$$\forall x \in X, f(x) \geqslant f(a)$$

On dit que f admet un minimum local en en $a \in A$ si

$$\exists \alpha > 0, \forall x \in X \cap B(a, \alpha), f(x) \geqslant f(a)$$

Remarque Les extremums globaux sont a fortiori des extremums locaux.

Définition

On dit qu'une application $f : \Omega \subset E \to \mathbb{R}$ différentiable admet un point critique en $a \in \Omega$ si $\mathrm{d}f(a) = \tilde{0}$.

Proposition

Soit $e = (e_1, \dots, e_n)$ une base de E, $f : \Omega \subset E \to \mathbb{R}$ différentiable et $a \in \Omega$.
On a équivalence entre :
(i) a est point critique de f ;
(ii) $\forall i \in \{1, \dots, n\}, \partial_i f(a) = 0$.

dém. :
(i) $\Rightarrow$ (ii) via $\partial_i f(a) = \mathrm{d}f(a) \cdot e_i$.
(ii) $\Rightarrow$ (i) via pour tout $h = h_1 e_1 + \cdots + h_n e_n \in E,\ \mathrm{d}f(a)h = \sum_{i=1}^{n} h_i \partial_i f(a)$.

□

Remarque Les points critiques correspondent aux points où le vecteur gradient est nul.

Théorème

Si $f : \Omega \subset E \to \mathbb{R}$ différentiable admet un extremum local en $a \in \Omega$ alors a est point critique de f.

dém. :
Cas a minimum local :
Il existe $\alpha > 0$ tel que $B(a, \alpha) \subset U$ et

$$\forall x \in B(a, \alpha), f(x) \geqslant f(a)$$

Pour tout $v \in E$,

$$\mathrm{d}f(a) \cdot v = \mathrm{D}_v f(a) = \lim_{t \to 0} \frac{1}{t} \left(f(a + t.v) - f(a)\right)$$

Quand $t \to 0^+$,
Pour t suffisamment proche de 0, $a + t.v \in B(a, \alpha)$ et $(f(a + t.v) - f(a))/t \geqslant 0$ donc à la limite $\mathrm{d}f(a) \cdot v \geqslant 0$.
Quand $t \to 0^-$,
On obtient de façon semblable $\mathrm{d}f(a) \cdot v \leqslant 0$.
Ainsi $\mathrm{d}f(a) \cdot v = 0$ pour tout $v \in E$.
□

Attention : La réciproque n'est pas vraie.

Attention : Ce résultat ne s'applique qu'à une fonction différentiable définie sur un ouvert.

25.4.3.2 En pratique

Protocole :
Pour étudier les extremums locaux de $f : \Omega \subset E \to \mathbb{R}$ différentiable :
- on recherche les points critiques ;
- on étudie chacun en se ramenant en 0_E par translation si besoin.

Exemple Extremums de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 + xy + 1$.
f est différentiable sur l'ouvert $\mathbb{R}^2$.
Points critiques :
$\dfrac{\partial f}{\partial x}(x, y) = 2x + y$ et $\dfrac{\partial f}{\partial y}(x, y) = 2y + x$.

$$\begin{cases} 2x + y = 0 \\ x + 2y = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}$$

$(0, 0)$ est seul point critique.
Etude de $(0, 0)$.
$f(0, 0) = 1$, étudions le signe de $g(x, y) = f(x, y) - f(0, 0) = x^2 + y^2 + xy$.
En écrivant $x = r\cos\theta$ et $y = r\sin\theta$, $g(x, y) = r^2(1 + \cos\theta\sin\theta) \geqslant 0$.
$(0, 0)$ est un minimum global.

Exemple Extremums de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = x^2 + y^2 + 4xy - 1$.
f est différentiable sur l'ouvert $\mathbb{R}^2$.
Points critiques :

$\dfrac{\partial f}{\partial x}(x,y) = 2x + 4y$, $\dfrac{\partial f}{\partial y}(x,y) = 2y + 4x$.

$$\begin{cases} 4x + y = 0 \\ x + 4y = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}$$

$(0,0)$ est seul point critique.
Etude de $(0,0)$.
$f(0,0) = -1$. Etudions le signe de $g(x,y) = f(x,y) - f(0,0) = x^2 + y^2 + 4xy$.
En écrivant $x = r\cos\theta$ et $y = r\sin\theta$, $g(x,y) = r^2(1 + 4\cos\theta\sin\theta) = r^2(1 + 2\sin 2\theta)$ qui change de signe.
Concrètement :
$g\left(\dfrac{1}{n}, 0\right) = \dfrac{1}{n^2} > 0$ donc $(0,0)$ n'est pas un maximum local,
$g\left(-\dfrac{1}{n}, \dfrac{1}{n}\right) = -\dfrac{2}{n^2} < 0$ donc $(0,0)$ n'est pas un maximum local.

Exemple Extremums de $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3 + y^3 - 3xy$.
f est différentiable sur l'ouvert $\mathbb{R}^2$.
Points critiques :
$\dfrac{\partial f}{\partial x}(x,y) = 3x^2 - 3y$, $\dfrac{\partial f}{\partial y}(x,y) = 3y^2 - 3x$.

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 = y \\ y^2 = x \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ x^4 = x \end{cases} \Leftrightarrow \begin{cases} y = x^2 \\ x = 0 \text{ ou } 1 \end{cases}$$

$(0,0), (1,1)$ seuls points critiques
Etude en $(0,0)$:
$g(x,y) = f(x,y) - f(0,0) = x^3 + y^3 - 3xy$.
$g\left(\dfrac{1}{n}, 0\right) = \dfrac{1}{n^3} > 0$ et $g\left(-\dfrac{1}{n}, 0\right) = -\dfrac{1}{n^3} < 0$ donc $(0,0)$ n'est pas extremum local.
Etude en $(1,1)$:
$g(x,y) = f(x,y) - f(1,1) = x^3 + y^3 - 3xy + 1$.

$$\begin{cases} x = 1 + u \\ y = 1 + v \end{cases}$$

$g(x,y) = 3u^2 + 3v^3 - 3uv + u^3 + v^3$.

$$\begin{cases} u = r\cos\theta \\ v = r\sin\theta \end{cases}$$

$g(x,y) = r^2\left(3 - \dfrac{3}{2}\sin 2\theta + r\cos^3\theta + r\sin^3\theta\right)$.
Quand $(x,y) \to (1,1)$, on a $(u,v) \to (0,0)$ donc $r \to 0$ puis
$3 - \dfrac{3}{2}\sin 2\theta + r\cos^3\theta + r\sin^3\theta = 3 - \dfrac{3}{2}\sin 2\theta + o(1) \geqslant \dfrac{3}{2} + o(1) \geqslant 0$.
$(1,1)$ est un minimum local.
Cependant $f(t,0) = t^3 \xrightarrow[t \to -\infty]{} -\infty$ donc f n'est pas minorée et donc $(1,1)$ n'est pas un minimum global.

25.4.3.3 Calcul d'inf et de sup

Soit I, J des intervalles non vides de $\mathbb{R}$.

Remarque Pour $\varphi : I \to \mathbb{R}$, le calcul de $\inf_{t\in I} \varphi(t)$ est facile en dressant un tableau de variation.

Proposition

Si $f : I \times J \to \mathbb{R}$ est minorée alors

$$\inf_{(x,y)\in I\times J} f(x,y) = \inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right)$$

dém. :
Posons $m = \inf_{(x,y)\in I\times J} f(x,y)$.
Pour tout $x \in I$ et $y \in J$, $m \leqslant f(x,y)$ donc $m \leqslant \inf_{y\in J} f(x,y)$ puis

$$m \leqslant \inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right)$$

Inversement, pour $x_0 \in I$ et $y_0 \in J$,

$$\inf_{y\in J} f(x_0,y) \leqslant f(x_0,y_0)$$

or

$$\inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right) \leqslant \inf_{y\in J} f(x_0,y)$$

donc

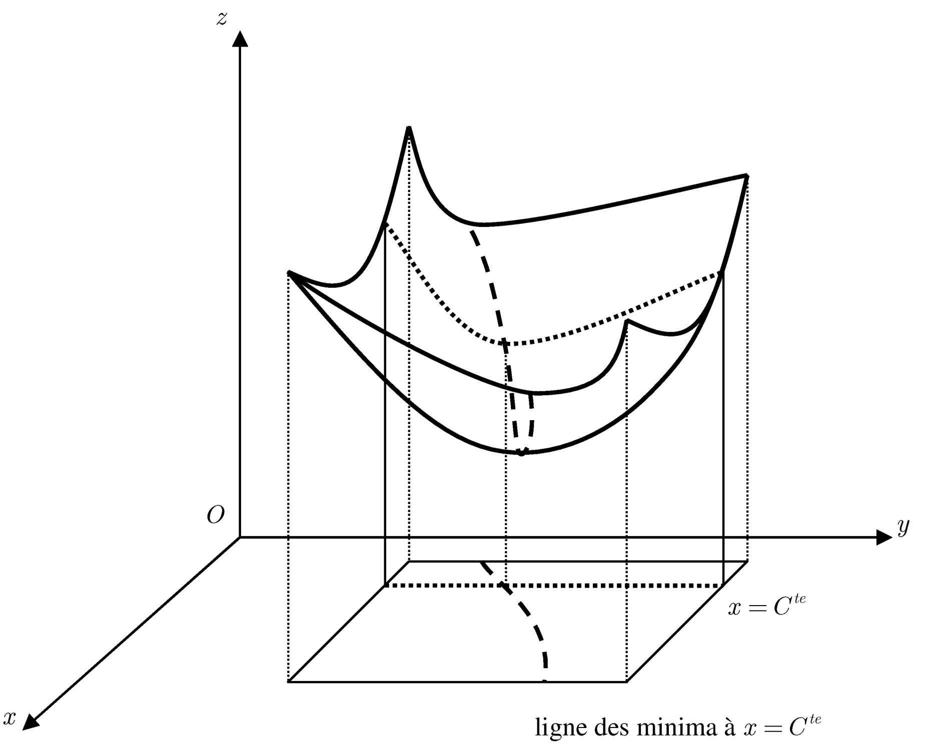
$$\inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right) \leqslant f(x_0,y_0)$$

Par suite $\inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right)$ minore f et donc

$$\inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right) \leqslant m$$

Finalement

$$\inf_{x\in I}\left(\inf_{y\in J} f(x,y)\right) = m$$



□

Exemple Calculons

$$M = \inf_{x,y>0} \left(x + y + \frac{1}{xy} \right)$$

$M = \inf_{x>0} m(x)$ avec $m(x) = \inf_{y>0} \varphi(y)$ où $\varphi(y) = x + y + 1/xy$.
Après étude des variations de φ $m(x) = \varphi\left(1/\sqrt{x}\right) = x + 2/\sqrt{x}$.
Après étude des variations de m, $M = m(1) = 3$.

25.4.3.4 Borne d'une fonction continue sur un compact

Exemple Calculons

$$M = \sup_{(x,y)\in T} xy(1-x-y) \text{ avec } T = \left\{(x,y) \in \mathbb{R}^2 / x, y \geqslant 0, x + y \leqslant 1\right\}$$

La partie T est compacte et non vide et la fonction $f : (x, y) \mapsto xy(1 - x - y)$ est continue sur T donc f admet un maximum en $a \in T$ et $M = f(a)$.
Puisque la fonction f est nulle sur le bord de T strictement positive sur l'intérieur de T on peut affirmer que a appartient à l'ouvert $U = T^\circ$. Or f est différentiable sur l'ouvert U donc a est point critique de f.

$\dfrac{\partial f}{\partial x}(x,y) = y(1-2x-y)$, $\dfrac{\partial f}{\partial y}(x,y) = x(1-2y-x)$,

$$\begin{cases} y(1-2x-y) = 0 \\ x(1-2y-x) = 0 \end{cases} \Leftrightarrow \begin{cases} 2x+y = 1 \\ x+2y = 1 \end{cases} \Leftrightarrow \begin{cases} x = 1/3 \\ y = 1/3 \end{cases}$$

car $x, y \neq 0$ pour $a \in U$.
Finalement

$$M = f(1/3, 1/3) = \frac{1}{27}$$

Remarque Cette borne supérieure peut aussi être déterminée en exploitant

$$M = \sup_{x\in[0,1]} \sup_{y\in[0,1-x]} xy(1-x-y)$$

25.4.4 Equations aux dérivées partielles

I et J désignent des intervalles de $\mathbb{R}$ ouverts et non vides.

25.4.4.1 Équation aux dérivées partielles d'ordre 1

Définition

Résoudre sur Ω une équation aux dérivées partielles d'ordre 1 en la fonction inconnue f, c'est déterminer toutes les fonctions $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^1$ vérifiant une relation donnée engageant f et/ou ses dérivées partielles.

Proposition

Les solutions sur $I \times J$ de l'équation

$$\frac{\partial f}{\partial x}(x,y) = 0$$

sont les fonctions

$$f : (x,y) \mapsto C(y) \text{ avec } C \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$$

dém. :
Soit $f : I \times J \to \mathbb{R}$ de classe $\mathcal{C}^1$ solution de l'équation aux dérivées partielles

$$\frac{\partial f}{\partial x}(x,y) = 0$$

Soit $y \in J$ fixé. L'application partielle $x \mapsto f(x,y)$ a pour dérivée $\dfrac{\partial f}{\partial x}(x,y)$.
L'application partielle $x \mapsto f(x,y)$ est donc de dérivée nulle sur l'intervalle I, c'est donc une fonction constante. Ainsi, il existe $C_y \in \mathbb{R}$ telle que

$$\forall x \in I, f(x,y) = C_y$$

Considérons alors $C : J \to \mathbb{R}$ définie par $C(y) = C_y$.
On définit ainsi une application $C : J \to \mathbb{R}$ vérifiant

$$\forall (x, y) \in I \times J, f(x, y) = C(y)$$

Soit $x_0 \in I$ fixé. La composition $y \mapsto (x_0, y) \mapsto f(x_0, y)$ est de classe $\mathcal{C}^1$, donc C est une fonction $\mathcal{C}^1$.
Résumons :
Si f est solution sur $I \times J$ de l'équation $\dfrac{\partial f}{\partial x}(x, y) = 0$ alors il existe $C : J \to \mathbb{R}$ de classe $\mathcal{C}^1$ vérifiant

$$\forall (x, y) \in I \times J, f(x, y) = C(y)$$

Inversement, les fonctions proposées sont évidemment solutions.
□

Exemple Résolvons sur $\mathbb{R}^2$ l'équation aux dérivées partielles

$$\frac{\partial f}{\partial x}(x, y) = xy$$

En intégrant par rapport à x

$$f(x, y) = \frac{1}{2}x^2 y + C(y) \text{ avec } C : \mathbb{R} \xrightarrow{\mathcal{C}^1} \mathbb{R}$$

Exemple Résolvons sur $\mathbb{R}^3$ l'équation aux dérivées partielles

$$\frac{\partial f}{\partial x}(x, y, z) = xy + z$$

En intégrant par rapport à x

$$f(x, y) = \frac{1}{2}x^2 y + C(y) \text{ avec } C : \mathbb{R}^2 \xrightarrow{\mathcal{C}^1} \mathbb{R}$$

Exemple Résolvons sur $\mathbb{R}^2$ l'équation aux dérivées partielles $\dfrac{\partial f}{\partial y}(x, y) = xf(x, y)$
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1$ solution.
Pour $x \in \mathbb{R}$ fixé, l'application partielle $y \mapsto f(x, y)$ a pour dérivée $\dfrac{\partial f}{\partial y}(x, y)$.
L'application partielle $y \mapsto f(x, y)$ est donc solution de l'équation différentielle

$$z'(y) = xz(y)$$

dont la solution générale est de la forme

$$z(y) = C\mathrm{e}^{xy}$$

Par suite, il existe une constante $C(x) \in \mathbb{R}$ telle que

$$\forall (x, y) \in \mathbb{R}^2, f(x, y) = C(x)\mathrm{e}^{xy}$$

$C : x \mapsto (x, 0) \mapsto f(x, 0)$ est de classe $\mathcal{C}^1$ par composition.
Inversement, de telle fonctions sont solutions.

Exemple Résolvons sur $\mathbb{R}^2$ l'équation

$$(E) : 2\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial y}(x,y) = 0$$

via le changement de variables

$$\begin{cases} u = x + y \\ v = x + 2y \end{cases}$$

Commençons par étudier le changement de variables de sorte d'exprimer les anciennes variables en fonction des nouvelles variables :

$$\begin{cases} u = x + y \\ v = x + 2y \end{cases} \Leftrightarrow \begin{cases} x = 2u - v \\ y = v - u \end{cases}$$

L'application $\Phi : (u,v) \mapsto (2u - v, v - u)$ traduit le changement de variable.
Φ est une bijection de classe $\mathcal{C}^1$ de $\mathbb{R}^2$ vers $\mathbb{R}^2$.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $g : \mathbb{R}^2 \to \mathbb{R}$ définie par « $g(u,v) = f(x,y)$ » i.e.

$$g : (u,v) = f(2u - v, v - u)$$

$g = f \circ \Phi$ est de classe $\mathcal{C}^1$.

$$\frac{\partial g}{\partial u}(u,v) = 2\frac{\partial f}{\partial x}(2u - v, v - u) - \frac{\partial f}{\partial y}(2u - v, v - u) = \left[2\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial y}(x,y)\right]_{\substack{x=2u-v \\ y=v-u}}$$

f est solution sur $\mathbb{R}^2$ de l'équation aux dérivées partielles proposée
$\Leftrightarrow \forall (x,y) \in \mathbb{R}^2, 2\dfrac{\partial f}{\partial x}(x,y) - \dfrac{\partial f}{\partial y}(x,y) = 0,$
$\Leftrightarrow \forall (u,v) \in \mathbb{R}^2, \dfrac{\partial g}{\partial u}(u,v) = 0$
$(\Rightarrow)$ immédiat et $(\Leftarrow)$ car Φ est surjective.
$\Leftrightarrow \exists C : \mathbb{R} \overset{\mathcal{C}^1}{\to} \mathbb{R}, \forall (u,v) \in \mathbb{R}^2, g(u,v) = C(v),$
$\Leftrightarrow \exists C : \mathbb{R} \overset{\mathcal{C}^1}{\to} \mathbb{R}, \forall (x,y) \in \mathbb{R}^2, f(x,y) = C(x + 2y).$
$(\Rightarrow)$ car $f = g \circ \Phi^{-1}$ et $(\Leftarrow)$ car $g = f \circ \Phi$
Finalement la solution générale de (E) est $f(x,y) = C(x + 2y)$ avec $C : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$.

Exemple Résolvons sur $\mathbb{R}^2 \setminus \{(0,0)\}$ l'équation aux dérivées partielles

$$(E) : x\frac{\partial f}{\partial y}(x,y) - y\frac{\partial f}{\partial x}(x,y) = 0$$

en passant en coordonnées polaires.

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

Puisqu'on se limite à $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, on peut se contenter de $r \in \mathbb{R}^{+\star}$ auquel cas $r = \sqrt{x^2 + y^2}$.
En revanche on ne peut pas exprimer θ mais au final ce ne sera pas utile.
Soit $\Phi : \mathbb{R}^{+\star} \times \mathbb{R} \to \mathbb{R}^2 \setminus \{(0,0)\}$ définie par $\Phi(r,\theta) \mapsto (r\cos\theta, r\sin\theta)$.

Φ est une surjection de classe $\mathcal{C}^1$ de $\mathbb{R}^{+\star} \times \mathbb{R}$ sur $\mathbb{R}^2 \setminus \{(0,0)\}$.
Soit $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $g : \mathbb{R}^{+\star} \times \mathbb{R} \to \mathbb{R}$ définie de sorte que « $g(r,\theta) = f(x,y)$ » i.e.

$$g(r,\theta) = f(r\cos\theta, r\sin\theta)$$

$g = f \circ \Phi$ est de classe $\mathcal{C}^1$.

$$\begin{aligned}\frac{\partial g}{\partial \theta}(r,\theta) &= -r\sin\theta \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\\ &= \left[-y\frac{\partial f}{\partial x}(x,y) + x\frac{\partial f}{\partial y}(x,y)\right]_{\substack{x=r\cos\theta\\ y=r\sin\theta}}\end{aligned}$$

f est solution sur $\mathbb{R}^2 \setminus \{(0,0)\}$ de l'équation aux dérivées partielles proposée E
$\Leftrightarrow \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}, x\dfrac{\partial f}{\partial y}(x,y) - y\dfrac{\partial f}{\partial x}(x,y) = 0,$
$\Leftrightarrow \forall (r,\theta) \in \mathbb{R}^{+\star} \times \mathbb{R}, \dfrac{\partial g}{\partial \theta}(r,\theta) = 0$
($\Rightarrow$) immédiat et ($\Leftarrow$) car Φ est surjective.
$\Leftrightarrow \exists C : \mathbb{R}^{+\star} \overset{\mathcal{C}^1}{\to} \mathbb{R}, \forall (r,\theta) \in \mathbb{R}^{+\star} \times \mathbb{R}, g(r,\theta) = C(r),$
$\Leftrightarrow \exists C : \mathbb{R}^{+\star} \overset{\mathcal{C}^1}{\to} \mathbb{R}, \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}, f(x,y) = C(\sqrt{x^2+y^2}),$
($\Leftarrow$) car $g = f \circ \Phi$ et ($\Rightarrow$) car Φ est surjective et $\Phi(r,\theta) = (x,y) \Rightarrow r = \sqrt{x^2+y^2}$.
$\Leftrightarrow \exists \tilde{C} : \mathbb{R}^{+\star} \overset{\mathcal{C}^1}{\to} \mathbb{R}, \forall (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}, f(x,y) = \tilde{C}(x^2+y^2).$
($\Rightarrow$) via $\tilde{C} = C \circ \sqrt{.}$ et ($\Leftarrow$) via $\tilde{C} = C \circ .^2$.
Finalement, la solution générale sur $\mathbb{R}^2 \setminus \{(0,0)\}$ de l'équation aux dérivées partielles (E) est
$f(x,y) = C(x^2+y^2)$ avec $C : \mathbb{R} \overset{\mathcal{C}^1}{\to} \mathbb{R}$.

25.4.4.2 Équations aux dérivées partielles d'ordre 2

Définition

Résoudre sur Ω une équation aux dérivées partielles d'ordre 2 en la fonction inconnue f, c'est déterminer toutes les fonctions $f : \Omega \to \mathbb{R}$ de classe $\mathcal{C}^2$ vérifiant une relation donnée engageant f et/ou ses dérivées partielles d'ordre 1 et 2.

Exemple L'équation de la chaleur

$$\frac{\partial f}{\partial t}(x,t) = D\frac{\partial^2 f}{\partial x^2}(x,t) \text{ avec } D > 0$$

Lorsque des conditions aux limites sont imposées, on peut avancer dans sa résolution par une décomposition en séries de fonctions.

Exemple L'équation de propagation des ondes

$$\frac{\partial^2 f}{\partial x^2}(x,t) - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}(x,t) = 0$$

On procède à sa résolution par changement de variables (voir plus bas).

Proposition

La solution générale sur $I \times J$ de l'équation aux dérivées partielles

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 0$$

est

$$f : (x,y) \mapsto xC(y) + D(y) \text{ avec } C, D : J \xrightarrow{\mathcal{C}^2} \mathbb{R}$$

Proposition

La solution générale sur $\mathbb{R}^2$ de l'équation aux dérivées partielles

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = 0$$

est

$$f : (x,y) \mapsto C(x) + D(y) \text{ avec } C : I \xrightarrow{\mathcal{C}^2} \mathbb{R} \text{ et } D : J \xrightarrow{\mathcal{C}^2} \mathbb{R}$$

Exemple Soit $c > 0$.
Résolvons sur $\mathbb{R}^2$ l'équation

$$(E) : \frac{\partial^2 f}{\partial x^2}(x,t) - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}(x,t) = 0$$

via le changement de variables :

$$\begin{cases} u = x + ct \\ v = x - ct \end{cases}$$

On a

$$\begin{cases} u = x + ct \\ v = x - ct \end{cases} \Leftrightarrow \begin{cases} x = (u+v)/2 \\ t = (u-v)/2c \end{cases}$$

L'application $\Phi : (u,v) \mapsto ((u+v)/2, (u-v)/2c)$ est une bijection de classe $\mathcal{C}^2$ de $\mathbb{R}^2$ vers $\mathbb{R}^2$.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^2$ et $g : \mathbb{R}^2 \to \mathbb{R}$ définie par « $g(u,v) = f(x,t)$ » i.e.

$$g(u,v) = f\left(\frac{u+v}{2}, \frac{u-v}{2c}\right)$$

$g = f \circ \Phi$ est de classe $\mathcal{C}^2$.
Après calculs,

$$\frac{\partial^2 g}{\partial u \partial v}(u,v) = \frac{1}{4}\left[\frac{\partial^2 f}{\partial x^2}(x,t) - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}(x,t)\right]_{\substack{x=(u+v)/2 \\ y=(u-v)/2c}}$$

f est solution sur $\mathbb{R}^2$ de l'équation des ondes
$\Leftrightarrow \forall (u,v) \in \mathbb{R}^2, \dfrac{\partial^2 g}{\partial u \partial v}(u,v) = 0$
$\Leftrightarrow \exists C, D : \mathbb{R} \xrightarrow{\mathcal{C}^2} \mathbb{R}, \forall (u,v) \in \mathbb{R}^2, g(u,v) = C(u) + D(v),$
$\Leftrightarrow \exists C, D : \mathbb{R} \xrightarrow{\mathcal{C}^2} \mathbb{R}, \forall (x,t) \in \mathbb{R}^2, f(x,t) = C(x+ct) + D(x-ct).$

Exemple Résolvons sur $\mathbb{R}^{+\star}\times\mathbb{R}$ l'équation aux dérivées partielles

$$(E) : x^2\frac{\partial^2 f}{\partial x^2}+2xy\frac{\partial^2 f}{\partial x\partial y}+y^2\frac{\partial^2 f}{\partial y^2}=xy$$

en passant aux coordonnées polaires.

$$\begin{cases} x=r\cos\theta \\ y=r\sin\theta \end{cases},\ \begin{cases} r=\sqrt{x^2+y^2} \\ \theta=\arctan(y/x) \end{cases}$$

L'application $\Phi : (r,\theta)\mapsto(r\cos\theta,r\sin\theta)$ est une bijection de classe $\mathcal{C}^2$ de $\mathbb{R}^{+\star}\times\,]-\pi/2,\pi/2[$ vers $\mathbb{R}^{+\star}\times\mathbb{R}$.
Soit $f:\mathbb{R}^{+\star}\times\mathbb{R}\to\mathbb{R}$ de classe $\mathcal{C}^2$ et $g:\mathbb{R}^{+\star}\times\,]-\pi/2,\pi/2[\,\to\mathbb{R}$ définie de sorte que « $g(r,\theta)=f(x,y)$ » i.e.

$$g(r,\theta)=f(r\cos\theta,r\sin\theta)$$

$g=f\circ\Phi$ est de classe $\mathcal{C}^2$.
Après calculs,

$$r^2\frac{\partial^2 g}{\partial r^2}(r,\theta)=\left[x^2\frac{\partial^2 f}{\partial x^2}+2xy\frac{\partial^2 f}{\partial x\partial y}+y^2\frac{\partial^2 f}{\partial y^2}\right]_{\substack{x=r\cos\theta\\ y=r\sin\theta}}$$

f est solution sur $\mathbb{R}^{+\star}\times\mathbb{R}$ de l'équation E
$\Leftrightarrow\forall(r,\theta)\in\mathbb{R}^{+\star}\times\,]-\pi/2,\pi/2[\,,r^2\dfrac{\partial^2 g}{\partial r^2}(r,\theta)=r^2\cos\theta\sin\theta,$
$\Leftrightarrow\exists C,D:\,]-\pi/2,\pi/2[\xrightarrow{\mathcal{C}^2}\mathbb{R},\forall(r,\theta)\in\mathbb{R}^{+\star}\times\,]-\pi/2,\pi/2[\,,g(r,\theta)=\dfrac{1}{2}r^2\cos\theta\sin\theta+rC(\theta)+D(\theta),$
$\Leftrightarrow\exists C,D:\,]-\pi/2,\pi/2[\xrightarrow{\mathcal{C}^2}\mathbb{R},\forall(x,y)\in\mathbb{R}^{+\star}\times\mathbb{R},f(x,y)=$
$\dfrac{1}{2}xy+\sqrt{x^2+y^2}C(\arctan(y/x))+D(\arctan(y/x)),$
$\Leftrightarrow\exists\tilde{C},\tilde{D}:\mathbb{R}\xrightarrow{\mathcal{C}^2}\mathbb{R},\forall(x,y)\in\mathbb{R}^{+\star}\times\mathbb{R},f(x,y)=\dfrac{1}{2}xy+\sqrt{x^2+y^2}\tilde{C}(y/x)+\tilde{D}(y/x),$
$\Leftrightarrow\exists\hat{C},\tilde{D}:\mathbb{R}\xrightarrow{\mathcal{C}^2}\mathbb{R},\forall(x,y)\in\mathbb{R}^{+\star}\times\mathbb{R},f(x,y)=\dfrac{1}{2}xy+x\hat{C}(y/x)+\tilde{D}(y/x)$
car $\sqrt{x^2+y^2}\underset{x>0}{=}x\psi(t)$ avec $\psi(t)=\sqrt{1+t^2}$, ψ de classe $\mathcal{C}^2$ ne s'annulant pas.

25.5 Eléments d'analyse vectorielle

On suppose le plan géométrique muni d'un repère orthonormé direct $\mathcal{R}=(O;\vec{i},\vec{j})$.

25.5.1 Gradient géométrique

Soit f une fonction réelle définie sur une partie du plan.
Si (x,y) sont les coordonnées cartésiennes de M, on pose $f_c(x,y)=f(M)$.
Exemple $f(M)=OM^2$, $f(M)=C/OM$,...

Définition

f_c est appelée représentation cartésienne de f dans le repère $\mathcal{R}$

Sous réserve d'existence, on pose

$$\frac{\partial f}{\partial x}(M) = \frac{\partial f_c}{\partial x}(x, y) \text{ et } \frac{\partial f}{\partial y}(M) = \frac{\partial f_c}{\partial y}(x, y)$$

Exemple Si $f(M) = OM^2$ alors $f_c(x, y) = x^2 + y^2$ et donc

$$\frac{\partial f}{\partial x}(M) = 2x \text{ et } \frac{\partial f}{\partial y}(M) = 2y$$

Définition

On appelle vecteur gradient de f en M le vecteur

$$\overrightarrow{\text{grad}}\, f(M) = \frac{\partial f}{\partial x}(M).\vec{i} + \frac{\partial f}{\partial y}(M).\vec{j}$$

On vérifie

$$f(M + \vec{h}) = f(M) + (\overrightarrow{\text{grad}}\, f(M) \mid \vec{h}) + o(\vec{h}) \text{ quand } \vec{h} \to \vec{0}$$

Cette relation caractérise le vecteur $\overrightarrow{\text{grad}}\, f(M)$ et assure que celui-ci est indépendant du choix du repère orthonormé $\mathcal{R}$. Elle peut être mise en résonance avec l'écriture physicienne

$$\mathrm{d}f = \overrightarrow{\text{grad}} f.\overrightarrow{\mathrm{d}M}$$

25.5.2 Gradient en coordonnées polaires

Si (r, θ) est un système de coordonnées polaires de M dans $\mathcal{R}$, on pose $f_p(r, \theta) = f(M)$.

Définition

f_p est appelée représentation polaire de f dans le repère $\mathcal{R}$.

Sous réserve d'existence, on pose

$$\frac{\partial f}{\partial r}(M) = \frac{\partial f_p}{\partial r}(r, \theta) \text{ et } \frac{\partial f}{\partial \theta}(M) = \frac{\partial f_p}{\partial \theta}(r, \theta)$$

Exemple Si $f(M) = OM^2$ alors $f_p(M) = r^2$ et

$$\frac{\partial f}{\partial r}(M) = 2r \text{ et } \frac{\partial f}{\partial \theta}(M) = 0$$

Proposition

On a

$$\overrightarrow{\text{grad}}\, f(M) = \frac{\partial f}{\partial r}(M)\vec{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}(M)\vec{u}_\theta$$

en notant $\vec{u}_r = \cos\theta\vec{i} + \sin\theta\vec{j}$ et $\vec{u}_\theta = -\sin\theta\vec{i} + \cos\theta\vec{j}$.

dém. :
Si (r,θ) est un système de coordonnées polaires de M alors ses coordonnées cartésiennes sont $(r\cos\theta, r\sin\theta)$.
Par suite $f_p(r,\theta) = f_c(r\cos\theta, r\sin\theta)$.
On en déduit

$$\frac{\partial f_p}{\partial r}(r,\theta) = \cos\theta \frac{\partial f_c}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta \frac{\partial f_c}{\partial y}(r\cos\theta, r\sin\theta)$$

$$\frac{\partial f_p}{\partial \theta}(r,\theta) = -r\sin\theta \frac{\partial f_c}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta \frac{\partial f_c}{\partial y}(r\cos\theta, r\sin\theta)$$

ce qui se réécrit

$$\frac{\partial f}{\partial r}(M) = \cos\theta \frac{\partial f}{\partial x}(M) + \sin\theta \frac{\partial f}{\partial y}(M) \text{ (1)}$$

$$\frac{\partial f}{\partial \theta}(M) = -r\sin\theta \frac{\partial f}{\partial x}(M) + r\cos\theta \frac{\partial f}{\partial y}(M) \text{ (2)}$$

$\cos\theta \times (1) - \dfrac{1}{r}\sin\theta \times (2)$ donne

$$\frac{\partial f}{\partial x}(M) = \cos\theta \frac{\partial f}{\partial r}(M) - \frac{1}{r}\sin\theta \frac{\partial f}{\partial \theta}(M)$$

$\sin\theta \times (1) + \dfrac{1}{r}\cos\theta \times (2)$ donne

$$\frac{\partial f}{\partial y}(M) = \sin\theta \frac{\partial f}{\partial r}(M) + \frac{1}{r}\cos\theta \frac{\partial f}{\partial \theta}(M)$$

On en déduit

$$\overrightarrow{\text{grad}}\, f(M) = \frac{\partial f}{\partial x}(M)\vec{\imath} + \frac{\partial f}{\partial y}(M)\vec{\jmath} = \frac{\partial f}{\partial r}(M)\vec{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}(M)\vec{u}_\theta$$

□

Remarque Le physicien retrouve les relations (1) et (2) de la démonstration ci-dessus en écrivant

$$\frac{\partial f}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial f}{\partial y} \text{ et } \frac{\partial f}{\partial \theta} = \frac{\partial x}{\partial \theta}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial f}{\partial y}$$

25.5.3 Intégration d'un champ de vecteurs

Soit $\vec{F}$ un champ de vecteurs défini sur une partie du plan. On peut écrire

$$\vec{F}(M) = F_x(M).\vec{\imath} + F_y(M).\vec{\jmath}$$

Soit Γ une courbe inscrite dans le domaine de définition de $\vec{F}$ joignant un point A à un point B. On suppose que la courbe Γ peut être paramétrée par

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \text{ avec } t \in [a,b]$$

Définition

On appelle circulation du champ de vecteur $\vec{F}$ le long de l'arc Γ le réel

$$\int_\Gamma \overrightarrow{F(M)}.\overrightarrow{\mathrm{d}M} \underset{\text{déf}}{=} \int_a^b \left(F_x(M(t))x'(t) + F_y(M(t))y'(t)\right)\,\mathrm{d}t$$

Remarque On peut montrer que cette valeur est géométrique dans le sens où, si l'on détermine un autre paramétrage de Γ, le résultat du calcul est inchangé.

Théorème

Si $\vec{F} = -\overrightarrow{\mathrm{grad}}V$ alors

$$\int_\Gamma \overrightarrow{F(M)}.\overrightarrow{\mathrm{d}M} = V(A) - V(B)$$

En particulier, si $M(a) = M(b)$ alors

$$\int_\Gamma \overrightarrow{F(M)}.\overrightarrow{\mathrm{d}M} = 0$$

dém. :
Par hypothèse

$$F_x = -\frac{\partial V}{\partial x} \text{ et } F_y = -\frac{\partial V}{\partial y}$$

donc

$$\int_\Gamma \overrightarrow{F(M)}.\overrightarrow{\mathrm{d}M} = -\int_a^b x'(t)\frac{\partial V}{\partial x}(x(t),y(t)) + y'(t)\frac{\partial V}{\partial y}(x(t),y(t))\,\mathrm{d}t$$

Or

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(V(x(t),y(t)\right) = x'(t)\frac{\partial V}{\partial x}(x(t),y(t)) + y'(t)\frac{\partial V}{\partial y}(x(t),y(t))$$

donc

$$\int_\Gamma \overrightarrow{F(M)}.\overrightarrow{\mathrm{d}M} = -\left[V(x(t),y(t))\right]_{t=a}^b$$

□

25.5.4 Laplacien

Soit f une fonction réelle définie sur une partie du plan.

Définition

On appelle laplacien d'une fonction f définie sur une partie du plan la quantité

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Remarque On peut montrer que cette quantité ne dépend pas du choix du repère orthonormé (c'est la trace de la matrice Hessienne).

Exemple L'équation de la chaleur en dimension 2 s'exprimer

$$\frac{\partial f}{\partial t}(x,t) = D.\Delta f(x,t)$$

Proposition

En coordonnées polaires

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}$$

Troisième partie

Probabilité

Chapitre 26

Probabilités

26.1 Espace probabilisé

26.1.1 Univers

Définition

L'ensemble des résultats possibles décrivant une expérience aléatoire est appelé univers. Il est généralement noté Ω. Les éléments ω de Ω sont les issues observées de l'expérience aléatoire, on les appelle éventualités. La réalisation de l'expérience aléatoire revient au choix d'une éventualité dans l'univers i.e. d'un élément ω à l'intérieur de l'ensemble Ω.

Exemple On lance une pièce pour obtenir Pile ou Face.
Il est naturel de choisir $\Omega = \{P, F\}$ pour modéliser les issues de l'expérience.
On lance la pièce n fois, on choisira $\Omega = \{P, F\}^n$.
On lance la pièce indéfiniment : on choisira $\Omega = \{P, F\}^{\mathbb{N}^\star}$.

Exemple On lance un dé : on choisit $\Omega = [\![1, 6]\!]$.
On lance deux dés : on choisit $\Omega = [\![1, 6]\!] \times [\![1, 6]\!]$ ou $\Omega = [\![2, 12]\!]$ selon l'ambition de l'étude menée.
Si l'on prend $\Omega = [\![1, 6]\!] \times [\![1, 6]\!]$, c'est aussi que l'on suppose les deux dés discernables.

Exemple On compte le nombre de jets d'un dé avant d'obtenir un premier 6, on choisira $\Omega = \mathbb{N}^\star$.

Exemple Une urne contient 1 boule blanche et 4 boules rouges.
On tire successivement deux boules avec remise :

$$\Omega = \{(B, B), (B, R), (R, B), (R, R)\}$$

On tire successivement deux boules sans remise :

$$\Omega = \{(B, R), (R, B), (R, R)\}$$

On tire simultanément deux boules :

$$\Omega = \{\{B, R\}, \{R, R\}\}$$

Remarque Le choix de l'univers Ω dépend de la modélisation choisie pour l'expérience aléatoire
- il ne doit pas être trop petit pour pouvoir étudier toutes les issues souhaitées ;
- il ne doit pas être inutilement grand en prenant en compte des phénomènes inutiles.

26.1.2 Tribu

Les sous-ensembles de l'univers Ω serviront pour décrire des événements dont on veut mesurer la probabilité d'occurrence. Contrairement à ce qui a été vu en première année dans le cas où l'ensemble Ω est fini, toute partie de Ω ne définira pas nécessairement un événement : on se limitera aux parties éléments d'une tribu.

Définition

On appelle tribu sur un ensemble Ω toute partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ vérifiant :
1) $\Omega \in \mathcal{A}$;
2) $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$;
3) $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}, \bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$
La dernière propriété s'appelle la stabilité par réunion dénombrable.

Exemple $\mathcal{A} = \mathcal{P}(\Omega)$ est une tribu de Ω.

Exemple $\mathcal{A} = \{\emptyset, \Omega\}$ est une tribu de Ω.

Exemple Soit A une partie de Ω. $\mathcal{A} = \left\{\emptyset, A, \bar{A}, \Omega\right\}$ est une tribu de Ω.

Théorème

Si $\mathcal{A}$ est une tribu sur un ensemble Ω alors
a) $\emptyset \in \mathcal{A}$;
b) $\forall A, B \in \mathcal{A}, A \cup B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$ et $A \backslash B \in \mathcal{A}$
c) $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}, \bigcap_{n=0}^{+\infty} A_n \in \mathcal{A}$

dém. :
a) $\Omega \in \mathcal{A}$ donc $\bar{\Omega} = \emptyset \in \mathcal{A}$.
b) Soit $A, B \in \mathcal{A}$. En choisissant $A_0 = A$, $A_1 = B$ et $A_n = \emptyset$ pour $n \geqslant 2$, $A \cup B = \bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$.
Aussi $\overline{A \cap B} = \bar{A} \cup \bar{B} \in \mathcal{A}$ donc $A \cap B \in \mathcal{A}$ et $A \backslash B = A \cap \bar{B} \in \mathcal{A}$.
c) $\overline{\bigcap_{n=0}^{+\infty} A_n} = \bigcup_{n=0}^{+\infty} \overline{A_n} \in \mathcal{A}$ donc $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$.
□

Remarque Une tribu est donc stable :
- par passage au complémentaire ;
- par réunion et intersection finie ou dénombrable.

Définition

On appelle espace probabilisable tout couple $(\Omega, \mathcal{A})$ constitué d'un ensemble Ω et d'une tribu $\mathcal{A}$ sur Ω.

Exemple $(\Omega, \mathcal{P}(\Omega))$ est un espace probabilisable.

26.1.3 Evénements

Définition

Si $(\Omega, \mathcal{A})$ est un espace probabilisable, les parties A de Ω éléments de la tribu $\mathcal{A}$ sont appelées événement de l'univers Ω.

Exemple On lance un dé et l'on considère $\Omega = [\![1, 6]\!]$ et $\mathcal{A} = \mathcal{P}(\Omega)$.
L'événement élémentaire $\Omega = \{6\}$ traduit « on a obtenu un 6 » .
L'événement $\Omega = \{2, 4, 6\}$ traduit « le tirage est un nombre pair » .

Exemple Une famille à deux enfants dont on étudie le genre en fonction du rang de naissance.

$$\Omega = \{(F, F), (F, G), (G, F), (G, G)\} \text{ et } \mathcal{A} = \mathcal{P}(\Omega)$$

L'événement l'aîné est un garçon est

$$A = \{(G, G), (G, F)\}$$

Définition

L'événement $\emptyset$ est appelé événement impossible.
L'événement Ω est appelé événement certain.
Les événements de la forme $\{\omega\}$ sont appelés événements élémentaires.

Définition

Si A et B sont deux événements de l'espace probabilisable $(\Omega, \mathcal{A})$ alors
- $\bar{A}$ est l'événement contraire de A ;
- $A \cap B$ est l'événement conjonction de A et B ;
- $A \cup B$ est l'événement disjonction de A et B.

Définition

Soit A et B deux événements de l'espace probabilisable $(\Omega, \mathcal{A})$.
On dit que l'événement A implique B si $A \subset B$.
On dit que les événements A et B sont incompatibles si $A \cap B = \emptyset$.

Exemple Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de l'espace probabilisable $(\Omega, \mathcal{A})$.
L'événement $\displaystyle\bigcap_{n=0}^{+\infty} A_n$ correspond à la réalisation de tous les A_n.
L'événement $\displaystyle\bigcup_{n=0}^{+\infty} A_n$ correspond à la réalisation d'au moins un A_n.

L'événement $\bigcup_{N=0}^{+\infty} \bigcap_{n=N}^{+\infty} A_n$ correspond à la réalisation de tous les A_n à partir d'un certain rang.
L'événement $\bigcap_{N=0}^{+\infty} \bigcup_{n=N}^{+\infty} A_n$ correspond à la réalisation d'une infinité de A_n.

Remarque Notons que les ensembles décrits dans l'exemple au dessus sont bien éléments de la tribu $\mathcal{A}$.

26.2 Probabilités

$(\Omega, \mathcal{A})$ désigne un espace probabilisable

26.2.1 Définition

Définition

On appelle probabilité sur l'espace probabilisable $(\Omega, \mathcal{A})$ toute application $P : \mathcal{A} \to \mathbb{R}^+$ vérifiant :
- $P(\Omega) = 1$;
- Pour toute suite $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ d'événements deux à deux incompatibles

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) \text{ [}\sigma\text{-additivité]}$$

Exemple Soit Ω un ensemble fini et $\mathcal{A} = \mathcal{P}(\Omega)$.
On définit la probabilité uniforme sur Ω par

$$P(A) = \frac{\text{Card}A}{\text{Card}\Omega}$$

Exemple Soit ω un élément de Ω. On définit une probabilité sur $(\Omega, \mathcal{A})$ par

$$P(A) = \begin{cases} 0 & \text{si } \omega \notin A \\ 1 & \text{si } \omega \in A \end{cases}$$

Définition

On appelle espace probabilisé tout triplet $(\Omega, \mathcal{A}, P)$ formé d'un ensemble Ω, d'une tribu $\mathcal{A}$ sur Ω et d'une probabilité P sur $(\Omega, \mathcal{A})$.

26.2.2 Propriétés élémentaires

Soit P une probabilité sur $(\Omega, \mathcal{A})$.

Théorème

a) $P(\emptyset) = 0$
b) Si $A_0, \dots, A_n$ sont des événements deux à deux incompatibles

$$P\left(\bigcup_{k=0}^{n} A_k\right) = \sum_{k=0}^{n} P(A_k)$$

c) $\forall A \in \mathcal{A}, P(\bar{A}) = 1 - P(A)$
d) $\forall A \in \mathcal{A}, P(A) \in [0,1]$

dém. :
a) En prenant $A_n = \emptyset$ pour tout $n \in \mathbb{N}$, on obtient

$$P(\emptyset) = \sum_{n=0}^{+\infty} P(\emptyset)$$

et donc $P(\emptyset) = 0$.
b) On choisit $A_k = \emptyset$ pour $k > n$ et on exploite

$$P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \sum_{k=0}^{+\infty} P(A_k)$$

c) Ω est la conjonction des événements incompatibles A et $\bar{A}$ donc

$$1 = P(\Omega) = P(A) + P(\bar{A})$$

d) $P(A) \geqslant 0$ et $P(\bar{A}) = 1 - P(A) \geqslant 0$.
□

Théorème

Soit A et B deux événements
a) $A \subset B \Rightarrow P(A) \leqslant P(B)$
b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

dém. :
a) Si $A \subset B$ alors B est la réunion disjointe de A et de $B \backslash A$. L'égalité $P(B) = P(A) + P(B \backslash A)$ donne alors $P(B) \geqslant P(A)$.
b) $A \cup B$ est la réunion disjointe de A et de $B \backslash A$. On a donc $P(A \cup B) = P(A) + P(B \backslash A)$.
Or B est la réunion disjointe de $B \backslash A$ et de $A \cap B$ donc $P(B) = P(B \backslash A) + P(A \cap B)$ ce qui permet de conclure.
□

Corollaire

Si $A_0, \dots, A_n$ sont des événements alors

$$P\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \sum_{k=0}^{n} P(A_k)$$

dém. :
Par récurrence sur $n \in \mathbb{N}$.
□

Remarque On peut énoncer une égalité connue sous le nom de formule du crible, mais celle-ci est hors-programme.

Corollaire

Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leqslant \sum_{n=0}^{+\infty} P(A_n)$$

26.2.3 Continuité monotone

Théorème

Si (A_n) est une suite croissante d'événements alors

$$P(A_n) \xrightarrow[n\to+\infty]{} P\left(\bigcup_{n=0}^{+\infty} A_n\right)$$

dém. :
Posons $B_0 = A_0$ puis, pour tout $n \geqslant 1$, $B_n = A_n \backslash A_{n-1}$.
Puisque la suite (A_n) est croissante pour l'inclusion, les événements de la suite (B_n) sont deux à deux disjoints. De plus

$$A_n = \bigcup_{k=0}^{n} B_k \text{ et } \bigcup_{n=0}^{+\infty} A_n = \bigcup_{n=0}^{+\infty} B_n$$

Par conséquent

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = P\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} P(B_n) = \lim_{n\to+\infty} \sum_{k=0}^{n} P(B_k)$$

avec

$$\sum_{k=0}^{n} P(B_k) = P(A_n)$$

□

Remarque Ce résultat est utile pour calculer la probabilité d'une union dénombrable.

Corollaire

On a

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P\left(\bigcup_{k=0}^{n} A_k\right)$$

Théorème

Si (A_n) est une suite décroissante d'événements alors

$$P(A_n) \xrightarrow[n\to+\infty]{} P\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

dém. :
Posons $B_n = \overline{A_n}$. (B_n) est une suite croissante d'événements avec

$$\overline{\bigcup_{n=0}^{+\infty} B_n} = \bigcap_{n=0}^{+\infty} \overline{B_n} = \bigcap_{n=0}^{+\infty} A_n$$

Par continuité croissante

$$P(B_n) \xrightarrow[n\to+\infty]{} P\left(\bigcup_{n=0}^{+\infty} B_n\right)$$

et donc

$$P(A_n) = 1 - P(B_n) \xrightarrow[n\to+\infty]{} 1 - P\left(\bigcup_{n=0}^{+\infty} B_n\right) = P\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

□

Corollaire

On a

$$P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} P\left(\bigcap_{k=0}^{n} A_k\right)$$

Exemple On lance indéfiniment un dé équilibré. Montrer que l'événement « on n'obtient jamais de 6 » est de probabilité nulle.
On note A l'événement : « on n'obtient jamais de 6 » On note A_n l'événement

« on n'a pas obtenu de 6 lors des n premiers lancers »

En supposant les lancers indépendants

$$P(A_n) = (5/6)^n$$

Puisque la suite (A_n) est décroissante, on a par continuité

$$P(A) = P\left(\bigcap_{n=1}^{+\infty} A_n\right) = \lim_{n\to+\infty} P(A_n) = 0$$

26.2.4 Evénements presque sûrs

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

Définition

On dit qu'un événement A est négligeable si $P(A) = 0$.

Exemple L'événement impossible est négligeable.

Exemple Ne jamais obtenir de six en lançant indéfiniment un dé équilibré est négligeable.

Proposition

Un événement inclus dans un événement négligeable est négligeable

dém. :
Cas

$$A \subset B \Rightarrow P(A) \leqslant P(B)$$

□

Proposition

Une réunion finie ou dénombrable d'événements négligeables est négligeable.

dém. :
Car

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leqslant \sum_{n=0}^{+\infty} P(A_n)$$

□

Définition

On dit qu'un événement A est presque sûr si $P(A) = 1$.
Ceci signifie encore que l'événement $\bar{A}$ est négligeable.

Exemple L'événement certain est presque sûr.

Exemple Obtenir un six en lançant indéfiniment un dé équilibré est un événement presque sûr.

Proposition

Un événement contenant un événement presque sûr est presque sûr.

Proposition

Une intersection finie ou dénombrable d'événements presque sûrs est presque sûre.

26.2.5 Probabilité sur un univers au plus dénombrable

Soit Ω un ensemble fini ou dénombrable, $\mathcal{A} = \mathcal{P}(\Omega)$ et P une probabilité sur $(\Omega, \mathcal{A})$.

Définition

Pour tout $\omega \in \Omega$, on introduit les probabilités élémentaires

$$p_\omega = P(\{\omega\})$$

Théorème

La famille $(p_\omega)_{\omega\in\Omega}$ est une famille de réels positifs, sommable et de somme égale à 1.

dém. :
$p_\omega = P(\{\omega\}) \in [0,1]$ donc $p_\omega \in \mathbb{R}^+$.
Cas Ω fini : $\Omega = \{\omega_1, \ldots, \omega_n\}$ avec $\omega_1, \ldots, \omega_n$ deux à deux distincts

$$\sum_{\omega \in \Omega} p_w = \sum_{i=1}^{n} P(\{\omega_i\}) = P\left(\bigcup_{i=1}^{n} \{\omega_i\}\right) = P(\Omega) = 1$$

Cas Ω dénombrable : $\Omega = \{\omega_n / n \in \mathbb{N}\}$ avec les ω_n deux à deux distincts

$$\sum_{\omega \in \Omega} p_\omega = \sum_{n=0}^{+\infty} P(\{\omega_n\}) = P\left(\bigcup_{n=0}^{+\infty} \{\omega_n\}\right) = P(\Omega) = 1$$

□

Théorème

Si $(p_\omega)_{\omega \in \Omega}$ est une famille de réels positifs, sommable et de somme égale à 1 alors il existe une unique probabilité P sur $(\Omega, \mathcal{A})$ vérifiant

$$\forall \omega \in \Omega, P(\{\omega\}) = p_\omega$$

De plus, celle-ci est déterminée par

$$\forall A \subset \Omega, P(A) = \sum_{\omega \in A} p_\omega$$

dém. :
Analyse : Supposons P probabilité solution.
Pour tout $A \subset \Omega$, on a la réunion disjointe

$$A = \bigcup_{\omega \in A} \{\omega\}$$

et donc, que A soit fini ou dénombrable

$$P(A) = \sum_{\omega \in A} p_\omega$$

La probabilité P est donc déterminée de façon unique.
Synthèse : Supposons $P : \mathcal{P}(\Omega) \to \mathbb{R}^+$ définie par

$$\forall A \subset \Omega, P(A) = \sum_{\omega \in A} p_\omega$$

L'application P est bien définie à valeurs dans $\mathbb{R}^+$.
$P(\Omega) = 1$ car par hypothèse la somme de p_ω vaut 1.

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements deux à deux incompatibles et $A = \bigcup_{n=0}^{+\infty} A_n$. Par sommation par paquets

$$\sum_{\omega \in A} p_\omega = \sum_{n=0}^{+\infty} \sum_{\omega \in A_n} p_\omega$$

et donc

$$P(A) = \sum_{n=0}^{+\infty} P(A_n)$$

□

Exemple Cas Ω fini : $\Omega = \{\omega_1, \dots, \omega_n\}$
Une probabilité sur Ω est entièrement déterminée par le choix de $p_1, \dots, p_n \in \mathbb{R}^+$ avec

$$p_1 + \cdots + p_n = 1$$

En prenant $p_k = 1/n$, on définit l'équiprobabilité sur $[\![1, n]\!]$.
En prenant $p_k = \binom{n}{k} p^k (1-p)^{n-k}$ avec $p \in \,]0,1[$, on définit une probabilité sur $[\![0, n]\!]$

Exemple Cas $\Omega = \mathbb{N}$
Une probabilité sur Ω est déterminée par le choix de $(p_n)_{n\in\mathbb{N}} \in \mathbb{R}^{+\mathbb{N}}$ avec

$$\sum_{n=0}^{+\infty} p_n = 1$$

L'équiprobabilité sur $\mathbb{N}$ est impossible.
Plus généralement, elle est impossible sur Ω infini dénombrable.
En revanche

$$p_n = \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} \text{ avec } \lambda \in \mathbb{R}^+$$

définit une probabilité sur $\mathbb{N}$.
Aussi

$$p_n = p(1-p)^{n-1} \text{ avec } p \in \,]0,1[$$

définit une probabilité sur $\mathbb{N}^\star$

26.3 Probabilités conditionnelles

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

26.3.1 Définition

Définition

Soit B un événement de Ω vérifiant $P(B) > 0$.
Pour tout événement A de Ω, la probabilité conditionnelle de A sachant B est définie par

$$P(A \mid B) \underset{\text{déf}}{=} \frac{P(A \cap B)}{P(B)}$$

Si $P(B) = 0$, on convient de poser $P(A \mid B) = 0$.

Exemple On lance un dé équilibré. $\Omega = \{1, 2, 3, 4, 5, 6\}$.

On considère les événements

$$A = \text{« on obtient 6 » et } B = \text{« le tirage est pair »}$$

Déterminons $P(A \mid B)$ et $P(A \mid \bar{B})$
Par retour à la définition

$$P(A \mid B) = \frac{1/6}{1/2} = \frac{1}{3} \text{ et } P(A \mid \bar{B}) = \frac{0}{1/3} = 0$$

Théorème

Si B est événement de Ω vérifiant $P(B) > 0$ alors l'application $P_B : \mathcal{P}(\Omega) \to \mathbb{R}^+$ donnée par

$$P_B(A) = P(A \mid B)$$

définit une probabilité sur $(\Omega, \mathcal{A})$.

dém. :
D'une part

$$P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = 1$$

et d'autre part, pour (A_n) suites d'événements deux à deux incompatibles

$$P_B\left(\bigcup_{n=0}^{+\infty} A_n\right) = \frac{P\left(\bigcup_{n=0}^{+\infty} (A_n \cap B)\right)}{P(B)} = \frac{\sum_{n=0}^{+\infty} P(A_n \cap B)}{P(B)} = \sum_{n=0}^{+\infty} P_B(A_n)$$

□

Corollaire

Les propriétés calculatoires relatives aux probabilités sont aussi vraies pour les probabilités conditionnelles.

26.3.2 Formule des probabilités composées

Théorème

Soit A, B deux événements de Ω. On a

$$P(A \cap B) = P(A \mid B) P(B)$$

dém. :
C'est immédiat compte tenu de la définition de $P(A \mid B)$ quand $P(B) > 0$. L'identité est aussi vraie quand $P(B) = 0$ car $A \cap B \subset B$.
□

Corollaire

Soit $A_1, \ldots, A_n$ des événements de Ω. On a

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2 \mid A_1) \ldots P(A_n \mid A_1 \cap \ldots \cap A_{n-1})$$

dém. :
Par récurrence sachant que le théorème ci-dessus avec $A = A_{n+1}$ et $B = A_1 \cap \ldots \cap A_n$ fournit

$$P(A_1 \cap \ldots \cap A_{n+1}) = P(A_1 \cap \ldots \cap A_n)P(A_{n+1} \mid A_1 \cap \ldots \cap A_n)$$

□

Exemple Une urne contient n boules blanches et n boules rouges.
On tire successivement et sans remise n boules dans cette urne.
Déterminons la probabilité qu'une boule rouge figure dans ce tirage.
Nous allons en fait mesurer l'événement contraire.
Notons A_k l'événement

« la boule obtenue lors du k-ième tirage est blanche »

$$P(A_1) = \frac{n}{2n} = \frac{1}{2} \text{ et } P(A_k \mid A_1 \cap \ldots \cap A_{k-1}) = \frac{n-(k-1)}{2n-(k-1)}$$

Par probabilités composées

$$P(A_1 \cap \ldots \cap A_n) = \frac{n}{2n} \times \frac{n-1}{2n-1} \times \cdots \times \frac{1}{n+1} = \frac{(n!)^2}{(2n)!}$$

et la probabilité cherchée est donc

$$P\left(\bar{A}_1 \cup \ldots \cup \bar{A}_n\right) = 1 - \frac{(n!)^2}{(2n)!}$$

Exemple Une urne contient une boule blanche et une boule rouge.
On tire successivement des boules dans cette urne. A chaque boule tirée, on note la couleur de celle-ci et on la remet dans l'urne accompagnée d'une boule de la même couleur.
Montrons qu'il est presque sûr que la boule rouge initiale sera tirée.
Notons A_n l'événement « la boule tirée au -ième tirage est blanche » Par probabilités composées

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)P(A_2 \mid A_1) \ldots P(A_n \mid A_1 \cap \ldots \cap A_{n-1})$$

avec

$$P(A_1) = \frac{1}{2}, P(A_2 \mid A_1) = \frac{2}{3}, \ldots, P(A_n \mid A_1 \cap \ldots \cap A_{n-1}) = \frac{n}{n+1}$$

On a donc

$$P(A_1 \cap \ldots \cap A_n) = \frac{1}{n+1}$$

Par continuité décroissante

$$P\left(\bigcap_{n=1}^{+\infty} A_n\right) = \lim_{n \to +\infty} P\left(A_1 \cap \ldots \cap A_n\right) = 0$$

Ainsi, l'événement « toutes les boules tirées sont blanches » est négligeable et l'événement complémentaire « la boule rouge initiale est tirée » est presque sûr.

26.3.3 Formule des probabilités totales

Définition

On appelle système complet d'événements toute famille $(A_i)_{i\in I}$ d'événements avec ensemble fini ou dénombrable vérifiant :
1) $\forall i,j\in I, i\neq j\Rightarrow A_i\cap A_j=\emptyset$;
2) $\bigcup_{i\in I} A_i=\Omega$
Autrement dit, la famille $(A_i)_{i\in I}$ est une famille au plus dénombrable d'événements deux à deux incompatibles et de réunion Ω.

Exemple Si A est un événement de Ω alors $(A,\bar{A})$ est un système complet d'événements.

Exemple Si Ω est dénombrable avec $\Omega=\{\omega_n/n\in\mathbb{N}\}$ (où les ω_n sont deux à deux distincts) et si $\mathcal{A}=\mathcal{P}(\Omega)$ alors les $A_n=\{\omega_n\}$ définissent un système complet d'événements.

Théorème

Si $(A_i)_{i\in I}$ est un système complet d'événements de l'espace probabilisé $(\Omega,\mathcal{A},P)$ alors pour tout événement B de Ω

$$P(B)=\sum_{i\in I}P\left(B\mid A_i\right)P(A_i)$$

dém. :
On a

$$B=B\cap\Omega=B\cap\left(\bigcup_{i\in I}A_i\right)=\bigcup_{i\in I}(B\cap A_i)$$

$$I$$

Les événements $B\cap A_i$ étant deux à deux incompatibles, que l'ensemble soit fini ou dénombrable, on obtient

$$P(B)=\sum_{i\in I}P(B\cap A_i)$$

Enfin, par probabilités composées

$$P(B)=\sum_{i\in I}P(B\mid A_i)P(A_i)$$

□

Exemple On dispose de six urnes numérotées de 1 à 6.
L'urne numéro k comporte k boules blanches et une boule rouge.
Un joueur lance un dé équilibré puis choisit une boule dans l'urne correspondant au résultat du dé.
Déterminons la probabilité que la boule tirée soit blanche.
On considère le système complet d'événements $(A_1,\dots,A_6)$ avec $A_k=$ « le dé donne la valeur k » et on étudie l'événement

$$B=\text{« la boule tirée est blanche »}$$

On a

$$P(A_k) = 1/6 \text{ et } P(B \mid A_k) = k/(k+1)$$

Par formule des probabilités totales

$$P(B) = \frac{1}{6}\sum_{k=1}^{6}\frac{k}{k+1} = \frac{617}{840}$$

Exemple Une urne contient une boule rouge.
Un joueur lance un dé équilibré.
S'il obtient un six, il tire une boule dans l'urne.
Sinon, il rajoute une boule blanche dans l'urne et répète la manipulation.
Sachant qu'il est presque sûr que le joueur fera un six, quelle est la probabilité que la boule tirée soit rouge ?
Le système complet d'événements choisi est $(A_n)_{n\in\mathbb{N}^\star}$ avec $A_n =$ « le joueur fait son premier six lors du n-ième lancer »
L'événement étudié est $B =$ « la boule tirée est rouge »
On a

$$P(A_n) = \frac{1}{6}\times\left(\frac{5}{6}\right)^{n-1} \text{ et } P(B \mid A_n) = \frac{1}{n}$$

Par la formule des probabilités totales

$$P(B) = \sum_{n=1}^{+\infty}\frac{1}{6n}\left(\frac{5}{6}\right)^{n-1} = \frac{1}{6}\sum_{n=1}^{+\infty}\frac{1}{n}\left(\frac{5}{6}\right)^{n} = -\frac{1}{6}\ln\left(1-\frac{5}{6}\right) = \frac{1}{6}\ln 6$$

26.3.4 Formule de Bayes

Théorème

Si A et B sont deux événements de probabilités non nulles alors

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

dém. :
C'est immédiat puisque

$$P(A \mid B)P(B) = P(A\cap B) = P(B \mid A)P(A)$$

□

Corollaire

Si $(A_i)_{i\in I}$ est un système complet d'événements alors pour tout événement B de probabilité non nulle et tout $k\in I$

$$P(A_k \mid B) = \frac{P(B \mid A_k)P(A_k)}{\sum_{i\in I} P(B \mid A_i)P(A_i)}$$

dém. :
Il suffit d'employer la formule précédente en exploitant celle des probabilités totales

$$P(B) = \sum_{i \in I} P(B \mid A_i) P(A_i)$$

□

Remarque La formule de Bayes est utile pour les raisonnements « rétroactifs » . Si on sait mesurer la conséquence B d'un événement A et que l'on sait l'événement B réalisé, la formule de Bayes permet de savoir si l'événement A l'a été. On parle parfois de la formule de probabilité des causes.

Exemple Une urne contient deux dés : l'un est équilibré et l'autre donne systématiquement un 6.
On choisit un dé dans l'urne et on le lance. On suppose que le dé lancé donne un 6, déterminons la probabilité que ce dé soit équilibré.
Notons A l'événement « le dé choisi est équilibré » On a $P(A) = P(\bar{A}) = 1/2$.
Notons B l'événement « le dé lancé donne un 6 » On veut mesurer $P(A \mid B)$.
Par la formule de Bayes

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

avec

$$P(B \mid A)P(A) = 1/6 \times 1/2$$

et

$$P(B) = P(B \mid A)P(A) + P(B \mid \bar{A})P(\bar{A}) = 1/12 + 1 \times 1/2$$

Ainsi

$$P(A \mid B) = \frac{1}{7}$$

26.4 Indépendance

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

26.4.1 Couple d'événements indépendants

Définition

On dit que deux événements A et B de l'espace probabilisé (Ω, P) sont indépendants si

$$P(A \cap B) = P(A)P(B)$$

Remarque Si $P(B) > 0$, on a alors

$$P(A \mid B) = P(A)$$

L'indépendance des événements A et B entraîne que la connaissance de la réalisation de B n'apporte rien pour savoir si A est aussi réalisé.

Exemple On lance deux fois le même dé (équilibré ou non). Les événements « le premier lancer donne un six » et « le second lancer donne un six » sont généralement modélisés indépendants.

Exemple On tire successivement et sans remise deux boules dans une urne contenant 5 boules blanches et 2 boules rouges. Les événements « la première boule tirée est blanche » et « la seconde boule tirée est blanche » ne sont pas indépendants.
En revanche, si l'on procède à un tirage avec remise, ces événements deviennent indépendants.

Attention : Ne pas confondre indépendance et incompatibilité : deux événements incompatibles sont rarement indépendants !

Proposition

Si A et B sont des événements indépendants alors A et $\bar{B}$ le sont aussi

dém. :
Puisque $\Omega = B \cup \bar{B}$

$$P(A) = P\left(A \cap (B \cup \bar{B})\right) = P\left((A \cap B) \cup (A \cap \bar{B})\right)$$

Or $A \cap B$ et $A \cap \bar{B}$ sont incompatibles

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(A)P(B) + P(A \cap \bar{B})$$

Ainsi

$$P(A \cap \bar{B}) = P(A)\,(1 - P(B)) = P(A)P(\bar{B})$$

□

Remarque Aussi $\bar{A}$ et B sont indépendants ainsi que $\bar{A}$ et $\bar{B}$.

26.4.2 Famille d'événements mutuellement indépendants

Définition

On dit que les événements d'une famille quelconque $(A_i)_{i \in I}$ d'événements de l'espace probabilisé $(\Omega, \mathcal{A}, P)$ sont mutuellement indépendants si

$$\forall J \text{ finie} \subset I, P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$$

Exemple On lance indéfiniment une pièce.
Soit A_i l'événement

« on obtient face lors du i-ème lancer »

Les événements de la famille $(A_n)_{n \geqslant 1}$ sont modélisés mutuellement indépendants.

Si la probabilité d'obtenir face lors de chaque lancer vaut $p \in]0,1[$, alors la probabilité que face apparaît pour la première fois lors du n-ième lancer vaut

$$P(A_n \cap \overline{A_{n-1}} \cap \ldots \cap \overline{A_1}) = p(1-p)^{n-1}$$

En effet, on peut montrer que les événements $\overline{A_1}, \ldots, \overline{A_{n-1}}$ et A_n sont mutuellement indépendants (voir ci-dessous).

Exemple A, B, C sont mutuellement indépendants si

$$P(A \cap B) = P(A)P(B), P(A \cap C) = P(A)P(C), P(B \cap C) = P(B)P(C)$$

et aussi

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

Attention : Il ne faut pas confondre l'indépendance mutuelle et l'indépendance deux à deux.

Exemple On lance deux dés discernables et l'on considère les événements

$A =$ « le premier dé lancé donne un résultat pair » $B =$ « le second dé lancé donne un résultat pair »

et

$C =$ « la somme des deux dés est un résultat pair »

Les événements A, B et C sont deux à deux indépendants, mais pas mutuellement indépendants.
En effet

$$P(A \cap B \cap C) = P(A \cap B) = \frac{1}{4} \text{ et } P(A)P(B)P(C) = \frac{1}{8}$$

Proposition

> Si $(A_i)_{i\in I}$ est une famille d'événements mutuellement indépendants alors, pour toute partie $J \subset I$, la sous-famille $(A_i)_{i\in J}$ est, elle aussi, constituée d'événements mutuellement indépendants.

dém. :
Immédiat par retour à la définition.
□

Proposition

> Soit $(A_i)_{i\in I}$ une famille d'événements et $(\varepsilon_i)_{i\in I}$ une famille de réels avec $\varepsilon_i = 0$ ou 1.
> On pose
> $$A_i^{\varepsilon_i} = \begin{cases} A_i & \text{si } \varepsilon_i = 0 \\ \overline{A_i} & \text{si } \varepsilon_i = 1 \end{cases}$$
> Si la famille $(A_i)_{i\in I}$ est constituée d'événements mutuellement indépendant alors la famille $(A_i^{\varepsilon_i})_{i\in I}$ aussi.

dém. :
Etape 1 : On montre

$$P(A_1 \cap \ldots \cap A_n) = P(A_1)\ldots P(A_n) \Rightarrow P(A_1 \cap \ldots \cap A_{n-1} \cap \overline{A_n}) = P(A_1)\ldots P(A_{n-1})P(\overline{A_n})$$

Etape 2 : On généralise

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n) \Rightarrow P(A_1^{\varepsilon_1} \cap \ldots \cap A_n^{\varepsilon_n}) = P(A_1^{\varepsilon_1}) \ldots P(A_n^{\varepsilon_n})$$

Etape 3 : On établit le résultat
Soit J finie $\subset I$. Par énumération de l'ensemble J

$$P\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} P(A_j)$$

puis par l'étude qui précède

$$P\left(\bigcap_{j\in J} A_j^{\varepsilon_j}\right) = \prod_{j\in J} P(A_j^{\varepsilon_j})$$

et l'on peut conclure que la famille $(A_i^{\varepsilon_i})_{i\in I}$ est constituée d'événements mutuellement indépendants.

□

Chapitre 27

Variables aléatoires discrètes

$(\Omega, \mathcal{A}, P)$ désigne un espace probabilisé.

27.1 Variables aléatoires discrètes

27.1.1 Définition

Définition

On appelle variable aléatoire discrète définie sur l'espace probabilisé Ω et à valeurs dans un ensemble E toute application $X : \Omega \to E$ vérifiant
1) l'ensemble des valeurs prises $X(\Omega)$ est fini ou dénombrable ;
2) $\forall x \in X(\Omega)$, $X^{-1}(\{x\}) = \{\omega \in \Omega / X(\omega) = x\}$ est élément de la tribu $\mathcal{A}$.
Lorsque $E = \mathbb{R}$, on parle de variable aléatoire réelle.

Remarque L'appellation variable aléatoire est usuelle bien que malheureuse. En effet, X n'est pas une variable, mais bien une fonction et celle-ci n'est pas aléatoire, mais plutôt parfaitement déterminée. Ce sont les valeurs de X qui correspondent à des quantités qui vont varier selon le résultat de l'expérience aléatoire.

Exemple On tire avec remise n boules dans une urne contenant des boules blanches et rouges en proportion p et $q = 1 - p$. On note X le nombre de boules blanches obtenues dans un tirage, X est une variable aléatoire discrète.

Exemple On lance indéfiniment un dé et l'on note X_n la valeur obtenue lors du n-ième lancer. $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires discrètes. On pose

$$T = \min\{n \in \mathbb{N}^\star / X_n = 6\} \text{ ou } T = +\infty \text{ si le min porte sur l'ensemble vide}$$

T est une variable aléatoire discrète (c'est le temps d'attente du premier 6).

Remarque Comme dans les exemples ci-dessus, il est fréquent de manipuler des variables aléatoires sans même avoir précisé l'espace probabilisé d'étude.

27.1.2 Evénements valeurs

Définition

Soit $X : \Omega \to E$ une variable aléatoire discrète.
Pour tout $x \in E$, on note $(X = x)$ ou $\{X = x\}$ l'événement

$$X^{-1}(\{x\}) = \{\omega \in \Omega / X(\omega) = x\}$$

Il s'agit bien d'un événement par définition d'une variable aléatoire discrète et l'on peut en calculer la probabilité

$$P(X = x)$$

Exemple On lance deux dés et X désigne la somme de leur valeur.
L'événement $(X = 12)$ correspond au cas où les deux dés valent 6.

Définition

Soit $X : \Omega \to E$ une variable aléatoire discrète. Pour toute partie A de E on note $(X \in A)$ ou $\{X \in A\}$ l'événement $X^{-1}(A)$. Autrement dit

$$(X \in A) = \{\omega \in \Omega / X(\omega) \in A\}$$

Remarque $(X \in A)$ est bien un événement. En effet, $X(\Omega)$ étant au plus dénombrable,

$$(X \in A) = \bigcup_{x \in X(\Omega) \cap A} (X = x)$$

est un événement car réunion au plus dénombrable d'événements.
Cela autorise le calcul de sa probabilité $P(X \in A)$

Remarque La notation $(X \in A)$ est compatible avec les opérations ensemblistes

$$(X \in A) \cap (X \in B) = (X \in A \cap B)$$

$$(X \in A) \cup (X \in B) = (X \in A \cup B)$$

$$\left(X \in \bar{A}\right) = \overline{(X \in A)}$$

Définition

Si X est une variable aléatoire discrète réelle et si $a \in \mathbb{R}$, on introduit l'événement

$$(X \leqslant a) = X^{-1}(]-\infty, a]) = \{\omega \in \Omega / X(\omega) \leqslant a\}$$

On peut aussi définir $(X < a)$, $(X \geqslant a)$,... et calculer leur probabilité.

27.1.3 Loi d'une variable aléatoire discrète

Soit $X : \Omega \to E$ une variable aléatoire discrète et $X(\Omega)$ son univers valeurs (au plus dénombrable).

Définition

On appelle loi de la variable $X : \Omega \to E$ l'application

$$P_X : \wp(X(\Omega)) \to [0,1]$$

définie par

$$\forall A \in \wp(X(\Omega)), P_X(A) = P(X \in A)$$

Théorème

La loi P_X définit une probabilité sur l'espace probabilisable $(X(\Omega), \wp(X(\Omega)))$

dém. :
$P_X(X(\Omega)) = P(X \in X(\Omega)) = 1$.
Soit $(A_n)_{n\in\mathbb{N}}$ une suite de parties deux à deux disjointes de $X(\Omega)$.
Les événements $(X \in A_n)$ sont deux à deux disjoints et

$$\bigcup_{n\in\mathbb{N}} (X \in A_n) = \left(X \in \bigcup_{n\in\mathbb{N}} A_n\right)$$

On en déduit

$$P_X\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} P(X \in A_n) = \sum_{n=0}^{+\infty} P_X(A_n)$$

□

Corollaire

La loi P_X est entièrement déterminée par les valeurs

$$p_x = P_X(x) = P(X = x)$$

pour chaque $x \in X(\Omega)$.

dém. :
L'espace $X(\Omega)$ étant au plus dénombrable, une probabilité sur celui-ci est entièrement déterminée par ses probabilités élémentaires

$$p_x = P(X = x)$$

En effet, pour toute partie A de $X(\Omega)$, on a alors

$$P_X(A) = P(X = A) = \sum_{x\in A} p_x$$

la somme portant sur une famille finie ou dénombrable.
□

Remarque Les probabilités élémentaires p_x déterminent une famille de réels positifs $(p_x)_{x\in X(\Omega)}$ vérifiant

$$\sum_{x\in X(\Omega)} p_x = 1$$

Remarque Souvent, on résume la loi de X à la famille des probabilités p_x pour $x \in X(\Omega)$ puisque celles-ci suffisent à déterminer $P_X(A)$ pour toute partie A de $X(\Omega)$.

Remarque La loi P_X détermine la probabilité de chaque événement valeur lié à la variable X. Cependant, la loi P_X ne suffit pas à déterminer la variable aléatoire X

$$P_X = P_Y \text{ n'implique pas } X = Y$$

Exemple Considérons un lancer de deux équilibrés. Si X et Y désignent les valeurs de chaque dé, celles-ci suivent la même loi sans pour autant être égales !

Définition

Soit X et Y deux variables aléatoires discrètes sur Ω prenant les mêmes valeurs.
Si $P_X = P_Y$, on dit que X et Y suivent la même loi et l'on note

$$X \sim Y$$

Si la variable Y suit une loi usuellement notée $\mathcal{L}$, on écrit

$$X \sim \mathcal{L}$$

27.1.4 Lois finies usuelles

X désigne une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$.

27.1.4.1 Loi uniforme

Définition

On dit que la variable aléatoire X suit une loi uniforme sur un ensemble fini E si

$$X(\Omega) = E \text{ et } \forall x \in E, P(X = x) = 1/n \text{ avec } n = \text{Card}E$$

On note $\mathcal{U}(\llbracket a, b \rrbracket)$ la loi uniforme sur $\llbracket a, b \rrbracket$ et en particulier $\mathcal{U}(n)$ celle sur $\llbracket 1, n \rrbracket$.

Exemple Si X est la valeur du lancer d'un dé équilibré alors $X \sim \mathcal{U}(6)$

27.1.4.2 Loi de Bernoulli

Définition

On dit que la variable aléatoire X suit une loi de Bernoulli de paramètre p (avec $p \in \,]0, 1[\,$) si

$$X(\Omega) = \{0, 1\}, P(X = 0) = 1 - p \text{ et } P(X = 1) = p$$

On note $\mathcal{B}(p)$ cette loi.

Exemple Une urne contient des boules blanches en proportion p et des boules rouges en proportion $q = 1 - p$.
On tire une boule de cette urne.
Si X vaut 1 lorsque la boule est blanche et 0 sinon alors $X \sim \mathcal{B}(p)$

Remarque Les variables de Bernoulli sont utiles pour modéliser les situations à deux issues : succès (valeur 1) ou échec (valeur 0)

27.1.4.3 Loi binomiale

Définition

On dit que la variable aléatoire X suit une loi binomiale de paramètres n et p (avec $n \in \mathbb{N}^\star$ et $p \in \left]0,1\right[$) si

$$X(\Omega) = [\![0,n]\!] \text{ et } \forall k \in [\![0,n]\!]\,, P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

On note $\mathcal{B}(n,p)$ cette loi

Exemple Une urne contient des boules blanches en proportion p et des boules rouges en proportion $q = 1 - p$.
On tire n boules avec remise dans cette urne.
Si X désigne le nombre de boules blanches obtenues alors $X \sim \mathcal{B}(n,p)$.

Remarque La loi de Bernoulli est utile pour modéliser ce qui s'apparente à un tirage avec remise, elle permet aussi de mesurer le nombre de succès lorsqu'on répète indépendamment une expérience dont la probabilité de réussite égale p.

27.1.5 Variables aléatoires composées

Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$ à valeurs dans un ensemble E.

Définition

Si f est une application définie au moins sur $X(\Omega) \subset E$ à valeurs dans un ensemble E', on note $f(X)$ la variable aléatoire $Y = f \circ X$

$$Y : \Omega \to E' \text{ avec } Y(\omega) = f\left(X(\omega)\right)$$

Remarque On vérifie qu'il s'agit bien d'une variable aléatoire car

$$\forall y \in Y(\Omega), Y^{-1}\left(\{y\}\right) = \bigcup_{x \in X(\Omega), f(x) = y} X^{-1}\left(\{x\}\right) \in \mathcal{A}$$

Remarque Si la fonction f est une fonction présentant une notation usuelle particulière, on adapte celle-ci à la description de la variable aléatoire $f(X)$. C'est ainsi qu'on pourra écrire

$$X^2, \sqrt{X}, \ |X|, aX+b, \ldots$$

Théorème

Si $Y = f(X)$ alors la loi de Y est entièrement déterminée par celle de X :

$$\forall B \in Y(\Omega), P_Y(B) = P_X\left(f^{-1}(B)\right)$$

dém. :
Par définition

$$P_Y(B) = P(Y \in B) = P(f(X) \in B)$$

Or

$$(f(X) \in B) = \left(X \in f^{-1}(B)\right)$$

□

Remarque En pratique, connaître la loi de X suffira pour déterminer les lois des variables aléatoires composées déduites de X.

Exemple Si $X \sim \mathcal{B}(n,p)$ alors $Y = n - X \sim \mathcal{B}(n,q)$.
En effet

$$Y(\Omega) = [\![0,n]\!] \text{ et } P(Y=k) = P(X=n-k) = \binom{n}{n-k} p^{n-k}q^k = \binom{n}{k} q^k p^{n-k}$$

Définition

Plus généralement, si $X_1, \ldots, X_m$ sont des variables aléatoires discrètes sur $(\Omega, \mathcal{A}, P)$, on peut donner un sens à la variable aléatoire discrète $Y = f(X_1, \ldots, X_m)$ pour peu que f soit définie sur les valeurs prises par $\omega \mapsto (X_1(\omega), \ldots, X_m(\omega))$.

Remarque Pour connaître la loi de Y, connaître les lois des X_k ne suffit pas, il faut aussi connaître leurs comportements conjoints...

27.2 Couples de variables aléatoires discrètes

27.2.1 Loi conjointe

Définition

Soit X et Y deux variables aléatoires discrètes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$ et à valeurs dans des ensembles E et F respectivement. On appelle couple défini par les variables aléatoires X et Y la variable aléatoire $Z = (X, Y) : \Omega \to E \times F$ déterminée par

$$\forall \omega \in \Omega, Z(\omega) = (X(\omega), Y(\omega))$$

Remarque Il s'agit d'une variable aléatoire discrète car $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$ est au plus dénombrable.

Exemple On choisit une carte à l'intérieur d'un jeu de 32 cartes. On désigne par X la hauteur et Y la couleur de cette carte. La variable aléatoire $Z = (X, Y)$ détermine alors parfaitement la carte tirée.

Définition

On appelle loi conjointe de deux variables aléatoires X et Y la loi du couple $Z = (X, Y)$.

Remarque Celle-ci est entièrement déterminée à partir de la connaissance de

$$P(X = x_i, Y = y_j) \text{ avec } x_i \in X(\Omega) \text{ et } y_j \in Y(\Omega)$$

On pourra exploiter un tableau pour visualiser cette loi conjointe.

Exemple Une urne comporte 2 boules blanches, 1 rouge et 1 noire. On tire simultanément deux boules de cette urne et l'on note X le nombre de boules blanches et Y le nombre de boules noires tirées.

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	0	$1/3$	$1/6$
$Y = 1$	$1/6$	$1/3$	0

Remarque Evidemment la somme des valeurs du tableau donne 1.

27.2.2 Lois marginales

Soit Z une variable aléatoire discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$ à valeurs dans un produit cartésien $E \times F$. Pour chaque $\omega \in \Omega$, $Z(\omega)$ désigne un couple élément de $E \times F$. Notons $X(\omega) \in E$ et $Y(\omega) \in F$ les deux éléments de ce couple. La variable Z se comprend alors comme le couple (X, Y).

Définition

Les lois des deux variables aléatoires X et Y sont appelées les lois marginales de la variable Z .

Proposition

La loi de Z détermine entièrement ses lois marginales.

dém. :
Pour $x \in X(\Omega)$,

$$(X = x) = Z \in \{x\} \times F$$

et donc

$$P_X(x) = P_Z(\{x\} \times F) = \sum_{y \in F \cap Y(\Omega)} P(Z = (x, y))$$

□

Remarque Dans un tableau visualisant la loi conjointe, les lois marginales s'obtiennent en sommant sur les rangées

Exemple On reprend l'urne urne comporte 2 boules blanches, 1 rouge et 1 noire. On tire simultanément deux boules de cette urne et l'on note X le nombre de boules blanches et Y le nombre de boules noires tirées.

	$X = 0$	$X = 1$	$X = 2$	P_Y
$Y = 0$	0	$1/3$	$1/6$	$1/2$
$Y = 1$	$1/6$	$1/3$	0	$1/2$
P_X	$1/6$	$2/3$	$1/6$	

Remarque En revanche, les lois marginales ne suffisent pas à déterminer la loi conjointe.
Par exemple, les deux tableaux ci-dessous correspondent à de mêmes lois marginales pour des lois conjointes différentes

$X\backslash Y$	$X = 0$	$X = 1$	P_Y
$Y = 0$	$1/2$	0	$1/2$
$Y = 1$	0	$1/2$	$1/2$
P_X	$1/2$	$1/2$	

et

$X\backslash Y$	$X = 0$	$X = 1$	P_Y
$Y = 0$	0	$1/2$	$1/2$
$Y = 1$	$1/2$	0	$1/2$
P_X	$1/2$	$1/2$	

27.2.3 Lois conditionnelles

Soit X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

Définition

Soit $x \in X(\Omega)$. On appelle loi conditionnelle de Y sachant $X = x$ la loi de la variable aléatoire Y pour la probabilité conditionnelle $P(. \mid X = x)$.
Autrement dit, pour toute partie $B \subset Y(\Omega)$

$$P(Y \in B \mid X = x) = \begin{cases} \dfrac{P(Y \in B, X = x)}{P(X = x)} & \text{si } P(X = x) > 0 \\ 0 & \text{sinon} \end{cases}$$

Remarque Cette loi est entièrement déterminée par la connaissance de

$$P(Y = y \mid X = x) \text{ pour tout } y \in Y(\Omega)$$

Exemple Supposons X et Y variables aléatoires de loi conjointe donnée par

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	0	$1/3$	$1/6$
$Y = 1$	$1/6$	$1/3$	0

La loi de Y sachant $X = x$ est alors

	$X = 0$	$X = 1$	$X = 2$
$P(Y = 0 \mid X = x)$	0	$1/2$	1
$P(Y = 1 \mid X = x)$	1	$1/2$	0

Théorème

La connaissance :
- de la loi de X ;
- de la loi de Y sachant $X = x$ pour chaque $x \in X(\Omega)$

détermine entièrement la loi conjointe de $Z = (X, Y)$.

dém. :
Soit $(x, y) \in Z(\Omega)$. On a $x \in X(\Omega)$ et $y \in Y(\Omega)$.
Si $P(X = x) = 0$ alors $P(Z = (x, y)) = 0$ car $\{Z = (x, y)\} \subset \{X = x\}$.
Si $P(X = x) > 0$ alors $P(Z = (x, y)) = P(X = x, Y = y) = P(Y = y \mid X = x)P(X = x)$.
□

Remarque En particulier la loi de Y est alors connue

$$P(Y = y) = \sum_{x \in X(\Omega)} P(Y = y \mid X = x)P(X = x)$$

Définition

Plus généralement, si A est une partie de $X(\Omega)$, on peut définir la loi de Y sachant $X \in A$

$$P(Y = y \mid X \in A) = \begin{cases} \dfrac{P(X \in A, Y = y)}{P(X \in A)} & \text{si } P(X \in A) > 0 \\ 0 & \text{sinon} \end{cases}$$

Exemple Si X est à valeurs réelles, on peut introduire la loi de Y sachant $(X > x)$.

27.2.4 Vecteurs aléatoires

Soit $X_1, \dots, X_n$ des variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$.

Définition

On appelle vecteur aléatoire discret défini à partir des variables aléatoires $X_1, \dots, X_n$ la variable aléatoire discrète Z donnée par

$$\forall \omega \in \Omega, Z(\omega) = (X_1(\omega), \dots, X_n(\omega))$$

La loi de la variable Z est appelée loi conjointe des variables $X_1, \dots, X_n$ tandis que les lois de $X_1, \dots, X_n$ sont les lois marginales de Z.

Remarque La loi conjointe détermine les lois marginales, mais l'inverse n'est pas vrai.

27.3 Indépendance de variables aléatoires

27.3.1 Couple de variables indépendantes

Soit X et Y deux variables aléatoires discrètes sur l'espace probabilisé (Ω, P).

Définition

On dit que les deux variables X et Y sont indépendantes si pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $(X \in A)$ et $(Y \in B)$ sont indépendants.

Exemple On lance deux dés discernables. X détermine la valeur du premier et Y celle du second.
Il est usuel de modéliser X et Y en tant que variables indépendantes.

Exemple Une première urne contient 2 boules blanches et 3 boules noires et une seconde l'inverse.
On jette une pièce et si l'on obtient « face », on pioche une boule dans la première urne, sinon, on pioche cette boule dans la seconde urne.
On note X la valeur du lancer de la pièce et Y la couleur de la boule tirée.
Les variables X et Y ne sont pas indépendantes !

Remarque Si X et Y sont indépendantes, la loi de Y sachant $X \in A$ se résume à la loi de Y.

Théorème

On a équivalence entre :
(i) les variables aléatoires X et Y sont indépendantes ;
(ii) $\forall (x,y) \in X(\Omega) \times Y(\Omega), P(X = x, Y = y) = P(X = x)P(Y = y)$

dém. :
(i) $\Rightarrow$ (ii) Supposons (i)
Soit $(x,y) \in X(\Omega) \times Y(\Omega)$. Les événements $(X = x)$ et $(Y = y)$ sont indépendants donc

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

(ii) $\Rightarrow$ (i) Supposons (ii)
Soit $A \subset X(\Omega)$ et $B \subset Y(\Omega)$. Par probabilités totales (avec $A \times B$ au plus dénombrable)

$$P(X \in A \cap Y \in B) = \sum_{(x,y) \in A \times B} P(X = x, Y = y)$$

donc

$$P(X \in A \cap Y \in B) = \sum_{(x,y) \in A \times B} P(X = x)P(Y = y)$$

En sommant par paquets

$$P(X \in A \cap Y \in B) = \sum_{x \in A} \sum_{y \in B} P(X = x)P(Y = y)$$

puis

$$P(X \in A \cap Y \in B) = \sum_{x \in A} P(X = x) \sum_{y \in B} P(Y = y) = P(X \in A)P(Y \in B)$$

□

Exemple Supposons X et Y variables aléatoires de loi conjointe donnée par

	$X = 0$	$X = 1$
$Y = 0$	$1/12$	$2/12$
$Y = 1$	$3/12$	$6/12$

La loi de Y sachant $X = x$ est alors

	$X = 0$	$X = 1$
$P(Y = 0 \mid X = x)$	$1/4$	$1/4$
$P(Y = 1 \mid X = x)$	$3/4$	$3/4$

Les variables X et Y sont indépendantes.

Théorème

Si X et Y sont deux variables indépendantes alors pour toutes fonctions f, g définies sur les domaines de valeurs de X et Y, les variables $f(X)$ et $g(Y)$ sont indépendantes.

dém. :
Soit $x' \in f(X(\Omega))$ et $y' \in g(Y(\Omega))$. On a

$$P(f(X) = x', g(Y) = y') = P\left(X \in f^{-1}(\{x'\}) \cap Y \in g^{-1}(\{y'\})\right)$$

Les variables X et Y étant indépendantes

$$P(f(X) = x', g(Y) = y') = P\left(X \in f^{-1}(\{x'\})\right) P\left(Y \in g^{-1}(\{y'\})\right)$$

ce qui donne

$$P(f(X) = x', g(Y) = y') = P(f(X) = x')\, P(g(Y) = y')$$

□

27.3.2 Famille finie de variables mutuellement indépendantes

Soit $(X_i)_{1 \leqslant i \leqslant n}$ une famille de n variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$.

Définition

On dit que celle-ci sont mutuellement indépendantes si pour toute famille $(A_i)_{1 \leqslant i \leqslant n}$ avec $A_i \subset X_i(\Omega)$ les événements $(X_i = A_i)$ sont mutuellement indépendants.

Théorème

On a équivalence entre :
(i) les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes ;
(ii) $\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega), P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \dots P(X = x_n)$

dém. :
Il suffit d'adapter la démonstration présentée pour les couples de variables sachant que pour (ii) $\Rightarrow$ (i) on étudiera l'indépendance en considérant les sous familles finies de la famille des événements $(X_i = A_i)$.
□

Remarque On répète n fois la même expérience aléatoire et l'on note $X_1, \dots, X_n$ les résultats successifs.
En supposant que le résultat d'une expérience est sans incidence sur les autres, il est usuel de modéliser l'expérience en supposant les variables $X_1, \dots, X_n$ mutuellement indépendantes.
C'est le cas lorsqu'on lance plusieurs fois une même pièce de monnaie que celle-ci soit ou non équilibrée.

Exemple On tire des boules dans une urne contenant des boules blanches et rouges.
On note X_i la couleur obtenue lors du i-ème tirage.
Si l'on suppose que le tirage a lieu avec remise, il est usuel de supposer les variables $X_1, \dots, X_n$ mutuellement indépendantes.
Si l'on ne suppose pas la remise, les variables X_i ne sont plus indépendantes !

Attention : L'indépendance mutuelle ne doit pas être confondues avec l'indépendance deux à deux.
Si on lance deux dés discernables que l'on note X et Y les parités de chaque dé et Z la parité de la somme alors les variables X, Y, Z sont deux à deux indépendantes, mais pas mutuellement indépendantes.

Théorème

Si les variables $X_1, \ldots, X_n$ sont mutuellement indépendantes alors pour tout m compris entre 1 et $n-1$ et toutes fonctions f et g définies sur des domaines convenables, les variables

$$X = f(X_1, \ldots, X_m) \text{ et } Y = g(X_{m+1}, \ldots, X_n)$$

sont indépendantes.

dém. :
Soit $x \in X(\Omega)$ et $y \in Y(\Omega)$.

$$P(X = x \cap Y = y) = \sum_{(x_1,\ldots,x_n)\in f^{-1}(\{x\}),(x_{m+1},\ldots,x_n)\in f^{-1}(\{y\})} P\left(X_1 = x_1, \ldots, X_n = x_n\right)$$

Par indépendance

$$P\left(X_1 = x_1, \ldots, X_n = x_n\right) = P\left(X_1 = x_1\right) \ldots P\left(X_n = x_n\right)$$

puis

$$P\left(X_1 = x_1, \ldots, X_n = x_n\right) = P\left(X_1 = x_1, \ldots, X_m = x_m\right) P\left(X_{m+1} = x_{m+1}, \ldots, X_n = x_n\right)$$

En réorganisant la somme par paquets

$$P(X = x \cap Y = y) = \sum_{(x_1,\ldots,x_n)\in f^{-1}(\{x\})} P\left(X_1 = x_1, \ldots, X_m = x_m\right) \sum_{(x_{m+1},\ldots,x_n)\in f^{-1}(\{y\})} P\left(X_{m+1} = x_{m+1}, \ldots, X_n = x_n\right)$$

et finalement

$$P(X = x \cap Y = y) = P(X = x)P(Y = y)$$

□

Remarque Si Z est indépendant de X et de Y, il se peut que Z ne soit pas indépendant de $X + Y$. C'est le cas lors d'un lancer de dés où X et Y teste la parité de chaque dé et Z la parité de la somme. Dans l'énoncé qui précède, l'hypothèse d'indépendance mutuelle est donc essentielle.

27.3.3 Famille infinie de variables mutuellement indépendantes

Soit $(X_i)_{i\in I}$ une famille infinie de variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, P)$.

Définition

On dit que les variables aléatoires de la famille $(X_i)_{i\in I}$ sont mutuellement indépendantes si toutes ses sous-familles finies sont mutuellement indépendantes.

Exemple On lance indéfiniment une pièce de monnaie et l'on note X_n la variable de Bernoulli égale à 1 lorsqu'on obtient face au n-ième lancer.
Il est usuel de modéliser l'expérience en supposant la famille $(X_n)_{n\geqslant 1}$ constituée de variables aléatoires mutuellement indépendantes.

27.3.4 Suites infinies d'épreuves

Afin d'assurer l'existence de cadre probabiliste permettant l'étude de la répétition indépendante et infinie d'une même expérience, nous admettons le résultat (difficile) suivant

Théorème

Soit $\mathcal{L}$ la loi d'une certaine variable aléatoire discrète.
Il existe un espace probabilisé $(\Omega, \mathcal{A}, P)$ sur lequel existe une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires mutuellement indépendantes et qui sont toutes de loi $\mathcal{L}$.

Exemple Il existe un cadre probabiliste permettant de modéliser un jeu de « pile ou face »infinie où
- chaque X_n suit une même loi de Bernoulli de paramètre p ;
- la famille $(X_n)_{n\geqslant 1}$ est constituée de variables mutuellement indépendantes.

27.4 Espérance

Les variables aléatoires introduites seront toutes supposées réelles, discrètes et définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

27.4.1 Définition

Définition

On dit que la variable X admet une espérance si la famille $(xP(X = x))_{x\in X(\Omega)}$ est sommable.
Sa somme définit alors l'espérance de X

$$E(X) \underset{\text{déf}}{=} \sum_{x\in X(\Omega)} xP(X = x)$$

Celle-ci ne dépend que la loi de la variable X.

Remarque Si la variable X ne prend qu'un nombre fini de valeurs $x_1, \ldots, x_n$ alors celle-ci est assurément d'espérance finie et

$$E(X) = \sum_{k=1}^{n} x_k P(X = x_k)$$

Exemple Rappelons :
Si $X \sim \mathcal{B}(p)$ alors

$$E(X) = 0 \times (1 - p) + 1 \times p = p$$

Si $X \sim \mathcal{B}(n, p)$ alors

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1 - p)^{n-k} = np$$

Exemple Si $A \in \mathcal{A}$ alors

$$E(1_A) = 0 \times P(\bar{A}) + 1 \times P(A) = P(A)$$

Remarque Si la variable X prend une infinité (nécessairement dénombrable) de valeurs alors, en introduisant $(x_n)_{n\in\mathbb{N}}$ une énumération de celles-ci, la variable X admet une espérance si, et seulement si, il y a convergence absolue de la série $\sum x_n P(X = x_n)$. On a alors

$$E(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n)$$

La valeur obtenue ne dépend pas de l'énumération choisie.

Remarque Si la variable X ne prend que des valeurs positives

$$\forall \omega \in \Omega, X(\omega) \in \mathbb{R}^+$$

on peut encore définir son espérance dans $\mathbb{R}^+ \cup \{+\infty\}$ par la relation

$$E(X) \underset{\text{déf}}{=} \sum_{x\in X(\Omega)} xP(X = x)$$

Exemple Soit X une variable aléatoire avec

$$X(\Omega) = \mathbb{N} \text{ et } P(X = n) = \frac{1}{2^{n+1}}$$

La variable X est à valeurs positives et

$$E(X) = \sum_{n=0}^{+\infty} \frac{n}{2^{n+1}}$$

Pour calculer cette somme, exploitons la série entière

$$\sum_{n=1}^{+\infty} nx^{n-1} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=0}^{+\infty} x^n\right) = \frac{1}{(1-x)^2}$$

donnant

$$\sum_{n=1}^{+\infty} n\frac{1}{2^{n-1}} = 4$$

puis

$$E(X) = \sum_{n=0}^{+\infty} \frac{n}{2^{n+1}} = 1$$

Exemple Soit X une variable aléatoire avec

$$X(\Omega) = \mathbb{N}^\star \text{ et } P(X = n) = \frac{1}{n} - \frac{1}{n+1}$$

La variable X est à valeurs positives et

$$E(X) = \sum_{n=1}^{+\infty} n\left(\frac{1}{n} - \frac{1}{n+1}\right) = \sum_{n=1}^{+\infty} \frac{1}{n+1} = +\infty$$

Exemple Si la variable X est constante égale à C alors

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x) = C \times P(X = C) = C$$

Le résultat est encore vraie si l'égalité $X = C$ est presque sûre (i.e. $P(X = C) = 1$)

Définition

Si la variable X admet une espérance et si celle-ci est nulle, on dit que la variable X est centrée.

27.4.2 Propriétés

Théorème

Si les variables X et Y admettent des espérances alors pour tout $\lambda \in \mathbb{R}$ les variables λX et $X + Y$ admettent une espérance et

$$E(\lambda X) = \lambda E(X) \text{ et } E(X + Y) = E(X) + E(Y)$$

dém. :
Etude de $E(\lambda X) = \lambda E(X)$
Le cas $\lambda = 0$ est immédiat. Pour $\lambda \neq 0$,

$$\lambda E(X) = \lambda \sum_{x \in X(\Omega)} xP(X = x) = \sum_{x \in X(\Omega)} \lambda x P(\lambda X = \lambda x)$$

puis, sachant que x parcourt $X(\Omega)$ si, et seulement si, λx parcourt $(\lambda X)(\Omega)$,

$$\lambda E(X) = \sum_{y \in (\lambda X)(\Omega)} yP(\lambda X = y) = E(\lambda X)$$

Etude de $E(X + Y) = E(X) + E(Y)$
Par la formule des probabilités totales

$$P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y)$$

En sommant par paquets

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xP(X = x, Y = y)$$

De même

$$E(Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} yP(X = x, Y = y)$$

et donc, avec sommabilité

$$E(X) + E(Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (x + y)P(X = x, Y = y)$$

En sommant par paquets selon la valeur de $z = x + y$

$$E(X) + E(Y) = \sum_{z \in (X+Y)(\Omega)} z \sum_{(x,y) \in X(\Omega) \times Y(\Omega), x+y=z} P(X = x, Y = y)$$

soit

$$E(X) + E(Y) = \sum_{z \in (X+Y)(\Omega)} zP(X + Y = z)$$

□

Corollaire

L'ensemble des variables aléatoires réelles discrètes définies sur $(\Omega, \mathcal{A}, P)$ admettant une espérance est un espace vectoriel et l'espérance y définit une forme linéaire.

dém. :
C'est un sous-espace vectoriel de l'espace des variables aléatoires.
□

Exemple Si a et b sont deux réels

$$E(aX + b) = aE(X) + b$$

Exemple Si X admet une espérance alors la variable $Y = X - E(X)$ est centrée.

Théorème

Si X est à valeurs positives alors $E(X) \geqslant 0$.
Si de plus $E(X) = 0$ alors $X = 0$ presque sûrement.

dém. :
Si X est à valeur positives

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x) \in \mathbb{R}^+ \cup \{+\infty\}$$

car somme d'une famille de réels tous positifs.
Si de plus $E(X) = 0$ alors

$$\forall x \in X(\Omega), xP(X = x) = 0$$

et donc

$$\forall x \in X(\Omega) \setminus \{0\}, P(X = x) = 0$$

On en déduit $P(X = 0) = 1$.
□

Corollaire

Si X et Y admettent une espérance et si $X \leqslant Y$ alors

$$E(X) \leqslant E(Y)$$

dém. :
$Z = Y - X$ est une variable positive.
□

Théorème

Si $|X| \leqslant Y$ et si Y admet une espérance alors X aussi.

dém. :
Par probabilités totales

$$E(|X|) = \sum_{x \in |X|(\Omega)} xP(|X| = x) = \sum_{x \in |X|(\Omega)} \sum_{y \in Y(\Omega)} xP(|X| = x, Y = y)$$

Or $|X| \leqslant Y$ donc

$$xP(|X| = x, Y = y) \leqslant yP(|X| = x, Y = y)$$

En effet, si le terme de probabilité est nul, l'inégalité est vraie, sinon il existe $\omega \in \Omega$ tel que $|X(\omega)| = x$ et $Y(\omega) = y$ donc $x \leqslant y$ et l'inégalité est encore vraie.
En réordonnant la somme

$$E(|X|) \leqslant \sum_{y \in Y(\Omega)} \sum_{x \in |X|(\Omega)} yP(|X| = x, Y = y)$$

et par probabilité totales

$$E(|X|) \leqslant \sum_{y \in Y(\Omega)} yP(Y = y) = E(Y) < +\infty$$

□

Exemple Si la variable aléatoire X est bornée, elle admet assurément une espérance.

27.4.3 Formule de transfert

Théorème

Soit X une variable et f une fonction définie au moins sur $X(\Omega)$ et à valeurs dans $\mathbb{R}$.
On a équivalence entre :
(i) la variable $f(X)$ admet une espérance ;
(ii) la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable.
De plus, si tel est le cas

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x)$$

dém. :
$E(f(X)) = \sum_{y \in f(X)(\Omega)} yP(f(X) = y)$
Par probabilités totales

$$E(f(X)) = \sum_{y \in f(X)(\Omega)} \sum_{x \in X(\Omega)} yP(f(X) = y, X = x)$$

Or

$$yP(f(X) = y, X = x) = f(x)P(f(X) = y, X = x)$$

car l'égalité est vraie quand la probabilité est nulle, mais aussi quand elle est non nulle car il existe un événement ω vérifiant $f(X(\omega)) = y$ et $X(\omega) = x$ donc $f(x) = y$.
En réordonnant les sommes

$$E\left(f(X)\right) = \sum_{x\in X(\Omega)} \sum_{y\in f(X)(\Omega)} f(x)P\left(f(X) = y, X = x\right)$$

Par probabilités totales

$$E\left(f(X)\right) = \sum_{x\in X(\Omega)} f(x)P(X = x)$$

□

Exemple Sous réserve de sommabilité

$$E\left(X^k\right) = \sum_{x\in X(\Omega)} x^k P(X = x), E\left(\mathrm{e}^X\right) = \sum_{x\in X(\Omega)} \mathrm{e}^x P(X = x),\ldots$$

27.4.4 Inégalité de Markov

Théorème

Soit X une variable à valeurs positives admettant une espérance.
Pour tout $a \geqslant 0$, on a

$$aP\left(X \geqslant a\right) \leqslant E(X)$$

dém. :
Par définition

$$E(X) = \sum_{x\in X(\Omega)} xP(X = x)$$

On sépare la somme en deux

$$E(X) = \sum_{x\in X(\Omega), x<a} xP(X = x) + \sum_{x\in X(\Omega), x\geqslant a} xP(X = x)$$

D'une part

$$\sum_{x\in X(\Omega), x<a} xP(X = x) \geqslant 0$$

car la variable aléatoire est à valeurs positives.
D'autre part

$$\sum_{x\in X(\Omega), x\geqslant a} xP(X = x) \geqslant \sum_{x\in X(\Omega), x\geqslant a} aP(X = x) = aP(X \geqslant a)$$

□

Exemple L'inégalité de Markov possède de nombreuses déclinaisons

$$P\left(|X| \geqslant \varepsilon\right) \leqslant \frac{E\left(|X|\right)}{\varepsilon}$$

$$P(|X - E(X)| \geqslant \varepsilon) \leqslant \frac{E(|X - E(X)|)}{\varepsilon}$$

et

$$P(|X| \geqslant \varepsilon) \leqslant \frac{E\left(X^2\right)}{\varepsilon^2}$$

27.4.5 Variables indépendantes

Théorème

Si les variables X et Y sont indépendantes et admettent une espérance alors XY admet une espérance et

$$E(XY) = E(X)E(Y)$$

dém. :
Avec sommabilité

$$E(X)E(Y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x)P(Y = y)$$

Par indépendance

$$E(X)E(Y) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X = x, Y = y)$$

En regroupant par paquets selon la valeur de $z = xy$

$$E(X)E(Y) = \sum_{z\in(XY)(\Omega)} z \sum_{(x,y)\in X(\Omega)\times Y(\Omega), z=xy} P(X = x, Y = y)$$

puis

$$E(X)E(Y) = \sum_{z\in(XY)(\Omega)} zP(XY = z) = E(XY)$$

□

Remarque La réciproque est fausse : on peut avoir $E(XY) = E(X)E(Y)$ sans pour autant indépendance de X et Y.

Corollaire

Si $f(X)$ et $g(Y)$ admettent des espérances avec X et Y variables indépendantes alors

$$E(f(X)g(Y)) = E(f(X))\,E(g(Y))$$

27.5 Variance d'une variable aléatoire

Les variables aléatoires introduites seront toutes supposées réelles, discrètes et définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.

27.5.1 Moments

Définition

On dit que la variable X admet un moment d'ordre $k \in \mathbb{N}$ si la variable X^k admet une espérance. Celle-ci est alors appelée moment d'ordre k de X et on note

$$m_k = E(X^k) = \sum_{x \in X(\Omega)} x^k P(X = x)$$

Exemple X admet assurément un moment d'ordre 0 et

$$m_0 = 1$$

X admet un moment d'ordre 1 si, et seulement si, X admet une espérance et alors

$$m_1 = E(X)$$

27.5.2 Espace des variables possédant un moments d'ordre 2

Théorème

Si la variable X admet un moment d'ordre 2 alors X admet une espérance.

dém. :
Pour tout $x \in \mathbb{R}$, on a

$$2|x| \leqslant 1 + x^2$$

donc

$$|X| \leqslant \frac{1}{2}\left(1 + X^2\right)$$

Puisque les variables 1 et X^2 admettent une espérance, la variable X aussi.
□

Remarque Ce résultat se généralise : si X admet un moment d'ordre n, X admet un moment d'ordre k pour tout $k \leqslant n$.

Théorème

Si les variables X et Y admettent chacune un moment d'ordre 2 alors XY est d'espérance finie et

$$E(XY)^2 \leqslant E(X^2)E(Y^2)$$

dém. :
Pour tout $x, y \in \mathbb{R}$, on a

$$2|xy| \leqslant x^2 + y^2$$

donc

$$|XY| \leqslant \frac{1}{2}\left(X^2 + Y^2\right)$$

Puisque les variables X^2 et Y^2 admettent une espérance, la variable XY aussi.
Soit $\lambda \in \mathbb{R}$. Introduisons la variable $Z = (\lambda X + Y)^2 = \lambda^2 X^2 + 2\lambda XY + Y^2$. Par combinaison linéaire, Z admet une espérance et puisque Z est positive

$$\lambda^2 E(X^2) + 2\lambda E(XY) + E(Y^2) \geqslant 0$$

Cette identité vaut pour tout $\lambda \in \mathbb{R}$.
Cas $E(X^2) \neq 0$: le trinôme associé au premier membre ne peut posséder deux racines réelles et donc

$$\Delta = 4E\left(XY\right)^2 - 4E\left(X^2\right)E\left(Y^2\right) \leqslant 0$$

Cas $E(X^2) = 0$: on a nécessairement $E(XY) = 0$ car sinon la constance de signe est impossible.
□

Théorème

L'ensemble des variables admettant un moment d'ordre 2 est un sous-espace vectoriel de l'espace des variables admettant un moment d'ordre 1.

dém. :
L'inclusion a déjà été vue.
Si X et Y admettent des moments d'ordre 2 alors $Z = \lambda X + \mu Y$ aussi car

$$Z^2 = \lambda^2 X^2 + 2\lambda\mu XY + \mu^2 Y^2$$

admet une espérance par combinaison linéaire.
□

27.5.3 Variance et écart-type

Définition

Si X admet un moment d'ordre 2, on appelle variance de la variable X le réel

$$V(X) = E\left(\left(X - E(X)\right)^2\right)$$

On introduit aussi son écart type

$$\sigma(X) = \sqrt{V(X)}$$

Remarque La variance un bien définie car X et la constante $E(X)$ admettent des moments d'ordre 2.

Remarque Variance et écart-type permettent de mesurer la dispersion de la variable X autour de sa moyenne.
Si la variable X se comprend avec une unité (des mètres, des années, des points,...) espérance et écart-type s'exprime avec la même unité.

Théorème

Si X admet un moment d'ordre 2 alors

$$V(X) = E\left(X^2\right) - E\left(X\right)^2$$

dém. :
En développant

$$(X - E(X))^2 = X^2 - 2E(X)X + E(X)^2$$

et par linéarité de l'espérance

$$V(X) = E\left(X^2\right) - 2E\left(X\right)^2 + E\left(X\right)^2 = E\left(X^2\right) - E\left(X\right)^2$$

□

Exemple Si $X \sim \mathcal{B}(p)$ alors $V(X) = p(1-p)$.

Exemple Si $X \sim \mathcal{B}(n, p)$ alors $V(X) = np(1-p)$.

Théorème

Si X est admet un moment d'ordre 2 alors pour tout $a, b \in \mathbb{R}$,

$$V(aX + b) = a^2 V(X)$$

dém. :

$$V(aX+b) = E\left(a^2X^2 + 2abX + b^2\right) - (aE(X)+b)^2 = a^2\left(E\left(X^2\right) - E(X)^2\right) = a^2V(X)$$

□

Remarque Il est naturel que la translation de b ne modifie pas la valeur de la variance car, si cette translation modifie la moyenne, elle ne modifie pas la dispersion de la variable autour de celle-ci.

Définition

Lorsqu'une variable aléatoire est de variance égale à 1, on la qualifie de réduite.

Exemple Si X est une variable admettant un moment d'ordre 2 alors en introduisant son espérance m et son écart type σ (supposé non nul), la variable

$$Y = \frac{X - m}{\sigma}$$

est centrée réduite.

27.5.4 Covariance

Définition

Si les variables X et Y admettent des moments d'ordre 2, on introduit leur covariance

$$\text{Cov}(X,Y) = E\left((X - E(X))\left(Y - E(Y)\right)\right)$$

Exemple $\text{Cov}(X,X) = V(X)$.

Proposition

La covariance définit une application bilinéaire symétrique sur l'espace des variables admettant un moment d'ordre 2.

dém. :
La symétrie est évidente.
De plus, on peut simplifier

$$\text{Cov}(X,Y) = E\left((X - E(X)Y\right)$$

et la linéarité de $Y \mapsto \text{Cov}(X,Y)$ est alors évidente.
□

Théorème

Si X et Y sont deux variables aléatoires réelles sur l'espace probabilisé (Ω, P) alors

$$\text{Cov}(X,Y) = E\left(XY\right) - E(X)E(Y)$$

dém. :
En développant

$$(X - E(X))\left(Y - E(Y)\right) = XY - E(X)Y - E(Y)X + E(X)E(Y)$$

puis par linéarité de l'espérance

$$\text{Cov}(X,Y) = E\left(XY\right) - 2E(X)E(Y) + E(X)E(Y) = E\left(XY\right) - E(X)E(Y)$$

□

Corollaire

Si les variables X et Y sont indépendantes

$$\text{Cov}(X,Y) = 0$$

La réciproque est fausse.

Remarque Par l'inégalité de Cauchy-Schwarz, on a

$$|\text{Cov}(XY)| \leqslant V(X)V(Y)$$

Si $V(X) > 0$ et $V(Y) > 0$ on peut introduire

$$\text{cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}} \in [-1, 1]$$

appelé coefficient de corrélation de X et Y.
Si les variables X et Y sont indépendantes, ce coefficient est nul.
Si les variables X et Y ont des « comportements analogues », ce coefficient est proche de 1.
Si les variables X et Y ont des « comportements opposés », ce coefficient est proche de -1.

27.5.5 Variance d'une somme

Proposition

Si X et Y admettent un moment d'ordre 2 alors

$$V(X + Y) = V(X) + 2\text{Cov}(X, Y) + V(Y)$$

dém. :
Par la formule de Huygens

$$V(X + Y) = E\left((X + Y)^2\right) - (E(X + Y))^2$$

En développant et par linéarité de l'espérance

$$V(X + Y) = E\left(X^2\right) + 2E\left(XY\right) + E\left(Y^2\right) - E\left(X\right)^2 - 2E(X)E(Y) - E\left(Y\right)^2$$

puis immédiatement

$$V(X + Y) = V(X) + 2\text{Cov}(X, Y) + V(Y)$$

□

Théorème

Si les variables $X_1, \ldots, X_n$ admettent des moments d'ordre 2 alors

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i<j} \text{Cov}(X_i, X_j)$$

dém. :
On a

$$V\left(\sum_{i=1}^{n} X_i\right) = \text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i\right)$$

Par bilinéarité

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \text{Cov}(X_i, X_j)$$

On obtient l'identité voulue en réorganisant via

$$\text{Cov}(X_i, X_i) = V(X_i) \text{ et } \text{Cov}(X_j, X_i) = \text{Cov}(X_i, X_j)$$

□

Corollaire

Si les variables X_1,..., X_n sont deux à deux indépendantes

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V\left(X_i\right)$$

Exemple On peut exploiter ce résultat pour retrouver la variance d'une variable X suivant une loi binomiale de paramètres n et p.
En effet, celle-ci peut être simulée par la somme de $X_1 + \cdots + X_n$ de n variables mutuellement indépendantes suivant une loi de Bernoulli de paramètre p et alors

$$V(X) = V(X_1) + \cdots + V(X_n) = np(1-p)$$

27.5.6 Inégalité de Bienaymé-Tchebychev

Théorème

Si la variable X admet un moment d'ordre 2 alors pour tout $\varepsilon > 0$,

$$P\left(|X - E(X)| \geqslant \varepsilon\right) \leqslant \frac{V(X)}{\varepsilon^2}$$

dém. :
On a

$$\left(|X - E(X)| \geqslant \varepsilon\right) = \left((X - E(X))^2 \geqslant \varepsilon^2\right)$$

et par l'inégalité de Markov appliquée à la variable positive $Y = (X - E(X))^2$

$$V(X) = E(Y) \geqslant \varepsilon^2 P\left(Y \geqslant \varepsilon^2\right) = \varepsilon^2 P\left(|X - E(X)| \geqslant \varepsilon\right)$$

□

Remarque Cette inégalité permet de mesurer dans quelle mesure l'expérimentation peut s'écarter de sa moyenne.

27.5.7 Loi faible des grands nombres

Théorème

Soit $(X_n)_{n\geqslant 0}$ une suite de variables aléatoires deux à deux indépendantes et suivant une même loi.
Si celles-ci admettent un moment d'ordre 2 alors en introduisant m leur espérance commune et

$$S_n = \sum_{k=1}^{n} X_k$$

on a

$$P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0$$

dém. :
Introduisons σ la variance commune aux variables X_n. On a

$$E(S_n) = nm \text{ et } V(S_n) = n\sigma^2$$

Par l'inégalité de Bienaymé-Tchebychev

$$P\left(|S_n - nm| \geqslant a\right) \leqslant \frac{V(S_n)}{a^2}$$

En prenant $a = n\varepsilon$

$$P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = P\left(|S_n - nm| \geqslant a\right) \leqslant \frac{V(S_n)}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

□

Exemple On veut estimer l'équilibre d'une pièce. On note p la probabilité (inconnue) que la pièce donne « face »lors d'un lancer.
On lance n fois la pièce et l'on pose S égale au nombre de lancers ayant donné « face ».
En posant X_k la variable de Bernoulli testant si le k-ième lancer donne « face », on

$$S = \sum_{k=1}^{n} X_k$$

Sachant $E(X_k) = p$ et $V(X_k) = p(1-p) \leqslant 1/4$, l'inégalité de Bienaymé-Tchebychev donne

$$P\left(|S/n - p| \geqslant \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}$$

Pour $\varepsilon = 0,01$, on obtient que S/n est une valeur approchée de p à ε près avec une probabilité supérieure à 5 % sous réserve de prendre $n \geqslant 50\,000$!

27.6 Variables aléatoires à valeurs naturelles

27.6.1 Loi de Poisson

Définition

On dit qu'une variable aléatoire X suit une loi de Poisson de paramètre λ (avec $\lambda > 0$) si

$$X(\Omega) = \mathbb{N} \text{ et } P(X = k) = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}$$

On note $\mathcal{P}(\lambda)$ cette loi.

Remarque On vérifie

$$\sum_{k=0}^{+\infty} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = 1 \text{ avec } \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \geqslant 0$$

Il est donc possible qu'une telle loi existe...

Théorème

Si $X \sim \mathcal{P}(\lambda)$ alors

$$E(X) = \lambda \text{ et } V(X) = \lambda$$

dém. :

$$E(X) = \sum_{n=0}^{+\infty} k\mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=1}^{+\infty} \mathrm{e}^{-\lambda} \frac{\lambda^k}{(k-1)!} = \lambda \mathrm{e}^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = \lambda$$

et

$$V(X) = E(X^2) - E(X)^2 = E(X(X-1)) + E(X) - E(X)^2$$

avec

$$E(X(X-1)) = \sum_{k=0}^{+\infty} k(k-1)\mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = \lambda^2 \mathrm{e}^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = \lambda^2$$

et donc

$$V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

□

Exemple Si durant un laps de temps T un phénomène se produit en moyenne λ fois, il est fréquent de dire que le nombre d'occurrences de ce phénomène durant ce laps de temps suit une loi de Poisson de paramètre λ.
Par exemple, le nombre de désintégrations radioactives par seconde, le nombre de passages journalier le long d'une route, le nombre d'accidents annuel, etc.
Cette interprétation s'explique par un passage à la limite de la loi binomiale dans le cadre des événements rares.

Théorème

Soit $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires avec $X_n \sim \mathcal{B}(n, p_n)$.
Si $np_n \xrightarrow[n\to+\infty]{} \lambda$ alors

$$P(X_n = k) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

dém. :
Par définition d'une loi binomiale

$$P(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}$$

Or

$$\binom{n}{k} \underset{n\to+\infty}{\sim} \frac{n^k}{k!} \text{ et } n^k p_n^k (1-p_n)^{n-k} = \left(\frac{np_n}{1-p_n}\right)^k \exp\left(n\ln(1-p_n)\right)$$

avec

$$\frac{np_n}{1-p_n} \xrightarrow[n\to+\infty]{} \lambda \text{ et } (1-p_n)^n = \mathrm{e}^{n\ln(1-p_n)} = \mathrm{e}^{-np_n+o(1)} \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\lambda}$$

donc

$$P(X_n = k) \xrightarrow[n\to+\infty]{} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

□

Exemple Dans une certaine quantité de matière, il y a une grande quantité n d'atomes radioactifs. Chacun à une probabilité p très faible de se désintégrer mais l'on sait qu'en moyenne il y a λ désintégration durant un laps de temps T : $np = \lambda$. En supposant l'indépendance des désintégrations atomiques, il serait rigoureux de modéliser le nombre X de désintégration par une loi de Bernoulli $\mathcal{B}(n, p)$.
En pratique, les calculs numériques seraient difficiles alors que l'approximation avec une loi de Poisson de paramètre λ est bien plus commode.

Exemple Soit X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ et $\mu > 0$. Etudions la loi de la variable $Z = X + Y$.$X + Y$ est à valeurs dans $\mathbb{N}$ et

$$P(X+Y=k) = \sum_{\ell=0}^{k} P(X=\ell, Y=k-\ell)$$

Par indépendance

$$P(X+Y=k) = \sum_{\ell=0}^{k} P(X=\ell)P(Y=k-\ell)$$

puis

$$P(X+Y=k) = \sum_{\ell=0}^{k} \mathrm{e}^{-\lambda}\frac{\lambda^\ell}{\ell!}\mathrm{e}^{-\mu}\frac{\mu^{k-\ell}}{(k-\ell)!}$$

On réorganise

$$P(X+Y=k)=\frac{\mathrm{e}^{-(\lambda+\mu)}}{k!}\sum_{\ell=0}^{k}\binom{k}{\ell}\lambda^{\ell}\mu^{k-\ell}$$

Par la formule du binôme

$$P(X+Y=k)=\mathrm{e}^{-(\lambda+\mu)}\frac{(\lambda+\mu)^k}{k!}$$

La variable $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

27.6.2 Loi géométrique

Définition

On dit que la variable aléatoire X suit une loi géométrique de paramètre p (avec $p\in\,]0,1[$) si

$$X(\Omega)=\mathbb{N}^\star \text{ et } P(X=k)=p(1-p)^{k-1}$$

On note $\mathcal{G}(p)$ cette loi.

Remarque On vérifie

$$\sum_{k=1}^{+\infty}p(1-p)^{k-1}=1 \text{ avec } p(1-p)^{k-1}\geqslant 0$$

Il est donc possible qu'une telle loi existe.

Exemple On lance successivement un dé équilibré jusqu'à obtention d'un six.
On pose X le nombre de lancers nécessaires. On a

$$P\left(X=n\right)=\left(\frac{5}{6}\right)^{n-1}\left(\frac{1}{6}\right)$$

et donc $X\sim\mathcal{G}(p)$ avec $p=1/6$.

Remarque Plus généralement, la loi géométrique est utile pour évaluer le temps d'attente du premier succès dans une suite d'épreuves de Bernoulli mutuellement indépendantes de même paramètre p.

Théorème

Si $X\sim\mathcal{G}(p)$ alors

$$E(X)=\frac{1}{p} \text{ et } V(X)=\frac{1-p}{p^2}$$

dém. :

$$E(X)=\sum_{k=1}^{+\infty}k(1-p)^{k-1}p=\frac{1}{p}$$

et

$$V(X) = E\left(X(X-1)\right) + E(X) - E(X)^2$$

avec

$$E\left(X(X-1)\right) = \sum_{k=2}^{+\infty} k(k-1)(1-p)^{k-1}p$$

Or

$$\sum_{k=2}^{+\infty} k(k-1)x^{k-2} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{1}{1-x}\right) = \frac{2}{(1-x)^3}$$

donc

$$E\left(X(X-1)\right) = p(1-p)\frac{2}{p^3}$$

puis

$$V(X) = 2\frac{1-p}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

□

Théorème

Si X est une variable aléatoire à valeurs dans $\mathbb{N}^\star$ vérifiant la condition d'absence de mémoire

$$\forall n, k \in \mathbb{N}, P\left(X > n+k \mid X > n\right) = P\left(X > k\right)$$

alors X suit une loi géométrique.

dém. :
Posons $q = P(X > 1)$. La condition imposée donne

$$P(X > n+1 \mid X > n) = q$$

Or

$$P(X > n+1) = P(X > n+1 \mid X > n)P(X > n)$$

et donc

$$P(X > n+1) = qP(X > n)$$

Par une récurrence immédiate et sachant $P(X > 0) = 1$, on obtient

$$\forall n \in \mathbb{N}, P(X > n) = q^n$$

puis

$$\forall n \in \mathbb{N}^\star, P(X = n) = P(X > n-1) - P(X > n) = (1-q)q^{n-1}$$

Ainsi, la variable X suit une loi géométrique de paramètre $p = 1 - q$.
□

27.6.3 Fonctions génératrices

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$.

Définition

On appelle fonction génératrice de la variable X la série entière

$$\sum P(X=n)t^n$$

On note $G_X(t)$ sa somme là où elle est définie

$$G_X(t)=\sum_{n=0}^{+\infty} P(X=n)t^n = E\left(t^X\right)$$

Théorème

Cette série entière est de rayon de convergence R_X au moins égale à 1 et converge normalement sur $[-1,1]$.

dém. :
Pour $t=1$, la série numérique $\sum P(X=n)1^n=\sum P(X=n)$ converge.
Puisque la série entière converge en $t=1$, son rayon de convergence est au moins égale à 1.
Posons $u_n(t)=P(X=n)t^n$ définie sur $[-1,1]$.
Pour tout $t\in[-1,1]$, $|u_n(t)|\leqslant P(X=n)$.
C'est une majoration uniforme et $\sum P(X=n)$ converge donc la série de fonctions $\sum u_n$ converge normalement sur $[-1,1]$.
□

Corollaire

La fonction génératrice G_X est au moins définie et continue sur $[-1,1]$.

Remarque La fonction génératrice est entièrement déterminée par la loi de X. Inversement, la fonction génératrice caractérise la loi de X puisque

$$P(X=n)=\frac{G_X^{(n)}(0)}{n!}$$

Exemple Si $X\sim\mathcal{B}(p)$ alors

$$G_X(t)=(1-p)+pt \text{ et } R_X=+\infty$$

Exemple Si $X\sim\mathcal{B}(n,p)$ alors

$$G_X(t)=\sum_{k=0}^{n}\binom{n}{k}(pt)^k(1-p)^{n-k}=(1-p+pt)^n \text{ et } R_X=+\infty$$

Exemple Si $X \sim \mathcal{P}(\lambda)$ alors

$$G_X(t) = \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda} \frac{(\lambda t)^n}{n!} = \mathrm{e}^{\lambda(t-1)} \text{ et } R_X = +\infty$$

Exemple Si $X \sim \mathcal{G}(p)$ alors

$$G_X(t) = \sum_{n=1}^{+\infty} (1-p)^{n-1} p t^n = \frac{pt}{1-(1-p)t} \text{ et } R_X = 1/p$$

27.6.4 Calcul d'espérances et de variances

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice. On remarque

$$G_X(1) = \sum_{n=0}^{+\infty} P(X=n) = 1 = E(1)$$

Par dérivation de série entière sur $]-1,1[$

$$G_X^{(k)}(t) = \sum_{n=0}^{+\infty} n(n-1)\dots(n-k+1)P(X=n)t^{n-k}$$

et donc

$$G_X^{(k)}(t) = E\left(X(X-1)\dots(X-k+1)t^X\right)$$

Sous réserve d'existence

$$G_X^{(k)}(1) = E\left(X(X-1)\dots(X-k+1)\right)$$

ce qui donne accès aux moments de X... Approfondissons dans le cadre de l'espérance et de la variance.

Théorème

On a équivalence entre :
(i) la variable X admet une espérance ;
(ii) la fonction génératrice G_X est dérivable en 1.
De plus, on a alors

$$E(X) = G'_X(1)$$

dém. :
Sur $[-1,1]$, G_X est la somme de la série de fonctions $\sum u_n$ où

$$u_n(t) = P(X=n)t^n$$

Celles-ci sont de classe $\mathcal{C}^1$ sur $[-1,1]$ et

$$u'_n(t) = nP(X=n)t^{n-1}$$

(i) $\Rightarrow$ (ii) Si X admet une espérance alors $\sum nP(X=n)$ converge. Or on a la majoration uniforme $|u'_n(t)| \leqslant nP(X=n)$ et donc la série $\sum u'_n$ converge normalement sur $[-1,1]$.

On en déduit que G_X est de classe $\mathcal{C}^1$ sur $[-1,1]$. En particulier, G_X est dérivable en 1.
De plus
$$G'_X(1) = \sum_{n=0}^{+\infty} u'_n(1) = \sum_{n=0}^{+\infty} nP(X=n) = E(X)$$
(ii) $\Rightarrow$ (i) Supposons G_X dérivable en 1. Le taux d'accroissement
$$\frac{G_X(t) - G_X(1)}{t-1} = \sum_{n=0}^{+\infty} P(X=n)\frac{t^n - 1}{t-1}$$
admet une limite quand $t \to 1^-$. Exploitons l'écriture
$$\frac{G_X(t) - G_X(1)}{t-1} = \sum_{n=0}^{+\infty} P(X=n)(1 + t + \cdots + t^{n-1})$$
Soit $N \in \mathbb{N}$,
$$\sum_{n=0}^{N} nP(X=n) = \lim_{t\to 1} \sum_{n=0}^{N} P(X=n)(1 + t + \cdots + t^{n-1})$$
Par positivité des termes sommés
$$\sum_{n=0}^{N} nP(X=n) \leqslant \lim_{t\to 1} \frac{G_X(t) - G_X(1)}{t-1} = G'_X(1)$$
La série $\sum nP(X=n)$ est donc convergente car c'est une série à termes positifs aux sommes partielles majorées.
□

Théorème

On a équivalence entre :
(i) la variable X admet un moment d'ordre 2 ;
(ii) la fonction génératrice G_X est deux fois dérivable en 1.
De plus, on a alors
$$V(X) = G''_X(1) + G'_X(1) - \left(G'_X(1)\right)^2$$

dém. :
(i) $\Rightarrow$ (ii) Si la variable X admet un moment d'ordre 2, il y a convergence de $\sum n^2 P(X=n)$ mais aussi de $\sum n(n-1)P(X=n)$. On peut alors adapter la démonstration précédente et obtenir G_X de classe $\mathcal{C}^2$ sur $[-1,1]$ avec
$$G''_X(1) = \sum_{n=0}^{+\infty} n(n-1)P(X=n) = E\left(X(X-1)\right)$$
La relation
$$V(X) = E(X^2) - E(X)^2 = E\left(X(X-1)\right) + E(X) - E(X)^2$$
fournit alors
$$V(X) = G''_X(1) + G'_X(1) - \left(G'_X(1)\right)^2$$

(ii) ⇒ (i) Supposons G_X deux fois dérivable en 1. La fonction G_X est au moins dérivable en 1 et donc X admet une espérance. On sait alors exprimer $G'_X(t)$ sur $[-1,1]$ par

$$G'_X(t) = \sum_{n=1}^{+\infty} nP(X=n)t^{n-1}$$

La poursuite de la démonstration est alors la même que celle précédente afin d'établir la convergence de

$$\sum n(n-1)P(X=n)$$

□

Exemple Si $X \sim \mathcal{B}(p)$ alors $G_X(t) = (1-p) + pt$.
$E(X) = G'_X(1) = p$ et $V(X) = 0 + p - p^2 = p(1-p)$

Exemple Si $X \sim \mathcal{B}(n,p)$ alors $G_X(t) = (1-p+pt)^n$.
$E(X) = np$ et $V(X) = n(n-1)p^2 + np - (np)^2 = np(1-p)$

Exemple Si $X \sim \mathcal{P}(\lambda)$ alors $G_X(t) = \mathrm{e}^{\lambda(t-1)}$.
$E(X) = \lambda$ et $V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

Exemple Si $X \sim \mathcal{G}(p)$ alors $G_X(t) = \dfrac{pt}{1-(1-p)t} = \dfrac{p}{p-1} + \dfrac{\frac{p}{1-p}}{1-(1-p)t}$.
$E(X) = \dfrac{1}{p}$ et $V(X) = \dfrac{2(1-p)}{p^2} + \dfrac{1}{p} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$

27.6.5 Fonctions génératrices d'une somme

Théorème

Soit X et Y sont deux variables aléatoires discrètes à valeurs dans $\mathbb{N}$.
Si X et Y sont indépendantes alors

$$G_{X+Y} = G_X \times G_Y$$

dém. :
Pour $t \in [-1,1]$,

$$G_{X+Y}(t) = E\left(t^{X+Y}\right) = E\left(t^X \times t^Y\right)$$

Or les variables t^X et t^Y sont indépendantes car X et Y le sont donc

$$G_{X+Y}(t) = E\left(t^X\right)E\left(t^Y\right) = G_X(t)G_Y(t)$$

□

Corollaire

Si $X_1, \ldots, X_n$ sont des variables mutuellement indépendantes

$$G_{X_1+\cdots+X_n}(t) = G_{X_1}(t) \times \cdots \times G_{X_n}(t)$$

Exemple Sachant qu'une loi $\mathcal{B}(n,p)$ peut être simulée par la somme de n loi $\mathcal{B}(p)$ indépendantes, on retrouve que si $X \sim \mathcal{B}(n,p)$ alors $G_X(t) = (1-p+pt)^n$.

27.6.6 Musculation : somme aléatoire

Théorème

Soit N une variable aléatoire à valeurs dans $\mathbb{N}$ et $(X_n)_{n\in\mathbb{N}^\star}$ une suite de variables aléatoires suivant toutes une même loi de fonction génératrice G_X.
Si ces variables sont mutuellement indépendantes alors la fonction génératrice de la variable $S = \sum_{k=1}^{N} X_k$ est donnée par

$$G_S(t) = G_N\left(G_X(t)\right)$$

dém. :
Par formule des probabilités totales

$$P(S=n) = \sum_{k=0}^{+\infty} P(N=k)P(X_1+\cdots+X_k=n)$$

donc

$$G_S(t) = \sum_{n=0}^{+\infty}\sum_{k=0}^{+\infty} P(N=k)P(X_1+\cdots+X_k=n)t^n$$

En réordonnant la somme de cette famille sommable

$$G_S(t) = \sum_{k=0}^{+\infty} P(N=k)\sum_{n=0}^{+\infty} P(X_1+\cdots+X_k=n)t^n$$

soit

$$G_S(t) = \sum_{k=0}^{+\infty} P(N=k)G_{X_1+\cdots+X_k}(t)$$

Or $G_{X_1+\cdots+X_k}(t) = [G_X(t)]^k$ donc

$$G_S(t) = \sum_{k=0}^{+\infty} P(N=k)\,[G_X(t)]^k = G_N\left(G_X(t)\right)$$

□

Corollaire

Si N et X possèdent une espérance

$$E(S) = E(N)E(X)$$

dém. :
Car

$$G'_S(1) = G'_X(1)G'_N(G_X(1)) = G'_X(1)G'_N(1)$$

□

Exemple On lance une pièce équilibrée. Tant que l'on obtient « face », on jette un dé et on avance le personnage d'un jeu de plateau du nombre correspondant de cases.
En moyenne, le personnage avance de $E(S) = E(N) \times E(X) = 3,5$ cases.

Table des matières