CPGE MPSI-MP Sainte Croix - Saint Euverte Sylvain Arnt

Cours de Mathématiques CPGE MP

Programme 2021

Table des matières

Chapitre VIII : Fonctions vectorielles 392

Chapitre IX : Endomorphismes d'un espace euclidien 406

Chapitre X : Suites, séries d'intégrales et intégrales à paramètres 455

Chapitre XI : Probabilités discrètes 500

Chapitre I

Structures algébriques usuelles

Table des matières

Partie $*$

Rappels de Sup' sur les groupes

Cette partie a été vue en classe de Sup' et n'est présente dans ce cours qu'à titre de révisions. *Le lecteur pourra faire les exercices de cette partie pour s'assurer de bien maîtriser ses bases !*

1. Structure de groupe

Définition $*$1. Groupe

Soit G un ensemble et $*$ une loi de composition interne sur G. On dit que le couple $(G, *)$ est **une structure de groupe**, ou plus simplement G est un **groupe** (muni de la loi $*$), si :

i) *Associativité* : $\forall x, y, z \in G$,
$$(x * y) * z = x * (y * z);$$
ii) *Élément neutre* : $\exists e \in G,\ \forall x \in G$,
$$e * x = x = x * e;$$
iii) *Symétrique* : $\forall x \in G,\ \exists y \in G$,
$$x * y = e = y * x.$$

On dit de plus qu'un groupe G est **commutatif** si :

iv) *Commutativité* : $\forall x, y \in G$,
$$x * y = y * x.$$

Remarque $*$1.

On rappelle deux des principales notations pour la loi d'un groupe :

— La notation additive $(G, +)$ qui est utilisée exclusivement dans le cas commutatif. Dans ce cas, l'élément neutre est noté 0 et le symétrique de $x \in G$ est noté $-x$.

— La notation multiplicative $(G, .)$ (ou $(G, \times)$) qui peut s'employer dans les cas commutatifs ou non. Dans ce cas, l'élément neutre est souvent noté e ou 1 et le symétrique de $x \in G$ est noté x^{-1}.

Exemple $*$1.

— Les ensembles de nombres suivants munis de l'addition sont des groupes : $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

— Les ensembles de nombres suivants munis de la multiplication sont des groupes : $\mathbb{Q}^*$, $\mathbb{Q}_+^*$, $\mathbb{R}^*$, $\mathbb{R}_+^*$, $\mathbb{C}^*$, $\mathbb{U}$, $\mathbb{U}_n$ (pour $n \in \mathbb{N}^*$).

> On utilise ici les notations :
>
> $$\mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\} \quad \text{et pour } n \in \mathbb{N}^*, \quad \mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}.$$

— Pour X un ensemble, l'ensemble $\mathcal{S}_X$ des permutations de X (i.e. des bijections de X dans X) est un groupe pour la composition. Ce groupe est appelé le groupe **symétrique** de l'ensemble X.

— Pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $n \in \mathbb{N}^*$, L'ensemble $GL_n(\mathbb{K})$ des matrices inversibles est un groupe pour le produit matriciel.

— Pour $n \in \mathbb{N}^*$, l'ensemble $O_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) \mid {}^tMM = I_n\}$ des matrices orthogonales est un groupe pour le produit matriciel.

Exercice $*$1.

1. Déterminer, pour chaque groupe G parmi ceux de l'exemple précédent, l'élément neutre de G, le symétrique d'un élément dans G et si G est commutatif.
2. Les ensembles munis des lois suivantes sont-ils des groupes ? $(\mathbb{N}, +)$, $(\mathbb{Z}, \times)$, $(SL_n(\mathbb{R}), \times)$, $(GL_n(\mathbb{R}), +)$.

Correction.

1. — Pour $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ munis de l'addition : l'élément neutre est 0 ; le symétrique de x est son opposé $-x$ et ils sont tous commutatifs.
   — Pour $\mathbb{Q}^*$, $\mathbb{Q}_+^*$, $\mathbb{R}^*$, $\mathbb{R}_+^*$, $\mathbb{C}^*$, $\mathbb{U}$, $\mathbb{U}_n$ munis de la multiplication : l'élément neutre est 1 ; le symétrique de x est son inverse $\frac{1}{x}$ et ils sont tous commutatifs.
   — Pour $\mathcal{S}_X$ muni de la composition : l'élément neutre est l'application identité $\mathrm{id} : x \mapsto x$; le symétrique de σ est son application réciproque σ^{-1}. En général, $(\mathcal{S}_X, \circ)$ n'est pas commutatif.

   > *Exercice pour le lecteur :* montrer que $(\mathcal{S}_X, \circ)$ est commutatif si, et seulement si, X contient 1 ou 2 éléments.
   > *Indication :* Déterminer $\mathcal{S}_X$ pour $X = \{1\}, \{1, 2\}$ et $\{1, 2, 3\}$ et écrire leurs tables d'opérations.

   — Pour $GL_n(\mathbb{K})$ et $O_n(\mathbb{R})$ munis de la multiplication matricielle : l'élément neutre est la matrice identité I_n ; le symétrique de M est sa matrice inverse M^{-1}. Si $n \geq 2$, ils ne sont pas commutatif.
2.

Exercice ∗2.

Soit $(G, .)$ un groupe et $g \in G$. Alors les applications φ_g et ψ_g de G dans G telles que

$$\varphi_g : x \mapsto gx \quad \text{et} \quad \psi_g : x \mapsto xg$$

sont bijectives.

Correction.

Soit $x, y \in G$ tel que $\varphi_g(x) = \varphi_g(y)$. Alors $g^{-1}gx = g^{-1}gy$, donc $x = ex = ey = y$. Par suite, φ_g est injective.
Soit $y \in G$. Alors $\varphi_g(g^{-1}y) = gg^{-1}y = ey = y$. Donc y possède un antécédent par φ_g. Par suite, φ_g est surjective.
Il en résulte que φ_g est bijective.
On emploie un raisonnement similaire pour ψ_g.

Proposition ∗1.

(Structure de groupe produit) Soit $(G_1, .), (G_2, .)$ des groupes et on note $G = G_1 \times G_2$. On considère la loi de composition suivante sur G : pour $(x_1, x_2), (y_1, y_2) \in G$,

$$(x_1, x_2).(y_1, y_2) := (x_1.y_1, x_2.y_2).$$

Alors G muni de cette loi est un groupe et :

- L'élément neutre de G est $e = (e_1, e_2)$ où e_1 est l'élément neutre de G_1 et e_2 l'élément neutre de G_2.
- Le symétrique de $(x_1, x_2) \in G$, est (x_1^{-1}, x_2^{-1}).

Démonstration.

Pour tous $x_1, y_1 \in G_1$ et $x_2, y_2 \in G_2$, $x_1, y_1 \in G_1$ et $(x_2, y_2) \in G_2$ donc $(x_1y_1, x_2, y_2) \in G_1 \times G_2$.
Donc . est bien une loi de composition interne.
De plus, par associativité des lois sur G_1 et G_2, la loi . est associative.
Soit $(x_1, x_2) \in G$.

— Élément neutre : on a

$$(x_1, x_2).(e_1, e_2) = (x_1e_1, x_2e_2) = (x_1, x_2) = (e_1x_1, e_2x_2) = (e_1, e_2).(x_1, x_2);$$

donc $e = (e_1, e_2)$ est un élément neutre pour la loi . .

— Symétrique : on a

$$(x_1, x_2).(x_1^{-1}, x_2^{-1}) = (x_1x_1^{-1}, x_2x_2^{-1}) = (e_1, e_2) = (x_1^{-1}x_1, x_2^{-1}x_2) = (x_1^{-1}, x_2^{-1}).(x_1, x_2);$$

donc (x_1^{-1}, x_2^{-1}) est le symétrique de (x_1, x_2) pour la loi . .

Il en résulte que G est un groupe. □

Remarque ∗2.

Par récurrence, on peut ainsi munir un produit fini de groupes d'une structure de groupe.

Exercice ∗3.

Montrer que $G = G_1 \times G_2$ est commutatif si, et seulement si, G_1 et G_2 sont commutatifs.

Correction.

Soit $x_1, y_1 \in G_1$, $x_2, y_2 \in G_2$. On a :

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

si, et seulement si,

$$(x_1x_2, y_1y_2) = (x_2x_1, y_2y_1),$$

si, et seulement si,

$$x_1x_2 = x_2x_1 \quad \text{et} \quad y_1y_2 = y_2y_1.$$

2. Sous-groupes

a. Généralités

Définition ∗2. Sous-groupe

Soit G un groupe et $H \subset G$. On dit que H est un sous-groupe de G si :
- H est non vide ;
- pour tous $x, y \in H$, $x.y \in H$;
- pour $x \in H$, $x^{-1} \in H$.

Proposition ∗2.

(Caractérisation des sous-groupes) Soit G un groupe et $H \subset G$. Alors H est un sous-groupe de G, si, et seulement si :

i) L'élément neutre e de G appartient à H ;

ii) pour tous $x, y \in H$, $x.y^{-1} \in H$.

Démonstration.

- ($\Rightarrow$). On suppose que H est un sous-groupe de G. Alors,
  i) H est non vide, donc il existe $x \in H$, et d'après les hypothèses, $x^{-1} \in H$. Par suite, $e = x.x^{-1} \in H$.

  ii) Pour tous $x, y \in H$, $y^{-1} \in H$ d'après les hypothèses, donc

  $$x.y^{-1} \in H.$$

- ($\Leftarrow$). On suppose i) et ii) vérifiés. Alors
  - H est non vide, car $e \in H$.
  - Soit $x \in H$. Alors d'après i) et ii),
  $$x^{-1} = e.x^{-1} \in H.$$
  - Soit $x, y \in H$. Alors $y^{-1} \in H$ d'après ce qui précède et $y = (y^{-1})^{-1}$, donc, d'après ii)
  $$x.y = x.(y^{-1})^{-1} \in H.$$

  Il en résulte que H est un sous-groupe de G.

□

Exemple ∗2.

— Si G est un groupe, $\{e\}$ et G sont des sous-groupes de G. On les appelle les sous-groupes triviaux de G.

— La chaîne d'inclusions suivante est également une chaîne de sous-groupes :
$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

— Soit $n \in \mathbb{N}^*$. $\mathbb{U}_n$ est un sous-groupe de $\mathbb{U}$.

— Soit $n \in \mathbb{N}^*$. $O_n(\mathbb{R})$ est un sous-groupe de $GL_n(\mathbb{R})$.

3. Morphismes de groupes

a. Définition

Définition ∗3. Morphisme de groupes

Soit $(G_1, *), (G_2, \star)$ des groupes et $f : G_1 \to G_2$.
On dit que f est un **morphisme de groupes** si, pour tous $x, y \in G_1$:
$$f(x * y) = f(x) \star f(y).$$

Exemple ∗3.

— L'exponentielle est un morphisme de groupes de $(\mathbb{R}, +)$ dans $(\mathbb{R}_+^*, \times)$.

— Le déterminant est un morphisme de groupes de $(GL_n(\mathbb{K}), \times)$ dans $(\mathbb{K}^*, \times)$.

— Soit $n \in \mathbb{N}^*$. La signature ε est un morphisme de groupes de $(\mathcal{S}_n, \circ)$ dans $(\{-1, 1\}, \times)$

Exercice ∗4.

Soit $n \in \mathbb{N}$. Montrer que l'application $\varphi : \left| \begin{array}{ccc} \mathbb{Z} & \to & \mathbb{C}^* \\ k & \mapsto & e^{i\frac{2k\pi}{n}} \end{array} \right.$ est un morphisme de $(\mathbb{Z}, +)$ dans

$(\mathbb{C}^*, \times)$.

Correction.

Soit $k, k' \in \mathbb{Z}$. Alors

$$\varphi(k+k') = e^{i\frac{2(k+k')\pi}{n}} = e^{i\frac{2k\pi}{n}} e^{i\frac{2k'\pi}{n}} = \varphi(k)\varphi(k').$$

Donc φ est un morphisme de groupes.

b. Noyau, Image et sous-groupes

Définition ∗4. Noyau, Image d'un morphisme de groupes

Soit G_1, G_2 des groupes d'éléments neutres respectifs e_1,e_2 et $f : G_1 \to G_2$ un morphisme de groupes.

— On appelle **noyau de** f l'ensemble $\mathrm{Ker}(f) = \{x \in G_1 \mid f(x) = e_2\}$.

— On appelle **image de** f l'ensemble $\mathrm{Im}(f) = f(G_1) = \{f(x) \mid x \in G_1\}$.

Lemme ∗1. -NoValue-

Soit G_1, G_2 des groupes d'éléments neutres respectifs e_1,e_2 et $f : G_1 \to G_2$ un morphisme de groupes. Alors :

$$f(e_1) = e_2; \quad \forall x \in G_1,\ f(x^{-1}) = f(x)^{-1} \text{ et } \forall x \in G_1, \forall n \in \mathbb{N}^*, f(x^n) = f(x)^n.$$

Démonstration.

Soit $x \in G_1$.

- On a, $f(e_1)f(e_1) = f(e_1e_1) = f(e_1) = f(e_1)e_2$. Donc, en composant à gauche cette égalité par $(f(e_1))^{-1}$, on obtient le résultat :

$$f(e_1) = e_2.$$

- On a $f(x^{-1})f(x) = f(x^{-1}x) = f(e_1) = e_2$, donc en composant à droite cette égalité par $(f(x))^{-1}$, on obtient le résultat :

$$f(x^{-1}) = f(x)^{-1}.$$

- On raisonne par récurrence sur $n \in \mathbb{N}^*$.
  Initialisation : On $f(x^1) = f(x) = f(x)^1$.
  Hérédité : Soit $n \in \mathbb{N}^*$. On suppose que $f(x^n) = f(x)^n$. Alors on a :

$$f(x^{n+1}) = f(x^nx) = f(x^n)f(x) = f(x)^nf(x) = f(x)^{n+1}.$$

Ce qui achève le raisonnement par récurrence.
Il en résulte que, pour tout $n \in \mathbb{N}^*$,

$$f(x^n) = f(x)^n.$$

□

Proposition *3.

Soit G_1, G_2 des groupes, H_1, H_2 des sous-groupes de G_1, G_2 respectivement et $f : G_1 \to G_2$ un morphisme de groupes. Alors :
- $f^{-1}(H_2)$ est un sous-groupe de G_1 ;
- $f(H_1)$ est un sous-groupe de G_2.

Démonstration.

- i) On a $f(e_1) = e_2 \in H_2$, donc $e_1 \in f^{-1}(H_2)$.
  ii) Soit $x, y \in f^{-1}(H_2)$. Montrons que $xy^{-1} \in f^{-1}(H_2)$. On a :
  $$f(xy^{-1}) = f(x)f(y^{-1}) = f(x)f(y)^{-1} \in H_2,$$
  car $f(x), f(y) \in H_2$ et H_2 est un sous-groupe. Donc $xy^{-1} \in f^{-1}(H_2)$.
  Il en résulte que $f^{-1}(H_2)$ est un sous-groupe de G_1.
- i) On a $e_1 \in H_1$ et $f(e_1) = e_2$, donc $e_2 \in f(H_1)$.
  ii) Soit $x, y \in f(H_1)$. Montrons que $xy^{-1} \in f(H_1)$. Il existe $z, t \in H_1$ tels que $x = f(z)$ et $y = f(t)$. On a alors :
  $$xy^{-1} = f(z)f(t)^{-1} = f(z)f(t^{-1}) = f(zt^{-1}) \in f(H_1)$$
  car $z, t \in H_2$ et H_2 est un sous-groupe. Donc $xy^{-1} \in f(H_1)$.
  Il en résulte que $f(H_1)$ est un sous-groupe de G_2.

□

Corollaire *1.

Soit G_1, G_2 des groupes et $f : G_1 \to G_2$ un morphisme de groupes. Alors :
- $\mathrm{Ker}(f)$ est un sous-groupe de G_1 ;
- $\mathrm{Im}(f)$ est un sous-groupe de G_2.

Démonstration.

- On a :
  $$\mathrm{Ker}(f) = \{x \in G_1 \mid f(x) = e_2\} = f^{-1}(\{e_2\});$$
  Or $\{e_2\}$ est un sous-groupe de G_2, donc, d'après la proposition précédente, $\mathrm{Ker}(f)$ est

un sous-groupe de G_1 comme image réciproque d'un sous-groupe par un morphisme de groupes.

— On a :
$$\mathrm{Im}(f) = f(G_1);$$
Or G_1 est un sous-groupe de G_1, donc, d'après la proposition précédente, $\mathrm{Im}(f)$ est un sous-groupe de G_2 comme image directe d'un sous-groupe par un morphisme de groupes.

□

Proposition ∗4.

Soit G_1, G_2 des groupes et $f : G_1 \to G_2$ un morphisme de groupes. On note e_1 l'élément neutre de G_1.
Alors l'application f est injective si, et seulement si, $\mathrm{Ker}(f) = \{e_1\}$

Démonstration.

- $(\Rightarrow)$. On suppose f injective. Soit $x \in \mathrm{Ker}(f)$. Alors $f(x) = e_2 = f(e_1)$. Par injectivité de f, il en résulte que $x = e_1$. Par suite, $\mathrm{Ker}(f) = \{e_1\}$.
- $(\Leftarrow)$. On suppose $\mathrm{Ker}(f) = \{e_1\}$. Soit $x, x' \in G_1$ tels que $f(x) = f(x')$. Alors on a :
$$f(xx'^{-1}) = f(x)f(x'^{-1}) = f(x)f(x')^{-1} = e_2,$$
donc $xx'^{-1} \in \mathrm{Ker}(f) = \{e_1\}$; d'où $xx'^{-1} = e_1$. Par suite $x = x'$.
Il en résulte que f est injective.

□

Exemple ∗4.

— Soit $n \in \mathbb{N}^*$. L'application $z \mapsto z^n$ de $\mathbb{C}^*$ dans $\mathbb{C}^*$ est un morphisme de groupes surjectif et son noyau est $\mathbb{U}_n$.

> En effet : soit $z, z' \in \mathbb{C}^*$. On note $f : z \mapsto z^n$. Alors :
> $$f(zz') = (zz')^n = z^n z'^n = f(z)f(z'),$$
> car la multiplication sur $\mathbb{C}^*$ est commutative. Donc f est un morphisme de groupes. De plus, pour $\zeta = re^{i\theta} \in \mathbb{C}^*$, $z = r^{\frac{1}{n}} e^{i\frac{\theta}{n}}$ est un antécédent de ζ par f. Donc f est surjective de $\mathbb{C}^*$ dans lui-même.

— Le groupe spécial orthogonal $SO_n(\mathbb{R})$ des matrices orthogonales de déterminant 1 est un sous-groupe de $O_n(\mathbb{R})$: il est l'image réciproque de $\{1\}$ par le morphisme de groupes $\det : O_n(\mathbb{R}) \to \mathbb{R}^*$.

Exercice ∗5.

Montrer que exp : $z \mapsto e^z$ est un morphisme de groupes surjectif de $(\mathbb{C}, +)$ dans $(\mathbb{C}^*, \times)$ et déterminer son noyau.

Remarque : on définira "proprement" l'exponentielle complexe dans le chapitre dédié aux séries entières. Pour cet exercice, on prendra la définition de exp et ses propriétés vues en Sup', à savoir : $\exp(0) = 1$ et pour $z = re^{i\theta} \in \mathbb{C}^*$, $\exp(z) = e^r e^{i\theta}$ où e^r est l'exponentielle réelle de r et $e^{i\theta} = \cos(\theta) + i \sin(\theta)$.

Correction.

Soit $z, z' \in \mathbb{C}$. On a $|e^z| = e^{|z|} > 0$ donc $e^z \in \mathbb{C}^*$ et $\exp(z + z') = \exp(z)\exp(z')$. Donc exp est un morphisme de groupes de $(\mathbb{C}, +)$ dans $(\mathbb{C}^*, \times)$.

Image : Pour $\zeta = re^{i\theta} \in \mathbb{C}^*$ avec $r \in \mathbb{R}_+^*$ et $\theta \in [0, 2\pi[$, $z = \ln(r) + i\theta$ est un antécédent de ζ par exp. Par suite, $\mathrm{Im}(\exp) = \mathbb{C}^*$. Il en résulte que exp est surjective de $\mathbb{C}$ dans $\mathbb{C}^*$.

Noyau : On a $z = x + iy \in \mathrm{Ker}(\exp)$ si, et seulement si, $e^x e^{iy} = e^z = 1$. Par suite, $x = 0$ (car $e^x = |e^z| = 1$) et $y \in \{2k\pi \mid k \in \mathbb{Z}\}$. (Et réciproquement, un élément de cette forme est dans le noyau).
Il en résulte que $\mathrm{Ker}(\exp) = \{i2k\pi \mid k \in \mathbb{Z}\}$.

Exercice ∗6.

Soit $n \in \mathbb{N}^*$. On note :

$$SL_n(\mathbb{R}) = \{M \in GL_n(\mathbb{R}) \mid \det(M) = 1\}$$

Montrer que $SL_n(\mathbb{R})$ est un sous-groupe de $GL_n(\mathbb{R})$.

Correction.

Pour tous $A, B \in M_n(\mathbb{R})$, on a $\det(AB) = \det(A)\det(B)$, donc l'application det restreinte à $GL_n(\mathbb{R})$ est un morphisme de groupes de $GL_n(\mathbb{R})$ dans $\mathbb{R}^*$. De plus, on remarque que $SL_n(\mathbb{R}) = \mathrm{Ker}(\det)$ donc $SL_n(\mathbb{R})$ est un sous-groupe de $GL_n(\mathbb{R})$ comme noyau d'un morphisme de groupes.

c. Isomorphismes de groupes

Définition ∗5. Isomorphisme de groupes

Soit G_1, G_2 des groupes et $f : G_1 \to G_2$. Si f est un morphisme de groupes bijectif, on dit que f est un **isomorphisme de groupes**.

Si f est un isomorphisme de groupes et $G_1 = G_2$, on dit que f et un **automorphisme de groupes**. On note $\mathrm{Aut}(G)$ l'ensemble des automorphismes de G.

Exemple ∗5.

— L'exponentielle est un isomorphisme de $(\mathbb{R}, +)$ dans $(\mathbb{R}_+^*, \times)$.

> On a montré précédemment que exp est un morphisme de groupes de $(\mathbb{R},+)$ dans $(\mathbb{R}_+^*,\times)$ et de plus exp est bijective de $(\mathbb{R},+)$ dans $(\mathbb{R}_+^*,\times)$ de réciproque ln.
> Donc exp est un isomorphisme de $(\mathbb{R},+)$ dans $(\mathbb{R}_+^*,\times)$.

— Soit G un groupe et $g \in G$. L'application $x \mapsto gxg^{-1}$ est un automorphisme de G (on appelle automorphismes intérieurs de G de telles applications).

> Soit $g \in G$. Notons $f_g : x \mapsto gxg^{-1}$.
> *Morphisme :* Soit $x, x' \in G$. On a :
>
> $$f_g(xx') = g^{-1}xx'g = g^{-1}xgg^{-1}x'g = f_g(x)f_g(x').$$
>
> *Image :* On a, pour tout $y \in G$, $y = gg^{-1}ygg^{-1} = f_g(g^{-1}yg)$; d'où $x = g^{-1}yg$ est un antécédent de y par f_g. Par suite, f_g est surjective de G dans G.
> *Noyau :* Soit $x \in \mathrm{Ker}(f_g)$. Alors $g^{-1}xg = f_g(x) = e$, donc, en composant cette égalité à gauche par g et à droite par g^{-1}, on obtient $x = e$. Par suite, f_g est injective.

Proposition ∗5.

Soit G_1, G_2 des groupes et $f : G_1 \to G_2$. Si f est un isomorphisme, alors f^{-1} est également un isomorphisme de groupes.

Démonstration.

On suppose que f est un isomorphisme. Alors f^{-1} existe et est bijective de G_2 dans G_1. Montrons que, de plus, f^{-1} est un morphisme de groupes.
Soit $y, y' \in G_2$. Comme f est bijective, il existe $x, x' \in G_1$ tels que $y = f(x)$, $y' = f(x')$ et $x = f^{-1}(y)$, $x' = f^{-1}(y')$. On a :

$$f^{-1}(yy') = f^{-1}(f(x)f(x')) = f^{-1}(f(xx')) = xx' = f^{-1}(y)f^{-1}(y').$$

Donc f^{-1} est un morphisme de groupes.
Il en résulte que f^{-1} est un isomorphisme de groupes. □

Exemple ∗6.

Le logarithme népérien est un isomorphisme de $(\mathbb{R}_+^*,\times)$ dans $(\mathbb{R},+)$.

On a montré dans les exercices précédents que exp est un isomorphisme de groupes de $(\mathbb{R},+)$ dans $(\mathbb{R}_+^*,\times)$. Or ln est la réciproque de cette fonction, donc, d'après la proposition précédente ln est un isomorphisme de groupes.

Exercice $*$7. Théorème de Cayley

1. Soit G un groupe. On considère $\mathcal{S}_G$ le groupe symétrique de G (groupe des permutations de G). On note φ l'application de G dans $\mathcal{F}(G,G)$ (ensemble des fonctions de G dans G) telle que, pour tout $g \in G$, $\varphi(g) : x \mapsto g.x$.
   (a) Montrer que $\mathrm{Im}(\varphi) \subset \mathcal{S}_G$.
   (b) Monter que φ est un morphisme injectif de G dans $\mathcal{S}_G$.
2. En déduire le résultat suivant :

> **Théorème de Cayley**
> *Tout groupe est isomorphe à un sous-groupe d'un groupe symétrique.*

Correction.

1. (a) D'après l'exercice $*2$, pour tout $g \in G$, $\varphi(g)$ est une bijection de G dans G i.e. $\varphi(g) \in \mathcal{S}_G$. D'où $\mathrm{Im}(\varphi) \subset \mathcal{S}_G$.
   (b) Montrons que φ est un morphisme de groupes de $(G,\cdot)$ dans $(\mathcal{S}_G,\circ)$. Soit $g,g' \in G$. On a, pour tout $x \in G$:
$$\begin{array}{rcl} \varphi(g.g')(x) & = & (g.g').x \\ & = & g.(g'.x) \text{ par associativité de } \cdot \\ & = & g.\varphi(g')(x) = \varphi(g)(\varphi(g')(x)) \\ \varphi(g.g')(x) & = & f(g)\circ\varphi(g')(x) \end{array}$$
   et donc $\varphi(g.g')(x) = f(g)\circ\varphi(g')$.
   Ainsi, φ est un morphisme de G dans $\mathcal{S}_G$.
   Montrons son injectivité. Soit $g \in \mathrm{Ker}(\varphi)$. Alors $\varphi(g) = \mathrm{id}$. Ainsi, on a :
$$g = g.e = \varphi(g)(e) = \mathrm{id}(e) = e$$
   Par suite, $\mathrm{Ker}(g) = \{e\}$ et don φ est injective.
2. Soit G un groupe. D'après la question précédente, l'application φ est un morphisme injectif de G dans $\mathcal{S}_G$ groupe symétrique de G donc φ est un isomorphisme de G dans le sous-groupe $\mathrm{Im}(\varphi)$ de $\mathcal{S}_G$.

Exercice ∗8.

Soit G un groupe et $\mathrm{Aut}(G)$ l'ensemble des automorphismes de G.

1. Montrer que, muni de la composition des applications, $\mathrm{Aut}(G)$ est un groupe.
2. On note $\mathrm{Int}(G) = \{\psi_g : x \mapsto gxg^{-1} \mid g \in G\}$. Montrer que $\mathrm{Int}(G)$ est un sous-groupe de $\mathrm{Aut}(G)$.

Correction.

1. — *1ère façon* (si on se rappelle du groupe symétrique de G noté $\mathcal{S}_G$) : montrons que $\mathrm{Aut}(G)$ est un sous-groupe de $(\mathcal{S}_G, \circ)$.
   On a bien $\mathrm{Aut}(G) \subset \mathcal{S}_G$ car tout automorphisme de G est en particulier une application bijective de G dans G.
   L'élément neutre Id_G de $(\mathcal{S}_G, \circ)$ est bien un automorphisme : comme il est dans $\mathcal{S}_G$, il est bijectif de G dans G et pour tout $x, x,' \in G$, $\mathrm{Id}_G(xx') = xx' = \mathrm{Id}_G(x)\mathrm{Id}_G(x')$ donc c'est un morphisme de groupes.
   De plus, pour tous $\varphi, \psi \in \mathrm{Aut}(G)$, ψ^{-1} est un automorphisme d'après la proposition précédente et on a, pour $g, g' \in G$:
   $$\begin{array}{rcl} \varphi \circ \psi^{-1}(gg') & = & \varphi(\psi^{-1}(gg')) \\ & = & \varphi(\psi^{-1}(g).\psi^{-1}(g')) \\ & = & \varphi(\psi^{-1}(g)).\varphi(\psi^{-1}(g')) \\ & = & \varphi \circ \psi^{-1}(g).\varphi \circ \psi^{-1}(g'). \end{array}$$
   Donc $\varphi \circ \psi^{-1}$ appartient à $\mathrm{Aut}(G)$.
   Par suite, $\mathrm{Aut}(G)$ est un sous-groupe de $(\mathcal{S}_G, \circ)$ et c'est donc un groupe.

   — *2ème façon* (si on ne se rappelle pas du groupe symétrique de G) : montrons le avec la définition !
   Soit $\varphi, \psi \in \mathrm{Aut}(G)$. Alors $\varphi \circ \psi : G \to G$ est bijective comme composée d'applications bijectives et on a, pour $g, g' \in G$:
   $$\varphi \circ \psi(gg') = \varphi(\psi(gg')) = \varphi(\psi(g).\psi(g')) = \varphi(\psi(g)).\varphi(\psi(g')) = \varphi \circ \psi(g).\varphi \circ \psi(g').$$
   Donc $\varphi \circ \psi$ appartient à $\mathrm{Aut}(G)$.
   Par suite, $\circ$ est une loi de composition interne sur $\mathrm{Aut}(G)$. Elle est associative et d'élément neutre la fonction identité $\mathrm{Id}_G : G \to G$. D'après la proposition précédente, si ψ est un automorphisme de G, alors ψ^{-1} l'est aussi, donc tout élément de $\mathrm{Aut}(G)$ possède un symétrique pour la loi $\circ$: il s'agit de sa réciproque.
   Ainsi, $(\mathrm{Aut}(G), \circ)$ est un groupe.

   Question : au fait, est-il commutatif ?

2. Remarquons tout d'abord deux faits. Soit $g, g' \in G$.
   - $\psi_g \circ \psi_{g'} = \psi_{gg'}$. En effet, pour tout $x \in G$, on a :
     $$\psi_g \circ \psi_{g'}(x) = \psi_g(g'xg'^{-1}) = g(g'xg'^{-1})g^{-1} = (gg')x(gg')^{-1} = \psi_{gg'}(x).$$
   - $(\psi_g)^{-1} = \psi_{g^{-1}}$. En effet, pour tout $x \in G$, on a, d'après le point précédent :
     $$\psi_g \circ \psi_{g^{-1}}(x) = \psi_e(x) = exe^{-1} = x = \mathrm{id}(x);$$
     et de même
     $$\psi_{g^{-1}} \circ \psi_g(x) = \psi_e(x) = exe^{-1} = x = \mathrm{id}(x).$$
     Donc $\psi_{g^{-1}} = (\psi_g)^{-1}$.

Montrons alors que $\mathrm{Int}(G)$ est un sous-groupe de $\mathrm{Aut}(G)$.

i) $\mathrm{id} = \psi_e \in \mathrm{Int}(G)$;

ii) Soit $\psi_g, \psi_{g'} \in \mathrm{Int}(G)$ avec $g, g' \in G$. On a, d'après les remarques précédentes :

$$\psi_g \circ (\psi_{g'})^{-1} = \psi_g \circ \psi_{g'^{-1}} = \psi_{gg'^{-1}} \in \mathrm{Int}(G).$$

Donc $\mathrm{Int}(G)$ est un sous-groupe de $\mathrm{Aut}(G)$.

Partie A
Compléments sur les groupes

1. Les sous-groupes de $\mathbb{Z}$

a. Sous-groupe engendré par une partie

Proposition 1. *Intersection de sous-groupes*

Soit G un groupe et $(H_i)_{i\in I}$ une famille quelconque de sous-groupes de G. Alors $\bigcap\limits_{i\in I} H_i$ est un sous-groupe de G.
Autrement dit, une intersection quelconque de sous-groupes est un sous-groupe.

Démonstration.

On note $H = \bigcap_{i\in I} H_i$.

i) On a, pour tout $i \in I$, $e \in H_i$ car H_i est un sous-groupe de G. Donc $e \in H$.
ii) Soit $x, y \in H$. Alors, pour tout $i \in I$, $x, y \in H_i$ qui est un sous-groupe de G, donc, pour tout $i \in I$,
$$x.y^{-1} \in H_i.$$
Par suite, $x.y^{-1} \in H$.

Il en résulte que H est un sous-groupe de G. □

Exercice 1.

Que dire d'une réunion de sous-groupes ?

Correction.

En général, une réunion de sous-groupes n'est pas un sous-groupe. Prendre par exemple les sous-groupes $i\mathbb{R}$ et $\mathbb{R}$ de $\mathbb{C}$.

Définition-Proposition 1.

Soit G un groupe et $A \subset G$. On note $\langle A\rangle$ l'intersection de tous les sous-groupes de G contenant A, i.e.
$$\langle A\rangle = \bigcap_{H\in\mathcal{H}_A} H \quad \text{où } \mathcal{H}_A = \{H \text{ sous-groupe de } G \mid A \subset H\}.$$

Alors $\langle A\rangle$ est le plus petit sous-groupe de G contenant A et on l'appelle le **sous-groupe engendré par** A.

Si de plus $\langle A\rangle = G$, on dit que A **engendre** G.

Démonstration.

Une intersection quelconque de sous-groupes de G est un sous-groupe de G, donc $\langle A\rangle = \bigcap_{\mathcal{H}_A} H$ est un sous-groupe de G. De plus, pour tout $H \in \mathcal{H}_A$, $A \subset H$, donc $A \subset \bigcap_{H\in\mathcal{H}_A} H = \langle A\rangle$.
Montrons alors que $\langle A\rangle$ est le plus petit sous groupe contenant A.
Soit $K \in \mathcal{H}_A$. Alors $\langle A\rangle = \bigcap_{H\in\mathcal{H}_A} H \subset K$. Donc $\langle A\rangle$ est le plus petit sous groupe contenant A. $\square$

La proposition suivante permet de se faire une meilleure idée de la notion de sous-groupe engendré : on y montre que le sous-groupe engendré par une partie est l'ensemble des éléments du groupe qui s'écrivent comme la composition d'un nombre fini d'éléments ou de symétriques d'éléments de cette partie.

Proposition 2.

Soit G un groupe et $A \subset G$. On a :

$$\langle A\rangle = \left\{a_1...a_n \mid n \in \mathbb{N}^*,\ a_1, ..., a_n \in A \cup A^{-1} \cup \{e\}\right\}.$$

où $A^{-1} = \{a^{-1} \mid a \in A\}$.

Démonstration.

On note $E(A) = \left\{a_1...a_n \mid n \in \mathbb{N}^*,\ a_1, ..., a_n \in A \cup A^{-1} \cup \{e\}\right\}$. On procède par double inclusion pour montrer que $\langle A\rangle = E(A)$.

$\subset$ Pour cette inclusion, il suffit de montrer que $E(A)$ est un sous-groupe de G contenant A car $\langle A\rangle$ est l eplus petit d'entre eux pour l'inclusion. Allons-y !
On remarque tout d'abord que $A \subset E(A)$; en effet, pour tout $a \in A$, $a \in A \cup A^{-1} \cup \{e\}$, donc $a \in E(A)$.
Montrons que $E(A)$ est un sous-groupe de G.

i) On a $e \in \{e\} \subset A \cup A^{-1} \cup \{e\}$ donc pour $n = 1$ et $a_1 = e$, on a $e = a_1 \in E(A)$.

ii) Soit $x, y \in E(A)$. Alors il existe $n, m \in \mathbb{N}^*$, $a_1, ..., a_n, a'_1, ..., a'_m \in A \cup A^{-1} \cup \{e\}$ tels que $x = a_1...a_n$ et $y = a'_1, ..., a'_m$. Alors

$$\begin{aligned} xy^{-1} &= a_1...a_n(a'_1...a'_m)^{-1} \\ &= a_1...a_n a'^{-1}_m...a'^{-1}_1 \\ &= a''_1...a''_{n+m} \end{aligned}$$

où

$$a''_i = \begin{cases} a_i \in A \cup A^{-1} \cup \{e\} & \text{si } i \in [\![1, n]\!]; \\ a'^{-1}_{i-n} \in A \cup A^{-1} \cup \{e\} & \text{si } i \in [\![n+1, n+m]\!]. \end{cases}$$

Par suite $xy^{-1} \in E(A)$.

Il en résulte que $E(A)$ est un sous-groupe de G.

Ainsi, $\langle A\rangle$ étant le plus petit sous-groupe de G contenant A et $E(A)$ étant un sous-groupe de G contenant A, on a $\langle A\rangle \subset E(A)$.

$\supset$ Soit $x \in E(A)$ et H un sous-groupe de G contenant A. Alors il existe $n \in \mathbb{N}^*$ tel que $x = a_1...a_n$ avec $a_1, ..., a_n \in A \cup A^{-1} \cup \{e\} \subset H$ car H contient A et H est un sous-groupe de G. Comme H est stable par composition $x = a_1...a_n \in H$.
Comme $\langle A\rangle$ est par définition l'intersection de tous les sous-groupes de G contenant A et chacun de ces sous-groupes contiennent $E(A)$ donc $E(A) \subset \langle A\rangle$.

Il en résulte que $E(A) = \langle A\rangle$. □

Exemple 1.

— $\mathbb{Z}$ est engendré par 1, i.e. $\mathbb{Z} = \langle 1\rangle$;

— Soit $n \in \mathbb{N}^*$. Le groupe $\mathcal{S}_n$ des permutations de $[\![1, n]\!]$ est engendré par les transpositions.

Exercice 2.

Déterminer le sous-groupe engendré par $A = \{2, 3\}$ dans $\mathbb{Z}$.

Correction.

En faisant un petit dessin, on se convaint que $\langle A\rangle = \mathbb{Z}$; montrons cette conjecture !

Comme $\langle A\rangle$ est un sous-groupe de $\mathbb{Z}$, on a en particulier $\langle A\rangle \subset \mathbb{Z}$. Montrons l'inclusion réciproque. Soit $n \in \mathbb{Z}$. Alors, comme 2 et 3 sont premiers entre eux, d'après le théorème de Bézout (on retrouvera l'énoncé de ce théorème vu en Sup' un peu plus loin), il existe $u', v' \in \mathbb{Z}$ tel que $2u+3v = 1$. On pose alors $u = nu'$ et $v = nv'$ et on obtient :

$$n = 2u + 3v = \begin{cases} \underbrace{2 + \ldots + 2}_{u \text{ termes}} + \underbrace{3 + \ldots + 3}_{v \text{ termes}} & \text{si } u, v \geq 0 \\ \underbrace{(-2) + \ldots + (-2)}_{|u| \text{ termes}} + \underbrace{3 + \ldots + 3}_{v \text{ termes}} & \text{si } u > 0 \text{ et } v \geq 0 \\ \underbrace{2 + \ldots + 2}_{u \text{ termes}} + \underbrace{(-3) + \ldots + (-3)}_{|v| \text{ termes}} & \text{si } u \geq 0 \text{ et } v < 0 \\ \underbrace{(-2) + \ldots + (-2)}_{|u| \text{ termes}} + \underbrace{(-3) + \ldots + (-3)}_{|v| \text{ termes}} & \text{si } u, v < 0 \end{cases}$$

donc $n \in \langle A\rangle$.
Notre conjecture est donc vraie !

b. les sous-groupes de $\mathbb{Z}$

Théorème 1. *Division euclidienne*

Soit $a \in \mathbb{Z}$ et $b \in \mathbb{N}^*$. Alors il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que :

$$a = bq + r \text{ et } 0 \leq r < b.$$

Démonstration.

Vue en Sup'. (Idée : pour $a \geq 0$, l'ensemble $\{p \in \mathbb{N} \mid bp \leq a\}$ est non vide et majoré, or toute partie non vide et majorée de $\mathbb{N}$ possède un plus grand élément : il s'agit du quotient q. Il ne reste plus qu'à encadrer le reste $r = a - bq$ et à démonter l'unicité du couple (q, r)). $\square$

Théorème 2.

Soit $H \subset \mathbb{Z}$. Alors H est un sous-groupe de $(\mathbb{Z}, +)$, si, et seulement si, il existe $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.

Démonstration.

- $(\Leftarrow)$. On suppose qu'il existe $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$. Alors $0 = n.0 \in H$. De plus, pour tous $x, y \in H$, il existe $p, q \in \mathbb{Z}$ tels que $x = np$ et $y = nq$ donc :

$$x - y = np - nq = n\underbrace{(p - q)}_{\in \mathbb{Z}} \in H.$$

Donc H est un sous-groupe de $\mathbb{Z}$.
- $(\Rightarrow)$.
1er cas : $H = \{0\}$. Alors $H = 0\mathbb{Z}$.
2eme cas : $H \neq \{0\}$. Pour $k \in H \setminus \{0\}$, $|k| \in H \cap \mathbb{N}^*$. Or tout sous-ensemble non vide de $\mathbb{N}$ possède un plus petit élément, donc $H \cap \mathbb{N}^*$ possède un plus petit élément n.
On a, pour tout $k \in \mathbb{Z}$,

$$nk = \begin{cases} \underbrace{n + \ldots + n}_{k \text{ termes}} \in H & \text{si } k \geq 0, \\ -(\underbrace{n + \ldots + n}_{-k \text{ termes}}) \in H & \text{si } k < 0; \end{cases}$$

donc $n\mathbb{Z} \subset H$.
Réciproquement, si $x \in H \subset \mathbb{Z}$, la division euclidienne de x par n nous fournit un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ avec $r < n$ tel que

$$x = nq + r$$

Comme $n\mathbb{Z} \subset H$, $nq \in H$ et donc $r = x - nq$ appartient à H. Or n est le plus petit élément positif et non nul de H et $r < n$, donc $r = 0$. Par suite, $x = nq \in n\mathbb{Z}$.
Il en résulte que $H = n\mathbb{Z}$.

$\square$

Proposition 3.

Si H est un sous-groupe de $\mathbb{Z}$, alors il existe un unique $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.

Correction.

Soit H un sous-groupe de $\mathbb{Z}$. L'existence est assurée par le théorème précédent. Montrons l'unicité. Soit $n, m \in \mathbb{N}$ tels que $n\mathbb{Z} = H = m\mathbb{Z}$. Ainsi $n \in n\mathbb{Z} \subset m\mathbb{Z}$, d'où $m|n$ donc $m = |m| \leq |n| = n$ (n, m étant positifs). De manière analogue, comme $m \in n\mathbb{Z}$, $n \leq m$.
Il en résulte que $n = m$. D'où le résultat.

Exercice 3.

Montrer que pour $p \in \mathbb{Z} \setminus \mathbb{N}$, $p\mathbb{Z}$ est un sous-groupe de $\mathbb{Z}$ (À quel sous-groupe de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$ est-il égal ?)

Correction.

Pour tout $k \in \mathbb{Z}$, $pk = (-p)(-k)$, donc $p\mathbb{Z} = (-p)\mathbb{Z}$ qui est un sous-groupe de $\mathbb{Z}$ d'après le théorème précédent.

2. Le groupe $\mathbb{Z}/n\mathbb{Z}$

a. Congruences

On rappelle la relation de congruence entre deux entiers relatifs pour un entier naturel non nul fixé :

Définition 2. *Relation de congruence*

Soit $n \in \mathbb{N}^*$. Pour $a, b \in \mathbb{Z}$, on dit que a **est congru à** b **modulo** n si

$$b - a \in n\mathbb{Z};$$

on note :

$$a \equiv b \bmod n.$$

Proposition 4.

Soit $n \in \mathbb{N}^*$. La relation de congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$. De plus, elle est *compatible* avec l'addition sur $\mathbb{Z}$, i.e. pour tous $a, b, c, d \in \mathbb{Z}$, si $a \equiv b \bmod n$ et $c \equiv d \bmod n$, alors

$$\begin{cases} a + c \equiv b + d \bmod n; \\ ac \equiv bd \bmod n. \end{cases}$$

Démonstration.

Montrons que $\cdot \equiv \cdot \bmod n$ est une relation d'équivalence sur $\mathbb{Z}$.

Soit $a, b, c \in \mathbb{Z}$.

- (*Réflexivité*) On a $a - a = 0 = 0.n \in n\mathbb{Z}$, donc $a \equiv a \bmod n$.
- (*Symétrie*) On suppose $a \equiv b \bmod n$. Alors il existe $k \in \mathbb{Z}$ tel que $b - a = kn$. On a :
$$a - b = -(b - a) = (-k)n \in n\mathbb{Z},$$
donc $b \equiv a \bmod n$.
- (*Transitivité*) On suppose $a \equiv b \bmod n$ et $b \equiv c \bmod n$. Alors il existe $k, k' \in \mathbb{Z}$ tel que $b - a = kn$ et $c - b = k'n$. Par suite,
$$c - a = (c - b) + (b - a) = (k + k')n \in n\mathbb{Z},$$
donc $a \equiv c \bmod n$.

Il en résulte que la relation de congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$. Montrons de plus qu'elle est compatible avec l'addition et la multiplication :

Soit $a, b, c, d \in \mathbb{Z}$ tels que $a \equiv b \bmod n$ et $c \equiv d \bmod n$. Alors il existe $k, k' \in \mathbb{Z}$ tel que $b - a = nk$ et $d - c = nk'$. On a alors :
$$(b + d) - (a + c) = (b - a) + (d - c) = n(k + k') \in n\mathbb{Z},$$
donc $a + c \equiv b + d \bmod n$.

Et on a :
$$bd - ac = (b - a)c + (d - c)b = n(kc + k'b) \in n\mathbb{Z},$$
donc $ac \equiv bd \bmod n$. □

b. L'ensemble $\mathbb{Z}/n\mathbb{Z}$

Notation 1. *Classes d'équivalence modulo n*

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

— Pour $k \in \mathbb{Z}$, on note $\overline{k} = \{x \in \mathbb{Z} \mid x \equiv k \bmod n\}$ la classe d'équivalence de k pour la relation de congruence modulo n ;

— On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes d'équivalence de la relation de congruence modulo n.

Remarque 1.

— On a $\overline{0} = n\mathbb{Z}, \quad \overline{1} = 1 + n\mathbb{Z}, \quad \ldots, \quad \overline{k} = k + n\mathbb{Z}$.

— Soit $\alpha \in \mathbb{Z}/n\mathbb{Z}$ est une classe d'équivalence pour la relation de congruence modulo n, si k est un entier tel que $k \in \alpha$, alors $\alpha = \overline{k}$. On dit alors que k est **représentant de la classe** α.

Exercice 4.

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

1. Écrire une description explicite de l'ensemble $\overline{k}$ pour $k \in \mathbb{Z}$ fixé.
2. Montrer que pour tous $x, y \in \mathbb{Z}$, $x \equiv y \mod n$ si, et seulement si, $\overline{x} = \overline{y}$.
3. Montrer que deux classes d'équivalences sont soit disjointes, soit égales.

Démonstration.

1.
$$\begin{aligned}\overline{k} &= \{x \in \mathbb{Z} \mid k \equiv x \mod n\}\\ &= \{x \in \mathbb{Z} \mid x - k \in n\mathbb{Z}\}\\ &= \{x \in \mathbb{Z} \mid \exists p \in \mathbb{Z},\ x = k + pn\}\\ &= \{k + pn \mid p \in \mathbb{Z}\}\\ &=: k + n\mathbb{Z}\end{aligned}$$

2. Si $x \equiv y \mod n$, alors il existe $p \in \mathbb{Z}$ tel que $y - x = pn$. Par suite, pour tout $q \in \mathbb{Z}$,
$$y + qn = x + pn + qn = x + (p+q)n \in x + n\mathbb{Z} = \overline{x},$$
et
$$x + qn = y - pn + qn = x + (q-p)n \in y + n\mathbb{Z} = \overline{y},$$
d'où $\overline{y} \subset \overline{x}$ et $\overline{x} \subset \overline{y}$. Donc $\overline{x} = \overline{y}$.
Réciproquement, si $\overline{x} = \overline{y}$, alors en particulier, $x = x + 0n \in \overline{x} = \overline{y}$, donc par définition, $x \equiv y \mod n$.

3. Soit $x, y \in \mathbb{Z}$. On suppose $\overline{x} \cap \overline{y} \neq \emptyset$. Alors il existe $k = \overline{x} \cap \overline{y}$, donc $k \equiv x \mod n$ et $k \equiv y \mod n$. Par symétrie et transitivité, on a alors
$$x \equiv y \mod n.$$
Par suite, d'après le résultat précédent, $\overline{x} = \overline{y}$.

□

Proposition 5.

Soit $n \in \mathbb{N}^*$. Alors $\mathbb{Z}/n\mathbb{Z}$ est un ensemble fini de cardinal n et on a :
$$\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, ..., \overline{n-1}\}.$$

Démonstration.

$\mathbb{Z}/n\mathbb{Z}$ étant l'ensemble des classes d'équivalence de la relation de congruence modulo n, pour tout $k \in \mathbb{Z}$, $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$, donc
$$\{\overline{0}, \overline{1}, ..., \overline{n-1}\} \subset \mathbb{Z}/n\mathbb{Z}.$$
Montrons que $\mathbb{Z}/n\mathbb{Z} \subset \{\overline{0}, \overline{1}, ..., \overline{n-1}\}$. Soit $\alpha \in \mathbb{Z}/n\mathbb{Z}$ et k un représentant de α (alors $\alpha = \overline{k}$).

On a, par division euclidienne, $k = nq + r$ où $q \in \mathbb{Z}$ et $r \in \mathbb{N}$ avec $r \in [\![0, n-1]\!]$. Alors

$$k \equiv r \mod n.$$

Donc $\alpha = \overline{k} = \overline{r} \in \mathbb{Z}/n\mathbb{Z}$. □

c. Le groupe $\mathbb{Z}/n\mathbb{Z}$

Théorème 3. *Structure de groupe sur $\mathbb{Z}/n\mathbb{Z}$*

Soit $n \in \mathbb{N}^*$. Il existe sur $\mathbb{Z}/n\mathbb{Z}$ une loi de composition interne notée $+$ et appelée **loi additive quotient** telle que, pour tous $x, y \in \mathbb{Z}$,

$$\overline{x} + \overline{y} = \overline{x + y}.$$

Muni de cette loi, $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe commutatif où :

— l'élément neutre est $\overline{0}$;

— pour $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$, $-\overline{k} = \overline{-k}$.

Démonstration.

Soit $\alpha, \beta \in \mathbb{Z}/n\mathbb{Z}$. Si $p, p' \in \mathbb{Z}$ sont des représentants de α et q, q' des représentants de β, alors $p \equiv p' \mod n$ et $q \equiv q' \mod n$, donc :

$$p + q \equiv p' + q' \mod n.$$

Ainsi, on peut poser $\alpha + \beta := \overline{p + q}$ car la classe de $\overline{p + q}$ ne dépend pas du choix des représentants p et q de α et β respectivement.

Vérifions alors que muni de cette opération, $\mathbb{Z}/n\mathbb{Z}$ est bien un groupe.
Soit $a = \overline{x}, \beta = \overline{y}, \gamma = \overline{z} \in \mathbb{Z}/n\mathbb{Z}$.

— (*Associativité*). On a :

$$\begin{aligned}(\overline{x} + \overline{y}) + \overline{z} &= \overline{x + y} + \overline{z} \\ &= \overline{(x + y) + z} \\ &= \overline{x + (y + z)} \\ &= \overline{x} + \overline{y + z} \\ &= \overline{x} + (\overline{y} + \overline{z})\end{aligned}$$

— (*Élément neutre*). On a :

$$\overline{0} + \overline{x} = \overline{0 + x} = \overline{x} = \overline{x + 0} = \overline{x} + \overline{0}.$$

— (*Symétrique*). On a :

$$\overline{x} + \overline{-x} = \overline{x + (-x)} = \overline{0} = \overline{-x + x} = \overline{-x} + \overline{x}.$$

□

Proposition 6.

Soit $n \in \mathbb{N}^*$. L'application $\pi_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ telle que pour $k \in \mathbb{Z}$,

$$\pi_n(k) = \overline{k}$$

est un morphisme surjectif de groupe.

Démonstration.

Soit $x, y \in \mathbb{Z}$. On a, par définition de l'addition sur $\mathbb{Z}/n\mathbb{Z}$:

$$\pi_n(x+y) = \overline{x+y} = \overline{x} + \overline{y} = \pi_(x) + \pi_n(y).$$

Donc π_n est un morphisme de groupes.
De plus, pour tout $\alpha \in \mathbb{Z}/n\mathbb{Z}$, si k est un représentant de α, alors k est un antécédent de α par π_n car $\alpha = \overline{k}$. Donc π_n est surjective. $\square$

Exercice 5.

Soit $n \in \mathbb{N}^*$. Déterminer le noyau de $\pi_n : k \mapsto \overline{k}$ de $\mathbb{Z}$ dans $\mathbb{Z}/n\mathbb{Z}$.

Correction.

On a, pour $k \in \mathbb{Z}$, $\overline{k} = \overline{0}$ si, et seulement si, $k \equiv 0 \mod n$, i.e. $k = k - 0 \in nZ$. Ainsi, $\mathrm{Ker}(\pi_n) = n\mathbb{Z}$.

3. Groupes monogènes

a. Généralités et exemples

Par mesure de simplicité, pour G un groupe et $x \in G$, on notera $\langle x \rangle$ en lieu et place de $\langle \{x\} \rangle$ pour désigner le sous-groupe engendré par le singleton $\{x\}$.

Définition 3. *Groupe monogène*

Soit G un groupe. On dit que G est monogène s'il est engendré par un seul élément i.e. s'il existe $x \in G$ tel que :

$$\langle x \rangle = G.$$

Dans ce cas, on dira que l'élément est un **générateur** de G ou encore que G est **engendré** par x.

Proposition 7.

Soit $(G, .)$ un groupe et $x \in G$. Alors :

$$\langle x \rangle = \{x^k \mid k \in \mathbb{Z}\}.$$

Démonstration.

On montre directement cette proposition mais on aura pu bien-sûr utiliser la proposition 2 pour conclure plus efficacement.

Montrons que $E(x) = \{x^k \mid k \in \mathbb{Z}\}$ est un sous-groupe de G. Comme G est un groupe, $E(x) \subset G$; de plus :

i) $e = x^0 \in E(x)$;

ii) Soit $y, z \in E(x)$. Alors il existe $k, k' \in \mathbb{Z}$ tels que $y = x^k$ et $z = x^{k'}$, et on a :

$$yz^{-1} = x^k x^{-k'} = x^{k-k'} \in E(x).$$

Donc $E(x)$ est un sous-groupe de G et il contient $\{x\}$: en effet, $x = x^1 \in E(x)$.
Par suite, $\langle x \rangle \subset E(x)$.

Réciproquement : soit $y \in E(x)$. Alors il existe $k \in \mathbb{Z}$ tel que $y = x^k$. Or $\langle x \rangle$ est un sous-groupe de G et $x \in \langle x \rangle$, donc $y = x^k \in \langle x \rangle$. Ainsi, $E(x) = \langle x \rangle$.

Il en résulte que $\langle x \rangle = \{x^k \mid k \in \mathbb{Z}\}$. □

Remarque 2.

Attention, pour la notation additive $(G, +)$ cette égalité devient :

$$\langle x \rangle = \{kx \mid k \in \mathbb{Z}\}.$$

Exemple 2. *Groupes monogènes classiques*

— pour $n \in \mathbb{N}^*$, $\mathbb{Z}/n\mathbb{Z}$ est engendré par $\overline{1}$, i.e. $\mathbb{Z}/n\mathbb{Z} = \langle \overline{1} \rangle$;

— pour $n \in \mathbb{N}^*$, le groupe $\mathbb{U}_n$ des racines n-ièmes de l'unité est engendré par $e^{i\frac{2\pi}{n}}$, i.e. $\mathbb{U}_n = \langle e^{i\frac{2\pi}{n}} \rangle$;

Exercice 6.

1. Montrer que $(\mathbb{Z}^2, +)$ est un groupe commutatif.
2. Est-il monogène ? Sinon, donner un ensemble minimal (de cardinal le plus petit possible) qui engendre $\mathbb{Z}^2$.

Correction.

1. $(\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}, +)$ un groupe comme le produit des groupes $\mathbb{Z}$ par $\mathbb{Z}$ chacun muni de l'addition. De plus, $(\mathbb{Z}, +)$ est commutatif, donc le produit de $\mathbb{Z}$, par $\mathbb{Z}$ l'est aussi.
2. Non, il n'est pas monogène. En effet, pour tout $(n, m) \in \mathbb{Z}^2$, on a par exemple $(n+1, m) \notin \langle (n, m) \rangle$ donc $\langle (n, m) \rangle \neq \mathbb{Z}^2$.
   La paire $\{(1, 0), (0, 1)\}$ engendre $\mathbb{Z}^2$.

Définition 4. *Groupe cyclique*

Soit G un groupe. On dit que G est **cyclique** s'il est monogène et fini.

Question 1.

Parmi les groupes de l'exemple précédent, lesquels sont cycliques ?

Correction.

$\mathbb{Z}/n\mathbb{Z}$ et $\mathbb{U}_n$.

b. Classification des groupes monogènes

Proposition-Notation 8.

Soit G un groupe et $x \in G$. Alors l'application notée $\varphi_x : \mathbb{Z} \to G$ telle que, pour $k \in \mathbb{N}$:

$$\varphi_x(k) = x^k$$

est un morphisme de groupes.

Démonstration.

Soit G un groupe et $x \in G$. Pour $k, k' \in \mathbb{Z}$, on a :

$$\varphi_x(kk') = x^{k+k'} = x^k x^{k'} = \varphi_x(k)\varphi_x(k')$$

donc φ_k est un morphisme de groupes. □

Théorème 4. *Classification des groupes monogènes*

— Tout groupe monogène *infini* est isomorphe à $(\mathbb{Z}, +)$;
— Tout groupe monogène *de cardinal* $n \in \mathbb{N}^*$ est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

Démonstration.

Soit G un groupe monogène. Alors il existe $x \in G$ tel que $G = \langle x \rangle = \{x^k \mid k \in \mathbb{Z}\}$. Considérons le morphisme de φ_x de la proposition précédente. On a

$$\mathrm{Im}(\varphi_x) = \{\varphi_x(k) \mid k \in \mathbb{Z}\} = \{x^k \mid k \in \mathbb{Z}\} = G;$$

donc φ_x est un morphisme surjectif.
Comme φ_x est un morphisme de groupes, alors $\mathrm{Ker}(\varphi_x)$ est un sous-groupe de $\mathbb{Z}$. Ainsi, il existe $n \in \mathbb{N}$ tel que $\mathrm{Ker}(\varphi_x) = n\mathbb{Z}$. On a donc l'alternative suivante :

- $n = 0$. Alors $\mathrm{Ker}(\varphi_x) = \{0\}$, d'où φ_x est injective, et donc φ_x est un isomorphisme de groupes. Ainsi, G est isomorphe à $(\mathbb{Z}, +)$.
- $n > 0$. Alors $\mathrm{Ker}(\varphi_x) = n\mathbb{Z} \neq \{0\}$, donc φ_x n'est pas injective.
  Soit $\alpha \in \mathbb{Z}/n\mathbb{Z}$ et $p, q \in \alpha$. Alors on a $p - q \in n\mathbb{Z} = \mathrm{Ker}(f_x)$ donc

  $$\varphi_x(p) = \varphi_x(q).$$

  Ainsi, φ_x est constante sur chaque classe d'équivalence de $\mathbb{Z}/n\mathbb{Z}$. Par suite, on peut définir l'application

  $$\tilde{\varphi_x} : \left| \begin{array}{ccll} \mathbb{Z}/n\mathbb{Z} & \to & \mathrm{G} & \\ \alpha & \mapsto & x^p = \varphi_x(p) & \text{où } p \in \alpha \end{array} \right.$$

  Alors
  - $\tilde{\varphi_x}$ est un morphisme de groupes. En effet, pour tous $p, q \in \mathbb{Z}$,

    $$\tilde{\varphi_x}(\overline{p} + \overline{q}) = \tilde{\varphi_x}(\overline{p+q}) = \varphi_x(p+q) = x^{p+q} = x^p x^q = \varphi_x(p)\varphi_x(q) = \tilde{\varphi_x}(\overline{p})\tilde{\varphi_x}(\overline{q}).$$

  - $\tilde{\varphi_x}$ est surjectif. En effet, pour tout $y \in \langle x \rangle$, il existe $k \in \mathbb{Z}$ tel que $y = x^k$ et

    $$y = x^k = f_x(k) = \tilde{\varphi_x}(\overline{k}).$$

    Donc $\overline{k}$ est un antécédent de y par $\tilde{\varphi_x}$.
  - $\tilde{\varphi_x}$ est injectif. En effet, si $\overline{k} \in \mathrm{Ker}(\tilde{\varphi_x})$, alors

    $$e = \tilde{\varphi_x}(\overline{k}) = \varphi_x(k),$$

    donc $k \in \mathrm{Ker}(\varphi_x) = n\mathbb{Z}$. Par suite, $k \equiv 0 \mod n$, d'où $\overline{k} = \overline{0}$.
    Il en résulte que $\mathrm{Ker}(\tilde{\varphi_x}) = \{\overline{0}\}$.

  Par suite, $\tilde{\varphi_x}$ est un isomorphisme de groupes. Ainsi, G est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

$\square$

Exemple 3.

Soit $n \in \mathbb{N}^*$. Le groupe $(\mathbb{U}_n, .)$ est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$.

> En effet, $\mathbb{U}_n = \langle e^{i\frac{2\pi}{n}} \rangle$ et $\#\mathbb{U}_n = n$; donc, d'après le théorème précédent, $(\mathbb{U}_n, .)$ est isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$. De plus, l'isomorphisme de $\mathbb{Z}/n\mathbb{Z}$ dans $\mathbb{U}_n$ est donné par :
>
> $$\overline{k} \mapsto e^{i\frac{2k\pi}{n}}.$$

Exercice 7.

Montrer que $G = \left\{ \begin{pmatrix} 3^n & n3^{n-1} \\ 0 & 3^n \end{pmatrix} \mid n \in \mathbb{Z} \right\}$ est un sous-groupe de $GL_2(\mathbb{R})$ isomorphe à $\mathbb{Z}$.

Correction.

On remarque que $G = \left\langle \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \right\rangle$. Donc G est monogène. De plus, G est infini, donc G est isomorphe à Z.

c. Les générateurs de $\mathbb{Z}/n\mathbb{Z}$

Théorème 5. *Théorème de Bézout*

Soit $n, m \in \mathbb{Z}$. Alors n et m sont premiers entre eux si, et seulement si, il existe $u, v \in \mathbb{Z}$ tels que

$$un + vm = 1.$$

Démonstration.

Vue en Sup. $\square$

Exercice 8.

Soit $n, m \in \mathbb{Z}$ et $d = \operatorname{pgcd}(n, m)$. Déduire du théorème de Bézout qu'il existe $u, v \in \mathbb{Z}$ tels que $un + vm = d$.

Correction.

d est le plus grand diviseur commun de n, m donc il existe $p, q \in \mathbb{Z}$ premiers entre eux tels que $n = dp$ et $m = dq$. D'après le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tels que $up + vq = 1$; donc on a, en multipliant la précédente égalité par d :

$$un + vm = d.$$

Proposition 9.

Soit $n \in \mathbb{N}$ et $k \in \mathbb{Z}$. Alors $\overline{k}$ est un générateur de $\mathbb{Z}/n\mathbb{Z}$ si, et seulement si, k et n sont premiers entre eux.

Démonstration.

- ($\Rightarrow$). On suppose que $\mathbb{Z}/n\mathbb{Z} = \langle \overline{k} \rangle$. Alors il existe $u \in \mathbb{Z}$ tel que

  $$u\overline{k} = \overline{1}.$$

  Par suite, $uk \equiv 1 \bmod n$, et donc, il existe $v \in \mathbb{Z}$ tel que

  $$uk + vn = 1.$$

  D'après le théorème de Bézout, il en résulte que k et n sont premiers entre eux.

- ($\Leftarrow$). On suppose que k et n sont premiers entre eux. D'après le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tels que :
$$uk + vn = 1.$$
Par suite, pour tout $p \in [\![0, n-1]\!]$,
$$p = puk + pvn,$$
d'où, si on note $q = pu$, $p \equiv puk = qk \mod n$ i.e.
$$\overline{p} = q\overline{k}.$$
Par suite, $\overline{k}$ engendre $\mathbb{Z}/n\mathbb{Z}$.

□

4. Ordre d'un élément

Définition 5. *Ordre d'un élément*

Soit G un groupe et $x \in G$. On dit que x est **d'ordre fini** si le cardinal de $\langle x \rangle$ est fini.
Dans ce cas, on appelle **ordre de** x et on note $o(x)$ le nombre entier naturel :
$$o(x) = \#\langle x \rangle.$$

Proposition 10.

Soit G un groupe et $x \in G$ un élément d'ordre fini $d \in \mathbb{N}^*$.

i) Le nombre d est le plus petit entier naturel non nul qui vérifie l'égalité $x^n = e$.

ii) Pour tout $n \in \mathbb{Z}$, $x^n = e$ si, et seulement si, d divise n.

Démonstration.

Comme $\langle x \rangle = \{e, x, ..., x^{d-1}\}$ est un groupe monogène de cardinal d, on peut considérer l'isomorphisme $\phi_x : \overline{k} \mapsto x^k$ de $\mathbb{Z}/d\mathbb{Z}$ dans $\langle x \rangle$.

i) On a
$$x^d = \phi_x(\overline{d}) = \phi_x(\overline{0}) = x^0 = e.$$
De plus, si $n \in \mathbb{N}^*$ est tel que $x^n = e$, alors $n > d-1$ car pour tout $i \in [\![1, d-1]\!]$, $x^i \neq e$. Donc $n \geq d$.

ii) Soit $n \in \mathbb{Z}$.
   - ($\Rightarrow$). On suppose $x^n = e$. Alors $\phi_x(n) = e$ donc $\overline{n} \in \mathrm{Ker}(\phi_x) = \{\overline{0}\}$. Ainsi $n \equiv 0 \mod d$ donc $d|n$.
   - ($\Leftarrow$). On suppose $d|n$. Alors il existe $k \in \mathbb{Z}$ tel que $n = dk$, donc
$$x^n = x^{dk} = (x^d)^k = e^k = e.$$

□

Exercice 9.

Soit G un groupe et x un élément d'ordre fini k.

1. Soit $n \in \mathbb{N}$ tel que $n|k$. Quel est l'ordre de x^n ?
2. Soit $p \in \mathbb{N}$. Quel est l'ordre de x^p ?

Correction.

1. Il existe $q \in \mathbb{N}$ tel que $k = nq$. Alors on a, pour tout $m \in \mathbb{Z}$:
$$(x^n)^m = e \Leftrightarrow x^{nm} = e \Leftrightarrow k|nm \Leftrightarrow nq|nm \Leftrightarrow q|m$$
Et de plus $(x^n)^q = e$ donc q est le plus petit entier naturel m non nul tel que $(x^n)^m = e$. Donc $o(x^n) = q = k/n$.
2. Considérons $d = \text{pgcd}(p, k)$. Montrons que $\langle x^d \rangle = \langle x^p \rangle$.
   — On a $\langle x^d \rangle \subset \langle x^p \rangle$. En effet, d'après le théorème de Bézout, il existe $u, v \in \mathbb{Z}$ tels que
   $$up + vk = d.$$
   Alors
   $$x^d = x^{up+vk} = (x^p)^u \underbrace{(x^k)^v}_{=e} = (x^p)^u \in \langle x^p \rangle.$$
   Donc $\langle x^d \rangle \subset \langle x^p \rangle$
   — On a $\langle x^p \rangle \subset \langle x^d \rangle$. En effet, $d|p$ donc il existe $q \in \mathbb{Z}$ tel que $p = dq$ et donc :
   $$x^p = x^{dq} = (x^d)^q \in \langle x^d \rangle.$$
   Donc $\langle x^p \rangle \subset \langle x^d \rangle$
   Ainsi, $o(x^p) = \#\langle x^p \rangle = \#\langle x^d \rangle = k/d (= ppcm(p, k)/p)$.

Proposition 11.

Soit G un groupe fini. Alors tout élément x de G est d'ordre fini et $o(x)$ divise $\#G$.

Démonstration.

Pour tout $x \in G$, on a $\langle x \rangle \subset G$, donc $o(x) = \#\langle x \rangle \leq \#G$ d'où x est d'ordre fini. Soit $x \in G$.

Montrons $o(x)$ divise $\#G$. *Démonstration dans le cas où G est commutatif* - voire TD pour la démonstration du cas général (non exigible).

Notons $n = \#G$. L'application $g \mapsto xg$ est une bijection de G dans G (en effet, $g \mapsto x^{-1}g$ est la réciproque de cette application) donc, on a, par le changement bijectif d'indice $g = xh$:
$$\prod_{g \in G} g = \prod_{h \in G} xh = x^n \prod_{h \in G} h,$$
Donc $x^n = e$. Par suite, $o(x)|n$. □

Partie $**$

Rappels de Sup' sur les anneaux

1. Structure d'anneau

a. Définitions et exemples

Définition $ $1.** Anneau

Soit A un ensemble muni de deux lois de composition interne sur $+$ et $\cdot$. On dit que le triplet $(A, +, \cdot)$ est **une structure d'anneau**, ou plus simplement A est un **anneau** (muni des lois $+$ et $\cdot$), si :

i) $(A, +)$ est un groupe *commutatif* d'élément neutre 0_A ;

ii) la loi $\cdot$ est associative ;

iii) *Distributivité* : pour tous $a, b, c \in A$,

$$a.(b + c) = a.b + a.c \quad \text{et} \quad (b + c).a = b.a + c.a$$

iv) *Unité* : la loi $\cdot$ possède un élément neutre noté 1_A et appelé **unité de** A.

On dit de plus qu'un anneau A est **commutatif** si la loi $\cdot$ est commutative.

Définition $ $2.** Corps

Un anneau $(A, +, \cdot)$ commutatif tel que $(A \setminus \{0_A\}, \cdot)$ est un groupe est appelé un **corps**.

Remarque $ $1.**

— Un anneau A est dit trivial si $0_A = 1_A$. Dans ce cas $A = \{0_A\}$.

— Il découle de la définition qu'un corps ne peut pas être un anneau trivial.

Exemple $ $1.**

— $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont des anneaux commutatifs munis de l'addition et de la multiplication des nombres. Les anneaux $\mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont même des corps munis de ces opérations.

— $\mathbb{R}[X]$, $\mathbb{C}[X]$ sont des anneaux commutatifs munis de l'addition et de la multiplication des polynômes.

— Soit $n \in \mathbb{N}$, $M_n(\mathbb{R})$ et $M_n(\mathbb{C})$ sont des anneaux non commutatifs (sauf pour $n = 1$) munis de l'addition et de la multiplication des matrices.

— Soit E un espace vectoriel. $\mathcal{L}(E)$ est un anneau non commutatif (sauf si E est de dimension inférieure ou égale à 1) muni de l'addition et de la composition des applications.

b. Anneaux intègres

Définition **3. Anneau intègre

Soit A un anneau. On dit que A est **intègre** si pour tous $a, b \in A$,

$$a.b = 0_A \quad \Rightarrow \quad a = 0_A \text{ ou } b = 0_A.$$

Remarque **2.

— Dans un anneau, un élément $a \neq 0_A$ est un **diviseur de zéro** s'il existe $b \neq 0_A$ tel que $a.b = 0_A$.

— Dans un anneau intègre, tout élément $a \neq 0_A$ est **régulier** pour la loi $\cdot$ i.e. pour tous $x, y \in A$, $ax = ay \Rightarrow x = y$.

Exercice **1.

1. Parmi les exemples d'anneaux précédents, lesquels sont intègres et lesquels ne le sont pas ?

2. Donner un exemple d'anneau commutatif non trivial qui n'est pas intègre.

Correction.

1. — $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ et $\mathbb{C}$ sont intègres.
   — $\mathbb{R}[X]$, $\mathbb{C}[X]$ sont intègres.
   — Soit $n \in \mathbb{N}$, $M_n(\mathbb{R})$ et $M_n(\mathbb{C})$ ne sont pas intègres.
   — Soit E un espace vectoriel. $\mathcal{L}(E)$ n'est pas intègre.
2. On considère $\mathcal{F}(\mathbb{R}, \mathbb{R})$ muni de la l'addition et de la multiplication des fonctions. Alors, $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +, \cdot)$ est un anneau commutatif mais n'est pas intègre : si on considère $A, B \subset \mathbb{R}$ non vides et disjoints, alors le produit des fonctions indicatrices de A et B est la fonction nulle.

2. Sous-anneaux

Définition **4. Sous-anneau

Soit $(A, +, \cdot)$ un anneau et $B \subset A$. On dit que B est un **sous-anneau** de A si :

i) B un sous-groupe de $(A, +)$;

ii) B est stable par $\cdot$ i.e. pour tout $a, b \in B$, $a.b \in B$.

iii) L'unité 1_A de A appartient à B.

Remarque ⁎⁎3.

Si $(A,+,\cdot)$ est un anneau et $B \subset A$ est un sous-anneau de A, alors $(B,+,\cdot)$ est un anneau.

Définition ⁎⁎5. Sous-corps

Soit $(K,+,\cdot)$ un corps et $L \subset K$. On dit que L est un **sous-corps de** K si L est un sous-anneau de K qui est un corps.

Proposition ⁎⁎1. Caractérisation des sous-anneaux

Soit $(A,+,\cdot)$ un anneau et $B \subset A$. Alors B est un sous-anneau de A si, et seulement si,

i) $1_A \in B$;

ii) pour tous $x, y \in B$, $x - y \in B$;

iii) pour tous $x, y \in B$, $x.y \in B$.

Démonstration.

- $(\Rightarrow)$. Immédiat ;
- $(\Leftarrow)$. Il suffit de montrer que $0_A \in B$. On a, d'après i), $1_A \in B$ et d'après ii)

$$0_A = 1_A - 1_A \in B.$$

□

Exemple ⁎⁎2.

— $\mathbb{R}$ est un sous-corps de $\mathbb{C}$, $\mathbb{Q}$ est un sous-corps de $\mathbb{R}$ et $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$.

— L'ensemble des matrices diagonales est un sous-anneau de $M_n(\mathbb{R})$.

Exercice ⁎⁎2.

1. Montrer que $\mathbb{Z} + i\mathbb{Z} = \{n + im \mid n, m \in \mathbb{Z}\}$ est un sous-anneau de $\mathbb{C}$;
2. Montrer que $\mathbb{Q} + \sqrt{2}\mathbb{Q} = \{p + \sqrt{2}q \mid p, q \in \mathbb{Q}\}$ est un sous-corps de $\mathbb{R}$.

Correction.

1. i) On a $1 = \underbrace{1}_{\in\mathbb{Z}} + i.\underbrace{0}_{\in\mathbb{Z}} \in \mathbb{Z} + i\mathbb{Z}$.

   ii) Soit $x = n + im, y = n' + im' \in \mathbb{Z} + i\mathbb{Z}$. On a :

$$\begin{aligned} x - y &= n + im - (n' + im') \\ &\underbrace{n - n'}_{\in\mathbb{Z}} + i\underbrace{(m - m')}_{\in\mathbb{Z}} \end{aligned}$$

   Donc $x - y \in \mathbb{Z} + i\mathbb{Z}$.

iii) Soit $x = n + im, y = n' + im' \in \mathbb{Z} + i\mathbb{Z}$. On a :

$$\begin{aligned} x.y &= (n+im).(n'+im') \\ & \underbrace{nn' - mm'}_{\in\mathbb{Z}} + i \underbrace{(nm' + n'm)}_{\in\mathbb{Z}} \end{aligned}$$

Donc $x.y \in \mathbb{Z} + i\mathbb{Z}$.

Il en résulte que $\mathbb{Z} + i\mathbb{Z}$ est un sous-anneau de $\mathbb{C}$.

2. i) On a $1 = \underbrace{1}_{\in\mathbb{Q}} + \sqrt{2}.\underbrace{0}_{\in\mathbb{Q}} \in \mathbb{Q} + \sqrt{2}\mathbb{Q}$.

ii) Soit $x = p + \sqrt{2}q, y = p' + \sqrt{2}q' \in \mathbb{Q} + \sqrt{2}\mathbb{Q}$. On a :

$$\begin{aligned} x - y &= p + \sqrt{2}q - (p' + \sqrt{2}q') \\ & \underbrace{p - p'}_{\in\mathbb{Q}} + \sqrt{2}\underbrace{(q - q')}_{\in\mathbb{Q}} \end{aligned}$$

Donc $x - y \in \mathbb{Q} + \sqrt{2}\mathbb{Q}$.

iii) Soit $x = p + \sqrt{2}q, y = p' + \sqrt{2}q' \in \mathbb{Q} + \sqrt{2}\mathbb{Q}$. On a :

$$\begin{aligned} x.y &= (p + \sqrt{2}q).(p' + \sqrt{2}q') \\ & \underbrace{pp' + 2qq'}_{\in\mathbb{Q}} + \sqrt{2}\underbrace{(pq' + p'q)}_{\in\mathbb{Q}} \end{aligned}$$

Donc $x.y \in \mathbb{Q} + \sqrt{2}\mathbb{Q}$.

Il en résulte que $\mathbb{Q}+\sqrt{2}\mathbb{Q}$ est un sous-anneau de $\mathbb{R}$. Montrons que $\mathbb{Q}+\sqrt{2}\mathbb{Q}$ est un corps. Soit $x = p+\sqrt{2}q \in \mathbb{Q}+\sqrt{2}\mathbb{Q}$ avec $p, q \neq 0$. Alors en particulier, $x \in \mathbb{R}^*$: en effet, $\mathbb{Q}+\sqrt{2}\mathbb{Q} \subset \mathbb{R}$ et $p+\sqrt{2}q \neq 0$ car sinon le rationnel p serait égal à l'irrationnel $-\sqrt{2}q$ ce qui est impossible. Ainsi on a :

$$x^{-1} = \frac{1}{x} = \frac{1}{p + \sqrt{2}q} = \frac{p - \sqrt{2}q}{p^2 - 2q^2} = \underbrace{\frac{p}{p^2 - 2q^2}}_{\in\mathbb{Q}} + \sqrt{2}\underbrace{\frac{-q}{p^2 - 2q^2}}_{\in\mathbb{Q}} \in \mathbb{Q} + \sqrt{2}\mathbb{Q}.$$

Par suite, tout élément non nul est inversible. Il en résulte que $\mathbb{Q} + \sqrt{2}\mathbb{Q}$ est un corps et donc un sous-corps de $\mathbb{R}$.

3. Inversibles d'un anneau

Définition $ $6.** Groupe des inversibles

Soit $(A, +, \cdot)$ un anneau. On appelle **groupe des inversibles** (ou **groupe des unités**) et on note $A^\times$ (ou $U(A)$) l'ensemble des éléments inversibles de $(A, \cdot)$.

Proposition $ $2.**

Soit $(A, +, \cdot)$ un anneau. Alors le groupe des inversibles $A^\times$ muni de la loi $\cdot$ est un groupe.

Démonstration.

Montrons que $\cdot$ est une loi de composition *interne* sur $A^\times$: pour tous $x, y \in A^\times$, x, y sont inversibles et

$$(xy).(y^{-1}x^{-1}) = 1_A,$$

donc xy est inversible, d'où $xy \in A^\times$. Par suite $\cdot$ est bien une loi de composition interne sur $x, y \in A^\times$.

i) La loi $\cdot$ est associative car $(A, +, \cdot)$ est un anneau ;

ii) 1_A est l'élément neutre de $\cdot$ donc il suffit de vérifier que $1_A \in A^\times$. En effet on $1_A.1_A = 1_A$ donc 1_A est inversible d'où $1_A \in A^\times$;

iii) Tout élément x de $A^\times$ est inversible par définition et x^{-1} étant également inversible, il appartient à $A^\times$.

Donc $(A^\times, \cdot)$ est un groupe □

Exemple ∗∗3.

— $\mathbb{Z}^\times = \{-1, 1\}$; $\mathbb{R}^\times = \mathbb{R}^*$; $\mathbb{C}^\times = \mathbb{C}^*$;

— $\mathbb{K}[X]^\times = \mathbb{K}^*$;

— $M_n(\mathbb{K})^\times = GL_n(\mathbb{K})$

4. Morphismes d'anneaux

a. Définition

Définition ∗∗7. Morphisme d'anneaux

Soit A, B deux anneaux et $f : A \to B$ une application. On dit que f est un **morphisme d'anneaux** si :

i) pour tous $x, y \in A$, $f(x + y) = f(x) + f(y)$;

ii) pour tous $x, y \in A$, $f(xy) = f(x)f(y)$;

iii) $f(1_A) = 1_B$.

Un morphisme d'anneau bijectif est appelé un **isomorphisme d'anneaux**.

Exercice ∗∗3.

Soit A un anneau non trivial et $u \in A^\times$. Montrer que $\varphi_u : x \mapsto uxu^{-1}$ est un isomorphisme d'anneaux.

b. Noyaux, images et sous-anneaux

Définition ∗∗8. Noyau/Image

Soit A, B des anneaux et $f : A \to B$ un morphisme d'anneaux.
— Le **noyau** de f est le sous-ensemble de A

$$\mathrm{Ker}(f) = \{a \in A \mid f(a) = 0_B\}.$$

— L'**image** de f est le sous-ensemble de B

$$\mathrm{Im}(f) = f(A) = \{f(a) \mid a \in A\}.$$

Proposition ∗∗3.

Soit A_1, A_2 des anneaux et $f : A_1 \to A_2$ un morphisme d'anneaux. Alors $\mathrm{Im}(f)$ est un sous-anneau de A_2.

Démonstration.

i) $1_{A_2} = f(1_{A_1}) \in \mathrm{Im}(f)$;

ii) pour $f(x), f(y) \in \mathrm{Im}(f)$ avec $x, y \in A_1$,

$$f(x) - f(y) = f(x - y) \in \mathrm{Im}(f);$$

iii) pour $f(x), f(y) \in \mathrm{Im}(f)$ avec $x, y \in A_1$,

$$f(x)f(y) = f(xy) \in \mathrm{Im}(f).$$

Donc $\mathrm{Im}(f)$ est un sous-anneau de A_2.

□

Remarque ∗∗4. ATTENTION !

Contrairement au cas des groupes où le noyau d'un morphisme est un sous-groupe, **le noyau d'un morphisme d'anneau n'est JAMAIS un sous-anneau de l'anneau de départ** (à moins que l'anneau d'arrivée ne soit trivial).

En effet, si $B \neq \{0_B\}$, $1_A \notin \mathrm{Ker}(f)$ car $f(1_A) = 1_B \neq 0_B$. Ainsi, $\mathrm{Ker}(f)$ ne peut pas être un sous-anneau puisqu'il ne contient pas 1_A.

Partie B
Compléments sur les anneaux ; idéaux

1. Structure d'anneau produit

Proposition 12. *Structure d'anneau produit*

Soit $(A_1,+,\cdot),(A_2,+,\cdot)$ des anneaux et on note $A=A_1\times A_2$. On considère les lois de composition suivantes sur A : pour $(x_1,x_2),(y_1,y_2)\in A$,

$$(x_1,x_2)+(y_1,y_2):=(x_1+y_1,x_2+y_2) \quad \text{et} \quad (x_1,x_2).(y_1,y_2):=(x_1.y_1,x_2.y_2).$$

Alors A muni de ces lois est un anneau et :

- L'élément nul de A est $0_A=(0_{A_1},0_{A_2})$.
- L'unité de A est $1_A=(1_{A_1},1_{A_2})$.

Démonstration.

— $(A,+)$ est un groupe commutatif d'élément neutre $0_A=(0_{A_1},0_{A_2})$ comme groupe produit des groupes commutatifs $(A_1,+)$ et $(A_2,+)$.

— La loi $\cdot$ est associative par associativité des lois multiplicatives de A_1 et A_2.

— *Distributivité :* Soit $x=(x_1,x_2),y=(y_1,y_2),z=(z_1,z_2)\in A$, alors :

$$\begin{aligned} x.(y+z) &= (x_1,x_2).(y_1+z_1,y_2+z_2) \\ &= (x_1.(y_1+z_1),x_2.(y_2+z_2)) \\ &= (x_1.y_1+x_1.z_1,x_2.y_2+x_2.z_2) \\ &= (x_1.y_1,x_2.y_2)+(x_1.z_1,x_2.z_2) \\ &= (x_1,x_2).(y_1,y_2)+(x_1,x_2).(z_1.z_2) \\ &= x.y+x.z \end{aligned}$$

et de même

$$(y+z).x=y.x+z.x$$

— Pour tous $x=(x_1,x_2)\in A$,

$$\begin{aligned} x.1_A &= (x_1,x_2).(1_{A_1},1_{A_2}) \\ &= (x_1.1_{A_1},x_2.1_{A_2}) \\ &= (x_1,x_2) \\ &= x \\ &= (1_{A_1}.x_1,1_{A_2}.x_2) \\ &= (1_{A_1},1_{A_2}).(x_1,x_2) \\ &= 1_A.x \end{aligned}$$

donc $x.1_A=x=1_A.x$ donc 1_A est l'élément neutre pour la multiplication.

Il en résulte que $(A,+,\cdot)$ est un anneau. $\square$

Remarque 3.

Par récurrence, on peut ainsi munir un produit fini d'anneaux d'une structure d'anneau.

Exercice 10.

Montrer que $A = A_1 \times A_2$ est commutatif si, et seulement si, A_1 et A_2 sont commutatifs.

Correction.

Soit $x_1, y_1 \in A_1$, $x_2, y_2 \in A_2$. On a :

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

si, et seulement si,

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

si, et seulement si,

$$x_1 x_2 = x_2 x_1 \quad \text{et} \quad y_1 y_2 = y_2 y_1.$$

2. Idéaux d'un anneau commutatif

a. Définition et premières propriétés

Définition 6. *Idéal d'un anneau commutatif*

Soit A un anneau commutatif et $I \subset A$. On dit que I est un **idéal** de A si :

i) I est un sous-groupe de $(A, +)$;

ii) I est stable par multiplication par les éléments de A, i.e. pour tout $x \in I$ et tout $a \in A$,

$$ax \in I.$$

Exemple 4.

Soit A, B des anneaux commutatifs.

— A et $\{0_A\}$ sont des idéaux de A.

— Pour $f : A \to B$ un morphisme d'anneaux, $\mathrm{Ker}(f)$ est un idéal de A.

> En effet,
>
> i) $\mathrm{Ker}(f)$ est un sous-groupe du groupe $(A, +)$ comme image réciproque de $\{0_B\}$ par f.
>
> ii) Soit $x \in \mathrm{Ker}(f)$ et $a \in A$; on a :
>
> $$f(ax) = f(a)f(x) = f(a)0_B = 0_B,$$
>
> donc $ax \in \mathrm{Ker}(f)$.

Exercice 11.

Soit A un anneau et I un idéal de A.

1. Montrer que si $1_A \in I$, alors $I = A$.
2. Soit $u \in A^\times$. En déduire que si $u \in I$, alors $I = A$.

Correction.

1. On suppose $1_A \in I$. On a $I \subset A$. Montrons $A \subset I$. Soit $a \in A$. Alors
$$\underbrace{a}_{\in A} \cdot \underbrace{1_A}_{\in I} \in I.$$
Donc $a \in I$; d'où $A \subset I$.
Il en résulte que $I = A$.
2. Si $u \in I$, alors
$$1_A \underbrace{u^{-1}}_{\in A} \cdot \underbrace{u}_{\in I} \in I.$$
Donc d'après la question précédente $I = A$.

Proposition 13. *Image réciproque d'un idéal*

Soit A, B des anneaux commutatifs, $f : A \to B$ un morphisme et J un idéal de B. Alors $f^{-1}(J)$ est un idéal de A.

Démonstration.

i) $f^{-1}(J)$ est un sous-groupe du groupe $(A, +)$ comme image réciproque du sous-groupe J de $(B, +)$ par f.

ii) Soit $x \in f^{-1}(J)$ et $a \in A$; on a :
$$f(ax) = \underbrace{f(a)}_{\in B}\underbrace{f(x)}_{\in J} \in J$$
donc $ax \in f^{-1}(J)$.

Donc $f^{-1}(J)$ est un idéal de A. □

b. Opérations sur les idéaux

Proposition 14. *Somme d'idéaux*

Soit A un anneau commutatif et I, J des idéaux de A. Alors l'ensemble $I + J = \{x + y \mid x \in I\ , y \in J\}$ est un idéal de A

Démonstration.

Montrons que $I+J$ est un sous-groupe de $(A,+)$. On a

$$0_A = \underbrace{0_A}_{\in I} + \underbrace{0_A}_{\in J} \in I+J$$

car I, J sont des sous-groupes de $(A,+)$; et pour tous $x = x_I + x_J, y = y_I + y_J \in I+J$,

$$x - y = x_I + x_J - (y_I + y_J) = \underbrace{(x_I - y_I)}_{\in I} + \underbrace{(x_J - y_J)}_{\in J} \in I+J,$$

car I, J sont des sous-groupes de $(A,+)$;
Donc $I+J$ est un sous-groupe de $(A,+)$

Soit $x = x_I + x_J \in I+J$ et $a \in A$. Par distributivité, on a :

$$ax = a(x_I + x_J) = \underbrace{(ax_I)}_{\in I} + \underbrace{(ax_J)}_{\in J} \in I+J,$$

car I, J sont stables par multiplication par les éléments de A.
Il en résulte que $I+J$ est un idéal de A. □

Remarque 4.

On peut généraliser ce résultat par récurrence : une somme finie d'idéaux est un idéal.

Proposition 15. *Intersection d'idéaux*

Soit A un anneau commutatif et $(I_k)_{k\in K}$ une famille quelconque d'idéaux de A. Alors $\bigcap\limits_{k\in K} I_k$ est un idéal de A.

Démonstration.

$I = \bigcap\limits_{k\in K} I_k$ est un sous-groupe de $(A,+)$ comme intersections de sous-groupes de $(A,+)$.
Soit $x \in I$ et $a \in A$. Alors pour tout $k \in K$, $x \in I_k$ qui est un idéal de A donc $ax \in I_k$ pour tout $k \in K$. Par suite, $ax \in I$.
Il en résulte que I est un idéal de A. □

Définition-Proposition 7.

Soit A un anneau commutatif et $X \subset A$. On appelle **idéal engendré par** X l'ensemble $I = \bigcap_{J\in\mathcal{I}_H} J$ où $\mathcal{I}_X = \{J \text{ idéal de } A \mid X \subset J\}$; autrement dit, I est l'intersection de tous les idéaux contenant X.
L'idéal engendré par X est le plus petit idéal contenant X.

Démonstration.

I est un idéal comme intersection d'idéaux et comme $X \subset J$ pour tout $J \in \mathcal{I}_H$, $X \subset I$.
Par suite, I est le plus petit idéal contenant X ; en effet, I est inclus dans tous les idéaux contenant X car il est défini comme leur intersection. $\square$

Définition-Proposition 8.

Soit A un anneau commutatif et $x \in A$. L'idéal engendré par le singleton $\{x\}$ est égal à l'ensemble :

$$Ax := \{ax \mid a \in A\} \quad (= xA := \{xa \mid a \in A\}).$$

L'élément x est appelé **générateur** de l'idéal Ax engendré par $\{x\}$.

Démonstration.

On note $I_x = \bigcap_{J \in \mathcal{I}_{\{x\}}} J$ l'idéal engendré par $\{x\}$. Montons que $I_x = Ax$.

$\underline{I_x \subset Ax}$:

Comme I_x est contenu dans tous les idéaux contenant x, montrons que Ax est un idéal contenant x :

On a $x = 1_A.x \in Ax$ donc Ax contient x.

On vient de voir que Ax est non vide (on aurait pu également voir que $0_A = 0_A.x \in Ax$) et pour tous $ax, bx \in Ax$, on a, par distributivité de $\cdot$ par rapport à $+$:

$$ax - bx = (a - b)x \in Ax$$

donc Ax est un sous-groupe de $(A, +)$.

De plus, pour tout $b \in A$ et tous $ax \in Ax$, par associativité de $\cdot$:

$$b.(ax) = (ba).x \in Ax$$

Il en résulte que Ax est un sous-anneau de A qui contient x donc $I_x \subset Ax$.

$\underline{Ax \subset I_x}$:

Soit $ax \in Ax$ où $a \in A$. Comme I est un idéal et que x appartient à I_x, par stabilité de I_x par multiplication par les éléments de A, on a $ax \in I_x$. D'où $Ax \subset I_x$.

Conclusion : on a $I_x = Ax = \{ax \mid a \in A\}$. $\square$

Définition 9.

Soit A un anneau commutatif.

— On dit qu'un idéal de A est **principal** s'il est engendré par un singleton.

— On dit que l'anneau A est **principal** si tous les idéaux de A sont principaux.

Exemple 5.

Pour $n \in \mathbb{Z}$, l'ensemble $n\mathbb{Z}$ des multiples de n dans $\mathbb{Z}$ est l'idéal principal engendré par $\{n\}$.

c. Divisibilité dans un anneau commutatif intègre

Définition 10. *Diviseur et multiple*

Soit A un anneau commutatif intègre et $x, y \in A$. On dit que x **divise** y et on note $x|y$ s'il existe $a \in A$ tel que $y = ax$.
Dans ce cas, on dira également que x est **un diviseur de** y ou encore que y est **un multiple de** x.

Proposition 16. *Caractérisation de la divisibilité en terme d'idéaux*

Soit A un anneau commutatif intègre et $x, y \in A$. Alors $x|y$ si, et seulement si, $Ay \subset Ax$.

Démonstration.

- $(\Rightarrow)$. On suppose $x|y$. Alors il existe $a \in A$ tel que $y = ax$. Soit $a'y \in Ay$. Alors
$$a'y = a'ax \in Ax,$$
car A est stable par $\cdot$. Donc $Ay \subset Ax$.
- $(\Leftarrow)$. On suppose $Ay \subset Ax$. Alors en particulier, $y = 1_A.y \in Ax$ donc il existe $a \in A$ tel que $y = ax$. D'où $x|y$.

□

Exercice 12.

Soit A un anneau commutatif intègre et $x, y \in A$.
1. Montrer que $x|y$ et $y|x$ si, et seulement s'il existe $u \in A^\times$ tel que $y = ux$. Dans ce cas, on dit que x et y sont **associés**.
2. Montrer que $Ax = Ay$ si, et seulement si, x et y sont associés.

Correction.

1. - $(\Rightarrow)$. On suppose $x|y$ et $y|x$. Alors il existe $u, v \in \mathbb{Z}$ tels que $y = ux$ et $x = vy$. Ainsi, par exemple, $x = vux$ d'où $x(1_A - vu) = 0_A$. Comme A est intègre, alors
$$x = 0_A \text{ ou } 1_A - vu = 0_A.$$
1er cas : $x = 0_A$. Alors $y = 0_A$ et par exemple, $y = 1_A x$.
2eme cas : $1_A - vu = 0_A$. Alors $vu = 1_A$ donc u est inversible d'inverse v.
Dans tous les cas il existe $u \in A^\times$ tel que $y = ux$.
   - $(\Leftarrow)$. On suppose qu'il existe $u \in A^\times$ tel que $y = ux$. Alors $y = ux$ et $x = u^{-1}y$ donc $x|y$ et $y|x$
2. On a $Ax = Ay$ si, et seulement si, $Ax \subset Ay$ et $Ay \subset Ax$ si, et seulement si, $x|y$ et $y|x$ (d'après la proposition précédente).

d. Exemples : les idéaux de $\mathbb{Z}$

Théorème 6. *Idéaux de $\mathbb{Z}$*

Soit I un idéal de $\mathbb{Z}$. Alors il existe un unique $n \in \mathbb{N}$ tel que $I = n\mathbb{Z}$.

Démonstration.

Soit I un idéal de l'anneau $\mathbb{Z}$. Alors c'est un sous-groupe de $(\mathbb{Z}, +)$. Par suite, il est de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$ et ce n est unique d'après la proposition 3. □

Corollaire 1.

$\mathbb{Z}$ est un anneau principal.

Démonstration.

En effet, d'après le théorème précédent, tout idéal de $\mathbb{Z}$ est de la forme $n\mathbb{Z} = \mathbb{Z}n$ où $n \in \mathbb{N}$; or $n\mathbb{Z}$ est un idéal principal car engendré par le singleton $\{n\}$. □

Proposition 17.

Soit $a, b \in \mathbb{Z}$ non tous nuls.

— Le pgcd d de a et b est l'unique générateur positif de l'idéal $a\mathbb{Z} + b\mathbb{Z}$.
— Le ppcm m de a et b est l'unique générateur positif de l'idéal $a\mathbb{Z} \cap b\mathbb{Z}$.

Démonstration.

— Comme une somme d'idéaux est un idéal, d'après le théorème 6, il existe un unique $d \in \mathbb{N}$ tel que $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. Montrons que d est le pgcd de a et b i.e. le plus petit diviseur positif commun de a et b. On note d' ce pgcd.
On a $a = a \times 1 + b \times 0 \in a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$ et $b = a \times 0 + b \times 1 \in a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$ donc d est un diviseur commun de a et b. Or tout diviseur commun divise le pgcd donc $d|d'$.
De plus, comme $d \in d\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$, il existe $u', v' \in \mathbb{Z}$ tels que $d = au' + bv'$. Or d' étant un diviseur commun de a et b, par combinaison linéaire $d'|d$.
Ainsi, d, d' étant positifs, $d = d'$.
— Comme une intersection d'idéaux est un idéal, d'après le théorème 6, il existe un unique $m \in \mathbb{N}$ tel que $a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$. Montrons que m est le ppcm de a et b i.e. le plus petit multiple positif commun de a et b. On note m' ce ppcm.
Comme m' est un multiple commun de a et b, on $m' \in a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$ donc $m|m'$.
De plus, comme $m \in m\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$, m est un multiple commun de a et b. Or le ppcm divise tout multiple commun donc $m'|m$.
Ainsi, m, m' étant positifs, $m = m'$.

□

Remarque 5.

La proposition précédente nous invite à revoir certaines définitions de bases de l'arithmétique dans $\mathbb{Z}$: on pourrait oublier nos "vieilles" définitions du pgcd et de ppcm et les redéfinir en termes d'idéaux ! Et c'est ce qu'on fera pour les polynômes. Un des avantages de partir des idéaux pour la définition du pgcd est que la relation de Bézout devient immédiate !

3. L'anneau $\mathbb{Z}/n\mathbb{Z}$

a. Structure d'anneau

Théorème 7. *Structure d'anneau sur $\mathbb{Z}/n\mathbb{Z}$*

Soit $n \in \mathbb{N}^*$. Il existe sur $\mathbb{Z}/n\mathbb{Z}$ des lois de composition internes notée $+$ et $\cdot$ appelées respectivement **loi additive quotient** et **loi multiplicative quotient** telles que, pour tous $x, y \in \mathbb{Z}$,

$$\overline{x} + \overline{y} = \overline{x+y} \quad \text{et} \quad \overline{x}.\overline{y} = \overline{x.y}.$$

Muni de ces lois, $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ est un anneau commutatif où l'élément nul est $\overline{0}$ et l'unité est $\overline{1}$.

Démonstration.

On a montré que $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe commutatif. Il s'agit ici de montrer que la multiplication $\cdot$ sur $\mathbb{Z}/n\mathbb{Z}$ est bien définie et qu'elle vérifie les axiomes requis pour la structure d'anneau. $\square$

Proposition 18.

Soit $n \in \mathbb{N}^*$. L'application :

$$\pi_n : \left| \begin{array}{ccc} \mathbb{Z} & \to & \mathbb{Z}/n\mathbb{Z} \\ k & \mapsto & \overline{k} \end{array} \right.$$

est un morphisme surjectif d'anneaux de noyau $\mathrm{Ker}(\pi_n) = n\mathbb{Z}$.

Démonstration.

On a déjà montré que π_n est un morphisme surjectif de groupes de $(\mathbb{Z}, +)$ dans $(\mathbb{Z}/n\mathbb{Z}, +)$ de noyau $\mathrm{Ker}(\pi_n) = n\mathbb{Z}$.
Il reste à montrer que $\pi_n(pq) = \pi_n(p)\pi_n(q)$ pour tous $p, q \in \mathbb{Z}$.
Soit $p, q \in \mathbb{Z}$, on a :

$$\pi_n(pq) = \overline{pq} = \overline{p}\,\overline{q} = \pi_n(p)\pi_n(q).$$

$\square$

b. Les inversibles de $\mathbb{Z}/n\mathbb{Z}$

Proposition 19. *Éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$*

Soit $n \in \mathbb{N}$ et $k \in \mathbb{Z}$. Alors $\overline{k}$ est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si, et seulement, si k est premier avec n.

Démonstration.

$$\overline{k} \text{ est inversible dans } \mathbb{Z}/n\mathbb{Z}$$

si, et seulement si,

$$\text{il existe } \overline{u} \in \mathbb{Z}/n\mathbb{Z} \text{ tel que } \overline{k}\overline{u} = \overline{1}$$

si, et seulement si,

$$\text{il existe } u \in \mathbb{Z} \text{ tel que } ku \equiv 1 \bmod n$$

si, et seulement si,

$$\text{il existe } u, v \in \mathbb{Z} \text{ tels que } ku + nv = 1$$

si, et seulement si,

$$k \text{ et } n \text{ sont premiers entre eux.}$$

□

Corollaire 2.

Soit $n \in \mathbb{N}^*$. On a équivalence entre :

i) n est premier ;

ii) $\mathbb{Z}/n\mathbb{Z}$ est un anneau intègre ;

iii) $\mathbb{Z}/n\mathbb{Z}$ est un corps.

Démonstration.

Démontrons i)⇒iii)⇒ii)⇒i)

- i)⇒iii) Si n est premier, alors pour tout $k \in [\![1, n-1]\!]$, k est premier avec n donc $\overline{k}$ est inversible. Donc $\mathbb{Z}/n\mathbb{Z}$ est un corps
- ii)⇒iii) Si $\mathbb{Z}/n\mathbb{Z}$ est un corps, alors tous ses éléments non nuls sont inversibles et donc sont réguliers. Ainsi, $\mathbb{Z}/n\mathbb{Z}$ est intègre.
- iii)⇒i) Raisonnons par contraposée. On suppose que n n'est pas premier. Si $n = 1$, l'anneau est trivial et donc n'est pas intègre. Supposons $n \geq 2$. Alors $n = pq$ avec $p, q \in [\![2, n-1]\!]$. On a alors $\overline{p} \neq \overline{0}$ et $\overline{q} \neq \overline{0}$ et
$$\overline{pq} = \overline{0},$$
par suite $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre.

□

Notation 2.

Soit p un nombre premier. On note $\mathbb{F}_p$ le corps $\mathbb{Z}/p\mathbb{Z}$.

Exercice 13.

1. Déterminer les éléments inversibles de $\mathbb{Z}/10\mathbb{Z}$ et calculer l'inverse de $\overline{9}$ et $\overline{7}$.
2. Déterminer l'inverse de $\overline{41}$ dans $\mathbb{Z}/152\mathbb{Z}$.

Correction.

1. On a $9 \times 9 = 81 = 1 + 8 \times 10$ donc $\overline{9}.\overline{9} = \overline{1}$. Donc $\overline{9}$ est sa propre inverse dans $\mathbb{Z}/10\mathbb{Z}$.
   On a $7 \times 3 = 21 = 1 + 2 \times 10$ donc $\overline{7}.\overline{3} = \overline{1}$. Donc $\overline{3}$ est l'inverse de $\overline{7}$ dans $\mathbb{Z}/10\mathbb{Z}$.
2. En appliquant l'algorithme d'Euclide pour le calcul du pgcd on obtient 1 et donc $\overline{41}$ est inversible dans $\mathbb{Z}/152\mathbb{Z}$. Ainsi, en remontant l'algorithme d'Euclide, on trouve les coefficients de Bézout suivants :
   $$(-63) \times 41 + 17 \times 152 = 1,$$
   et donc $\overline{-63}.\overline{41} = \overline{1}$. Par suite $\overline{-63} = \overline{89}$ est l'inverse de $\overline{41}$ dans $\mathbb{Z}/152\mathbb{Z}$.

c. Théorème Chinois

Théorème 8. *Théorème Chinois*

Soit $n, m \in \mathbb{N}$ deux entiers premiers entre eux. Alors les anneaux $\mathbb{Z}/(nm)\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ sont isomorphes et l'application :

$$\varphi \left| \begin{array}{ccc} \mathbb{Z}/(nm)\mathbb{Z} & \to & \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \\ \overline{k}^{nm} & \mapsto & \left(\overline{k}^n, \overline{k}^m\right) \end{array} \right.$$

est un isomorphisme d'anneaux.

Démonstration.

Montrons que φ est bien définie : si $p \equiv q \bmod nm$, alors $p - q \in nm\mathbb{Z}$. Or $nm\mathbb{Z} \subset n\mathbb{Z}$ et $nm\mathbb{Z} \subset m\mathbb{Z}$, donc $p \equiv q \bmod n$ et $p \equiv q \bmod m$. Ainsi $\varphi(\overline{p}^{nm}) = \varphi(\overline{q}^{nm})$ donc φ est bien définie.

φ est un morphisme d'anneaux car $\overline{k}^{nm} \mapsto \overline{k}^n$ et $\overline{k}^{nm} \mapsto \overline{k}^m$ sont des morphismes d'anneaux. On a :

$$\mathrm{Ker}(\varphi) = \{\overline{k}^{nm} \mid \overline{k}^n = \overline{0}^n, \overline{k}^m = \overline{0}^m\},$$

Or si $\overline{k}^n = \overline{0}^n$ et $\overline{k}^m = \overline{0}^m$, alors k est un multiple commun de n et m qui sont premiers entre eux, donc k est un multiple de nm i.e. $\overline{k}^{nm} = \overline{0}^{nm}$. D'où $\mathrm{Ker}(\varphi) = \{\overline{0}^{nm}\}$. Par suite φ est injective.

De plus, f est bijective car φ est injective et $\mathbb{Z}/(nm)\mathbb{Z}$ et $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ ont tout deux même cardinal (nm). $\square$

Corollaire 3. *Application du théorème Chinois*

Soit $n, m \in \mathbb{N}$ deux entiers premiers entre eux. Pour tout $a, b \in \mathbb{Z}$, il existe un entier k vérifiant le système :

$$\begin{cases} x \equiv a \bmod n \\ x \equiv b \bmod m \end{cases}$$

et les solutions de ce système sont exactement les entiers congrus à k modulo nm.

Méthode de résolution :

n et m étant premier entre eux, on cherche deux entiers $u, v \in \mathbb{Z}$ tels que $nu + mv = 1$. Ainsi, les entiers $x_1 = nu$ et $x_2 = mv$ vérifient

$$\begin{cases} x_1 \equiv 0 \bmod n \\ x_1 \equiv 1 \bmod m \end{cases} \quad \text{et} \quad \begin{cases} x_2 \equiv 1 \bmod n \\ x_2 \equiv 0 \bmod m \end{cases}$$

Ainsi, $x = bx_1 + ax_2$ est solution du système initial (*Toujours vérifier que ce x est bien solution pour éviter les erreurs dans les calculs précédents !*) ; par suite l'ensemble des solutions est :

$$x + nm\mathbb{Z} = \{x + nmk \mid k \in \mathbb{Z}\}.$$

Exercice 14.

Résoudre les systèmes suivants :

$$(S_1) \begin{cases} x \equiv 1 \bmod 6 \\ x \equiv 4 \bmod 7 \end{cases} \quad \text{et } (S_2) \begin{cases} 3x \equiv 2 \bmod 5 \\ 5x \equiv 1 \bmod 6 \end{cases}$$

Correction.

(S_1) On a $6 \times (-1) + 7 \times 1 = 1$ donc $x_1 = 7 \times 1$ et $x_2 = 6 \times (-1)$ vérifient :

$$\begin{cases} x_1 \equiv 1 \bmod 6 \\ x_1 \equiv 0 \bmod 7 \end{cases} \quad \text{et} \quad \begin{cases} x_2 \equiv 0 \bmod 6 \\ x_2 \equiv 1 \bmod 7 \end{cases}$$

donc $x = 1 \times x_1 + 4 \times x_2 = -17$ est une solution de (S_1). Ainsi l'ensemble des solutions de (S_1) est

$$\{-17 + 42k \mid k \in \mathbb{Z}\}.$$

(S_2) On a 3 et 5 premiers entre eux, donc $\overline{3}$ est inversible dans $\mathbb{Z}/5\mathbb{Z}$ et sont inverse est $\overline{2}$ car $3 \times 2 = 6 = 1 + 5$.

Ainsi, on a $3x \equiv 2 \bmod 5 \Leftrightarrow x \equiv 4 \bmod 5$.

On a 5 et 6 premiers entre eux, donc $\overline{5}$ est inversible dans $\mathbb{Z}/6\mathbb{Z}$ et sont inverse est $\overline{5}$ car $5 \times 5 = 24 = 1 + 4 \times 6$.

Ainsi, on a $5x \equiv 1 \bmod 6 \Leftrightarrow x \equiv 5 \bmod 6$.

d'où :

$$(S_2) \Leftrightarrow \begin{cases} x \equiv 4 \bmod 5 \\ x \equiv 5 \bmod 6 \end{cases}$$

Et on résout comme précédemment.

Exercice 15.

Résoudre l'équation $x^2 + x + 11 \equiv 0 \bmod 143$.

Correction.

On a $143 = 11 \times 13$ et 11 est premier avec 13 donc d'après le théorème chinois, $(*)$ $x^2 + x + 11 \equiv 0 \bmod 143$ si, et seulement si,

$$(1)\ x^2 + x + 11 \equiv 0 \bmod 11 \quad \text{et} \quad (2)\ x^2 + x + 11 \equiv 0 \bmod 13$$

Donc si x_1 est une solution de (1) et x_2 une solution de (2), alors $x = x_1 u + x_2 v$ est solution de $(*)$ où $u, v \in \mathbb{Z}$ sont tels que $11u + 13v = 1$, et toutes les solutions sont de cette forme.
On a

$$(x+1)x = x^2 + x \equiv x^2 + x + 11 \equiv 0 \bmod 11$$

et

$$(x-1)(x+2) = x^2 + x - 2 \equiv x^2 + x + 11 \equiv 0 \bmod 13$$

Donc les solutions de (1) sont $x \equiv 0 \bmod 11$ et $x \equiv -1 \bmod 11$ et les solutions de (2) sont $x \equiv 1 \bmod 13$ et $x \equiv -2 \bmod 13$.
Déterminons les coefficients u et v : on a $6 \times 11 - 5 \times 13 = 1$, d'où $u = 6$ et $v = 5$.
Ainsi, on a donc les solutions suivantes :

$$x \equiv 66 \bmod 143$$

$$x \equiv 11 \bmod 143$$

$$x \equiv -12 \bmod 143$$

$$x \equiv 76 \bmod 143$$

Méthode : Que faire dans le cas d'un système : $\begin{cases} x \equiv a \bmod p \\ x \equiv b \bmod q \end{cases}$ où p et q ne sont pas premiers entre eux ?
Soit $d = \text{pgcd}(p, q)$ et $M = \text{ppcm}(p, q)$. Alors on peut montrer qu'il existe une solution à ce système, si et seulement si $a \equiv b \bmod d$.
Dans ce cas, une solution est donnée par $x = a + qu\frac{b-a}{d}$ où u est un entier qui vérifie $pu + qv = d$ (où $v \in \mathbb{Z}$) et de plus, l'ensemble des solutions forme un classe modulo $M\mathbb{Z}$.

Exercice 16.

Résoudre les systèmes suivants :

$$(S_1)\ \begin{cases} x \equiv 1 \bmod 6 \\ x \equiv 4 \bmod 15 \end{cases} \quad \text{et } (S_2)\ \begin{cases} 3x \equiv 6 \bmod 18 \\ x \equiv 1 \bmod 21 \end{cases}$$

d. Indicatrice d'Euler

Définition 11. *Fonction indicatrice d'Euler*

On appelle **fonction indicatrice d'Euler** l'application $\varphi : \mathbb{N}^* \to \mathbb{N}$ définie, pour $n \in \mathbb{N}^*$, par :

$$\varphi(n) = \#\{k \in [\![1, n]\!] \mid k \text{ et } n \text{ sont premiers entre eux}\}.$$

Proposition 20. *Propriétés de l'indicatrice d'Euler*

i) $\varphi(1) = 1$,

ii) pour $n \geq 2$, $\varphi(n) = \#(\mathbb{Z}/n\mathbb{Z})^\times$,

iii) pour p premier,
$$\varphi(p) = p - 1,$$

iv) pour p premier et $\alpha \in \mathbb{N}^*$,
$$\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$$

v) pour $n, m \in \mathbb{N}^*$ avec n et m premiers entre eux,
$$\varphi(nm) = \varphi(n)\varphi(m).$$

Démonstration.

i) 1 est premier avec lui-même d'où $\varphi(1) = 1$,

ii) pour $n \geq 2$, on a $\overline{k} \in (\mathbb{Z}/n\mathbb{Z})^\times$ si, et seulement, si k premier avec n, d'où le résultat.

iii) pour p premier, alors chaque nombre compris entre 1 et $p - 1$ est premier avec p d'où le résultat.

iv) Soit p premier et $\alpha \in \mathbb{N}^*$. Pour $x \in [\![1, p^\alpha]\!]$, x est n'est pas premier avec p^α si, et seulement si $x \in p\mathbb{Z}$, i.e.
$$x \in [\![1, p^\alpha]\!] \cap \{kp \mid k \in \mathbb{Z}\} = \{kp \mid k \in [\![1, p^{\alpha-1}]\!]\}.$$
Donc
$$\varphi(n) = \#[\![1, p^\alpha]\!] - \#[\![1, p^{\alpha-1}]\!] = p^\alpha - p^{\alpha-1}.$$

v) Soit $n, m \in \mathbb{N}^*$ avec n et m premiers entre eux. D'après le théorème chinois, $\mathbb{Z}/nm\mathbb{Z}$ est isomorphe à $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. Donc $\overline{x}^{nm}$ est inversible dans $\mathbb{Z}/nm\mathbb{Z}$ si, et seulement si, $\overline{x}^n$

est inversible dans $\mathbb{Z}/n\mathbb{Z}$ et $\overline{x}^m$ est inversible dans $\mathbb{Z}/m\mathbb{Z}$ d'où :

$$\varphi(nm) = \#(\mathbb{Z}/nm\mathbb{Z})^\times = \#(\mathbb{Z}/n\mathbb{Z})^\times.\#(\mathbb{Z}/m\mathbb{Z})^\times = \varphi(n)\varphi(m).$$

$\square$

Exercice 17.

Soit $n \in \mathbb{N}^*$. Montrer que

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

Correction.

Considérons l'ensemble $F = \{\frac{k}{n} \mid k \in [\![1, n]\!]\}$. Alors $\#F = n$. De plus, chaque élément de F admet une forme irréductible $\frac{i}{d}$ où $d|n$ et i et d sont premiers entre eux. On a donc

$$F = \bigcup_{d|n} F_d \quad \text{où } F_d = \{\frac{i}{d} \mid \text{pgcd}(i, d) = 1 \text{et} 1 \leq i \leq d\}.$$

De plus les F_d sont disjoints par unicité du représentant irréductible d'un rationnel. Par suite, les F_d pour $d|n$ forment une partition de F et donc :

$$n = \#F = \sum_{d|n} \#F_d = \sum_{d|n} \varphi(d).$$

Corollaire 4.

Soit $n \in \mathbb{N}$ avec $n \geq 2$. On considère $n = p_1^{\alpha_1}...p_k^{\alpha_k} = \prod_{i=1}^{k} p_i^{\alpha_i}$ sa décomposition en facteurs premiers. Alors :

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)...\left(1 - \frac{1}{p_k}\right) = n\prod_{i=1}^{k}\left(1 - \frac{1}{p_i}\right).$$

Démonstration.

On applique par récurrence le point v) de la proposition précédente car chaque $p_i^{\alpha_i}$ est premier avec chaque $p_j^{\alpha_j}$ $(i \neq j)$ puis on applique le point iv) pour obtenir :

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

$\square$

Théorème 9. *Théorème d'Euler*

Soit $n \in \mathbb{N}$ avec $n \geq 2$. Alors, pour tout $a \in \mathbb{Z}$ tel que a et n sont premiers entre eux,

$$a^{\varphi(n)} \equiv 1 \bmod n.$$

Démonstration.

L'entier a est premier avec n donc $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$. Or $((\mathbb{Z}/n\mathbb{Z})^\times, \cdot)$ est un groupe fini, donc $o(\overline{a}) | \#(\mathbb{Z}/n\mathbb{Z})^\times = \varphi(n)$. Ainsi,

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

□

Corollaire 5. *Petit théorème de Fermat*

Soit p un nombre premier. Alors pour tout $a \in \mathbb{Z}$ tel que $a \notin p\mathbb{Z}$,

$$a^{p-1} \equiv 1 \bmod p$$

Démonstration.

Si p est premier, $\varphi(p) = p - 1$ et tout entier qui n'est pas multiple de p est premier avec p. On applique alors le théorème d'Euler. □

Partie C
Anneaux de polynômes

Dans toute cette partie, $\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$ et on considère l'anneau commutatif intègre (muni de ses opérations usuelles) $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$.

1. Propriétés arithmétiques élémentaires

a. Divisibilité

Théorème 10. *Division euclidienne dans $\mathbb{K}[X]$*

Soit $A, B \in \mathbb{K}[X]$ et $B \neq 0$. Alors il existe un unique couple $(Q, R) \in (\mathbb{K}[X])^2$ tel que :

$$A = BQ + R \quad \text{et} \quad \deg(R) \leq \deg(B) - 1.$$

On appelle Q le **quotient** et R le **reste** de la division euclidienne de A par B.

Démonstration.

Vue en sup. □

Remarque 6.

Ainsi $B|A$ si, et seulement si le reste R est nul.

b. Inversibles

On rappelle le fait suivant :

Proposition 21.

Les éléments inversibles de $\mathbb{K}[X]$ sont les éléments de $\mathbb{K}^*$ i.e. $\mathbb{K}[X]^\times = \mathbb{K}^*$.

c. Polynômes irréductibles

Définition 12. *Polynôme irréductible*

On dit que $A \in \mathbb{K}[X]$ est un **polynôme irréductible** dans $\mathbb{K}[X]$ si $\deg(A) \geq 1$ et :

$$B|A \quad \Rightarrow \quad B = \lambda \text{ ou } B = \lambda A \text{ avec } \lambda \in \mathbb{K}.$$

Remarque 7.

Ainsi, si $A = PQ$ avec $\deg(P) \geq 1$ et $\deg(Q) \geq 1$ alors A n'est pas irréductible dans $\mathbb{K}[X]$.

Exemple 6.

— $X^2 + X + 5$ et $X + 1$ sont irréductibles dans $\mathbb{R}[X]$ mais $X^2 + 2X + 1$ ne l'est pas.
— $X^2 - 2$ est irréductible dans $\mathbb{Q}[X]$.

Exercice 18.

1. Donner des exemples de polynômes irréductibles dans $\mathbb{C}[X]$ et dans $\mathbb{R}[X]$.
2. $P = X^3 + X + 1$ est-il irréductible dans $\mathbb{Q}[X]$?

Correction.

1. Dans $\mathbb{C}[X]$, $iX + 1$; dans $\mathbb{R}[X]$, $X^2 + 21$.
2. On suppose par l'absurde que P n'est pas irréductible. Alors il existe A, B non constants tels que $P = AB$. Alors quitte à échanger A et B, on peut supposer $\deg(A) = 2$ et $\deg(B) = 1$. Par suite, B admet une racine $\frac{p}{q} \in \mathbb{Q}$ mise ici sous forme irréductible i.e. p et q sont premiers entre eux, qui est donc une racine de P également.
   Par suite $0 = P(\frac{p}{q}) = \frac{p^3}{q^3} + \frac{p}{q} + 1$. Ainsi,
   $$p^3 + pq^2 + q^3 = 0;$$
   donc $q|p^3$ et $p|q^3$. Il en résulte que $p = \pm 1$ et $q = \pm 1$ car p et q sont premiers entre eux. Par suite, $\frac{p}{q} = \pm 1$. Or $P(1) = 3$ et $P(-1) = -1$, contradiction !

d. Polynômes premiers entre eux

Définition 13. *Polynômes premiers entre eux*

Soit $A, B \in \mathbb{K}[X]$. On dit que A et B sont **premiers entre eux** si, pour $P \in \mathbb{K}[X]$

$$P|A \text{ et } P|B \quad \Rightarrow \quad P \in \mathbb{K}^*,$$

i.e. si les seuls diviseurs communs de A et B sont les polynômes constants non nuls.

Proposition 22.

Soit $A, B \in \mathbb{K}[X]$ tel B est irréductible. Alors A et B sont premiers entre eux si, et seulement si, B ne divise pas A.

Démonstration.

- ($\Rightarrow$). On raisonne par contraposée : si B divise A alors A et B ne sont pas premiers entre eux car B est un diviseur commun non constant de A et B.
- ($\Leftarrow$). On raisonne également par contraposée : on suppose A et B ne sont pas premiers entre eux. Alors ils admettent un diviseur commun non constant P. En particulier, P divise B et B est irréductible, donc il existe $\lambda \in \mathbb{K}^*$ tel que $P = \lambda B$. Ainsi, comme $P|A$, il existe Q tel que :
$$A = PQ = \lambda BQ = B(\lambda Q),$$
donc $B|A$.

□

2. Idéaux de $\mathbb{K}[X]$

Théorème 11.

Les idéaux de $\mathbb{K}[X]$ sont les $P\mathbb{K}[X]$ pour $P \in \mathbb{K}[X]$.

Démonstration.

Pour $P \in \mathbb{K}[X]$, $P\mathbb{K}[X]$ est un idéal comme idéal engendré par P.

Soit I un idéal de $\mathbb{K}[X]$.

- *1er cas :* $I = \{0\}$. Alors $I = 0\mathbb{K}[X]$.
- *2eme cas :* $I \neq \{0\}$. Alors $\{\deg(P) \mid P \in I \setminus \{0\}\}$ est un ensemble non vide de $\mathbb{N}^*$ et donc possède un plus petit élément $p \in \mathbb{N}^*$. Soit $P \in I$ un polynôme de degré p.

  Montrons que $I = P\mathbb{K}[X]$.

  — $P\mathbb{K}[X] \subset I$. Comme P appartient à I qui est un idéal, on a l'inclusion voulue car pour tout $A \in \mathbb{K}[X]$, $PA \in I$.

  — $I \subset P\mathbb{K}[X]$. Soit $A \in I$. La division euclidienne de A par P nous donne l'existence de $Q, R \in \mathbb{K}[X]$ avec $\deg(R) < P$ tels que $A = PQ + R$. Or on a $R = A - PQ \in I$ car $A \in I$ et $PQ \in P\mathbb{K}[X] \subset I$ et comme $\deg(R) < P$ et P est de degré minimal dans I, $R = 0$. Donc $A = PQ \in P\mathbb{K}[X]$. Il en résulte que $I = P\mathbb{K}[X]$.

□

Remarque 8.

Ainsi, $\mathbb{K}[X]$ est un anneau principal.

3. Propriétés relatives au PGCD

a. PGCD et PPCM

Définition 14.

Soit $A, B \in \mathbb{K}[X]$ des polynômes non nuls.

— Le **PGCD** de A et B est le générateur unitaire de l'idéal $A\mathbb{K}[X] + B\mathbb{K}[X]$;

— Le **PPCM** de A et B est le générateur unitaire de l'idéal $A\mathbb{K}[X] \cap B\mathbb{K}[X]$.

Exercice 19.

1. Soit I un idéal de $\mathbb{K}[X]$. Montrer qu'il existe un unique polynôme unitaire P tel que $I = P\mathbb{K}[X]$.
2. En déduire que le PGCD et le PPCM de deux polynômes non nuls sont bien définis.

Correction.

1. Si I est un idéal de $\mathbb{K}[X]$ alors il existe $Q \in \mathbb{K}[X]$ tel que $I = Q\mathbb{K}[X]$. On a $Q\mathbb{K}[X] = Q'\mathbb{K}[X]$ si, et seulement si, $Q|Q'$ et $Q'|Q$ si, et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q' = \lambda Q$. Ainsi on pose le polynôme unitaire $P = q_n Q$ est l'unique générateur unitaire de I où q_n est l'inverse du coefficient de plus haut degré de Q.
2. $A\mathbb{K}[X] + B\mathbb{K}[X]$ et $A\mathbb{K}[X] \cap B\mathbb{K}[X]$ sont des idéaux de $\mathbb{K}[X]$ donc d'après la question précédente, le PGCD et le PPCM sont bien définis.

Notation 3.

Soit $A, B \in \mathbb{K}[X]$ des polynômes non nuls.

— On note $A \wedge B$ le PGCD de A et B ;

— On note $A \vee B$ le PPCM de A et B.

Proposition 23.

Soit $A, B, D, M, P \in \mathbb{K}[X]$ des polynômes non nuls avec D, M unitaires. Alors :

- $D = A \wedge B$ si, et seulement si, D vérifie les deux conditions :
  i) $D|A$ et $D|B$;
  ii) si $P|A$ et $P|B$ alors $P|D$.
- $M = A \vee B$ si, et seulement si, M vérifie les deux conditions :
  i) $A|M$ et $B|M$;
  ii) si $A|P$ et $B|P$ alors $M|P$.

Démonstration.

On note $\mathcal{D} = A\mathbb{K}[X] + B\mathbb{K}[X]$ et $\mathcal{M} = A\mathbb{K}[X] \cap B\mathbb{K}[X]$.

- — ($\Rightarrow$). On suppose $D = A \wedge B$. Alors

$$\mathcal{D} = D\mathbb{K}[X],$$

donc $A\mathbb{K}[X] \subset D\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset D\mathbb{K}[X]$, d'où

$$D|A \text{ et } D|B.$$

et de plus, si $P|A$ et $P|B$ alors $A\mathbb{K}[X] \subset P\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset P\mathbb{K}[X]$ et donc

$$D\mathbb{K}[X] = \mathcal{D} \subset P\mathbb{K}[X];$$

d'où $D|P$.

— ($\Leftarrow$). On suppose i) et ii). $D|A$ et $D|B$ donc $A\mathbb{K}[X] \subset D\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset D\mathbb{K}[X]$. Ainsi,

$$\mathcal{D} \subset D\mathbb{K}[X].$$

De plus, comme $\mathcal{D}$ est principal, il existe P tel que $\mathcal{D} = P\mathbb{K}[X]$. Alors $A\mathbb{K}[X] \subset P\mathbb{K}[X]$ et $B\mathbb{K}[X] \subset P\mathbb{K}[X]$ et donc $P|A$ et $P|B$ d'où, d'après ii), $P|D$. Par suite, $D\mathbb{K}[X] \subset P\mathbb{K}[X] = \mathcal{D}$.

Il résulte que $D\mathbb{K}[X] = \mathcal{D}$ et comme D est unitaire, $D = A \wedge B$.

- — ($\Rightarrow$). On suppose $M = A \vee B$. Alors

$$\mathcal{M} = M\mathbb{K}[X],$$

donc $M\mathbb{K}[X] \subset A\mathbb{K}[X]$ et $M\mathbb{K}[X] \subset B\mathbb{K}[X]$, d'où

$$A|M \text{ et } B|M.$$

et de plus, si $A|P$ et $B|P$ alors $P\mathbb{K}[X] \subset A\mathbb{K}[X]$ et $P\mathbb{K}[X] \subset B\mathbb{K}[X]$ et donc

$$P\mathbb{K}[X] \subset \mathcal{M} = M\mathbb{K}[X];$$

d'où $M|P$.

— ($\Leftarrow$). On suppose i) et ii). $A|M$ et $B|M$ donc $M\mathbb{K}[X] \subset A\mathbb{K}[X]$ et $M\mathbb{K}[X] \subset B\mathbb{K}[X]$. Ainsi,

$$M\mathbb{K}[X] \subset \mathcal{M}.$$

De plus, comme $\mathcal{M}$ est principal, il existe P tel que $\mathcal{M} = P\mathbb{K}[X]$. Alors $P\mathbb{K}[X] \subset A\mathbb{K}[X]$ et $P\mathbb{K}[X] \subset B\mathbb{K}[X]$ et donc $A|P$ et $B|P$ d'où, d'après ii), $M|P$. Par suite, $\mathcal{M} = P\mathbb{K}[X] \subset M\mathbb{K}[X]$.

Il résulte que $M\mathbb{K}[X] = \mathcal{M}$ et comme M est unitaire, $M = A \vee B$.

□

b. Relation de Bézout et algorithme d'Euclide

Proposition 24. *Relation de Bézout*

Soit $A, B \in \mathbb{K}[X]$ des polynômes non nuls et $D = A \wedge B$. Alors il existe $U, V \in \mathbb{K}[X]$ tels que

$$D = AU + BV.$$

Démonstration.

On a $D \in D\mathbb{K}[X] = A\mathbb{K}[X] + B\mathbb{K}[X]$, donc il existe $U, V \in \mathbb{K}[X]$ tels que $D = AU + BV$. □

Proposition 25.

Soit $A, B \in \mathbb{K}[X]$ des polynômes non nuls et R le reste de la division euclidienne de A par B. Alors
$$A \wedge B = B \wedge R.$$

Démonstration.

On note Q le quotient de la division et $D = A \wedge B$, $D' = B \wedge R$. Montrons que $D|D'$ et $D'|D$.

- $\underline{D|D'}$: On a $R = A - BQ \in A\mathbb{K}[X] + B\mathbb{K}[X] = D\mathbb{K}[X]$ donc D divise R, et D divise B donc $D|B \wedge R = D'$.
- $\underline{D'|D}$: On a $A = BQ + R \in B\mathbb{K}[X] + R\mathbb{K}[X] = D'\mathbb{K}[X]$ donc D' divise A, et D' divise B donc $D'|A \wedge B = D$.

Par suite, D et D' sont associés. Or D et D' sont unitaires, donc $D = D'$. □

Théorème 12. *Algorithme d'Euclide pour les polynômes*

Soit $A, B \in \mathbb{K}[X]$ des polynômes non nuls avec $\deg(A) \geq \deg(B)$ et $D = A \wedge B$. Alors la suite récurrente (R_n) :
$$\begin{cases} R_0 = A, \ R_1 = B; \\ R_{n+2} \text{ est le reste de la division euclidienne de } R_n \text{ par } R_{n+1} \end{cases}$$
est stationnaire en 0 et D est égal au polynôme unitaire associé à R_d où $d = \max\{n \in \mathbb{N} \mid R_n \neq 0\}$.

Démonstration.

Pour $n \in \mathbb{N}$, $\deg(R_{n+2}) < \deg(R_{n+1})$ et donc $\deg(R_{n+2}) \leq \deg(R_{n+1}) - 1$. Donc si $n = \deg(A)$, on a :
$$\deg(R_{n+2}) < \deg(R_{n+1}) \leq \deg(R_n) - 1 \leq \deg(R_{n-1}) - 2 \leq ... \leq \deg(R_1) - n = \deg(B) - n \leq \deg(A) = n - n = 0,$$
donc $\deg(R_{n+2}) = 0$.

De plus, par la proposition précédente, on a, pour $d = \max\{n \in \mathbb{N} \mid R_n \neq 0\}$ et U le polynôme unitaire associé à R_d :
$$U = R_d \wedge 0 = R_{d-1} \wedge R_d = R_{d-2} \wedge R_{d-1} = ... = R_0 \wedge R_1 = A \wedge B = D.$$
□

c. Lien avec les polynômes premiers entre eux

Proposition 26.

Soit $A, B \in \mathbb{K}[X]$ des polynômes non nuls. Alors A et B sont premiers entre eux si, et seulement si, $A \wedge B = 1$.

Démonstration.

On note $D = A \wedge B$.

- $(\Rightarrow)$. On suppose A, B premiers entre eux. On a $D|A$ et $D|B$ alors $D \in \mathbb{K}^*$. Or D est unitaire, donc $D = 1$.
- $(\Leftarrow)$. On suppose $D = 1$. Soit $P \in \mathbb{K}[X]$ tel que $P|A$ et $P|B$. D'après la relation de Bézout, il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$. Par suite, $P|AU + BV = 1$ et donc $P \in \mathbb{K}^*$.

□

Théorème 13. *Théorème de Bézout pour les polynômes*

Soit $A, B \in \mathbb{K}[X]$ des polynômes. Alors A et B sont premiers entre eux si, et seulement si, il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$.

Démonstration.

On note $D = A \wedge B$.

- $(\Rightarrow)$. On suppose A, B premiers entre eux. D'après la proposition précédente, $D = 1$. Donc, d'après la relation de Bézout, il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$.
- $(\Leftarrow)$. On suppose qu'il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$. On a $D\mathbb{K}[X] = A\mathbb{K}[X] + B\mathbb{K}[X]$ donc $1 \in D\mathbb{K}[X]$. Alors $D\mathbb{K}[X]$ est un idéal qui contient l'unité de l'anneau, donc $D\mathbb{K}[X] = \mathbb{K}[X]$. Comme 1 est unitaire, $D = 1$.

□

Théorème 14. *Lemme de Gauss*

Soit $A, B, C \in \mathbb{K}[X]$. Si A et B sont premiers entre eux et si $A|BC$ alors $A|C$.

Démonstration.

D'après le théorème de Bézout, il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$, donc $C = AUC + BCV$. Or $A|BC$, donc $BCV, AUV \in A\mathbb{K}[X]$; d'où $C \in A\mathbb{K}[X]$. Par suite, $A|C$. □

4. Décomposition d'un polynôme en facteurs irréductibles

a. Décomposition en facteurs irréductibles

Proposition 27.

On a les propriétés suivantes :

- Tout polynôme de degré supérieur ou égal à 1 possède un diviseur irréductible.
- Si A est irréductible et $A|P_1...P_n$, alors il existe $i \in [\![1, n]\!]$ tel que $A|P_i$.
- Tout polynôme de degré 1 est irréductible.
- Un polynôme de degré 2 ou 3 est irréductible si, et seulement si, il n'a pas de racine dans $\mathbb{K}$.

Démonstration.

- Soit P un polynôme de degré supérieur ou égal à 1. Alors l'ensemble des polynômes non constant qui divise P est non vide car il contient P. Ainsi, il existe un polynôme A non constant de degré minimal qui divise P. Or si $A = UV$ avec $U, V \in \mathbb{K}[X]$, alors $\deg(U), \deg(V) \leq \deg(A)$ et $U|P$, $V|P$. Par suite, par minimalité du degré de A, soit $\deg(U) = \deg(A)$ ou $\deg(V) = \deg(A)$ d'où U ou V est constant. Ainsi, pour tout $U \in \mathbb{K}[X]$ tel que $U|A$, $U = \lambda$ ou $U = \lambda A$ avec $\lambda \in \mathbb{K}^*$. Par suite, A est irréductible.

  Il en résulte que P possède un diviseur irréductible.
- On raisonne par récurrence sur $n \in \mathbb{N}^*$. Pour $n = 1$, si $A|P_1$ alors $A|P_1$! Soit $n \in \mathbb{N}^*$. On suppose la propriété vraie pour n. Soit $P_1, ..., P_{n+1} \in \mathbb{K}[X]$ tels que $A|P_1...P_{n+1}$. Alors $A|(P_1...P_n)P_{n+1}$. On a alors deux cas :
  - $A|P_{n+1}$.
  - $A \nmid P_{n+1}$. Alors A et P_{n+1} sont premiers entre eux car A est irréductible, donc, d'après le Lemme de Gauss, $A|P_1...P_n$. Par suite, par hypothèse de récurrence, il existe $i \in [\![1, n]\!]$ tel que $A|P_i$.

  Dans tous les cas, il existe $i \in [\![1, n+1]\!]$ tel que $A|P_i$. Donc la propriété est vraie pour $n + 1$. Ce qui achève le raisonnement par récurrence.
- Si $aX - b = PQ$ alors $\deg(P) + \deg(Q) = 1$ donc $\deg(P) = 1$ ou 0 et inversement pour Q. Si $\deg(Q) = 0$, alors $Q = \lambda \in \mathbb{K}^*$ et $P = \frac{1}{\lambda}X - a$ et inversement. Par suite, si $P|aX - b$, P est constant ou $P = \lambda(aX - b)$. Il en résulte que $aX - b$ est irréductible.
- Soit P un polynôme de degré 2 ou 3.
  - ($\Rightarrow$). Par contraposée. Si P admet une racine dans $\mathbb{K}$, alors $X - a|P$ et $\deg(P) > 1 = \deg(X - a)$, donc P n'est pas irréductible.
  - ($\Leftarrow$). Par contraposée. Si $P = AB$ avec A et B non associée à P alors $\deg(A), \deg(B) < \deg(P) = 2$ ou 3. Ainsi, soit $\deg(A) = 1$, soit $\deg(B) = 1$. Et donc A ou B possède une racine et donc P en possède une.

$\square$

Théorème 15. *Décomposition en facteurs irréductibles*

Soit $A \in \mathbb{K}[X]$ un polynôme non constant. Alors A s'écrit de façon unique comme le produit

$$A = \lambda \prod_{i=1}^{n} P_i^{\alpha_i},$$

où $\lambda \in \mathbb{K}^*$, $n \in \mathbb{N}$, et pour $i, j \in [\![1, n]\!]$, $i \neq j$, P_i est un polynôme unitaire irréductible, $\alpha_i \in \mathbb{N}^*$

et $P_i \neq P_j$.

Démonstration.

- **Existence de la décomposition :** On raisonne par récurrence sur le degré $n \in \mathbb{N}^*$ d'un polynôme.
  — Initialisation. Pour $n = 1$, $A = aX + b = a(X + \frac{b}{a})$.
  — Hérédité. Soit $n \in \mathbb{N}^*$. On suppose la propriété vraie pour $1 \leq k \leq n$. Soit A de degré $n+1$ et P un diviseur irréductible unitaire de A. Alors il existe $B \in \mathbb{K}[X]$ tel que $A = PB$. Or comme P est irréductible, $\deg(P) \geq 1$, d'où $\deg(B) \leq n$. On applique alors l'hypothèse de récurrence à B :
  $$B = \lambda P_1^{\alpha_i} ... P_i^{\alpha_i}$$
  d'où
  $$A = \lambda P_1^{\alpha_i} ... P_i^{\alpha_i}$$
  et si P est dans la liste, cela rajoute une puissance à un P_j, s'il ne l'est pas, la forme précédente est la forme voulue. Ce qui achève le raisonnement par récurrence.
- **Unicité de la décomposition :** Soit $A = \lambda \prod_{i=1}^n P_i^{\alpha_i} = \mu \prod_{i=1}^m Q_i^{\beta_i}$ deux décomposition de A. On a $\lambda = \mu$ car les P_i, Q_i étant unitaires, λ et μ sont égaux au coefficient dominant de A. Donc
  $$\prod_{i=1}^n P_i^{\alpha_i} = \prod_{i=1}^m Q_i^{\beta_i}.$$
  Donc $P_1 | \prod_{i=1}^m Q_i^{\beta_i}$ et P_1 est irréductible, alors il existe j tel que $P_1 | Q_j$. Quitte à changer l'ordre entre les Q_i, on peut suppose que $Q_j = Q_1$. $P_1 = Q_1$ est premier avec $Q_2^{\alpha_2} ... Q_m^{\alpha_m}$ donc $P_1^{\alpha_1} | Q_1^{\beta_1}$, d'où
  $$\alpha_1 \leq \beta_1$$
  . Et en raisonnant de même avec Q_1, on trouve $\beta_1 \leq \alpha_1$ d'où $\alpha_1 = \beta_1$.
  On obtient donc $\prod_{i=2}^n P_i^{\alpha_i} = \prod_{i=2}^m Q_i^{\beta_i}$ et on procède de la même manière pour $i = 2, 3, ...$.
  Ainsi, les deux décompositions sont les mêmes.

□

b. Irréductibles dans $\mathbb{R}[X]$ et $\mathbb{C}[X]$

Théorème 16. *Théorème de D'Alembert-Gauss*

Tout polynôme non constant de $\mathbb{C}$ admet au moins un racine dans $\mathbb{C}$.

Proposition 28.

— Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.
— Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif.

Partie D
Algèbres

Dans toute cette partie, $\mathbb{K}$ désigne un sous-corps de $\mathbb{C}$.

1. Structure d'algèbre

Définition 15.

Soit A un espace vectoriel sur $\mathbb{K}$ et $\cdot$ une loi de composition interne sur A. On dit que le couple $(A, \cdot)$ ou plus simplement que A est une **algèbre sur** $\mathbb{K}$ si :

i) la loi $\cdot$ est associative ;

ii) la loi $\cdot$ est bilinéaire ;

iii) la loi $\cdot$ possède un élément neutre 1_A.

Exemple 7.

— $(\mathbb{K}[X], \times)$ est une algèbre sur $\mathbb{K}$;

— si E est un espace vectoriel sur $\mathbb{K}$, $(\mathcal{L}(E), \circ)$ est une algèbre sur $\mathbb{K}$;

— $(M_n(\mathbb{K}), \times)$ est une algèbre sur $\mathbb{K}$;

— si X est un ensemble, $(\mathcal{F}(X, \mathbb{K}), \times)$ est une algèbre sur $\mathbb{K}$.

2. Sous-algèbres

Définition 16. *Sous-algèbre*

Soit $(A, \cdot)$ une algèbre sur $\mathbb{K}$ et $B \subset A$. On dit que B est une **sous-algèbre** de A si :

i) B est un sous-espace vectoriel de A.

ii) B est stable par $\cdot$;

iii) $1_A \in B$.

Exemple 8.

— $\mathrm{Vect}(1_A)$ et A sont des sous-algèbres de A ;

— L'ensemble $T_n^+(\mathbb{K})$ des matrices triangulaires supérieures est une sous-algèbre de $M_n(\mathbb{K})$.

3. Morphismes d'algèbres

Définition 17. *Morphisme d'algèbre*

Soit A, B deux algèbres sur $\mathbb{K}$ et $f : A \to B$. On dit que f est un **morphisme d'algèbres** si :

i) pour tous $\lambda, \mu \in \mathbb{K}$ et tous $x, y \in A$,
$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y);$$

ii) pour tous $x, y \in A$,
$$f(xy) = f(x)f(y);$$

iii) $f(1_A) = 1_B$.

Exemple 9.

— $\lambda \mapsto \lambda 1_A$ est un morphisme d'algèbres de $\mathbb{K}$ dans $\mathrm{Vect}(1_A)$;

— Pour $P \in GL_n(\mathbb{K})$, $M \mapsto PMP^{-1}$ est un morphisme d'algèbres de $M_n(\mathbb{K})$ dans lui-même.

Proposition 29.

Soit A, B deux algèbres sur $\mathbb{K}$ et $f : A \to B$ un morphisme d'algèbres. Alors

— Le noyau $\mathrm{Ker}(f)$ est un idéal de l'anneau $(A, +, \cdot)$

— L'image $\mathrm{Im}(f)$ est une sous-algèbre de B.

Démonstration.

— $\mathrm{Ker}(f)$ est un idéal de l'anneau $(A, +, \cdot)$ car c'est le noyau d'un morphisme d'anneaux.

—
- $\mathrm{Im}(f)$ est un sous-anneau de B comme image de l'anneau A par le morphisme d'anneaux f.
- Il reste à montrer que $\mathrm{Im}(f)$ est stable par multiplication externe : soit $\lambda \in \mathbb{K}$ et $f(x) \in \mathrm{Im}(f)$ avec $x \in A$. Alors :
$$\lambda f(x) = f(\lambda x) \in \mathrm{Im}(f).$$

Donc $\mathrm{Im}(f)$ est une sous-algèbre de B.

□

4. Algèbres et polynômes

a. Polynômes appliqués à un élément d'une algèbre

Notation 4. *Polynôme d'un élément*

Soit A une algèbre sur $\mathbb{K}$, $u \in A$ et $P \in \mathbb{K}[X]$ avec $P = \sum_{i=0}^{n} a_i X^i$. On note $P(u)$ l'élément de A

$$P(u) = \sum_{i=0}^{n} a_i u^i = a_0 1_A + a_1 u + ... + a_n u^n.$$

Exemple 10.

— Soit E un espace vectoriel sur $\mathbb{K}$. Pour $f \in \mathcal{L}(E)$ et $P = \sum_{i=0}^{n} a_i X^i \in \mathbb{K}[X]$,

$$P(f) = a_0 \mathrm{Id} + a_1 f + ... + a_n f^n;$$

— Soit $n \in \mathbb{N}^*$. Pour $M \in M_n(\mathbb{K})$ et $P = \sum_{i=0}^{n} a_i X^i \in \mathbb{K}[X]$,

$$P(M) = a_0 I_n + a_1 M + ... + a_n M^n.$$

Proposition-Notation 30.

Soit A une algèbre sur $\mathbb{K}$ et $u \in A$. L'application notée

$$f_u : \left| \begin{array}{ccc} \mathbb{K}[X] & \to & A \\ P & \mapsto & P(u) \end{array} \right.$$

est un morphisme d'algèbres.

Démonstration.

Soit $\lambda, \mu \in \mathbb{K}$, $P, Q \in \mathbb{K}[X]$. On a :

i)

$$\begin{aligned} f_u(\lambda P + \mu Q) &= (\lambda P + \mu Q)(u) \\ &= (\lambda P)(u) + (\mu Q)(u) \\ &= \lambda P(u) + \mu Q(u) \\ &= f_u(u) + f_u(u) \end{aligned}$$

ii)

$$\begin{aligned} f_u(PQ) &= (PQ)(u) \\ &= P(u)Q(u) \\ &= f_u(P) f_u(Q) \end{aligned}$$

iii) $f_u(1) = 1(u) = 1_A$

Il en résulte que f_u est un morphisme d'algèbres. $\square$

b. Polynômes annulateurs

Définition 18. *Idéal et polynôme annulateur*

Soit A une algèbre sur $\mathbb{K}$ et $u \in A$. On appelle **idéal annulateur de** u l'ensemble

$$\text{Ker}(f_u) = \{P \in \mathbb{K}[X] \mid P(u) = 0_A\}.$$

Un polynôme $P \in \mathbb{K}[X]$ est appelé **polynôme annulateur de** u s'il appartient à l'idéal annulateur de u i.e. si $P(u) = 0$.

On a montré dans la partie précédente que $\mathbb{K}[X]$ est un anneau principal. Ceci justifie la définition suivante :

Définition 19. *Polynôme minimal*

Soit A une algèbre sur $\mathbb{K}$ et $u \in A$ d'idéal annulateur non réduit à 0_A. On appelle **polynôme minimal de** u et on note π_u le générateur unitaire de l'idéal annulateur de u.

Exemple 11.

Soit A une algèbre sur $\mathbb{K}$.

— Un élément u de A est dit **nilpotent** s'il existe $n \in \mathbb{N}^*$ tel que $u^n = 0$.

  Dans ce cas, le polynôme X^n est un polynôme annulateur de u. Comme le polynôme minimal de u divise X^n donc il existe $k \leq n$ tel que $\pi_u = X^k$.

— Un élément u de A est un **idempotent** si $u^2 = u$.

  Dans ce cas, $X^2 - X$ est un polynôme annulateur de u. Comme $\pi_u | X^2 - X$, alors on a trois cas possibles :
  1) $\pi_u = X^2 - X$.
  2) $\pi_u = X$, auquel cas $u = 0_A$
  3) $\pi_u = X - 1$, auquel cas $u = 1_A$

Exercice 20.

1. Soit $f \in \mathcal{L}(\mathbb{R}^3)$ tel que $f(x, y, z) = (y, z, x)$. Calculer f^3 et en déduire un polynôme annulateur de f.
2. Soit $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in M_3(\mathbb{R})$. Calculer A^3 et déterminer un polynôme annulateur de A.

Correction.

1. On a, pour $(x, y, z) \in \mathbb{R}^3$:

$$f^3(x, y, z) = f^2(y, z, x) = f(z, x, y) = x, y, z$$

   Donc $f^3 = \text{id}$.
   Par suite, $X^3 - 1$ est un polynôme annulateur de f.
   Remarque : on a $P = X^3 - 1 = X - 1(X^2 + X + 1)$ donc P est le polynôme minimal de f car ni $X - 1$, ni $X^2 + X + 1$ ne sont des polynômes annulateur de f.

2. On a
$$A^3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0_3,$$
donc X^3 est un polynôme annulateur de A.
Remarque : X et X^2 sont les seuls diviseurs non triviaux de X^3 et aucun des deux n'est un polynôme annulateur de A donc X^3 est le polynôme minimal de A. Ainsi, A est une matrice nilpotente d'indice 3.

c. Algèbre engendrée par un élément

Notation 5.

Soit A une algèbre sur $\mathbb{K}$ et $u \in A$. On note
$$\mathbb{K}[u] = \{P(u) \mid P \in \mathbb{K}[X]\}.$$

Proposition 31.

Soit A une algèbre sur $\mathbb{K}$ et $u \in A$. Alors $\mathbb{K}[u]$ est une sous-algèbre *commutative* de A.

Démonstration.

On a $\mathbb{K}[u] = \mathrm{Im}(f_u)$ donc $\mathbb{K}[u]$ est une sous-algèbre de A comme image d'un morphisme d'algèbres. De plus, pour $P(u), Q(u) \in \mathbb{K}[u]$ avec $P, Q \in \mathbb{K}[X]$, on a :
$$P(u)Q(u) = (PQ)(u) = (QP)(u) = Q(u)P(u);$$
Donc $\cdot$ est commutative sur $\mathbb{K}[u]$. $\square$

Proposition 32.

Soit A une algèbre sur $\mathbb{K}$ et $u \in A$. Si u admet un polynôme minimal $\pi_u \in \mathbb{K}[X]$ avec $d = \deg(\pi_u)$, alors $\mathbb{K}[u]$ est un espace vectoriel de dimension finie d et
$$\left(u^k\right)_{0 \leq k \leq d-1}$$
est une base de $\mathbb{K}[u]$.

Démonstration.

On suppose que u admet un polynôme minimal π_u avec $d = \deg(\pi_u)$. Montrons que $\left(u^k\right)_{0 \leq k \leq d-1}$ est une base de $\mathbb{K}[u]$.

- *Famille libre :* soit $\lambda_0, ..., \lambda_{d-1} \in \mathbb{K}$ des scalaires tels que $\sum_{k=1}^{d-1} \lambda_k u^k = 0_A$. Alors le polynôme $P = \sum_{k=1}^{d-1} \lambda_k X^k$ est un polynôme annulateur de u de degré $\leq d-1 < d = \deg(\pi_u)$. Or π_u est de degré minimal parmi les polynômes annulateurs non nuls. Donc $P = 0$ et ainsi, pour tout $i \in [\![0, d-1]\!]$, $\lambda_i = 0$. Donc la famille $\left(u^k\right)_{0 \leq k \leq d-1}$ est libre.
- *Famille génératrice :* Soit $P(u) \in \mathbb{K}[u]$. Alors $P \in \mathbb{K}[X]$ et en faisant la division euclidienne de ce polynôme par π_u, on obtient qu'il existe $Q, R \in \mathbb{K}[X]$ tels que $P = \pi_u Q + R$ et $\deg(R) < d-1$. Par suite,
$$P(u) = \underbrace{\pi_u(u)}_{=0_A} Q(u) + R(u) = R(u).$$
et R est de degré $\leq d-1$ donc il existe $\lambda_0, ..., \lambda_{d-1} \in \mathbb{K}$ tels que $R = \sum_{k=1}^{d-1} \lambda_k X^k$. Il en résulte que :
$$P(u) = R(u) = \sum_{k=1}^{d-1} \lambda_k u^k \in \mathrm{Vect}\left(u^k\right)_{0 \leq k \leq d-1}.$$
Et ainsi, $\left(u^k\right)_{0 \leq k \leq d-1}$ est génératrice.

Donc $\left(u^k\right)_{0 \leq k \leq d-1}$ est une base de $\mathbb{K}[u]$ et de plus, cette base comporte d vecteurs donc $\dim(\mathbb{K}[u]) = d$. □

Question 2.

Que dire de $\mathbb{K}[u]$ lorsque u n'admet pas de polynôme annulateur non nul ?

Correction.

Si u n'admet pas de polynôme annulateur, alors pour tout $P \in \mathbb{K}[X]$ avec $P \neq 0$, $P(u) \neq 0$, ce qui permet de montrer que la famille $\left(u^k\right)_{k \in \mathbb{N}}$ est une famille libre de $\mathbb{K}[u]$. Comme cette famille est infinie, il en résulte que $\dim(\mathbb{K}[u]) = +\infty$.

Méthode : Connaissant le polynôme minimal π_u d'un élément u d'une algèbre A, on peut, pour $P \in \mathbb{K}[X]$, donner la décomposition de $P(u)$ dans la base $\left(u^k\right)_{0 \leq k \leq d-1}$ de $\mathbb{K}[u]$: il suffit de déterminer le reste R de la division euclidienne de P par π_u et d'évaluer R en u pour obtenir la décomposition voulue.

Ainsi, cette méthode donne un moyen pratique pour calculer les puissances successives u^n de u pour $n \in \mathbb{N}^*$!

Exercice 21.

Soit $A = \begin{pmatrix} 8 & -3 & -6 \\ -2 & 3 & 2 \\ 6 & -3 & -4 \end{pmatrix}$

1. Calculer A^2 et déterminer un polynôme annulateur de A.
2. Ce polynôme est-il le polynôme minimal de A ?
3. Montrer que A est inversible en utilisant son polynôme annulateur.
4. Calculer A^n pour $n \in \mathbb{N}$.

Correction.

1. On a
$$A^2 = \begin{pmatrix} 34 & -15 & -30 \\ -10 & 9 & 10 \\ 30 & -15 & -26 \end{pmatrix}$$
et on remarque que $A^2 - 5A = -6I_3$ i.e. $A^2 - 5A + 6I_3 = 0_3$. Ainsi, le polynôme
$$P = X^2 - 5X + 6$$
est un polynôme annulateur de A.
2. On a $P = (X-2)(X-3)$ donc on a trois possibilités pour π_A du fait que $\deg(\pi_A) \geq 1$ et $\pi_A | P$:
— $\pi_A = X - 2$: impossible car $A \neq 2I_3$
— $\pi_A = X - 3$: impossible car $A \neq 3I_3$
— et donc $\pi_A = (X-2)(X-3)$!
Par suite P est le polynôme minimal de A.
3. On a $A^2 - 5A + 6I_3 = 0_3$ donc $A(\frac{-1}{6}(A - 5I_3)) = I_3$. Par suite A est inversible et son inverse est :
$$\frac{-1}{6}(A - 5I_3).$$
4. On effectue la division euclidienne de X^n par P qui est de degré 2, alors il existe $Q, R \in \mathbb{K}[X]$ tels que $\deg(R) \leq 1$ et
$$X^n = QP + R \quad (*)$$
Comme R est de degré au plus 1, il existe $a, b \in \mathbb{K}$ tels que $R = aX + b$. les nombres 2 et 3 étant des racines de P i.e. $P(2) = 0$ et $P(3) = 0$, en évaluant $(*)$ en 2 et 3 on obtient :
$$\begin{cases} 2^n = 2a + b \\ 3^n = 3a + b \end{cases} \Leftrightarrow \begin{cases} a = 3^n - 2^n \\ b = 3.2^n - 2.3^n \end{cases}$$
et donc on obtient :
$$A^n = Q(A)\underbrace{P(A)}_{=0_3} + R(A) = aA + b = (3^n - 2^n)A + (3.2^n - 2.3^n)I_3.$$

Proposition 33.

Soit A une algèbre sur $\mathbb{K}$ de dimension finie. Alors tout élément de A admet un polynôme minimal.

Démonstration.

Soit n la dimension de A. Soit $u \in A$. Montrons que u possède un polynôme annulateur non nul. La famille $(1_A, u, ..., u^n)$ est liée car composée de $n+1$ vecteurs dans un espace de dimension n donc il existe $\lambda_0, ..., \lambda_n \in \mathbb{K}$ non tous nuls tels que

$$\sum_{i=0}^{n} \lambda_i u^i = 0_A.$$

Par suite $P = \sum_{i=0}^{n} \lambda_i X^i$ est un polynôme annulateur non nul de u d'où u admet un polynôme minimal. □

Chapitre II
Intégrale généralisée

Table des matières

Dans ce chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$; les intervalles considérés sont supposés d'intérieur non vide et a désigne un nombre réel.

Partie A

Intégrale généralisée sur $[a, +\infty[$

1. Généralités

a. Définitions et exemples

Définition 1. *Intégrale convergente*

Soit $f \in C_{pm}([a, +\infty[, \mathbb{K})$.

Si la limite $\lim\limits_{x \to +\infty} \displaystyle\int_a^x f$ existe et est finie, on note $\displaystyle\int_a^{+\infty} f$ ou encore $\displaystyle\int_a^{+\infty} f(t)dt$ cette quantité, i.e. :

$$\int_a^{+\infty} f(t)dt = \lim_{x \to +\infty} \int_a^x f(t)dt.$$

Dans ce cas, on dit que $\displaystyle\int_a^{+\infty} f$ **converge**.

Dans le cas contraire, on dit que $\displaystyle\int_a^{+\infty} f$ **diverge**.

Proposition-Notation 1.

Soit $f \in C([a, +\infty[, \mathbb{K})$ et F une primitive de f. L'intégrale $\displaystyle\int_a^{+\infty} f$ converge si, et seulement si F admet une limite finie en $+\infty$. Dans ce cas, on a :

$$\int_a^{+\infty} f(t)dt = \lim_{x \to +\infty} F(x) - F(a) = \underbrace{[F(t)]_a^{+\infty}}_{\text{notation}}$$

Démonstration.

D'après le théorème fondamental de l'analyse, on a, pour tout $x \in [a, +\infty[$:

$$\int_a^x f(t)dt = F(x) - F(a)$$

Donc l'intégrale $\displaystyle\int_a^{+\infty} f$ converge si, et seulement si, $x \mapsto \int_a^x f(t)dt$ admet une limite finie en $+\infty$ si, et seulement si, F admet une limite finie en $+\infty$.

Dans ce cas, on a :

$$\int_a^{+\infty} f(t)dt = \lim_{x\to+\infty} \int_a^x f(t)dt = \lim_{x\to+\infty} (F(x) - F(a)) = \lim_{x\to+\infty} F(x) - F(a).$$

□

Exemple 1.

1. Soit $\alpha > 0$. Alors $\displaystyle\int_0^{+\infty} e^{-\alpha t}dt$ converge et on a $\displaystyle\int_0^{+\infty} e^{-\alpha t}dt = \frac{1}{\alpha}$.
2. L'intégrale $\displaystyle\int_0^{+\infty} \cos(t)dt$ diverge.

> En effet, pour $x \in \mathbb{R}$, $\int_0^x \cos(t)dt = \sin(x)$; or sin ne possède pas de limite quand x tend vers $+\infty$.

Exercice 1.

Discuter de la convergence des intégrales suivantes et en cas de convergence, déterminer la valeur de l'intégrale :

$$\int_0^{+\infty} \frac{1}{1+t^2}dt \qquad \int_0^{+\infty} te^t dt \qquad \int_1^{+\infty} \ln\left(1+\frac{1}{t^2}\right) dt.$$

Correction.

1. La fonction $f : t \mapsto \frac{1}{1+t^2}$ est continue sur $I = [0, +\infty[$ et $F = \arctan$ est une primitive de f sur I. De plus, F admet $\frac{\pi}{2}$ comme limite en $+\infty$ donc $\int_0^{+\infty} \dfrac{1}{1+t^2}dt$ converge et on a :
$$\int_0^{+\infty} \frac{1}{1+t^2}dt = [\arctan(t)]_0^{+\infty} = \frac{\pi}{2}.$$
2. La fonction $f : t \mapsto te^t$ est continue sur $I = [0, +\infty[$ et $F : t \mapsto (t-1)e^t$ est une primitive de f sur I (on peut retrouver cette primitive, sans connaissance du résultat a priori, par une intégration par parties par exemple). De plus, F admet $+\infty$ comme limite en $+\infty$ donc $\int_0^{+\infty} te^t dt$ diverge.
3. La fonction $f : t \mapsto \ln\left(1+\frac{1}{t^2}\right)$ est continue sur $I = [1, +\infty[$. De plus, par une intégration

par paties, on trouve :

$$\begin{aligned}\int f(t)dt &= t\ln\left(1+\frac{1}{t^2}\right) - \int t\frac{\frac{-2}{t^3}}{1+\frac{1}{t^2}}dt\\ &= t\ln\left(1+\frac{1}{t^2}\right) + 2\int \frac{1}{1+t^2}dt\\ \int f(t)dt &= t\ln\left(1+\frac{1}{t^2}\right) + 2\arctan(t).\end{aligned}$$

Par suite, $F : t \mapsto t\ln\left(1+\frac{1}{t^2}\right) + 2\arctan(t)$ est une primitive de f sur I et comme $\ln\left(1+\frac{1}{t^2}\right) \underset{t\to+\infty}{\sim} \frac{1}{t^2}$,

$$\lim_{t\to+\infty} F(t) = 0 + 2\times\frac{\pi}{2} = \pi$$

Ainsi, $\int_1^{+\infty} \ln\left(1+\frac{1}{t^2}\right) dt$ converge et comme $F(1) = \ln(2) + \frac{\pi}{2}$, on a :

$$\int_1^{+\infty} \ln\left(1+\frac{1}{t^2}\right) dt = [F(t)]_0^{+\infty} = \pi - (\ln(2) + \frac{\pi}{2}) = \frac{\pi}{2} - \ln(2).$$

Exercice 2.

1. Donner un exemple de fonction $f : [1, +\infty[\to \mathbb{R}_+$ continue telle que $\lim_{t\to+\infty} f(t) = 0$ et telle que $\displaystyle\int_1^{+\infty} f$ diverge.
2. Donner un exemple de fonction $f : \mathbb{R}_+ \to \mathbb{R}_+$ continue telle que f n'est pas bornée et telle que $\displaystyle\int_1^{+\infty} f$ converge.

Correction.

1. On peut considérer la fonction $f : t \mapsto \dfrac{1}{t}$. En effet, on a $\lim_{t\to+\infty} f(t) = 0$ et
$$\int_1^x f(t)dt = [\ln(t)]_1^x = \ln(x) \xrightarrow[x\to+\infty]{} +\infty.$$
2. On peut considérer la fonction $f : \mathbb{R}_+ \to \mathbb{R}_+$ définie par :
— pour $n \in \mathbb{N}*$, f est affine sur :
    - l'intervalle $[n - \frac{1}{4^n}, n]$ avec $f(n - \frac{1}{4^n}) = 0$ et $f(n) = 2^n$;
    - l'intervalle $[n, n + \frac{1}{4^n}]$ avec $f(n) = 2^n$ et $f(n + \frac{1}{4^n}) = 0$;

— $f(t) = 0$ pour tout $t \in \mathbb{R}_+ \setminus \displaystyle\bigcup_{n\in\mathbb{N}^*} [n - \frac{1}{4^n}, n + \frac{1}{4^n}]$.

Alors f est continue et non bornée sur $\mathbb{R}_+$. De plus, on a, pour $n \in \mathbb{N}^*$:
$$\int_{n-\frac{1}{4^n}}^{n+\frac{1}{4^n}} f(t)dt = \frac{1}{2^n}.$$

On a, pour $x \in \mathbb{R}_+$ et $n \in \mathbb{N}^*$ tel que $n + \frac{1}{4^n} > x$:

$$\begin{aligned}\int_0^x f(t)dt &\leq \int_0^{n+\frac{1}{4^n}} f(t)dt = \sum_{k=1}^{n}\left(\int_{k-\frac{1}{4^k}}^{k+\frac{1}{4^k}} f(t)dt\right) = \sum_{k=1}^{n}\frac{1}{2^k} \\ &\leq \sum_{k=1}^{+\infty}\frac{1}{2^k} = 1.\end{aligned}$$

Ainsi la fonction $F : x \mapsto \int_0^x f(t)dt$ est majorée sur $\mathbb{R}_+$ et, de plus, la fonction f étant positive, la fonction F est croissante sur $\mathbb{R}_+$. Il en résulte que F admet une limite finie en $+\infty$ *(on peut montrer simplement que la limite est bien* 1*)* et donc $\displaystyle\int_0^{+\infty} f$ converge.

b. Intégrales de Riemann en $+\infty$

Proposition 2. *Intégrale de Riemann en $+\infty$*

Soit $\alpha \in \mathbb{R}$. Alors $\displaystyle\int_1^{+\infty} \frac{1}{t^\alpha}dt$ converge si, et seulement si, $\alpha > 1$.

Démonstration.

Pour $\alpha = 1$, $\displaystyle\int_1^{+\infty} \frac{1}{t}dt$ diverge. En effet, une primitive de $t \mapsto \frac{1}{t}$ sur $[1, +\infty[$ est la fonction ln qui admet pour limite $+\infty$ en $+\infty$.

Pour $\alpha \neq 1$, on a :

$$\int_1^x \frac{1}{t^\alpha}dt = \left[\frac{1}{(1-\alpha)t^{\alpha-1}}\right]_1^x = \frac{1}{(1-\alpha)x^{\alpha-1}} - \frac{1}{1-\alpha} \begin{cases} \xrightarrow[x\to+\infty]{} +\infty & \text{si } \alpha < 1; \\ \xrightarrow[x\to+\infty]{} \dfrac{1}{\alpha-1} & \text{si } \alpha > 1. \end{cases}$$

Ainsi, si $\alpha > 1$, $\displaystyle\int_1^{+\infty} \frac{1}{t}dt$ converge et si $\alpha < 1$ $\displaystyle\int_1^{+\infty} \frac{1}{t}dt$ diverge.

Conclusion : $\displaystyle\int_1^{+\infty} \frac{1}{t^\alpha}dt$ converge si, et seulement si, $\alpha > 1$; et dans le cas $\alpha > 1$, on a

$$\int_1^{+\infty} \frac{1}{t^\alpha}dt = \frac{1}{\alpha - 1}$$

□

c. Propriétés de l'intégrale généralisée sur $[a, +\infty[$

Proposition 3. *Relation de Chasles*

Soit $f \in C_{pm}([a, +\infty[, \mathbb{K})$ et $b \in [a, +\infty[$. Alors l'intégrale $\displaystyle\int_a^{+\infty} f$ est convergente si, et seulement si, $\displaystyle\int_b^{+\infty} f$ est convergente. Dans ce cas, on a :

$$\int_a^{+\infty} f(t)dt = \int_a^b f(t)dt + \int_b^{+\infty} f(t)dt.$$

Démonstration.

Soit $f \in C_{pm}([a, +\infty[, \mathbb{K})$ et $b \in [a, +\infty[$. Comme f est continue par morceaux sur $[a, +\infty[$ et donc $[a, b]$, $\int_a^b f(t)dt$ est bien définie et on a, pour tout $x \in [a, +\infty[$, d'après la relation de Chasles (pour les intégrales sur un segment) :

$$\int_a^x f(t)dt = \int_a^b f(t)dt + \int_b^x f(t)dt$$

Ainsi, $\int_a^{+\infty} f$ converge si, et seulement si, la fonction $x \mapsto \int_a^x f(t)dt$ admet une limite finie en $+\infty$, si et seulement si, $x \mapsto \int_b^x f(t)dt$ admet une limite finie en $+\infty$, si et seulement si, $\int_b^{+\infty} f$ converge.
Dans ce cas, on a :

$$\begin{aligned}
\int_a^{+\infty} f(t)dt &= \lim_{x \to +\infty} \int_a^x f(t)dt \\
&= \lim_{x \to +\infty} \left(\int_a^b f(t)dt + \int_b^x f(t)dt \right) \\
&= \int_a^b f(t)dt + \lim_{x \to +\infty} \int_b^x f(t)dt \\
&= \int_a^b f(t)dt + \int_b^{+\infty} f(t)dt.
\end{aligned}$$

$\square$

Proposition 4. *Linéarité de l'intégrale généralisée*

Soit $f, g \in C_{pm}([a, +\infty[, \mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$. Si les intégrales $\displaystyle\int_a^{+\infty} f$ et $\displaystyle\int_a^{+\infty} g$ convergent, alors l'intégrale $\displaystyle\int_a^{+\infty} (\lambda f + \mu g)$ converge et on a :

$$\int_a^{+\infty} (\lambda f + \mu g)(t)dt = \lambda \int_a^{+\infty} f(t)dt + \mu \int_a^{+\infty} g(t)dt.$$

Démonstration.

Soit $f, g \in C_{pm}([a, +\infty[, \mathbb{K})$ et $\lambda, \mu \in \mathbb{K}$. Pour tout $x \in [a, +\infty[$, on a, par linéarité de l'intégrale sur un segment :

$$\int_a^x (\lambda f + \mu g)(t)dt = \lambda \int_a^x f(t)dt + \mu \int_a^x g(t)dt$$

Par suite, si les fonctions $x \mapsto \int_a^x f(t)dt$ et $x \mapsto \int_a^x g(t)dt$ admettent des limites finies en $+\infty$, alors $\int_a^x (\lambda f + \mu g)(t)dt$ admet une limite finie en $+\infty$. Dans ce cas :

$$\begin{aligned}
\int_a^{+\infty} (\lambda f + \mu g)(t)dt &= \lim_{x \to +\infty} \int_a^x (\lambda f + \mu g)(t)dt \\
&= \lim_{x \to +\infty} \left(\lambda \int_a^x f(t)dt + \mu \int_a^x g(t)dt \right) \\
&= \lambda \lim_{x \to +\infty} \int_a^x f(t)dt + \mu \lim_{x \to +\infty} \int_a^x g(t)dt \\
&= \lambda \int_a^{+\infty} f(t)dt + \mu \int_a^{+\infty} g(t)dt.
\end{aligned}$$

□

Exercice 3.

Déduire de la proposition précédente que l'ensemble des fonctions f continues par morceaux telles que $\displaystyle\int_a^{+\infty} f$ converge est un espace vectoriel sur $\mathbb{K}$. Que dire de l'application $f \mapsto \displaystyle\int_a^{+\infty} f$?

Correction.

D'après la proposition pécédente, l'ensemble $E = \{f \in C_{pm}([a, +\infty[, \mathbb{K}) \mid \int_a^{+\infty} f \text{ converge}\}$ est stable par combinaison linéaire ; de plus, la fonction nulle appartient à E donc E est un sous-espace vectoriel de $C_{pm}([a, +\infty[, \mathbb{K})$: ainsi E est un espace vectoriel sur $\mathbb{K}$.

Toujours d'après la proposition précédente, on conclut que $f \mapsto \displaystyle\int_a^{+\infty} f$ est une forme linéaire sur E.

Proposition 5. *Positivité et croissance*

Soit $f, g \in C_{pm}([a, +\infty[, \mathbb{R})$ telles que $\displaystyle\int_a^{+\infty} f$ et $\displaystyle\int_a^{+\infty} g$ convergent.

— Si f est positive (i.e. pour tout $t \in [a, +\infty[$, $f(t) \geq 0$), alors $\displaystyle\int_a^{+\infty} f(t)dt \geq 0$.

— Si $f \leq g$ (i.e. pour tout $t \in [a, +\infty[$, $f(t) \leq g(t)$), alors $\displaystyle\int_a^{+\infty} f(t)dt \leq \int_a^{+\infty} g(t)dt$.

Démonstration.

— On suppose f positive. Alors, par positivité de l'intégrale sur un segment, pour tout $x \in [a+\infty[$, $\int_a^x f(t)dt \geq 0$. Par passage à la limite, on obtient donc que $\int_a^{+\infty} f(t)dt \geq 0$.

— Si $f \leq g$, alors $h = g - f \geq 0$. De plus, par combinaison linéaire, $\int_a^{+\infty} h$ converge donc $\int_a^{+\infty} \underbrace{h(t)}_{g(t)-f(t)} \, dt \geq 0$.

Ainsi, $\int_a^{+\infty} f$ et $\int_a^{+\infty} g$ étant convergentes, par linéarité de l'intégrale généralisée, $\int_a^{+\infty} f(t)dt \leq \int_a^{+\infty} g(t)dt$.

□

Théorème 1.

Soit $f \in C([a,+\infty[,\mathbb{R})$ une fonction **continue** et **positive** sur $[a,+\infty[$. Alors $\displaystyle\int_a^{+\infty} f(t)dt = 0$ si, et seulement si, $f = 0$.

Proposition 6. *Dérivation*

Soit $f \in C([a,+\infty[,\mathbb{R})$ une fonction **continue** sur $[a,+\infty[$ telle que $\displaystyle\int_a^{+\infty} f$ converge. Alors l'application de $\varphi : [a,+\infty[\to \mathbb{R}$ définie par :

$$\varphi : x \mapsto \int_x^{+\infty} f(t)dt$$

est de classe C^1 sur $[a,+\infty[$ et on a $\varphi' = -f$.

2. Intégrale généralisée de fonctions positives

Proposition 7.

Soit $f \in C_{pm}([a,+\infty[,\mathbb{R})$ une fonction positive.

— L'intégrale $\displaystyle\int_a^{+\infty} f$ converge si, et seulement si, la fonction $F : x \mapsto \displaystyle\int_a^x f(t)dt$ est majorée sur $[a,+\infty[$. Dans ce cas, on a :

$$\int_a^{+\infty} f(t)dt = \sup_{x\in[a,+\infty[} \int_a^x f(t)dt.$$

— L'intégrale $\displaystyle\int_a^{+\infty} f$ diverge si, et seulement si, $\displaystyle\lim_{x\to+\infty} \int_a^x f(t)dt = +\infty$.

Théorème 2. *Comparaison*

Soit $f, g \in C_{pm}([a, +\infty[, \mathbb{R})$ des fonctions positives.

i) On suppose qu'il existe $A \geq a$ tel que, pour tout $t \geq A$, $f(t) \leq g(t)$.

— Si $\displaystyle\int_a^{+\infty} g$ converge, alors $\displaystyle\int_a^{+\infty} f$ converge.

— Si $\displaystyle\int_a^{+\infty} f$ diverge, alors $\displaystyle\int_a^{+\infty} g$ diverge.

ii) On suppose $f = 0(g)$ (ou $f = o(g)$) en $+\infty$. Si $\displaystyle\int_a^{+\infty} g$ converge, alors $\displaystyle\int_a^{+\infty} f$ converge.

iii) On suppose $f \underset{x \to +\infty}{\sim} g$. Alors $\displaystyle\int_a^{+\infty} f$ et $\displaystyle\int_a^{+\infty} g$ sont de même nature.

Exercice 4.

Étudier la convergence des intégrales suivantes :

1. $\displaystyle\int_0^{+\infty} \frac{t^\alpha + 1}{3t^\beta + 5} dt$ pour $\alpha, \beta \in \mathbb{R}_+$.
2. $\displaystyle\int_1^{+\infty} t^\alpha e^{-t} dt$ pour $\alpha \in \mathbb{R}$.
3. $\displaystyle\int_2^{+\infty} \frac{1}{t^\alpha (\ln(t))^\beta} dt$ pour $\alpha, \beta \in \mathbb{R}$.

Correction.

1. Soit $f : t \mapsto \frac{t^\alpha+1}{3t^\beta+5}$. Alors f est définie (car $3t^\beta + 5 > 0$ pour tout $t \geq 0$), continue et positive sur $\mathbb{R}_+$ comme quotient de fonctions continues dont le dénominateur ne s'annule pas sur $\mathbb{R}_+$ et on a :
$$\frac{t^\alpha + 1}{3t^\beta + 5} \underset{t \to +\infty}{\sim} \frac{t^\alpha}{3t^\beta} = \frac{1}{3t^{\beta - \alpha}}$$
Or, d'après le critère de Riemann en $+\infty$, $\displaystyle\int_1^{+\infty} \frac{1}{t^{\beta-\alpha}} dt$ converge si, et seulement si, $\beta - \alpha > 1$.
Par suite, $\displaystyle\int_1^{+\infty} \frac{t^\alpha + 1}{3t^\beta + 5} dt$ converge si, et seulement si, $\beta - \alpha > 1$.
De plus, f étant continue sur $[0, 1]$, $\displaystyle\int_0^1 \frac{t^\alpha + 1}{3t^\beta + 5} dt$ converge et donc :
$\displaystyle\int_0^{+\infty} \frac{t^\alpha + 1}{3t^\beta + 5} dt$ converge si, et seulement si, $\displaystyle\int_1^{+\infty} \frac{t^\alpha + 1}{3t^\beta + 5} dt$ converge si, et seulement si, $\beta - \alpha > 1$.
2. Soit $f : t \mapsto t^\alpha e^{-t}$. Alors f est définie, continue et positive sur $[1, +\infty[$ et on a, par croissances comparées :
$$t^2 f(t) = t^{2+\alpha} e^{-t} \xrightarrow[t \to +\infty]{} 0$$

donc $f(t) = \underset{t\to+\infty}{o}\left(\frac{1}{t^2}\right)$; or, d'après le critère de Riemann en $+\infty$, $\int_1^{+\infty} \frac{1}{t^2}dt$ converge, donc, par comparaison :

$$\text{pour tout } \alpha > 0, \quad \int_1^{+\infty} t^\alpha e^{-t}dt \text{ converge.}$$

3. Soit $\alpha, \beta \in \mathbb{R}$. On pose $f : t \mapsto \frac{1}{t^\alpha(\ln(t))^\beta}$. Alors f est définie, continue et positive sur $[2, +\infty[$.
   — 1er cas : $\alpha > 1$.
   On pose $a = \frac{\alpha+1}{2}$ et ainsi $\alpha > a > 1$. Alors $\alpha - a > 0$ et donc, par croissances comparées lorsque $\beta < 0$ et par produit de limites lorsque $\beta \geq 0$:
   $$t^a f(t) = \frac{1}{t^{\alpha-a}} \ln(t)^{-\beta} \xrightarrow[n\to+\infty]{} 0$$
   donc $f(t) = o\left(\frac{1}{t^a}\right)$ en $+\infty$.
   Or, d'après le critère de Riemann en $+\infty$, comme $a > 1$, $\int_2^{+\infty} \frac{1}{t^a}dt$ converge, donc, par comparaison, $\int_2^{+\infty} f(t)dt$ converge.
   — 2eme cas : $\alpha < 1$.
   Si $\beta \leq 0$, pour tout $t \geq e$, on a $\ln(t)^{-\beta} \geq 1$ donc, pour tout $t \geq e$:
   $$f(t) \geq \frac{1}{t^\alpha}.$$
   Or, d'après le critère de Riemann en $+\infty$, comme $\alpha < 1$, $\int_e^{+\infty} \frac{1}{t^\alpha}dt$ diverge, donc, par comparaison, $\int_e^{+\infty} f(t)dt$ diverge et ainsi, $\int_2^{+\infty} f(t)dt$ diverge.
   On suppose $\beta > 0$. Comme pour tout $a > 0$, par croissances comparées, $\frac{\ln(t)^\beta}{t^a} \xrightarrow[t\to+\infty]{} 0$, il existe $t_0 > 1$ tel que, pour tout $t \geq t_0$, $\frac{\ln(t)^\beta}{t^a} \leq 1$ et donc, pour tout $t \geq t_0$:
   $$\frac{1}{\ln(t)^\beta} \geq \frac{1}{t^a}.$$
   On pose $a = \frac{1-\alpha}{2}$. Alors $\alpha + a = \frac{1+\alpha}{2} < 1$ et $a > 0$, donc il existe $t_0 > 1$ tel que, pour tout $t \geq t_0$:
   $$f(t) = \frac{1}{t^\alpha \ln(t)^\beta} \geq \frac{1}{t^{\alpha+a}}$$
   Or, d'après le critère de Riemann en $+\infty$, comme $\alpha + a < 1$, $\int_{t_0}^{+\infty} \frac{1}{t^{\alpha+a}}dt$ diverge, et donc, par comparaison, $\int_{t_0}^{+\infty} f(t)dt$ diverge et ainsi, $\int_2^{+\infty} f(t)dt$ diverge.
   — 3eme cas : $\alpha = 1$.
   Pour tout $t \geq 2$,
   $$f(t) = \frac{1}{t\ln(t)^\beta} = \begin{cases} \frac{d}{dt}\frac{1}{(1-\beta)\ln(t)^{\beta-1}} & \text{si } \beta \neq 1 \\ \frac{d}{dt}\ln(\ln(t)) & \text{si } \beta = 1 \end{cases}$$

Par suite, pour $x \geq 2$:

$$\int_2^x f(t)dt = \begin{cases} \frac{1}{(1-\beta)}(\ln(x)^{1-\beta} - \ln(2)^{1-\beta}) \xrightarrow[x\to+\infty]{} +\infty & \text{si } \beta < 1 \\ \ln(\ln(x)) - \ln(\ln(2)) \xrightarrow[x\to+\infty]{} +\infty & \text{si } \beta = 1 \\ \frac{1}{(\beta-1)}\left(\frac{1}{\ln(2)^{\beta-1}} - \frac{1}{\ln(x)^{\beta-1}}\right) \xrightarrow[x\to+\infty]{} \frac{1}{(\beta-1)\ln(2)^{\beta-1}} & \text{si } \beta > 1 \end{cases}$$

Donc, par définition, $\int_2^{+\infty} f(t)dt$ converge si, et seulement si, $\beta > 1$.

Conclusion : on a donc : $\int_2^{+\infty} f(t)dt$ converge si, et seulement si, $\alpha > 1$ ou ($\alpha = 1$ et $\beta > 1$).

3. Intégrabilité

Définition 2. *Intégrabilité*

Soit $f \in C_{pm}([a, +\infty[, \mathbb{R})$. On dit que f est **intégrable** sur $[a, +\infty[$ ou que $\displaystyle\int_a^{+\infty} f$ est **absolument convergente** si l'intégrale $\displaystyle\int_a^{+\infty} |f|$ est convergente.

Théorème 3.

Soit $f \in C_{pm}([a, +\infty[, \mathbb{R})$. Si f est intégrable sur $[a, +\infty[$, alors $\displaystyle\int_a^{+\infty} f$ converge.

Démonstration.

On suppose que f est intégrable sur $[a, +\infty[$. On note $f_+ = \max(f, \mathbf{0})$ qui est continue par morceaux et positive sur $[a, +\infty[$ et $f_- = \min(f, \mathbf{0})$ qui est continue par morceaux et négative sur $[a, +\infty[$. Alors $f = f_+ + f_-$, $|f| = f_+ - f_-$ et de plus :

$$0 \leq f_+ \leq |f| \text{ et } 0 \leq -f_- \leq |f|$$

Par suite, comme $\int_a^{+\infty} |f|$ converge et $f_+, -f_-$ sont positives, par comparaison, $\int_a^{+\infty} f_+$ et $\int_a^{+\infty} -f_-$ convergent. Ainsi, comme $f = f_+ + (-1)(-f_-)$, par combinaison linéaire, $\int_a^{+\infty} f$ converge. □

Exercice 5.

Déterminer la nature de $\displaystyle\int_1^{+\infty} \frac{t\sin(t)}{\sqrt{\operatorname{sh}(t)}}dt$.

Correction.

La fonction $f : t \mapsto \frac{t\sin(t)}{\sqrt{\mathrm{sh}(t)}}$ est continue sur $[1, +\infty[$ et on a, pour tout $t \in [1, +\infty[$:

$$\left|\frac{t\sin(t)}{\sqrt{\mathrm{sh}(t)}}\right| = \frac{t|\sin(t)|}{\sqrt{\mathrm{sh}(t)}} \leq \frac{t}{\sqrt{\mathrm{sh}(t)}} \underset{t\to+\infty}{\sim} \sqrt{2}te^{-\frac{t}{2}} = \underset{t\to+\infty}{o}\left(\frac{1}{t^2}\right)$$

Or, d'après le critère de Riemann en $+\infty$ $(2 > 1)$, $\int_1^{+\infty} \frac{1}{t^2}dt$ converge, donc par comparaison, $\int_1^{+\infty} |f(t)|dt$ converge i.e. f est intégrable sur $[1, +\infty[$.
Il en résulte que $\int_1^{+\infty} f(t)dt$ converge.

Partie B

Intégration sur un intervalle quelconque

Dans cette partie, a, b désignent des réels tels que $a < b$.

1. Intégrale généralisée sur $]-\infty, a]$

Définition 3. *Intégrale convergente sur* $]-\infty, a]$

Soit $f \in C_{pm}(]-\infty, a], \mathbb{K})$.
Si la limite $\lim\limits_{x \to -\infty} \int_x^a f$ existe et est finie, on note $\int_{-\infty}^a f$ ou encore $\int_{-\infty}^a f(t)dt$ cette quantité, i.e. :

$$\int_{-\infty}^a f(t)dt = \lim_{x \to -\infty} \int_x^a f(t)dt.$$

Dans ce cas, on dit que $\int_{-\infty}^a f$ **converge**.

Dans le cas contraire, on dit que $\int_{-\infty}^a f$ **diverge**.

Remarque 1.

Les propriétés de l'intégrale $\int_{-\infty}^a$ sont analogues à celle de l'intégrale $\int_a^{+\infty}$ et on les démontre sans difficulté en remarquant qu'en posant $\tilde{f} : t \mapsto f(-t)$, on obtient les propriétés suivantes :

— si $f \in C_{pm}(]-\infty, a], \mathbb{K})$ alors $\tilde{f} \in C_{pm}([-a, +\infty], \mathbb{K})$;

— les intégrales $\int_{-\infty}^a f$ et $\int_{-a}^{+\infty} \tilde{f}$ sont de même nature ;

— en cas de convergence, $\int_{-\infty}^a f(t)dt = \int_{-a}^{+\infty} \tilde{f}(t)dt$.

Exercice 6.

Soit $\alpha \in \mathbb{R}$. Étudier la convergence de l'intégrale $\int_{-\infty}^0 e^{\alpha t}dt$ en fonction de α.

Correction.

Soit $\alpha \in \mathbb{R}$. Considérons $f : t \mapsto e^{\alpha t}$. Alors

$$F : t \mapsto \begin{cases} \frac{e^{\alpha t}}{\alpha} & \text{si } \alpha \neq 0 \\ t & \text{si } \alpha = 0 \end{cases}$$

est une primitive de f sur $\mathbb{R}$ et donc sur $]-\infty, 0]$. Or F admet une limite finie en $-\infty$ si, et seulement si, $\alpha > 0$. Donc $\int_{-\infty}^{0} e^{\alpha t}dt$ converge si, et seulement si, $\alpha > 0$. Et dans ce cas,

$$\int_{-\infty}^{0} e^{\alpha t}dt = \left[\frac{e^{\alpha t}}{\alpha}\right]_{-\infty}^{0} = \frac{e^{\alpha 0}}{\alpha} - \lim_{t\to-\infty} \frac{e^{\alpha t}}{\alpha} = \frac{1}{\alpha}.$$

Remarque : on aurait également pu utiliser la remarque précédente en remarquant que $g : t \mapsto e^{-\alpha t} = f(-t)$ est intégrable sur $[0, +\infty[$ si, et seulement si, $\alpha > 0$.

2. Intégrale généralisée sur $]a, b]$ ou $[a, b[$

Définition 4. *Intégrale convergente sur $[a, b[$*

Soit $f \in C_{pm}([a, b[, \mathbb{K})$.

Si la limite $\lim\limits_{x\to b} \displaystyle\int_a^x f$ existe et est finie, on note $\displaystyle\int_a^b f$ ou encore $\displaystyle\int_a^b f(t)dt$ cette quantité, i.e. :

$$\int_a^b f(t)dt = \lim_{x\to b} \int_a^x f(t)dt.$$

Dans ce cas, on dit que $\displaystyle\int_a^b f$ **converge**.

Dans le cas contraire, on dit que $\displaystyle\int_a^b f$ **diverge**.

Définition 5. *Intégrabilité*

Soit $f \in C_{pm}(]a, b], \mathbb{R})$. On dit que f est **intégrable** sur $]a, b]$ ou que $\displaystyle\int_a^b f$ est **absolument convergente** si l'intégrale $\displaystyle\int_a^b |f|$ est convergente.

Remarque 2.

On définit de manière analogue l'intégrale sur un intervalle du type $]a, b]$.

Tous les résultats de la partie concernant l'intégrale sur $[?, +\infty[$ sont transposables aux cas des intégrales sur $]a, b]$ et sur $[a, b[$; citons notamment **les propriétés de comparaison** et le théorème "intrégrabilité d'une fonction implique convergence de l'intégrale de cette fonction" reproduit ci-après.

Théorème 4.

Soit $f \in C_{pm}([a,b[,\mathbb{R})$. Si f est intégrable sur $[a,b[$, alors $\int_a^b f$ converge.

Démonstration.

La démonstration est analogue au cas $[a,+\infty[$ en remplaçant simplement $+\infty$ par b. $\square$

Exemple 2.

— L'intégrale $\int_0^1 \ln(t)dt$ converge et $\int_0^1 \ln(t)dt = -1$

— L'intégrale $\int_0^1 \frac{1}{t}dt$ diverge

Proposition 8. *Intégrales de Riemann*

Soit $\alpha \in \mathbb{R}$. Les intégrales $\int_a^b \frac{1}{(b-t)^\alpha}dt$ et $\int_a^b \frac{1}{(t-a)^\alpha}dt$ convergent si, et seulement si, $\alpha < 1$.

Démonstration.

On traite la nature de $\int_a^b \frac{1}{(t-a)^\alpha}dt$:

- Si $\alpha = 1$, pour $a < x \leq b$, on a :

$$\int_x^b \frac{1}{t-a}dt = [\ln(t-a)]_x^b = \ln(b-a) - \ln(x-a) \xrightarrow[x\to a^+]{} +\infty$$

  donc $\int_a^b \frac{1}{t-a}dt$ diverge.

- Si $\alpha \neq 1$, on a, pour $a < x \leq b$:

$$\begin{aligned} \int_x^b \frac{1}{(t-a)^\alpha}dt &= \left[\frac{1}{(1-\alpha)(t-a)^{\alpha-1}}\right]_x^b \\ &= \frac{1}{(1-\alpha)(b-a)^{\alpha-1}} - \frac{1}{(1-\alpha)(x-a)^{\alpha-1}} \end{aligned}$$

$$\int_x^b \frac{1}{(t-a)^\alpha}dt \quad \begin{cases} \xrightarrow[x\to 0^+]{} \dfrac{1}{(1-\alpha)(b-a)^{\alpha-1}} & \text{si } \alpha < 1; \\ \xrightarrow[x\to 0^+]{} +\infty & \text{si } \alpha > 1. \end{cases}$$

Il en résulte que $\int_a^b \frac{1}{(t-a)^\alpha}dt$ converge si, et seulement si, $\alpha < 1$. $\square$

Corollaire 1. *Intégrale de Riemann en* 0

Soit $\alpha \in \mathbb{R}$. Les intégrales $\displaystyle\int_0^1 \frac{1}{t^\alpha} dt$ converge si, et seulement si, $\alpha < 1$.

Démonstration.

On applique la proposition précédente (deuxième intégrale) pour $a = 0$ et $b = 1$.

Voici tout de même la démonstration directe pour une potentielle question de colle :) :

- Si $\alpha = 1$, pour $0 < x \leq 1$, on a :

$$\int_x^1 \frac{1}{t} dt = [\ln(t)]_x^1 = -\ln(x) \xrightarrow[x \to 0^+]{} +\infty$$

donc $\displaystyle\int_0^1 \frac{1}{t} dt$ diverge.

- Si $\alpha \neq 1$, on a, pour $0 < x \leq 1$:

$$\int_x^1 \frac{1}{t^\alpha} dt = \left[\frac{1}{(1-\alpha)t^{\alpha-1}} \right]_x^1 = \frac{1}{1-\alpha} - \frac{1}{(1-\alpha)x^{\alpha-1}} \begin{cases} \xrightarrow[x \to 0^+]{} \dfrac{1}{1-\alpha} & \text{si } \alpha < 1; \\ \xrightarrow[x \to 0^+]{} +\infty & \text{si } \alpha > 1. \end{cases}$$

Il en résulte que $\displaystyle\int_0^1 \frac{1}{t^\alpha} dt$ converge si, et seulement si, $\alpha < 1$. □

Proposition 9.

Soit $f \in C_{pm}([a, b[, \mathbb{K})$.

— Si f est bornée sur $[a, b[$, alors f est intégrable sur $[a, b[$.

— Si f est prolongeable par continuité en b i.e. si f admet une limite finie en b, alors f est intégrable sur $[a, b[$.

Démonstration.

Soit $f \in C_{pm}([a, b[, \mathbb{K})$.

— On suppose f bornée sur $[a, b[$. Alors il existe $M \in \mathbb{R}_+$ tel que pour tout $t \in [a, b[$, $|f(t)| \leq M$. Or, la fonction constante en M sur $[a, b[$ est d'intégrale convergente, donc par comparaison (d'integrales de fonctions positives !), $\int_a^b |f(t)| dt$ converge i.e. f est intégrable sur $[a, b[$.

— On remarque qu'une fonction f continue par morceaux sur $[a, b[$ et prolongeable par continuité en b est la restriction d'une fonction g continue par morceaux sur $[a, b]$. Or, toute fonction continue par morceaux sur un **segment** est bornée sur ce segment (*Exercice : prouver cette affirmation ; puis trouver un exemple de fonction continue par morceaux sur* $[0, 1[$ *qui n'est pas bornée*), donc f est bornée sur $[a, b[$ comme restriction d'une fonction bornée. Ainsi, d'après le cas précédent, f est intégrable sur $[a, b[$.

□

Exercice 7.

Étudier la nature des intégrales suivantes :

$$\int_0^1 \frac{t^3}{t^{\frac{7}{2}}+1}dt \quad \int_0^1 \frac{\ln(1+t)}{t\sqrt{t}}dt \quad \int_0^1 \sin(\frac{1}{t})dt \quad \int_0^1 \frac{1}{e^{1-t}-1}dt.$$

Correction.

1. Aucun souci : la fonction $t \mapsto \dfrac{t^3}{t^{\frac{7}{2}}+1}$ est continue sur le segment $[0,1]$ donc intégrable sur ce segment ! Ainsi, l'intégrale converge.
2. La fonction $f : t \mapsto \dfrac{\ln(1+t)}{t\sqrt{t}}$ est continue sur $]0,1]$. Étudions f au voisinage de 0. On a :
$$|f(t)| = \frac{\ln(1+t)}{t\sqrt{t}} \underset{t\to 0}{\sim} \frac{t}{t\sqrt{t}} = \frac{1}{\sqrt{t}}$$
Or, d'après le critère de Riemann en 0, $\int_0^1 \frac{1}{\sqrt{t}}dt$ converge car $\frac{1}{2} < 1$; donc par comparaison $\int_0^1 |f(t)|dt = \int_0^1 |f(t)|dt$ converge.
3. Comme la fonction $|\sin|$ est bornée par 1 sur $\mathbb{R}$, la fonction $t \mapsto |\sin(\frac{1}{t})|$ est bornée sur 1 sur $]0,1]$. Or la fonction constante en 1 est d'intégrale convergente sur $]0,1]$ d'où $\int_0^1 |\sin(\frac{1}{t})|dt$ converge et donc $\int_0^1 \sin(\frac{1}{t})dt$ converge. (*On pouvait bien-sûr faire directement appel à la proposition précédente, mais ça ne coûte pas très cher de refaire ce raisonnement à chaque fois !*).
4. La fonction $f : t \mapsto \frac{1}{e^{1-t}-1}$ est positive et continue sur $[0,1[$.
Étudions f pour t au voisinage de 1^-. On a le développement limité à l'ordre 1 de la fonction exp suivant :
$e^u = 1 + u + \underset{u\to 0}{o}(u^2)$. Ainsi, en prenant $u = 1-t \xrightarrow[t\to 1^-]{} 0$, on obtient :
$$e^{1-t} = 1 + (1-t) + \underset{t\to 1}{o}((1-t)^2)$$
Ainsi,
$$f(t) = \frac{1}{e^{1-t}-1} = \frac{1}{(1-t) + \underset{t\to 1}{o}((1-t)^2)} \underset{t\to 1^-}{\sim} \frac{1}{1-t}$$
Or, d'après le critère de Riemann en 1, $\int_0^1 \frac{1}{1-t}dt$ diverge car $1 \geq 1$; donc par comparaison $\int_0^1 f(t)dt$ diverge.

3. Intégration sur un intervalle ouvert

Dans ce paragraphe a, b peuvent être égaux à $-\infty$ et $+\infty$ respectivement et on s'intéresse aux intégrales sur l'intervalle ouvert $]a,b[$.

Définition 6. *Intégrale convergente sur* $]a,b[$

Soit $f \in C_{pm}(]a,b[,\mathbb{K})$.

S'il existe $c \in]a,b[$ tel que $\displaystyle\int_a^c f$ **et** $\displaystyle\int_c^b f$ convergent, alors on note $\displaystyle\int_a^b f$ ou encore $\displaystyle\int_a^b f(t)dt$ la quantité :

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Dans ce cas, on dit que l'intégrale $\displaystyle\int_a^b f$ **converge**.

Dans le cas contraire, on dit que l'intégrale $\displaystyle\int_a^b f$ **diverge**.

Proposition 10.

Soit $f \in C_{pm}(]a,b[,\mathbb{K})$. Si $\displaystyle\int_a^b f$ converge, alors, pour tout $c \in]a,b[$, $\displaystyle\int_a^c f$ et $\displaystyle\int_c^b f$ convergent et on a :

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Notation 1.

Soit $F \in C(]a,b[,\mathbb{K})$. On note

$$[F(t)]_a^b = \lim_{t\to b} F(t) - \lim_{t\to a} F(t)$$

lorsque ces deux limites existent et sont finies.

Proposition 11.

Soit $f \in C(]a,b[,\mathbb{K})$ et F une primitive de f sur $]a,b[$. L'intégrale $\displaystyle\int_a^b f$ converge si, et seulement si, la fonction F admet des limites finies en a et b. Dans ce cas, on a :

$$\int_a^b f(t)dt = [F(t)]_a^b.$$

Exercice 8.

Étudier la convergence des intégrales suivantes :

$$\int_0^1 \frac{1}{\sqrt{t(1-t)}}dt \quad \int_1^{+\infty} \sin(t)\ln\left(\frac{t^2+1}{t^2-1}\right)dt \quad \int_0^{\frac{\pi}{2}} \frac{1}{\cos(t)\sqrt{e^t-1}}dt.$$

Correction.

1. La fonction $f : t \mapsto \frac{1}{\sqrt{t(1-t)}}$ est positive et continue sur $]0,1[$. On étudie donc f au voisinage de 0 puis de 1 :
   — <u>en 0 :</u> on a
   $$f(t) \underset{t\to 0^+}{\sim} \frac{1}{\sqrt{t}}$$
   Or $\int_0^{\frac{1}{2}} \frac{1}{\sqrt{t}}dt$ converge d'après le critère de Riemann en 0 ($\frac{1}{2} < 1$), donc, par comparaison, $\int_0^{\frac{1}{2}} \frac{1}{\sqrt{t(1-t)}}dt$ converge.
   — <u>en 1 :</u> on a
   $$f(t) \underset{t\to 1^-}{\sim} \frac{1}{\sqrt{1-t}}$$
   Or $\int_{\frac{1}{2}}^1 \frac{1}{\sqrt{1-t}}dt$ converge d'après le critère de Riemann en 0 ($\frac{1}{2} < 1$), donc, par comparaison, $\int_{\frac{1}{2}}^1 \frac{1}{\sqrt{t(1-t)}}dt$ converge.

   Comme $\int_0^{\frac{1}{2}} \frac{1}{\sqrt{t(1-t)}}dt$ et $\int_{\frac{1}{2}}^1 \frac{1}{\sqrt{t(1-t)}}dt$ convergent, par définition, $\int_0^1 \frac{1}{\sqrt{t(1-t)}}dt$ converge.

2. La fonction $f : t \mapsto \sin(t)\ln\left(\frac{t^2+1}{t^2-1}\right)$ est continue sur $]1,+\infty[$. On étudie donc f au voisinage de 1 puis de $+\infty$:
   — <u>en 1 :</u> on remarque que $\ln\left(\frac{t^2+1}{t^2-1}\right) = \ln(t^2+1) - \ln(t+1) - \ln(t-1)$ donc :
   $$|f(t)| \underset{t\to 1^+}{\sim} -\sin(1)\ln(t-1) = \underset{t\to 1^+}{o}\left(\frac{1}{\sqrt{t-1}}\right)$$
   Or $\int_1^2 \frac{1}{\sqrt{t-1}}dt$ converge d'après le critère de Riemann en 1 ($\frac{1}{2} < 1$), donc, par comparaison, $\int_1^2 |f(t)|dt$ converge.
   — <u>en $+\infty$:</u> on remarque que $\ln\left(\frac{t^2+1}{t^2-1}\right) = \ln(1+\frac{2}{t^2-1})$ donc, comme $\frac{2}{t^2-1} \xrightarrow[t\to+\infty]{} 0$:
   $$|f(t)| \leq \ln\left(\frac{t^2+1}{t^2-1}\right) \underset{t\to+\infty}{\sim} \frac{2}{t^2}$$
   Or $\int_2^{+\infty} \frac{1}{t^2}dt$ converge d'après le critère de Riemann en $+\infty$ ($2 > 1$), donc, par comparaison, $\int_2^{+\infty} |f(t)|dt$ converge.

   Comme $\int_1^2 \frac{1}{|}f(t)|dt$ et $\int_2^{+\infty} |f(t)|dt$ convergent, par définition, $\int_0^{+\infty} |f(t)|dt$ converge. Ainsi, f est intégrable sur $]1,+\infty[$ et donc $\int_1^{+\infty} f(t)dt$ converge.

3. La fonction $f : t \mapsto \frac{1}{\cos(t)\sqrt{e^t-1}}$ est continue et positive sur $]0,\frac{\pi}{2}[$. On étudie donc f au voisinage de 0 puis de $\frac{\pi}{2}$.
   — <u>en 0 :</u> comme $e^t - 1 \underset{t\to 0}{\sim} t$, on a :
   $$f(t) \underset{t\to 0}{\sim} \frac{1}{\sqrt{t}}$$
   Or $\int_0^1 \frac{1}{\sqrt{t}}dt$ converge d'après le critère de Riemann en 0 ($\frac{1}{2} < 1$), donc, par comparaison, $\int_0^1 f(t)dt$ converge.

— en $\frac{\pi}{2}$: on remarque que $\cos(t) = \sin(\frac{\pi}{2} - t) \underset{t \to \frac{\pi}{2}}{\sim} \frac{\pi}{2} - t$ donc :

$$f(t) \underset{t \to +\infty}{\sim} \frac{1}{(\frac{\pi}{2} - t)\sqrt{e^{\frac{\pi}{2}} - 1}}$$

Or $\int_1^{\frac{\pi}{2}} \frac{1}{\frac{\pi}{2} - t} dt$ diverge d'après le critère de Riemann en $\frac{\pi}{2}$ $(1 \geq 1)$, donc, par comparaison, $\int_1^{\frac{\pi}{2}} f(t)dt$ diverge.

Une des deux intégrales (au moins) $\int_0^1 f(t)dt$ et $\int_1^{\frac{\pi}{2}} f(t)dt$ diverge, donc $\int_0^{\frac{\pi}{2}} f(t)dt$ diverge.

Remarque : si on avait commencé par faire l'étude en $\frac{\pi}{2}$, on aurait pu conclure directement à la divergence de $\int_0^{\frac{\pi}{2}} f(t)dt$ sans faire l'étude en 0.

Partie C
Propriétés de l'intégrale généralisée

Dans cette partie, I désigne un intervalle quelconque (ouvert, semi-ouvert ou fermé, de longueur finie ou infinie) de $\mathbb{R}$ d'intérieur non vide et on note a et b ses extrémités (possiblement infinies).

1. Propriétés de l'intégrale

Dans ce paragraphe, on résume et généralise les proporiétés de l'intégrale généralisée ; les démonstrations sont laissées à titre d'exercices au lecteur.

a. Espace de fonctions continues par morceaux intégrables

Proposition 12.

L'ensemble $E = \{f \in C_{pm}(I, \mathbb{K}) \mid f \text{ est intégrable}\}$ est un espace vectoriel sur $\mathbb{K}$ et l'application $f \mapsto \displaystyle\int_I f$ est une forme linéaire sur E.
De plus, cette application est positive et croissante.

b. Inégalités

Proposition 13. *Inégalité triangulaire*

Soit $f, g \in C_{pm}(I, \mathbb{K})$ des fonctions intégrables. Alors

$$\int_I |f(t) + g(t)|dt \leq \int_I |f(t)|dt + \int_I |g(t)|dt.$$

Proposition 14.

Soit $f \in C_{pm}(I, \mathbb{K})$ une fonction intégrable. Alors

$$\left|\int_I f(t)dt\right| \leq \int_I |f(t)|dt.$$

c. Séparation

Proposition 15. *Séparation*

Si f est une fonction continue, positive et intégrable sur I, alors

$$\int_I f(t)dt = 0 \text{ si, et seulement si, } f = \mathbf{0}.$$

d. Relation de Chasles

Proposition 16.

Soit $f \in C_{pm}(I, \mathbb{K})$ une fonction intégrable. Alors pour tout $a, b, c \in I$ on a :

$$\int_a^c f(t)dt = \int_a^b f(t)dt + \int_b^c f(t)dt.$$

2. Calculs d'intégrales

a. Intégration par parties

Proposition 17. *Intégration par parties*

Soit $f, g \in C^1(I, \mathbb{K})$. Si la fonction produit fg admet des **limites finies** aux bornes a et b de I, alors les intégrales $\displaystyle\int_I fg'$ et $\displaystyle\int_I f'g$ sont de même nature. De plus, en cas de convergence, on a :

$$\int_a^b f'(t)g(t)dt = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)dt.$$

Démonstration.

Quitte à scinder l'intervalle I en deux et utiliser la définition 6, on peut supposer "qu'il n'y a de problème potentiel qu'en b" i.e. $I = [a, b[$.
Soit $x \in I$. Comme f et g sont de classe C^1 sur le segment $[a, x]$, alors d'après le théorème d'intégration par parties sur un segment (vu en Sup' - on rappelle que sa démonstration repose sur la formule de la dérivée du produit de fonctions !), on a :

$$\int_a^x f'(t)g(t)dt = [f(t)g(t)]_a^x - \int_a^x f(t)g'(t)dt. \quad (*)$$

On suppose que le produit fg admet une limite finie en b (d'après notre réduction initiale, c'est déjà le cas en a car f et g sont en particulier continues en a). Alors $[f(t)g(t)]_a^x$ tend vers la quantité finie $[f(t)g(t)]_a^b$ lorsque x tend vers b. Ainsi, d'après l'égalité $(*)$, $\int_a^x f'(t)g(t)dt$ admet une limite finie quand x tend vers b si, et seulement si, $\int_a^x f(t)g'(t)dt$ admet une limite finie quand x tend vers b d'où par définition, $\displaystyle\int_I fg'$ et $\displaystyle\int_I f'g$ sont de même nature.

De plus, en cas de convergence de $\displaystyle\int_I fg'$ ou $\displaystyle\int_I f'g$, toujours d'après (∗), on a :

$$\begin{aligned}\int_a^x f'(t)g(t)dt &= \lim_{x\to b}\int_a^x f'(t)g(t)dt \\ &= \lim_{x\to b}\left([f(t)g(t)]_a^x - \int_a^x f(t)g'(t)dt\right) \\ \int_a^x f'(t)g(t)dt &= [f(t)g(t)]_a^b - \int_a^b \text{`}f(t)g'(t)dt.\end{aligned}$$

□

Exercice 9.

1. Justifier la convergence de $\displaystyle\int_0^{+\infty}\frac{\sin(t)}{t}dt$
2. Justifier la convergence puis calculer $\displaystyle\int_0^1\frac{\ln(1-t^2)}{t^2}dt$.
3. On note, pour $n\in\mathbb{N}^*$, $I_n=\displaystyle\int_0^{+\infty}\frac{1}{(t^2+1)^n}dt$.
   i) Justifier l'existence de I_n pour $n\in\mathbb{N}^*$.
   ii) Soit $n\in\mathbb{N}^*$. Déterminer une relation entre I_{n+1} et I_n. En déduire I_n.

Correction.

1. On pose $f:t\mapsto\frac{\sin(t)}{t}$. Alors f est continue sur $]0,+\infty[$ et comme $|f(t)|\underset{t\to0^+}{\sim}1$, f est intégrable sur $]0,1]$ (elle est même prolongeable par continuité en 0 par $f(0)=1$). Par suite $\int_0^1 f(t)dt$ converge.

   On effectue une intégration par paties pour prouver la convergence de $\int_1^{+\infty}\frac{\sin(t)}{t}dt$. On pose $u:t\mapsto\frac{1}{t}$ et $v:t\mapsto-\cos(t)$. Alors u,v sont C^1 sur $[1,+\infty[$; le produit uv est défini en 1, vaut $u(1)v(1)=-\cos(1)$ en ce point et $u(t)v(t)=\frac{-\cos(t)}{t}\xrightarrow[t\to+\infty]{}0$.

   Ainsi, par intégration par parties, $\int_1^{+\infty}u(t)v'(t)dt=\int_1^{+\infty}\frac{\sin(t)}{t}dt$ et $\int_1^{+\infty}u'(t)v(t)dt=\int_1^{+\infty}\frac{\cos(t)}{t^2}dt$ sont de même nature.

   Etudions cette seconde intégrale : la fonction $g:t\mapsto\frac{\cos(t)}{t^2}$ est continue sur $[1,+\infty[$ et en $+\infty$, on a :

   $$|g(t)|=\frac{|\cos(t)|}{t^2}\leq\frac{1}{t^2}\ \text{intégrable sur } [1,+\infty[\text{ d'après le critère de Riemann en } +\infty$$

   Ainsi, par comparaison, g est intégrable sur $[1,+\infty[$ et donc les intégrales $\int_1^{+\infty}\frac{\cos(t)}{t^2}dt$ et $\int_1^{+\infty}\frac{\sin(t)}{t}dt$ convergent.

   Comme $\int_0^1\frac{\sin(t)}{t}dt$ et $\int_1^{+\infty}\frac{\sin(t)}{t}dt$ converge, il en résulte, par la relation de Chasles, que

$\int_0^{+\infty} \frac{\sin(t)}{t}dt$ converge.

2. Ici, on peut étudier directement la nature de cette intégrale avec les techniques vues précédemment. On pose $f : t \mapsto \frac{\ln(1-t^2)}{t^2}$. Alors f est continue sur $]0,1[$. On étudie donc f au voisinage de 0 et de 1 :
   — en 0 : on a $\ln(1-t^2) \underset{t\to 0}{\sim} t^2$ donc $f(t) \xrightarrow[t\to 0]{} 1$. Ainsi f est prolongeable par continuité en 0 (par 1) et donc intégrable au voisinage de 0. Ainsi, $\int_{\frac{1}{2}}^1 f(t)dt$ converge.
   — en 1 : on remarque que $\ln(1-t^2) = \ln(1-t) + \ln(1+t)$ donc :
   $$|f(t)| \underset{t\to 1}{\sim} -\ln(1-t) = \underset{t\to 0}{o}\left(\frac{1}{\sqrt{1-t}}\right)$$
   Or, d'après le critère de Riemann en 1, $\int_{\frac{1}{2}}^1 \frac{1}{\sqrt{1-t}}$ converge, donc par comparaison, f est intégrable au voisinage de 1. Ainsi, $\int_{\frac{1}{2}}^1 f(t)dt$ converge.

   Par suite, $\int_0^1 f(t)dt$ converge.

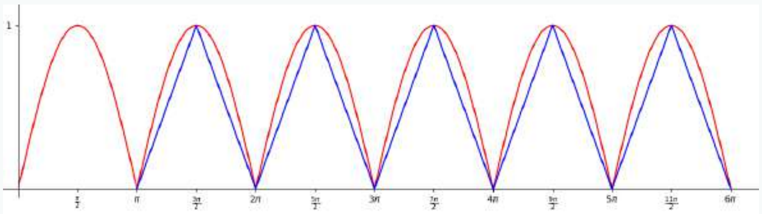
   Passons au calcul de $\int_0^1 f(t)dt$: utilisons une intégration par partie. On va se rendre compte qu'il faut parfois être subtil dans nos primitives :
   — *Premier essai, sans subtilité :* on considère les deux fonctions u, v de classe C^1 sur $]0,1[$ telles que, pour $t \in]0,1[$:
   $$u'(t) = \frac{1}{t^2} \qquad u(t) = -\frac{1}{t}$$
   $$v(t) = \ln(1-t^2) \qquad v'(t) = -\frac{2t}{1-t^2}.$$
   Pour appliquer le théorème d'intégration par parties, il faut vérifier que uv admet des limites en 0 et 1. Or ici, on a $\lim_{t\to 1} u(t)v(t) = +\infty$ "à cause" du $\ln(1-t^2)$... oups ! On ne peut donc pas faire notre IPP !
   Mais on se rend compte qu'on peut faire en sorte de compenser le fait que $\ln(1-t^2)$ tende vers $-\infty$ en choisissant plus intelligemment notre primitive de $\frac{1}{t^2}$: en effet, ici, $-\frac{1}{t}$ tend vers -1 alors qu'on voudrait "au minimum" que la primitive tende vers 0 pour potentiellement compenser le ∞ ! Ce qui nous amène au :
   — *Deuxième essai, avec subtilité :* on considère les deux fonctions u, v de classe C^1 sur $]0,1[$ telles que, pour $t \in]0,1[$:
   $$u'(t) = \frac{1}{t^2} \qquad u(t) = 1 - \frac{1}{t} = -\frac{1-t}{t}$$
   $$v(t) = \ln(1-t^2) \qquad v'(t) = -\frac{2t}{1-t^2}.$$
   Cette fois-ci, on a bien une limite finie en 1 :
   $$u(t)v(t) = -\frac{(1-t)\ln(1-t^2)}{t} \xrightarrow[t\to 1]{} 0$$
   et en 0 :
   $$u(t)v(t) = -\frac{(1-t)\ln(1-t^2)}{t} \xrightarrow[t\to 0]{} 0.$$

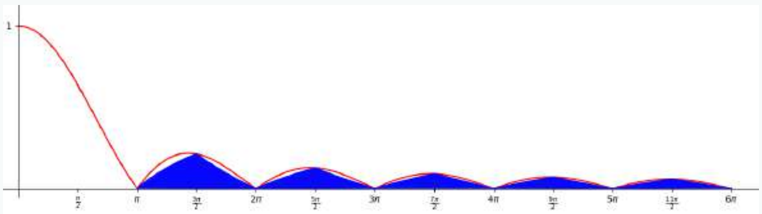
Ainsi, uv admet des limites finies en 0 et 1, donc d'après le théorème d'intégration par parties, $\int_0^1 u'v$ et $\int_0^1 uv'$ sont de même nature. Or on a prouvé que $\int_0^1 u'v$ converge, donc $\int_0^1 uv'$ converge et toujours d'après le théorème d'intégration par parties, on a :

$$\begin{aligned}
\int_0^1 f(t)dt &= \int_0^1 u'(t)v(t)dt \\
&= [u(t)v(t)]_0^1 - \int_0^1 u(t)v'(t)dt \\
&= -\int_0^1 \frac{1-t}{t} \times \frac{2t}{1-t^2}dt \\
&= -2\int_0^1 \frac{1}{1+t}dt \\
\int_0^1 f(t)dt &= -2\ln(2).
\end{aligned}$$

Remarque : On vient de montrer dans l'exercice précédent que $\int_0^{+\infty} \frac{\sin(t)}{t}dt$ converge, mais il est intéressant de remarquer que la fonction $f : t \mapsto \frac{\sin(t)}{t}$ n'est pas intégrable sur $[0, +\infty[$ i.e. $\int_0^{+\infty} \frac{|\sin(t)|}{t}dt$ diverge. Ce qui montre que la réciproque de "f intégrable sur I implique $\int_I f$ converge" est fausse.

Voici une démonstration de la non-intégrabilité de f sur $[0, +\infty[$. Tout d'abord, exposons l'idée de cette démonstration :





On va minorer la fonction $|sin|$ sur $[\pi, +\infty[$ par la fonction g "triangulaire" π-périodique dont le graphe est représenté en bleu dans le premier graphique. Ainsi, l'intégrale. Ainsi, $\int_\pi^{n\pi} \frac{|sin(x)|}{x}dx \geq \int_\pi^{n\pi} \frac{g(x)}{x}dx$ et cette dernière intégrale représentée par l'aire bleu sur le deuxième graphique se minore aisément par la somme partielle d'une série divergente. On pourra ainsi conclure la divergence de l'intégrale $\int_\pi^{+\infty} \frac{|sin(x)|}{x}dx$.

Voyons cela rigoureusement :

La fonction $|\sin|$ est π-périodique sur $\mathbb{R}$ et sur $[0, \pi]$ elle est égale à la fonction sin. Comme, pour

tout $x \in [0,\pi]$, $\sin''(x) = -\sin(x) \leq 0$, sin est concave sur $[0,\pi]$. Ainsi, sa courbe est au dessus de la corde entre les abscisses 0 et $\frac{\pi}{2}$ et de la corde entre les abscisses $\frac{\pi}{2}$ et π.
Ces dernières ont pour équations respectives $y = \frac{2}{\pi}x$ et $y = -\frac{2}{\pi}(x-\pi)$.
On considère alors une fonction g définie sur $[0,\pi]$, ayant pour image l'union de ces deux cordes, par exemple :

$$g : x \mapsto \begin{cases} \frac{2}{\pi}x & \text{si } x \in [0, \frac{\pi}{2}] \\ -\frac{2}{\pi}(x-\pi) & \text{si } x \in [\frac{\pi}{2}, \pi] \end{cases}$$

Alors g est continue sur $[0,\pi]$, et pour tout $x \in [0,\pi]$, $g(x) \leq \sin(x)$.
On pose, pour $k \in \mathbb{N}^*$, $x_k = (2k+1)\frac{\pi}{2} = \frac{\pi}{2} + k\pi$. Par π-périodicité de $|\sin|$, on a, pour tout $x \in [x_k - \frac{\pi}{2}, x_k - \frac{\pi}{2}] (= [0,\pi] + k\pi)$:

$$|\sin(x)| \geq g(x - k\pi).$$

Ainsi, pour tout $k \in \mathbb{N}^*$, en utilisant un changement de variable $t = x - k\pi$, on a :

$$\int_{x_k - \frac{\pi}{2}}^{x_k + \frac{\pi}{2}} \frac{|\sin(x)|}{x}dx \geq \int_{x_k - \frac{\pi}{2}}^{x_k + \frac{\pi}{2}} \frac{g(x-k\pi)}{x}dx = \int_0^{\pi} \frac{g(t)}{t + k\pi}dt \geq \frac{1}{(k+1)\pi}\int_0^{\pi} g(t)dt$$

Par suite, par symétrie du graphe de g par rapport à la droite $t = \frac{\pi}{2}$, on a :

$$\int_{x_k - \frac{\pi}{2}}^{x_k + \frac{\pi}{2}} \frac{|\sin(x)|}{x}dx \geq \frac{2}{(k+1)\pi}\int_0^{\frac{\pi}{2}} \frac{2}{\pi}tdt = \frac{4}{(k+1)\pi^2}\left[\frac{t^2}{2}\right]_0^{\frac{\pi}{2}} = \frac{1}{2(k+1)}.$$

Alors, pour $n \in \mathbb{N}^*$, comme $x_1 - \frac{\pi}{2} = \pi$ et $x_n + \frac{\pi}{2} = (n+1)\pi$ et pour tout $k \in \mathbb{N}^*$, $x_k + \frac{\pi}{2} = (k+1)\pi = x_{k+1} - \frac{\pi}{2}$, on a :

$$\int_{\pi}^{(n+1)\pi} \frac{|\sin(x)|}{x}dx = \sum_{k=1}^{n}\int_{x_k - \frac{\pi}{2}}^{x_k + \frac{\pi}{2}} \frac{|\sin(x)|}{x}dx \geq \sum_{k=1}^{n}\frac{1}{2(k+1)}$$

Pour $x \geq \pi$, on pose $n (= n_x) = E(\frac{x}{\pi} - 1)$; alors $x \geq (n+1)\pi$ et $n \xrightarrow[x\to+\infty]{} +\infty$ et donc, $\frac{1}{2(k+1)}$ étant le terme général d'une série à termes positifs divergente :

$$\int_{\pi}^{x} \frac{|\sin(x)|}{x}dx \geq \int_{\pi}^{(n+1)\pi} \frac{|\sin(x)|}{x}dx \geq \sum_{k=1}^{n}\frac{1}{2(k+1)} \xrightarrow[x\to+\infty]{} +\infty$$

Il en résulte que $\int_{\pi}^{+\infty} \frac{|\sin(x)|}{x}dx$ diverge et donc $\int_0^{+\infty} \frac{|\sin(x)|}{x}dx$ également. Ainsi, f n'est pas intégrable sur $[0,+\infty[$ mais son intégrale sur $[0,+\infty[$ converge.

b. Changement de variable

Dans ce paragraphe, on considère $-\infty \leq \alpha < \beta \leq +\infty$.

Proposition 18. *Changement de variable*

Soit $f \in C(]a,b[,\mathbb{K})$ et $\varphi :]\alpha,\beta[\to]a,b[$ une fonction de classe C^1, bijective et strictement monotone. Alors les intégrales $\displaystyle\int_a^b f(t)dt$ et $\displaystyle\int_{\alpha}^{\beta} (f\circ\varphi)(u).\varphi'(u)du$ sont de même nature. En cas de convergence, on a :

— si φ est croissante : $\displaystyle\int_a^b f(t)dt = \int_\alpha^\beta (f\circ\varphi)(u).\varphi'(u)du$;

— si φ est décroissante : $\displaystyle\int_a^b f(t)dt = \int_\beta^\alpha (f\circ\varphi)(u).\varphi'(u)du$.

Exercice 10.

Justifier la convergence de l'intégrale $\displaystyle\int_0^{+\infty} \frac{\ln(t)}{1+t^2}dt$ puis la calculer grâce au changement de variable $u = \frac{1}{t}$.

3. Intégration des relations de comparaison

Proposition 19.

Soit $f \in C([a,b[,\mathbb{K})$ et $g \in C([a,b[,\mathbb{R})$ une fonction **positive**. On suppose $f = O(g)$.

- Si $\displaystyle\int_a^b g$ converge, alors f est intégrable sur $[a,b[$ et on a :

$$\int_x^b f(t)dt = O\left(\int_x^b g(t)dt\right) \text{ quand } x \to b$$

- Si $\displaystyle\int_a^b g$ diverge, on a :

$$\int_a^x f(t)dt = O\left(\int_a^x g(t)dt\right) \text{ quand } x \to b$$

Ce résultat est toujours vrai lorsqu'on remplace les O par o.

Exercice 11.

1. Montrer que en $+\infty$, $\displaystyle\int_1^x \frac{e^{t^2}}{t^2}dt = o\left(\int_1^x e^{t^2}dt\right)$.
2. En déduire un équivalent simple de $\displaystyle\int_1^x e^{t^2}dt$ en $+\infty$ (utiliser une IPP).

Proposition 20.

Soit $f \in C([a, b[, \mathbb{K})$ et $g \in C([a, b[, \mathbb{R})$ une fonction **positive**. On suppose $f(x) \underset{x \to b}{\sim} g(x)$.

- Si $\displaystyle\int_a^b g$ converge, alors f est intégrable sur $[a, b[$ et on a :

$$\int_x^b f(t)dt \underset{x \to b}{\sim} \int_x^b g(t)dt.$$

- Si $\displaystyle\int_a^b g$ diverge, alors f n'est pas intégrable sur $[a, b[$ et on a :

$$\int_a^x f(t)dt \underset{x \to b}{\sim} \int_a^x g(t)dt.$$

Chapitre III

Topologie des espaces vectoriels normés

Table des matières

Partie A

Normes et espaces vectoriels normés

Dans cette partie, E désigne un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1. Définitions de base et premières propriétés

Définition 1. *Norme*

On appelle **norme** sur l'espace vectoriel, une application $N : E \to \mathbb{R}$ qui vérifie les axiomes suivants :

i) (Positivité) pour tout $x \in E$,

$$N(x) \geq 0;$$

ii) (Séparation) pour tout $x \in E$,

$$N(x) = 0 \text{ implique } x = 0_E;$$

iii) (Homogénéité) pour tous $x \in E$, $\lambda \in \mathbb{K}$,

$$N(\lambda x) = |\lambda| N(x);$$

iv) (Inégalité triangulaire) pour tous $x, y \in E$,

$$N(x + y) \leq N(x) + N(y).$$

Si N est une norme sur E, le couple (E, N) est appelé **espace vectoriel normé**.

Il est souvent d'usage de noter $\|\cdot\|$ d'un espace vectoriel normé et donc $\|x\|$ la norme d'un vecteur x de E.

Proposition 1.

Soit N une norme sur E. Alors $N(0_E) = 0$ et que pour tout $x \in E$, $N(-x) = N(x)$.

Démonstration.

D'après l'axiome iii) appliqué à $\lambda = 0$, on a :

$$N(0_E) = N(0.0_E) = 0N(0_E) = 0.$$

De plus, pour tout $x \in E$, $N(-x) = N((-1).x) = |-1|N(x) = N(x)$. □

Exercice 1.

Soit N une norme sur E. Montrer que l'axiome d'homogénéité est équivalent à :

$$\forall x \in E,\ \forall \lambda \in \mathbb{K}, \quad N(\lambda x) \leq |\lambda| N(x).$$

Correction.

La condition est clairement nécessaire. Montrons qu'elle est suffisante :
1er cas : $\lambda = 0$. Pour tout $x \in E$, on a $N(0.x) \leq 0N(x) = 0$. Or N est positive, donc $N(0.x) = 0 = 0N(x)$.
2eme cas : $\lambda \neq 0$. Pour tout $x \in E$,

$$N(x) = N(\frac{1}{\lambda}\lambda x) \leq \frac{1}{|\lambda|} N(\lambda x).$$

Par suite, $N(\lambda x) \geq |\lambda| N(x)$. D'où le résultat.

Exemple 1.

Soit E un espace vectoriel sur $\mathbb{R}$ et $< \cdot, \cdot >$ un produit scalaire sur E et $\| \cdot \|$ sa norme associée (i.e. $x \mapsto \|x\| = \sqrt{< x, x >}$). La terminologie utilisée est bien justifiée : en effet, $\| \cdot \|$ est bien une norme sur E au sens de la définition 1.

On rappelle qu'un produit scalaire sur E est une forme bilinéaire symétrique définie positive sur E. On utilise donc ces propriétés. Soit $x, y \in E$ et $\lambda \in \mathbb{R}$.

i) (Positivité) On a $< x, x > \geq 0$ par positivité du produit scalaire, donc $\| \cdot \|$ est bien définie et positive.

ii) (Séparation) . Si $\|x\| = 0_E$, alors, $\sqrt{< x, x >} = 0$. Ainsi, $< x, x >= 0$, d'où par définie positivité du produit scalaire, $x = 0_E$.

iii) (Homogénéité) On a, par linéarité du produit scalaire par rapport à chaque variable :

$$\|\lambda x\| = \sqrt{< \lambda x, \lambda x >} = \sqrt{\lambda^2 < x, x >} = \sqrt{\lambda^2}\sqrt{< x, x >} = |\lambda| \|x\|.$$

iv) (Inégalité triangulaire) On a, en utilisant la bilinéarité du produit scalaire et le théorème de Cauchy-Schwarz :

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

D'où le résultat.

Proposition 2. *Seconde Inégalité triangulaire*

Soit N une norme sur E. Pour tous $x, y \in E$ on a :

$$|N(x) - N(y)| \leq N(x - y) \leq N(x) + N(y).$$

Démonstration.

Soit $x, y \in E$. On a $N(x) = N(x-y+y) \leq N(x-y)+N(y)$. Par suite, $N(x)-N(y) \leq N(x-y)$. On a de même $N(y) - N(x) \leq N(x-y)$ en remarquant que $N(y) = N(y-x+x)$. □

Définition 2. *Vecteur unitaire*

Soit N une norme sur E.
On dit qu'un vecteur x de E est **unitaire** si $N(x) = 1$.

Notation 1. *Boule et sphère unité*

Soit N une norme sur E. On note :
- $S(0_E, 1) = \{x \in E \mid N(x) = 1\}$; on appelle cet ensemble **sphère unité** de E ;
- $B_f(0_E, 1) = \{x \in E \mid N(x) \leq 1\}$; on appelle cet ensemble **boule unité fermée** de E ;
- $B(0_E, 1) = \{x \in E \mid N(x) < 1\}$; on appelle cet ensemble **boule unité ouverte** de E ;

Définition 3.

Soit N une norme sur E et $x \in E$ un vecteur non nul.
On appelle **vecteur unitaire associé à** x le vecteur unitaire $\frac{1}{N(x)}x$.

Remarque 1.

Pour $x \neq 0_E$, le vecteur unitaire associé à x est clairement colinéaire à x.
Si $\mathbb{K} = \mathbb{R}$, il n'existe que deux vecteurs unitaires colinéaires à x :

$$\frac{1}{N(x)}x \text{ et } -\frac{1}{N(x)}x.$$

Dans le cas où $\mathbb{K} = \mathbb{C}$, il en a une infinité : ce sont exactement les

$$\frac{e^{it}}{N(x)}x \text{ pour } t \in [0, 2\pi[.$$

Exercice 2.

On considère l'espace euclidien $\mathbb{R}^2$ muni de sa norme canonique donnée par $\|(a,b)\| = \sqrt{a^2+b^2}$.
Dessiner la sphère unité et les boules unité fermée et ouverte de cet espace.

2. Premiers exemples d'espaces vectoriels normés

a. Quelques normes sur $\mathbb{K}^n$

Sur l'espace vectoriel $\mathbb{K}^n$ il existe de nombreuses normes qui permettent de le munir d'une structure d'espace vectoriel normé. En voici quelques unes parmi les plus fréquemment "rencontrées" :

Définition 4. *Normes usuelles sur $\mathbb{K}^n$*

Soit $x = (x_1, ..., x_n)$ un vecteur de $\mathbb{K}^n$. On définit :

— La **norme un**, notée $\|\cdot\|_1$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + ... + |x_n|.$$

— La **norme deux**, notée $\|\cdot\|_2$:

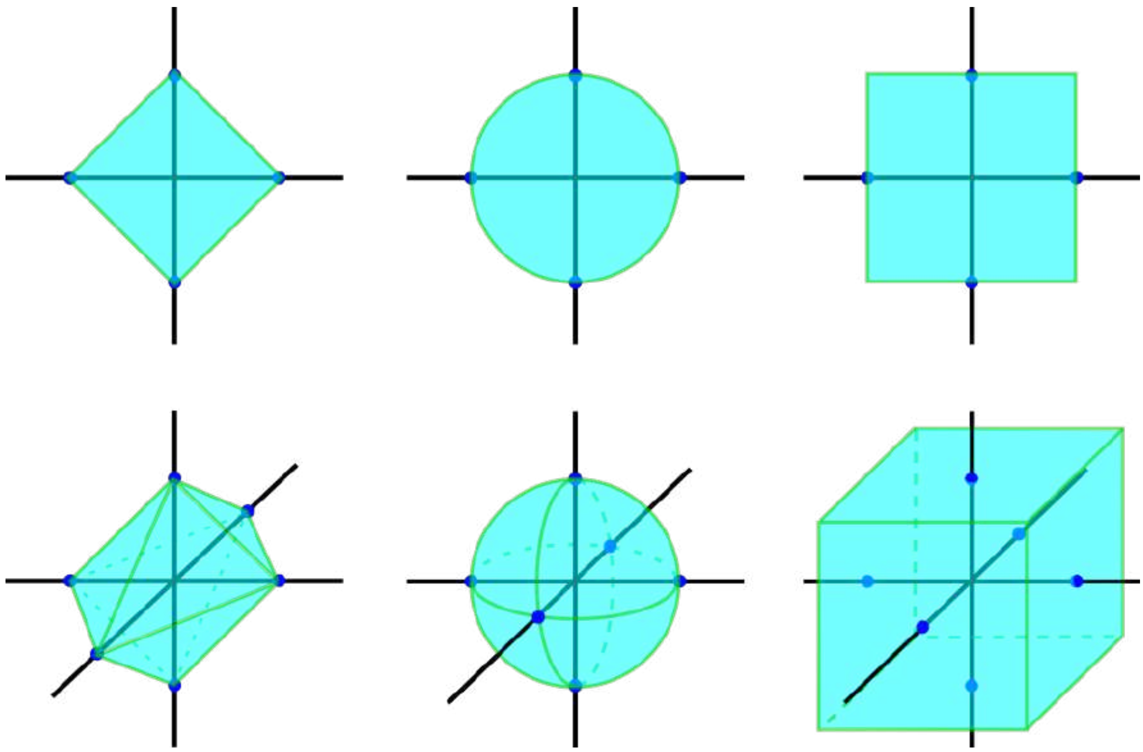
$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{|x_1|^2 + ... + |x_n|^2}.$$

— La **norme infinie**, notée $\|\cdot\|_\infty$:

$$\|x\|_\infty = \max_{1 \leq i \leq n} (|x_i|) = \max(|x_1|, ..., |x_n|).$$

On remarque immédiatement que dans le cas $\mathbb{K} = \mathbb{R}$, la norme deux correspond à la norme euclidienne sur $\mathbb{R}^n$.

Voici une illustration dans $\mathbb{R}^2$ puis $\mathbb{R}^3$ des boules des normes un, deux et infini (de gauche à droite).



Exercice 3.

Dessiner les boules des normes un, deux et infinie dans $\mathbb{R}^4$.

Correction.

includegraphics error, file not found.

Proposition 3.

Les normes un, deux, et infini sont bien des normes sur $\mathbb{R}^n$.

Démonstration.

Pour cette démonstration, nous aurons besoin de l'inégalité triangulaire de la fonction valeur absolue définie sur $\mathbb{R}$. Rappelons ici une démonstration de cette inégalité :

Inégalité triangulaire de la valeur absolue :

La fonction "valeur absolue" de $\mathbb{R}$ dans $\mathbb{R}_+$ est définie de la façon suivante :

$$|\cdot| : x \mapsto |x| = \begin{cases} x & \text{si } x \leq 0 \\ -x & \text{si } x < 0. \end{cases}$$

Montrons que, pour tout $x, y \in \mathbb{R}$, $|x+y| \leq |x| + |y|$.

Soit $x, y \in \mathbb{R}$. Quitte à échanger x et y, on peut supposer que $x \leq y$.

— 1er cas : x, y sont positifs. Alors $x + y \geq 0$ et ainsi, $|x| = x$, $|y| = y$ et $|x+y| = x+y$. Par suite, on a :
$$|x+y| = x+y = |x| + |y| \leq |x| + |y|$$
— 2eme cas : x, y sont négatifs. Alors $x + y \leq 0$ et ainsi, $|x| = -x$, $|y| = -y$ et $|x+y| = -(x+y) = -x-y$. Par suite, on a :
$$|x+y| = -x-y = |x| + |y| \leq |x| + |y|$$
— 3eme cas : x est négatif et y positif. Alors $|x| = -x$ et $|y| = y$. On distingue deux sous-cas :
  — Si $|x| \geq |y|$, alors $-x \geq y$ donc $x + y \leq 0$. Par suite, $|x+y| = -x-y$ et ainsi, comme $-|y| \leq 0 \leq |y|$:
  $$|x+y| = -x-y = |x| - |y| \leq |x| + |y|$$
  — Si $|x| < |y|$, alors $-x < y$ donc $x + y > 0$. Par suite, $|x+y| = x+y$ et ainsi, comme $-|x| \leq 0 \leq |x|$:
  $$|x+y| = x+y = -|x| + |y| \leq |x| + |y|$$

Dans tous les cas, on a $|x+y| \leq |x| + |y|$.

Il en résulte que, pour tout $x, y \in \mathbb{R}$, $|x+y| \leq |x| + |y|$.

Démontrons que les normes 1, 2 et infinie sont bien des normes sur $\mathbb{R}^n$. Ces fonctions sont bien définies : pour la norme 1, il s'agit d'une somme finie ; pour la norme 2, il s'agit de la racine d'une somme finie de nombres positifs (carrés) et donc positive ; pour la norme infinie, il s'agit du maximum d'un ensemble fini.

- $\|\cdot\|_1$: Soit $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$.
  i) (Positivité) pour tout $i \in [\![1, n]\!]$, $|x_i| \geq 0$, donc :
  $$\|x\|_1 = \sum_{i=1}^{n} |x_i| \geq 0.$$

ii) (Séparation) On suppose que $\|x\|_1 = 0$. Alors $\sum_{i=1}^n |x_i| = 0$. Or si une somme de nombres positifs est nulle, tous ses termes sont nuls (en effet, chaque terme d'une somme de nombres positifs est compris entre 0 et cette somme) ; ainsi, pour tout $i \in [\![1,n]\!]$, $|x_i| = 0$ et donc $x_i = 0$. Par suite, $x = (0, ..., 0) = 0_{\mathbb{R}^n}$.

iii) (Homogénéité) On a, en factorisant par $|\lambda|$:

$$\|\lambda x\|_1 = \sum_{i=1}^n \underbrace{|\lambda x_i|}_{=|\lambda|.|x_i|} = |\lambda| \sum_{i=1}^n |x_i| = |\lambda|.\|x\|_1.$$

iv) (Inégalité triangulaire) Par définition de l'addition dans $\mathbb{R}^n$, on a $x + y = (x_1 + y_1, ..., x_n + y_n)$ et ainsi, par inégalité triangulaire dans $\mathbb{R}$:

$$\|x+y\|_1 = \sum_{i=1}^n \underbrace{|x_i + y_i|}_{\leq |x_i|+|y_i|} \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = \|x\|_1 + \|y\|_1.$$

Il en résulte que $\|\cdot\|_1$ est une norme sur $\mathbb{R}^n$.

- $\|\cdot\|_2$: On rappelle que l'application $(\cdot|\cdot)$ de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$ définie par $(x|y) = \sum_{i=1}^n x_i y_i$ est un produit scalaire sur $\mathbb{R}^n$ (il s'agit du *produit scalaire canonique* de $\mathbb{R}^n$) et on remarque que $\|\cdot\|_2$ est la norme associée à ce produit scalaire. Ainsi, d'après l'exemple 1, $\|\cdot\|_2$ est un norme sur $\mathbb{R}^n$.

- $\|\cdot\|_\infty$: Soit $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$.

i) (Positivité) pour tout $i \in [\![1,n]\!]$, $|x_i| \geq 0$, donc en particulier :

$$\|x\|_\infty = \max_{1\leq i\leq n}(|x_i|) \geq 0.$$

ii) (Séparation) On suppose que $\|x\|_\infty = 0$. Alors $\max_{1\leq i\leq n}(|x_i|) = 0$. Ainsi, pour tout $j \in [\![1,n]\!]$, $0 \leq |x_j| \leq \max_{1\leq i\leq n}(|x_i|) = 0$, d'où $|x_j| = 0$ et donc $x_j = 0$. Par suite, $x = (0, ..., 0) = 0_{\mathbb{R}^n}$.

iii) (Homogénéité) On a $\|x\|_\infty = \max_{1\leq i\leq n}(|x_i|)$ qui est un maximum bien défini donc il existe $i_0 \in [\![1,n]\!]$ tel que $\|x\|_\infty = |x_{i_0}|$. Ainsi, on a, d'une part, comme $|\lambda|$ positif :

$$\|\lambda x\|_\infty = \max_{1\leq i\leq n} |\lambda x_i| \geq |\lambda x_{i_0}| = |\lambda|.|x_{i_0}| = |\lambda|\|x\|_\infty;$$

et, d'autre part, pour tout $i \in [\![1,n]\!]$, $|x_i| \leq |x_{i_0}|$ donc, toujours par positivité de $|\lambda|$:

$$|\lambda x_i| = |\lambda|.|x_i| \leq |\lambda|.|x_{i_0}| = |\lambda|\|x\|_\infty.$$

Donc, en particulier, $|\lambda|\|x\|_\infty$ est plus grand que le maximum des $|\lambda x_i|$ i.e. $|\lambda|\|x\|_\infty \geq \|\lambda x\|_\infty$. Par suite, $\|\lambda x\|_\infty = |\lambda|\|x\|_\infty$.

iv) (Inégalité triangulaire) Par définition de l'addition dans $\mathbb{R}^n$, on a $x + y = (x_1 + y_1, ..., x_n + y_n)$. Il existe $i_0 \in [\![1,n]\!]$ tel que $\|x+y\|_\infty = |x_{i_0} + y_{i_0}|$. Ainsi, par inégalité triangulaire dans $\mathbb{R}$:

$$\|x+y\|_\infty = |x_{i_0} + y_{i_0}| \leq |x_{i_0}| + |y_{i_0}| \leq \|x\|_\infty + \|y\|_\infty.$$

Il en résulte que $\|\cdot\|_\infty$ est une norme sur $\mathbb{R}^n$.

□

Remarque 2.

Soit $p \geq 1$ un réel. On peut généraliser l'idée des normes un et deux : on montrera en exercice que

$$x = (x_1, ..., x_n) \mapsto \|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}$$

est une norme sur $\mathbb{K}^n$. On l'appelle la **norme** p.

b. Exemples de normes sur des espaces de suites

On considère l'espace vectoriel $\mathbb{K}^{\mathbb{N}}$ des suites indexées par $\mathbb{N}$ et à valeurs dans $\mathbb{K}$. On définit les normes suivantes sur certains de ses sous-espaces :

Définition 5.

— On note $\ell^1(\mathbb{N}, \mathbb{K})$ l'ensemble des suites $u = (u_n)$ de $\mathbb{K}^{\mathbb{N}}$ telles que la série $\sum u_n$ est absolument convergente. On définit la **norme un**, notée $\|\cdot\|_1$ sur $\ell^1(\mathbb{N}, \mathbb{K})$ par :

$$\text{Pour } u = (u_n) \in \ell^1(\mathbb{N}, \mathbb{K}), \quad \|u\|_1 = \sum_{n=0}^{\infty} |u_n|.$$

— On note $\ell^2(\mathbb{N}, \mathbb{K})$ l'ensemble des suites $u = (u_n)$ de $\mathbb{K}^{\mathbb{N}}$ telles que la série $\sum u_n^2$ est absolument convergente. On définit la **norme deux**, notée $\|\cdot\|_2$ sur $\ell^2(\mathbb{N}, \mathbb{K})$ par :

$$\text{Pour } u = (u_n) \in \ell^2(\mathbb{N}, \mathbb{K}), \quad \|u\|_2 = \sqrt{\sum_{n=0}^{\infty} |u_n|^2}.$$

— On note $\ell^\infty(\mathbb{N}, \mathbb{K})$ l'ensemble des suites *bornées* de $\mathbb{K}^{\mathbb{N}}$. On définit la **norme infinie**, notée $\|\cdot\|_\infty$ sur $\ell^\infty(\mathbb{N}, \mathbb{K})$ par :

$$\text{Pour } u = (u_n) \in \ell^\infty(\mathbb{N}, \mathbb{K}), \quad \|u\|_\infty = \sup_{n \in \mathbb{N}}(|u_n|).$$

Exercice 4.

Montrer que $\ell^1(\mathbb{N}, \mathbb{K})$, $\ell^2(\mathbb{N}, \mathbb{K})$ et $\ell^\infty(\mathbb{N}, \mathbb{K})$ sont des sous-espaces vectoriels de $\mathbb{K}^{\mathbb{N}}$.
Que peut-on dire de l'ordre (au sens de l'inclusion) de ces espaces ?

Correction (exercice : rédiger cette démonstration correctement !).

Il est clair que la suite constante en 0 appartient à ces trois ensembles. Ensuite, on utilise le fait qu'une série dont le terme général est la somme des termes généraux de deux séries absolument convergentes est absolument convergente et que $(a+b)^2 \leq 2(a^2+b^2)$ pour montrer que $\ell^1(\mathbb{N}, \mathbb{K})$ et $\ell^2(\mathbb{N}, \mathbb{K})$ sont des sous-espaces vectoriels de $\mathbb{K}^{\mathbb{N}}$. Enfin, pour $\ell^\infty(\mathbb{N}, \mathbb{K})$, on utilise l'inégalité triangulaire et la sous-additivité du sup.

Soit $u = (u_n) \in \ell^1(\mathbb{N}, \mathbb{K})$. Alors u_n est le terme général d'une série absolument convergente donc

u converge vers 0 quand $n \to \infty$. Par suite, il existe $n_0 \in \mathbb{N}$ tel que, pour tout $n \geq n_0$, $|u_n| \leq 1$. On note $v = (v_n)$ avec $v_n = u_{n+n_0}$ pour $n \in \mathbb{N}$. Les séries de termes généraux u_n et v_n sont de même nature donc $\sum v_n$ est absolument convergente. De plus, comme pour tout $n \in \mathbb{N}$, $|v_n| \leq 1$, on a $v_n^2 \leq |v_n|$; par comparaison de séries à termes positifs, la série $\sum v_n^2$ est (absolument) convergente.
Or $\sum v_n^2$ et $\sum u_n^2$ sont de même nature donc u appartient à $\ell^2(\mathbb{N}, \mathbb{K})$. Il en résulte que $\ell^1(\mathbb{N}, \mathbb{K}) \subset \ell^2(\mathbb{N}, \mathbb{K})$.
Ensuite, si $u = (u_n)$ appartient à $\ell^2(\mathbb{N}, \mathbb{K})$ alors (u_n^2) converge vers 0 et donc (u_n) aussi. Par suite, (u_n) est bornée (exercice : montrer cette assertion) et donc u appartient $\ell^\infty(\mathbb{N}, \mathbb{K})$.

On obtient au final les inclusions suivantes :

$$\ell^1(\mathbb{N}, \mathbb{K}) \subset \ell^2(\mathbb{N}, \mathbb{K}) \subset \ell^\infty(\mathbb{N}, \mathbb{K}).$$

Proposition 4.

Les couples $(\ell^1(\mathbb{N}, \mathbb{K}), \|\cdot\|_1)$, $(\ell^2(\mathbb{N}, \mathbb{R}), \|\cdot\|_2)$, $(\ell^\infty(\mathbb{N}, \mathbb{K}), \|\cdot\|_\infty)$ sont des espaces vectoriels normés.

Démonstration.

N'oublions pas que la première chose à faire avant de se lancer dans l'étude des axiomes de norme pour une fonction, est de montrer que celle-ci est bien définie ! En effet, si on "essayait" $\|\cdot\|_1$ sur $\ell^\infty(\mathbb{N}, \mathbb{K})$, on pourrait vérifier que les axiomes de norme sont bien satisfaits mais que mais qu'en fait, la fonction est "très mal" définie : $\|\cdot\|_1$ n'est pas définie en la suite constante en 1 par exemple.

- $\|\cdot\|_1$: Montrons que la fonction $\|\cdot\|_1$ est bien définie sur $\ell^1 = \ell^1(\mathbb{N}, \mathbb{K})$. Soit $u = (u_n)_{n \in \mathbb{N}} \in \ell^1$. Alors $\sum u_n$ est une série numérique absolument convergente i.e. la série $\sum |u_n|$ converge donc la quantité $\|u\|_1 = \sum_{n=0}^{\infty} |u_n|$ est bien définie.
  Soit $u = (u_n)_{n \in \mathbb{N}}, v = (v_n)_{n \in \mathbb{N}} \in \ell^1$ et $\lambda \in \mathbb{R}$.
  i) (Positivité) La somme d'une série à termes positifs convergente est positive (car la limite d'une suite convergente à valeurs positives est positive ⇠ exercice) donc
  $$\|u\|_1 = \sum_{n=0}^{\infty} \underbrace{|u_n|}_{\geq 0} \geq 0.$$
  ii) (Séparation) On suppose que $\|u\|_1 = 0$. Alors $\sum_{n=0}^{+\infty} |u_n| = 0$. Comme $\sum |u_n|$ est une série à termes positifs, la suite $(S_n)_{n \in \mathbb{N}}$ de ses sommes partielles est croissante et ainsi, on a, pour tout $n \in \mathbb{N}$:
  $$0 \leq |u_n| \leq S_n \leq \lim_{k \to +\infty} S_k = \sum_{k=0}^{+\infty} |u_k| = 0$$
  Ainsi, pour tout $n \in \mathbb{N}$, $u_n = 0$. Par suite, $u = (0)_{n \in \mathbb{N}} = 0_{\ell^1}$.
  iii) (Homogénéité) On a, par linéarité de l'application $w = (w_n) \mapsto \sum_{n=0}^{+\infty} w_n$:
  $$\|\lambda u\|_1 = \sum_{n=0}^{+\infty} \underbrace{|\lambda u_n|}_{=|\lambda|.|u_n|} = |\lambda| \sum_{n=0}^{+\infty} |u_n| = |\lambda|.\|u\|_1.$$

iv) (Inégalité triangulaire) Par définition de l'addition dans $\mathbb{K}^{\mathbb{N}}$, on a $u+v=(u_n+v_n)_{n\in\mathbb{N}}$ et ainsi, par inégalité triangulaire dans $\mathbb{R}$ et par croissance et linéarité de l'application $w=(w_n)\mapsto\sum_{n=0}^{+\infty}w_n$:

$$\|u+v\|_1=\sum_{n=0}^{+\infty}\underbrace{|u_n+v_n|}_{\leq|u_n|+|v_n|}\leq\sum_{n=0}^{+\infty}|u_n|+\sum_{n=0}^{+\infty}|v_n|=\|u\|_1+\|v\|_1.$$

Il en résulte que $\|\cdot\|_1$ est une norme sur ℓ^1.

- $\|\cdot\|_2$: On note $\ell^2=\ell^2(\mathbb{N},\mathbb{R})$ et on considère l'application $(\cdot|\cdot)$ de $\ell^2\times\ell^2$ dans $\mathbb{R}$ définie par $(u|v)=\sum_{n=0}^{+\infty}u_nv_n$. Celle-ci est un produit scalaire sur ℓ^2 (nous allons le montrer ci-après) et on remarque que $\|\cdot\|_2$ est la norme associée à ce produit scalaire. Ainsi, d'après l'exemple 1, $\|\cdot\|_2$ est un norme sur ℓ^2.
Montrons que $(\cdot|\cdot)$ est un produit scalaire sur ℓ^2 i.e. une forme bilinéaire symétrique sur ℓ^2.
Encore une fois, il s'agit premièrement de montrer que $(\cdot|\cdot)$ est bien définie de $\ell^2\times\ell^2$ dans $\mathbb{R}$: pour $u,v\in\ell^2$, comme pour tout $(a,b)\in\mathbb{R}^2$, $2|ab|\leq a^2+b^2$, on a, pour tout $n\in\mathbb{N}$:

$$|u_nv_n|\leq\frac{u_n^2+v_n^2}{2}$$

Or, par hypothèses, u_n^2 et v_n^2 sont des termes généraux de séries convergentes d'où $\frac{u_n^2+v_n^2}{2}$ l'est aussi et ainsi, par comparaison $\sum|u_nv_n|$ converge.
Par suite, la série numérique $\sum u_nv_n$ est absolument convergente et donc convergente. Il en résulte que la quantité $(u|v)=\sum_{n=0}^{+\infty}u_nv_n$ est bien définie.
Soit $u=(u_n),v=(v_n),w=(w_n)\in\ell^2$ et $\lambda\in\mathbb{R}$.

i) (Symétrie) Par commutativité de la multiplication dans $\mathbb{R}$, on a :

$$(v|u)=\sum_{n=0}^{+\infty}v_nu_n=\sum_{n=0}^{+\infty}u_nv_n=(u|v).$$

ii) (Bilinéarité) Comme $(\cdot|\cdot)$ est symétrique, il suffit de montrer la linéarité par rapport à la première variable. Par linéarité de l'application $(s_n)\mapsto\sum_{n=0}^{+\infty}s_n$, on a :

$$(\lambda u+v|w)=\sum_{n=0}^{+\infty}\underbrace{(\lambda u_n+v_n)w_n}_{=\lambda u_nw_n+v_nw_n}=\lambda\sum_{n=0}^{+\infty}u_nw_n+\sum_{n=0}^{+\infty}v_nw_n=\lambda(u|w)+(v|w).$$

iii) (Définie positivité) On a, d'une part

$$(u|u)=\sum_{n=0}^{+\infty}\underbrace{u_n^2}_{\geq 0}\geq 0$$

et d'autre part, si $(u|u)=0$, alors, pour tout $k\in\mathbb{N}$:

$$0\leq u_k^2\leq\sum_{n=0}^{+\infty}u_n^2=(u|u)=0,$$

donc $u_k=0$. Par suite, $u=(0)_{n\in\mathbb{N}}=0_{\ell^2}$.

Ceci montre que $(\cdot|\cdot)$ est bien un produit scalaire sur ℓ^2.

- $\|\cdot\|_\infty$: Montrons que la fonction $\|\cdot\|_\infty$ est bien définie sur $\ell^\infty = \ell^\infty(\mathbb{N},\mathbb{K})$. Soit $u = (u_n)_{n\in\mathbb{N}} \in \ell^\infty$. Comme u est bornée, l'ensemble $U = \{|u_n| \mid n \in \mathbb{N}\}$ est une partie non vide et majorée de $\mathbb{R}$ (en effet, $|u_0| \in U$ et on utilise le caractère borné de u pour montrer que U est majoré) donc U possède une borne supérieure. Ainsi, la quantité $\|u\|_\infty = \sup U$ est bien définie.

  Soit $u = (u_n)_{n\in\mathbb{N}}, v = (v_n)_{n\in\mathbb{N}} \in \ell^\infty$ et $\lambda \in \mathbb{R}$. On note $U = \{|u_n| \mid n \in \mathbb{N}\}$, $V = \{|v_n| \mid n \in \mathbb{N}\}$, $W = \{|\lambda u_n| \mid n \in \mathbb{N}\}$ et $X = \{|u_n + v_n| \mid n \in \mathbb{N}\}$.

  i) (Positivité) pour tout $n \in \mathbb{N}$, $|u_n| \geq 0$, donc en particulier, $\sup U$ étant un majorant de U :
  $$\|u\|_\infty = \sup U \geq |u_0| \geq 0.$$
  ii) (Séparation) On suppose que $\|u\|_\infty = 0$. Alors $\sup U = 0$. Ainsi, pour tout $n \in \mathbb{N}$, $\sup U$ étant un majorant de U :
  $$0 \leq |u_n| \leq \sup U = 0.$$
  Par suite, $u = (0)_{n\in\mathbb{N}} = 0_{\ell^\infty}$.

  iii) (Homogénéité) Comme $|\lambda|$ est positif, $\sup U$ étant le plus petit des majorants de U, $|\lambda| \sup U$ est le plus petit des majorants de $\{|\lambda|.|u_n| \mid n \in \mathbb{N}\}$. Or, pour tout $n \in \mathbb{N}$, $|\lambda|.|u_n| = |\lambda u_n|$ donc $\{|\lambda|.|u_n| \mid n \in \mathbb{N}\} = W$. Par suite :
  $$\|\lambda u\|_\infty = \sup W = |\lambda| \sup U = |\lambda|.\|u\|_\infty$$
  iv) (Inégalité triangulaire) Par définition de l'addition dans $\mathbb{K}^\mathbb{N}$, on a $u+v = (u_n+v_n)_{n\in\mathbb{N}}$. Pour tout $n \in \mathbb{N}$, par inégalité triangulaire dans $\mathbb{R}$:
  $$|u_n + v_n| \leq |u_n| + |v_n| \leq \|u\|_\infty + \|v\|_\infty.$$
  Ainsi, $\|u\|_\infty + \|v\|_\infty$ est un majorant de X ; or $\|u+v\|_\infty = \sup X$ étant le plus petit des majorants de X, on obtient :
  $$\|u+v\|_\infty \leq \|u\|_\infty + \|v\|_\infty.$$
  Il en résulte que $\|\cdot\|_\infty$ est une norme sur ℓ^∞.

□

c. Exemples de normes sur des espaces de fonctions

Soit X un ensemble non vide. On considère l'ensemble $\mathcal{F}_b(X,\mathbb{K})$ (egalement noté $\ell^\infty(X,\mathbb{K})$ des fonctions bornées de X à valeurs dans $\mathbb{K}$. C'est un espace vectoriel sur $\mathbb{K}$: $\mathcal{F}_b(X,\mathbb{K})$ est un sous-espace vectoriel de $\mathcal{F}(X,\mathbb{K})$, l'ensemble de toutes les fonctions de X dans $\mathbb{K}$.

Voici la norme naturelle sur $\mathcal{F}_b(X,\mathbb{K})$:

Définition 6. *Norme sur l'espace des fonctions bornées*

Soit $f \in \mathcal{F}_b(X, \mathbb{K})$. On appelle **norme infinie** de f, et on note $\|\cdot\|_\infty$ la quantité :

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

Proposition 5.

La norme infinie de $\mathcal{F}_b(X, \mathbb{K})$ est bien définie et c'est une norme sur $\mathcal{F}_b(X, \mathbb{K})$.

Démonstration.

L'ensemble X est non vide et f est bornée, donc l'ensemble $F = \{|f(x)| \mid x \in X\}$ est une partie non vide et majorée de $\mathbb{R}$. Elle possède donc une borne supérieure. Par suite, $\|f\|_\infty$ est bien définie et égale à cette borne supérieure.

Soit $f, g \in \mathcal{F}_b(X, \mathbb{K})$ et $\lambda \in \mathbb{K}$.

— (Positivité) Comme X est non vide, il existe $x_0 \in X$. Alors on a $\|f\|_\infty \geq |f(x_0)| \geq 0$.

— (Séparation) : Si $\|f\|_\infty = 0$, alors, pour tout $x \in X$, $0 \leq |f(x)| \leq \|f\|_\infty = 0$ donc f est la fonction nulle sur X.

— (Homogénéité) : On considère l'ensemble $G = \{|\lambda f(x)| \mid x \in X\}$. Alors $\|\lambda f\|_\infty = \sup G$. Montrons l'égalité demandée en utilisant la caractérisation séquentielle de la borne supérieure i.e. $a = \sup A$ si, et seulement si, a est un majorant de A et il existe une suite (a_n) à valeurs dans A telle que $a_n \xrightarrow[n \to +\infty]{} a$.

Pour tout $u \in F$, il existe $x \in X$ tel que $u = |\lambda f(x)| = |\lambda||f(x)|$; et ainsi,

$$u = |\lambda|.|f(x)| \leq |\lambda|.\|f\|_\infty.$$

Par suite, $|\lambda|.\|f\|_\infty$ est un majorant de G.
De plus, comme $\|f\|_\infty = \sup F$, d'après la caractérisation séquentielle de la borne supérieure de F, il existe un suite (u_n) à valeurs dans F (avec $u_n = |f(x_n)|$ où $x_n \in X$) telle que $u_n \xrightarrow[n \to +\infty]{} \|f\|_\infty$.
Ainsi, pour tout $n \in \mathbb{N}$, $|\lambda|u_n = |\lambda|.|f(x_n)| = |\lambda f(x_n)| \in G$ et $|\lambda|u_n \xrightarrow[n \to +\infty]{} |\lambda|.\|f\|_\infty$.
Il en résulte, par caractérisation séquentielle de la borne supérieure de G que :

$$|\lambda|.\|f\|_\infty = \sup G = \|\lambda f\|_\infty.$$

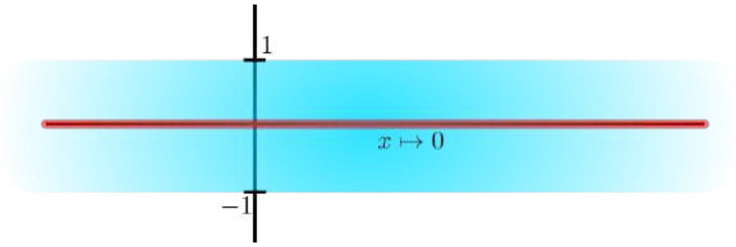
— (Inégalité triangulaire) : On a, pour tout $x \in X$, par inégalité triangulaire dans $\mathbb{R}$,

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty.$$

Ainsi, $\|f\|_\infty + \|g\|_\infty$ est un majorant de l'ensemble $\{|f(x) + g(x)| \mid x \in X\}$ dont la borne supérieure est $\|f + g\|_\infty$ qui est le plus petit de ses majorants. Par suite :

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

$\square$



Le graphe d'une fonction qui appartient à la boule unité ouverte de l'espace vectoriel normé $(\mathcal{F}_b(\mathbb{R},\mathbb{R}), \|\cdot\|_\infty)$ se trouve dans la bande bleue de la figure ci-dessus.

Exercice 5.

- Formaliser l'affirmation précédente et la démontrer.
- Montrer que les fonctions sin et cos appartiennent à la sphère unité de $\mathcal{F}_b(\mathbb{R},\mathbb{R})$.

Soit $a, b \in \mathbb{R}$ avec $a < b$. On considère l'espace vectoriel $C([a,b],\mathbb{K})$ des fonctions continues de $[a,b]$ dans $\mathbb{K}$. On définit trois normes usuelles sur cet espace par analogie du cas de $\mathbb{K}^n$:

Définition 7. *Normes sur $C([a,b],\mathbb{K})$*

Soit $f \in C([a,b],\mathbb{K})$. On définit les **normes de la convergence** : **en moyenne**, **en moyenne quadratique** et **uniforme** respectivement notées $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$, par :

$$\|f\|_1 = \int_a^b |f(t)|dt, \quad \|f\|_2 = \sqrt{\int_a^b |f(t)|^2 dt}$$

$$\text{et } \|f\|_\infty = \sup_{t\in[a,b]} |f(t)|.$$

Proposition 6.

Les normes de la convergence en moyenne, en moyenne quadratique et uniforme sont bien des normes sur $C([a,b],\mathbb{K})$.

Démonstration.

- $\|\cdot\|_1$. Soit $f, g \in E = C([a,b],\mathbb{K})$ et $\lambda \in \mathbb{K}$.
  La norme de f est bien définie car $|f|$ est continue sur le **segment** $[a,b]$ et donc intégrable sur $[a,b]$.
  i) **Positivité**. L'intégrale d'une fonction positive est positive donc $\|f\|_1 = \int_a^b |f(t)|dt \geq 0$.
  ii) **Séparation**. Si $\|f\|_1 = 0$, alors $\int_a^b |f(t)|dt = 0$, or une fonction **continue**, **positive**, d'intégrale nulle est nulle, d'où $f = 0_E$.

iii) **Homogénéité**. Par linéarité de l'intégrale, on a :

$$\|\lambda f\|_1 = \int_a^b \underbrace{|\lambda f(t)|}_{=|\lambda|.|f(t)|} \, dt = |\lambda| \int_a^b |f(t)|dt = |\lambda|.\|f\|_1.$$

iv) **Inégalité triangulaire**. Par linéarité de l'intégrale, on a :

$$\|f+g\|_1 = \int_a^b \underbrace{|f(t)+g(t)|}_{\leq |f(t)|+|g(t)|} dt \leq \int_a^b |f(t)|dt + \int_a^b |g(t)|dt = \|f\|_1 + \|g\|_1.$$

Il en résulte que $\|\cdot\|_1$ est une norme sur E.

— $\|\cdot\|_2$: il suffit de remarquer que la norme de la convergence en moyenne quadratique est la norme associée au produit scalaire

$$(f,g) \mapsto < f,g >= \int_a^b f(t)g(t)dt$$

sur $C([a,b],\mathbb{K})$.

— $\|\cdot\|_\infty$: il suffit de reprendre la démonstration de la norme infinie sur $\mathcal{F}_b(X,\mathbb{K})$.

□

Remarque 3.

On verra dans la suite du chapitre, que la norme infinie sur $C([a,b],\mathbb{K})$ peut-être obtenue comme norme *induite* par la norme infinie sur $\mathcal{F}_b([a,b],\mathbb{K})$ via l'inclusion $C([a,b],\mathbb{K}) \subset \mathcal{F}_b([a,b],\mathbb{K})$.

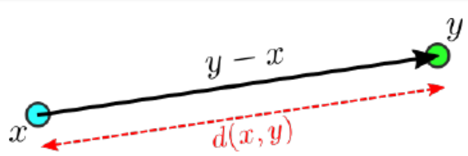
3. Distance associée à une norme

La norme d'un espace vectoriel normé E est une manière d'associer une *longueur* à chaque vecteur de E. Du point de vue *affine* i.e. si on se place dans un espace affine dirigé par E, une manière de mesurer la distance entre deux points M et N devient alors clair : il suffit de regarder la longueur - la norme - du vecteur $\overrightarrow{MN}$ qui "sépare" ces deux points.

Définition 8. *Distance associée à une norme*

Soit N une norme sur E. On appelle **distance associée à** N l'application $d : E \times E \to \mathbb{R}_+$ définie par :

$$\text{pour } (x,y) \in E^2, \quad d(x,y) = N(y-x).$$



Remarque 4.

Soit N une norme sur E. Alors la distance d associée à N vérifie les propriétés suivantes, pour tous $x, y, z \in E$:

i) (Symétrie) $d(x, y) = d(y, x)$

ii) (Séparation) $d(x, y) = 0 \Rightarrow x = y$.

iii) (Inégalité triangulaire) $d(x, z) \leq d(x, y) + d(y, z)$.

Pour un ensemble quelconque X, une application $d : X \times X \to \mathbb{R}_+$ qui vérifie ces trois axiomes est appelé une *distance* sur X.

Nous ne nous attarderons pas sur ce concept mais dans la suite, nous utiliserons librement ces propriétés dans l'étude des distances associées aux normes. En plus de ces propriétés, nous utiliserons également la suivante dans la suite :

— (Invariance par translation) $d(x + z, y + z) = d(x, y)$.

Exercice 6.

Soit N une norme sur E. Démontrer que la distance associée à N vérifie les quatre axiomes de la remarque précédente.

Proposition 7. *Seconde inégalité triangulaire*

Soit N une norme sur E. Alors la distance d associée à N vérifie, pour tous $x, y, z \in E$:

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

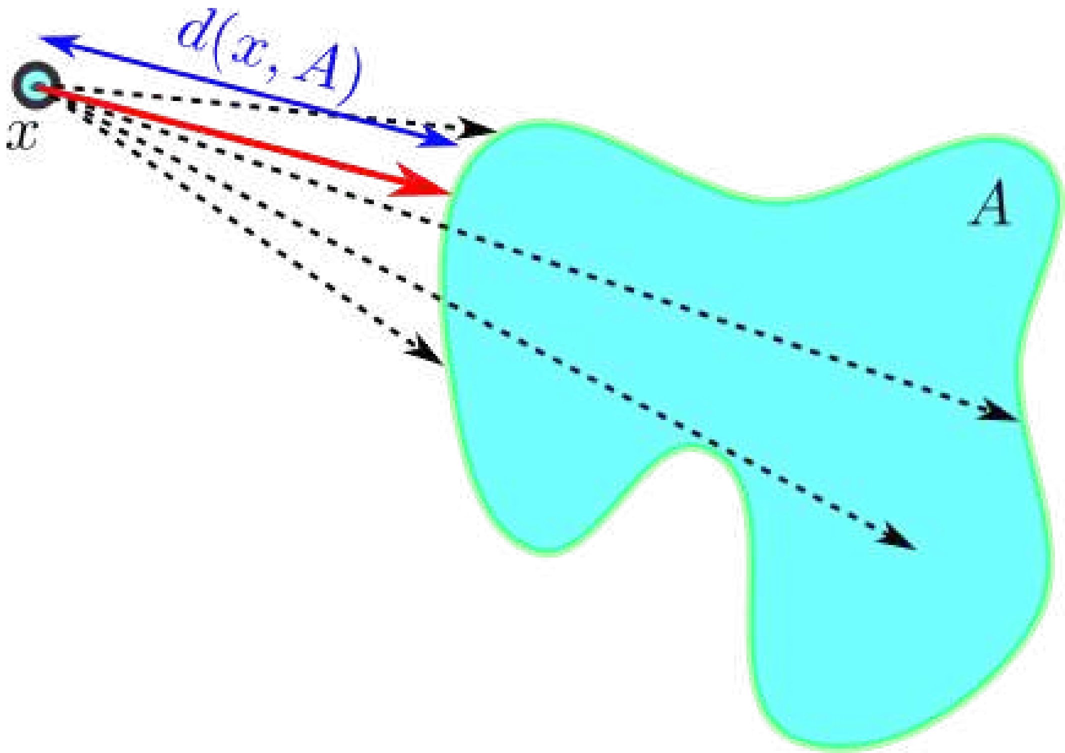
Démonstration.

Il suffit d'appliquer la définition de distance associée à une norme et d'utiliser la seconde inégalité triangulaire pour les normes. (On peut aussi vérifier que pour la notion de distance "générale" au sens de la remarque précédente, cette inégalité est encore valable.) $\square$

Définition 9. *Distance à une partie*

Soit N une norme sur E, d sa distance associée, A une partie non vide de E et $x \in E$. On appelle **distance de x à A** et on note $d(x, A)$, la quantité :

$$d(x, A) = \inf\{d(x, a) \mid a \in A\}.$$



On peut également définir la distance entre deux parties de E :

Définition 10. *Distance entre deux parties*

Soit N une norme sur E, d sa distance associée et A, B des parties non vides de E. On appelle **distance de** A **à** B et on note $d(A, B)$, la quantité :

$$d(A, B) = \inf\{d(a, b) \mid (a, b) \in A \times B\}.$$

Exercice 7.

Soit d la distance associée à une norme sur E et A une partie non vide de E. Montrer que, pour tous $x, y \in E$, on a :

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

Correction.

Soit $a \in A$. On a, par définition de $d(x, A)$ et par inégalité triangulaire :

$$d(x, A) \leq d(x, a) \leq d(x, y) + d(y, a).$$

Par suite, $d(x, A) - d(x, y) \leq d(y, a)$. Cette dernière inégalité est vraie pour tout $a \in A$.

Exercice 8.

Dans l'espace vectoriel $\mathbb{R}^2$, calculer la distance entre $(-1,1)$ et la droite D d'équation $y = 2x$ pour les normes un, deux et infini de $\mathbb{R}^2$. Représenter graphiquement chacune des situations.

Correction.

On note d_1 la distance associée à $\|\cdot\|_1$, d_2 la distance associée à $\|\cdot\|_2$, d_∞ la distance associée à $\|\cdot\|_\infty$ et $u = (-1,1)$.
La droite D est le sous-ensemble $D = \{(a,2a) \mid a \in \mathbb{R}\}$. Pour $(a,2a) \in D$ avec $a \in \mathbb{R}$, on a :

$$d_i(u,(a,2a)) = \|(a+1,2a-1)\|_i \text{ pour } i = 1,2 \text{ ou } \infty.$$

1. On a :

$$d_1(u,(a,2a)) = |a+1| + |2a-1| = \begin{cases} -(a+1)-(2a-1) = -3a & \text{pour } a \leq -1 \\ (a+1)-(2a-1) = -a+2 & \text{pour } -1 \leq a \leq \frac{1}{2} \\ (a+1)+(2a-1) = 3a & \text{pour } a \geq \frac{1}{2}. \end{cases}$$

On étudie $a \mapsto d_1(u,(a,2a))$ sur chacun des trois intervalles ci-dessus pour obtenir le minimum de cette fonction sur $\mathbb{R}$ qui est égal à $\frac{3}{2}$ (atteint pour $a = \frac{1}{2}$). Par suite,

$$d_1(u,D) = \inf_{(a,2a)\in D} d_1(u,(a,2a)) = \inf_{a\in\mathbb{R}} d_1(u,(a,2a)) = \frac{3}{2} \left(= d_1(u,(\frac{1}{2},1))\right).$$

2. On a, pour $a \in \mathbb{R}$:

$$d_2(u,(a,2a)) = \sqrt{(a+1)^2+(2a-1)^2} = \sqrt{5a^2-2a+2},$$

or $5a^2-2a+2$ est de discriminant strictement négatif, donc $5a^2-2a+2 > 0$ (car $5 > 0$) et le minimum de $a \mapsto 5a^2-2a+2$ est atteint pour $a = \frac{2}{2\times 5} = \frac{1}{5}$ et vaut $\frac{9}{5}$. De plus, la fonction $\sqrt{\ }$ est croissante, donc le minimum de la fonction $a \mapsto \sqrt{5a^2-2a+2}$ est égal à $\sqrt{\frac{9}{5}} = \frac{3}{\sqrt{5}} = \frac{3\sqrt{5}}{5}$. Par suite,

$$d_2(u,D) = \inf_{a\in\mathbb{R}} d_2(u,(a,2a)) = \frac{3\sqrt{5}}{5} \left(= d_1(u,(\frac{1}{5},\frac{2}{5}))\right).$$

Remarque : Comme la norme 2 est la norme associée au produit scalaire canonique sur $\mathbb{R}^2$ et D est un sous-espace vectoriel de $\mathbb{R}^2$, la distance de u à D aurait pu être calculée en déterminant la distance de u avec son projeté orthogonal sur D (qui est bien $(\frac{1}{5},\frac{2}{5})$).

∞. On a, pour $a \in \mathbb{R}$, $d_1(u,(a,2a)) = \max(|a+1|,|2a-1|)$, et on a :

$$|2a-1| - |a+1| = \begin{cases} -(2a-1)+(a+1) = -a+2 & \text{pour } a \leq -1 \\ -(2a-1)-(a+1) = -3a & \text{pour } -1 \leq a \leq \frac{1}{2} \\ (2a-1)-(a+1) = a-2 & \text{pour } a \geq \frac{1}{2} \end{cases}$$

donc $|2a-1| \leq |a+1|$ si, et seulement si, $a \in [0,2]$. Par suite,

$$d_\infty(u,(a,2a)) = \begin{cases} -2a+1 & \text{pour } a \leq 0 \\ a+1 & \text{pour } 0 \leq a \leq 2 \\ 2a-1 & \text{pour } a \geq 2. \end{cases}$$

On étudie $a \mapsto d_\infty(u,(a,2a))$ sur chacun des trois intervalles ci-dessus pour obtenir le minimum de cette fonction sur $\mathbb{R}$ qui est égal à 1 (atteint pour $a=0$). Par suite,

$$d_\infty(u,D) = \inf_{a\in\mathbb{R}} d_\infty(u,(a,2a)) = 1\,(= d_\infty(u,(0,0)))\,.$$

4. Boules et sphères associées à une distance

On généralise dans ce paragraphe la notion de cercle/sphère, disque/boule bien connue dans les espaces euclidiens de dimension 2 et 3 qui nous apparaissent naturels. Comme pour la sphère et la boule unité d'une norme, la forme des sphères et boules que nous allons considérer peut être tout-à-fait contre-intuitive vis-a-vis de la terminologie et de nos habitudes.

Notation 2. *Boule et sphère unité*

Soit N une norme sur E, d sa norme associée, $x_0 \in E$ et r un réel strictement positif. On note :

— $S(x_0,r) = \{x\in E \mid d(x,x_0) = r\}$; on appelle cet ensemble **sphère** de centre x_0 et de rayon r ;

— $B_f(x_0,r) = \{x\in E \mid d(x,x_0) \leq r\}$; on appelle cet ensemble **boule fermée** de centre x_0 et de rayon r ;

— $B(x_0,r) = \{x\in E \mid d(x,x_0) < r\}$; on appelle cet ensemble **boule ouverte** de centre x_0 et de rayon r ;

Remarque 5.

On peut remarquer que $S(x_0,r) = B_f(x_0,r) \setminus B(x_0,r)$.

Proposition 8.

Soit N une norme sur E. Toute boule (ouverte ou fermée) est une partie convexe de E.

Démonstration.

Soit $x_0\in E$ et $r>0$. Soit $x,y\in B(x_0,r)$ et $t\in[0,1]$. On pose $z = tx+(1-t)y$.
Alors on a :

$$\begin{array}{rcll} d(x_0,z) & = & N(z-x_0) & \\ & = & N(t(x-x_0)+(1-t)(y-x_0)) & \\ & \leq & N(t(x-x_0)) + N((1-t)((y-x_0)) & \text{(inégalité triangulaire)} \\ & \leq & tN(x-x_0)+(1-t)N(y-x_0) & \text{(homogénéité)} \\ & < & tr+(1-t)r = r & (x,y\in B(x_0,r)) \end{array}$$

Donc z appartient à $B(x_0,r)$. Par suite, $B(x_0,r)$ est convexe.

Il en est de même pour $B_f(x_0,r)$: il suffit de remplacer l'inégalité stricte par une inégalité large dans la dernière ligne du raisonnement précédent. □

5. Parties bornées, applications bornées

a. Parties bornées

Définition 11. *Partie bornée*

Soit N une norme sur E et A une partie de E. On dit que A est **bornée** s'il existe un réel positif R tel que :

$$\text{pour tout } x \in A, \quad N(x) \leq R.$$

Proposition 9.

Soit N une norme sur E, d sa distance associée et A une partie de E. Les propositions suivantes sont équivalentes :

i) A est bornée ;

ii) il existe $R \geq 0$ tel que $A \subset B_f(0_E, R)$;

iii) il existe $x_0 \in E$ et $R \geq 0$ tels que $A \subset B_f(x_0, R)$;

iv) il existe $R \geq 0$ tel que pour tous $x, y \in A$, $d(x, y) \leq R$.

Correction.

Schéma de la preuve : i) $\Rightarrow$ ii) $\Rightarrow$ iii) $\Rightarrow$ iv) $\Rightarrow$ i)

- i) $\Rightarrow$ ii) : On suppose A bornée. Alors il existe $R \geq 0$ tel que pour tout $x \in A$, $N(x) \leq R$, d'où $d(x, 0_E) \leq R$ et donc $x \in B_f(0_E, R)$. Par suite, $A \subset B_f(0_E, R)$.
- ii) $\Rightarrow$ iii) : Évident.
- iii) $\Rightarrow$ iv) : On suppose iii). Soit $x, y \in A$. Alors x, y appartiennent à $B_f(x_0, R)$ et on a :

$$d(x, y) \leq d(x, x_0) + d(x_0, x) \leq 2R.$$

- iv) $\Rightarrow$ i) : On suppose iv). On fixe x_0 dans A.
  Soit $x \in A$. Alors

$$N(x) = d(x, 0_E) \leq d(x, x_0) + d(x_0, 0_E) = R + N(x_0).$$

  Par suite, A est bornée.

Remarque 6.

Attention, la notion de partie bornée dépend bien-sûr de la norme utilisée ! Une partie bornée pour une certaine norme peut ne pas l'être pour une autre. Nous en verrons un exemple dans la partie consacrée à la comparaison des normes.

Définition 12. *Diamètre d'une partie*

Soit N une norme sur E et A une partie non vide et bornée de E. On appelle **diamètre** de A et on note $\operatorname{diam}(A)$, le réel positif :

$$\operatorname{diam}(A) = \sup_{(x,y)\in A^2} d(x,y).$$

Exercice 9.

Soit N une norme sur E et A une partie et bornée non vide de E.

1. Justifier l'existence du réel $\operatorname{diam}(A)$.
2. On suppose que E n'est pas réduit à $\{0_E\}$. Soit $r > 0$. Déterminer le diamètre d'une boule fermée de rayon r. De même pour une boule ouverte de rayon r.

Correction.

1. L'ensemble $\{d(x,y) \mid (x,y) \in A^2\}$ est une partie non vide - car A non vide - et majorée de $\mathbb{R}$ - d'après la proposition 9 iv) - donc elle possède une borne supérieure.
2. Il est clair qu'une boule (ouverte ou fermée) de rayon r à un diamètre inférieur ou égal à $2r$: il suffit d'appliquer l'inégalité triangulaire en utilisant le centre de la boule. Montrons que c'est exactement $2r$.

   Commençons par le cas d'une boule fermée $B_f = B_f(x_0, r)$ (qui nous donnera des idées pour le cas ouvert) :
   Soit u un vecteur unitaire de E (il en existe car $E \neq \{0_E\}$). Alors les éléments $x_0 \pm ru$ appartiennent à B_f, en effet :

   $$d(x_0, x_0 \pm ru) = d(0_E, \pm ru) = N(\pm ru) = rN(u) = r.$$

   Or on a :
   $$d(x_0 - ru, x_0 + ru) = N(ru - (-ru)) = N(2ru) = 2r.$$

   Donc $\operatorname{diam}(B_f) = 2r$.
   Passons maintenant au cas d'une boule ouverte $B = B(x_0, r)$. Le raisonnement précédent ne convient pas car $x_0 \pm ru$ ne sont pas dans B... Mais on peut s'en rapprocher autant que l'on veut !
   Soit u un vecteur unitaire de E et pour $n \in \mathbb{N}^*$, on considère les éléments $x_n^{\pm} = x_0 \pm r(1-\frac{1}{n})u$.
   Pour tout $n \in \mathbb{N}^*$, $x_n^{\pm} \in B$, en effet :

   $$d(x_0, x_n^{\pm}) = d(0_E, \pm r(1-\frac{1}{n})u) = N(\pm r(1-\frac{1}{n})u) = r(1-\frac{1}{n}) < r.$$

   De plus, on a :

   $$d(x_n^-, x_n^+) = d(x_0 - r(1-\frac{1}{n})u, x_0 + r(1-\frac{1}{n})u) = N(2r(1-\frac{1}{n})u) = 2r(1-\frac{1}{n}).$$

   Par suite, $d(x_n^-, x_n^+) \xrightarrow[n\to\infty]{} 2r$ et donc $\operatorname{diam}(B) = 2r$.

Remarque 7.

Si une partie A d'un espace vectoriel normé E n'est pas bornée, alors, par convention, on dira que $\text{diam}(A) = +\infty$.

b. Applications et suites bornées

Définition 13. *Application bornée*

Soit N une norme sur E et X un ensemble non vide. Une fonction $f : X \to E$ est dite **bornée** si $f(X)$ est une partie bornée de E i.e. s'il existe $R \in \mathbb{R}_+$ tel que, pour tout $x \in X$:

$$N(f(x)) \leq R.$$

Remarque 8. *Suite bornée*

Une suite $(u_n)_{n\in\mathbb{N}}$ à valeurs dans E peut être vue comme l'application $f : \mathbb{N} \to E$ telle que $f(n) = u_n$ pour $n \in \mathbb{N}$. La définition précédente s'applique donc également aux suites i.e. une suite $(u_n)_{n\in\mathbb{N}}$ est bornée s'il existe $R \in \mathbb{R}_+$ tel que, pour tout $n \in \mathbb{N}$:

$$N(u_n) \leq R.$$

On peut généraliser la structure de l'espace vectoriel normé $\mathcal{F}_b(X, \mathbb{K})$ vue dans la paragraphe 2 au cas des fonctions à valeurs dans E bornées :

Définition-Proposition 14.

Soit N une norme sur E et X un ensemble non vide. L'ensemble $\mathcal{F}_b(X, E)$ est un espace vectoriel sur $\mathbb{K}$ et l'application $\|\cdot\|_\infty$ appelée **norme infinie** définie par :

$$\text{pour } f \in \mathcal{F}_b(X, E), \quad \|f\|_\infty = \sup_{x\in X} N(f(x)),$$

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

Démonstration.

Il s'agit de la même démonstration que pour la norme infinie de $\mathcal{F}_b(X, \mathbb{K})$, en remplaçant le module/valeur absolue par la norme N. □

6. Constructions d'espaces vectoriels normés

a. Opérations sur les normes

Proposition 10.

Soit N, N' deux normes sur E et $\alpha \in \mathbb{R}_+^*$. Alors :

— $N + N' : x \mapsto N(x) + N'(x)$ est une norme sur E.

— $\alpha N : x \mapsto \alpha N(x)$ est une norme sur E.

Exercice 10.

Soit $n \in \mathbb{N}^*$, $N_1, ..., N_n$ des normes sur E et $\lambda_1, ..., \lambda_n \in \mathbb{R}_+^*$.

1. Montrer que
$$N = \sum_{i=1}^{n} \lambda_i N_i.$$
est une norme sur E.

2. On suppose de plus que $\sum_{i=1}^{n} \lambda_i = 1$. Que peut-on dire de l'intersection des boules unités (fermées ou ouvertes) des normes $N_1, ..., N_n$ par rapport à la boule unité (fermée ou ouverte) de N.

Correction.

1. On raisonne par récurrence et on utilise la proposition précédente.

2. On note, pour $i = 1, ..., n$, B_i la boule unité fermée de N_i et B la boule unité fermée de N. Soit $x \in \bigcap_{i=1}^{n} B_i$. Alors, pour tout $i \in \{1, ..., n\}$, $N_i(x) \leq 1$, donc on a :
$$N(x) = \sum_{i=1}^{n} \lambda_i N_i(x) \leq \sum_{i=1}^{n} \lambda_i = 1.$$
Par suite, $x \in B$. Donc $\bigcap_{i=1}^{n} B_i \subset B$.
L'inclusion réciproque est fausse en général (re-exercice : trouver un contre-exemple !).

Exercice 11.

On considère $\mathbb{K}$ muni de sa structure d'espace vectoriel sur lui-même. Montrer que toute norme de $\mathbb{K}$ est de la forme $x \mapsto \alpha|x|$ avec $\alpha \in \mathbb{R}_+^*$.

Correction.

D'après la proposition précédente, $x \mapsto \alpha|x|$ est une norme sur $\mathbb{K}$. Réciproquement, soit N une norme sur $\mathbb{K}$. On pose $\alpha = N(1)$ qui est strictement positif d'après l'axiome de séparation. D'après l'axiome d'homogénéité, on a pour tout $x \in \mathbb{K}$,
$$N(x) = N(x.1) = |x|N(1) = \alpha|x|.$$

b. Composition par une fonction injective

Proposition 11.

Soit F un espace vectoriel sur $\mathbb{K}$, N une norme sur E et u une application linéaire *injective* de F dans E. Alors l'application $N' : F \to \mathbb{R}_+$ définie par :

$$N' = N \circ u : x \mapsto N(u(x)),$$

est une norme sur F.

Démonstration.

Soit $x, y \in F$ et $\lambda \in \mathbb{K}$.

i) (Positivité) Comme N est positive, $N' = N \circ u$ l'est aussi.

ii) (Séparation) On suppose que $N'(x) = 0$. Alors $N(u(x)) = 0$ donc $u(x) = 0_E$. Or u est une application linéaire injective donc $x = 0_F$.

iii) (Homogénéité) Comme u est linéaire, on a

$$N'(\lambda x) = N(\lambda u(x)) = |\lambda| N(u(x)) = |\lambda| N'(x).$$

iv) (Inégalité triangulaire) Toujours par linéarité de u :

$$N'(x+y) = N(u(x) + u(y)) \leq N(u(x)) + N(u(y)) = N'(x) + N'(y).$$

□

c. Application : normes sur $\mathbb{K}[X]$

On considère l'espace vectoriel $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$ à une indéterminée. Grâce à la proposition précédente, on va munir $\mathbb{K}[X]$ de plusieurs structures d'espace vectoriel normé à partir des exemples que l'on a étudiés dans le paragraphe 2).

Exemple 2.

Soit P un polynôme de $\mathbb{K}[X]$. On note $(a_i)_{i\in\mathbb{N}}$ la suite de ses coefficients (on rappelle que cette suite est stationnaire en 0) et on notera $t \mapsto P(t)$ la fonction polynomiale associée à P.

Normes provenant des espaces de suites : Les applications $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ sont des normes sur $\mathbb{K}[X]$ où :

$$\|P\|_1 = \sum_{i=1}^{+\infty} |a_i| \quad \|P\|_2 = \sqrt{\sum_{i=1}^{+\infty} |a_i|^2} \text{ et } \|P\|_\infty = \sup_{i\in\mathbb{N}} |a_i|.$$

Normes provenant des espaces de fonctions : Soit $a < b$ des réels. Les applications $\mathcal{N}_1$, $\mathcal{N}_2$ et $\mathcal{N}_\infty$ sont des normes sur $\mathbb{K}[X]$ où :

$$\mathcal{N}_1(P) = \int_a^b |P(t)| dt \quad \mathcal{N}_2(P) = \sqrt{\int_a^b |P(t)|^2 dt} \text{ et } \mathcal{N}_\infty(P) = \sup_{t\in[a,b]} |P(t)|.$$

Exercice 12.

Expliciter les espaces vectoriels normés et les applications linéaires injectives qui permettent, grâce à la proposition précédente, de munir $\mathbb{K}[X]$ des normes ci-dessus.

Correction.

Pour les trois premières normes, on utilise l'application :

$$u : \left| \begin{array}{ccc} \mathbb{K}[X] & \to & \ell^{1,2,\infty}(\mathbb{K}) \\ P & \mapsto & (a_i)_{i\in\mathbb{N}} \end{array} \right.$$

où $\ell^1(\mathbb{N},\mathbb{K})$, $\ell^2(\mathbb{N},\mathbb{K})$ et $\ell^\infty(\mathbb{N},\mathbb{K})$ sont munis de leur normes canoniques respectives.

Pour les trois suivantes, on utilise l'application :

$$u : \left| \begin{array}{ccc} \mathbb{K}[X] & \to & C([a,b],\mathbb{K}) \\ P & \mapsto & \{t \mapsto P(t)\} \end{array} \right.$$

où $C([a,b],\mathbb{K})$ est tour à tour muni des normes de la convergence en moyenne, en moyenne quadratique et uniforme.

Exercice 13.

Soit $n \in \mathbb{N}^*$ et $x_1, ..., x_{n+1}$ des éléments de $\mathbb{K}$ deux à deux distincts. Montrer que, pour $P \in \mathbb{K}[X]$ de degré inférieur ou égal à n, la quantité :

$$\|P\| = \max(|P(x_1)|, ..., |P(x_{n+1})|),$$

définit une norme sur l'espace vectoriel $\mathbb{K}_n[X]$ des polynômes de degré inférieur ou égal à n.

Correction.

Voici deux façons de procéder pour cet exercice :

- **À partir de la définition :**

  Soit $P, Q \in K_n[X]$ et $\lambda \in \mathbb{K}$.

  — (Positivité). Pour tout $i \in [\![1, n+1]\!]$, $|P(x_i)| \geq 0$ donc $\|P\| = \max\limits_{1\leq i\leq n+1}(|P(x_i)|) \geq 0$.

  — (Séparation). On suppose $\|P\| = 0$. Alors, pour tout $i \in [\![1, n+1]\!]$, $P(x_i) = 0$. Par suite, P est un polynôme de degré au plus n qui possède $n+1$ racines (car les x_i sont tous distincts). Il en résulte que P est le polynôme nul.

  — (Homogénéité). On a :

  $$\|\lambda P\| = \max_{1\leq i\leq n+1}(\underbrace{|\lambda P(x_i)|}_{=|\lambda||P(x_i)|}) = |\lambda| \max_{1\leq i\leq n+1}(|P(x_i)|) = |\lambda|\|P\|.$$

  — (Inégalité triangulaire). On a :

  $$\|P+Q\| = \max_{1\leq i\leq n+1}(\underbrace{|P(x_i)+Q(x_i)|}_{\leq|P(x_i)|+|Q(x_i)|}) \leq \max_{1\leq i\leq n+1}(|P(x_i)|) + \max_{1\leq i\leq n+1}(|Q(x_i)|) = \|P\|+\|Q\|.$$

- **En utilisant une application linéaire injective :**

  On remarque que, pour $\|P\| = \|(P(x_1), ..., P(x_{n+1}))\|_\infty$ où $\|\cdot\|_\infty$ est la norme infinie sur $\mathbb{K}^{n+1}$. On définit alors l'application $u : \mathbb{K}_n[X] \to \mathbb{K}^{n+1}$ de la manière suivante : pour $P \in \mathbb{K}_n[X]$,
  $$u(P) = (P(x_1), ..., P(x_{n+1})).$$
  Montrons que u est une application linéaire injective.

  — (Linéarité). Soit $P, Q \in \mathbb{K}_n[X]$ et $\lambda, \mu \in \mathbb{K}$. On a :
  $$\begin{aligned} u(\lambda P + \mu Q) &= ((\lambda P + \mu Q)(x_1), ..., (\lambda P + \mu Q)(x_{n+1})) \\ &= (\lambda P(x_1) + \mu Q(x_1), ..., \lambda P(x_{n+1}) + \mu Q(x_{n+1})) \\ &= \lambda(P(x_1), ..., P(x_{n+1})) + \mu(Q(x_1), ..., Q(x_{n+1})) \\ &= \lambda u(P) + \mu u(Q). \end{aligned}$$
  Donc u est linéaire.

  — (Injectivité). Soit $P \in \mathrm{Ker}(u)$. Alors $u(P) = (0, ..., 0)$. Par suite, pour tout $i \in [\![1, n+1]\!]$, x_i est une racine de P. Or les x_i sont tous différents, donc P est un polynôme de degré au plus n qui possède $n+1$ racines. Il s'agit donc du polynôme nul (*Un polynôme non nul de degré au plus n a au plus n racines*). Ainsi, $\mathrm{Ker}(u) = \{0\}$ et donc u est injective.

  Remarque : on peut montrer que u est en fait bijective (et donc que c'est un isomorphisme) grâce au théorème du rang ou en montrant directement que u est surjective.

  Ainsi, comme u est linéaire injective et que
  $$P \mapsto \|P\| = \|u(P)\|_\infty,$$
  il résulte de la Proposition 11 que $\|\cdot\|$ est une norme sur $\mathbb{K}_n[X]$.

d. Normes induites

Étant donné une sous-espace F de l'espace vectoriel E muni d'une norme, on utilise la restriction à F de cette dernière pour munir d'une structure d'espace vectoriel normé :

Définition-Proposition 15. *Norme induite*

Soit N une norme sur E et F un sous-espace vectoriel de E. Alors la restriction de N à F est une norme sur F.

Cette restriction est appelée **norme induite** sur F par N.

Démonstration.

On utilise la proposition précédente : en effet, l'injection canonique $\iota : F \to E$ est une application linéaire injective de F dans E. La norme induite sur F par N est alors la norme $N \circ \iota$. □

Remarque 9.

Pour désigner la norme induite sur F, on utilisera en général la même notation que pour la norme sur E.

Exemple 3.

- La valeur absolue sur $\mathbb{R}$ est la norme induite par le module sur $\mathbb{C}$.
- Soit $a, b \in \mathbb{R}$ avec $a < b$. La norme de la convergence uniforme sur $C([a,b],\mathbb{K})$ est la norme induite par la norme infinie sur $\mathcal{F}_b([a,b],\mathbb{K})$.

Exercice 14.

Quelle norme usuelle peut-t-on induire sur l'espace vectoriel $c_0(\mathbb{K})$ des suites à valeurs dans $\mathbb{K}$ qui convergent vers 0.

Correction.

Parmi les espaces vectoriel normés de suites que l'on a étudié, seul $\ell^\infty(\mathbb{N},\mathbb{K})$ contient $c_0(\mathbb{K})$. On induit alors la norme infinie de $\ell^\infty(\mathbb{N},\mathbb{K})$ sur $c_0(\mathbb{K})$.

e. Normes produits

On peut définir une norme naturelle sur un produit fini d'espaces vectoriels normés :

Définition-Proposition 16. *Norme produit*

Soit $n \in \mathbb{N}^*$ et $(E_1, N_1), ..., (E_n, N_n)$ des espaces vectoriels normés sur $\mathbb{K}$.
On pose $E = E_1 \times ... \times E_n$. Alors l'application $N : E \to \mathbb{R}_+$, appelée **norme produit**, et définie, pour $x = (x_1, ..., x_n) \in E$, par :

$$N(x) = \max(N_1(x_1), ..., N_n(x_n)),$$

est une norme sur E.

Correction.

L'application N est clairement positive.
Soit $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in E$ et $\lambda \in \mathbb{K}$.

- Séparation. On suppose $N(x) = 0$. Alors $\max(N_1(x_1), ..., N_n(x_n)) = 0$, donc $N_i(x_i) = 0$ pour $i = 1, ..., n$ car max est une norme. Par suite, $x_i = 0_{E_i}$ pour $i = 1, ...n$ car chaque N_i est une norme.
- Homogénéité. On a :

$$\begin{aligned} N(\lambda x) &= \max(N_1(\lambda x_1), ..., N_n(\lambda x_n)) \\ &= \max(|\lambda| N_1(x_1), ..., |\lambda| N_n(x_n)) \\ &= \max(|\lambda|(N_1(x_1), ..., N_n(x_n))) \\ &= |\lambda| \max(N_1(x_1), ..., N_n(x_n)) \\ &= |\lambda| N(x). \end{aligned}$$

— Inégalité triangulaire. On a :

$$\begin{aligned} N(x+y) &= \max(N_1(x_1+y_1), ..., N_n(x_n+y_n)) \\ &\leq \max(N_1(x_1)+N_1(y_1), ..., N_n(x_n)+N(y_n)) \\ &= \max(N_1(x_1), ..., N_n(x_n)) + \max(N_1(y_1), ..., N_n(y_n)) \\ &= N(x)+N(y). \end{aligned}$$

Exercice 15.

Comment peut-on obtenir la norme infinie sur $\mathbb{K}^n$ en utilisant une norme produit ?

Correction.

On considère, pour chaque $i \in [\![1, n]\!]$, l'espace vectoriel normé (E_i, N_i) où $E_i = \mathbb{K}$ et $N_i = |\cdot|$. Alors la norme produit sur $\mathbb{K}^n = E_1 \times ... \times E_n$ est égale à la norme infinie sur $\mathbb{K}^n$.

Exercice 16.

En s'inspirant de la définition-proposition précédente, définir deux autres normes sur un produit fini d'espaces vectoriels normés.

Correction.

Soit $n \in \mathbb{N}^*$ et $(E_i, N_i)_{1 \leq i \leq n}$ une famille d'espaces vectoriels normés. Alors, pour $p \geq 1$, on peut définir une norme N sur $E = E_1 \times ... \times E_n$ de la façon suivante : pour $x = (x_1, ..., x_n) \in E$,

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

Partie B
Suites dans un espace vectoriel normé

Dans cette partie, $(E, \|\cdot\|)$ désigne un espace vectoriel normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1. Suites convergentes

Définition 17. *Suite convergente*

Soit $u = (u_n)_{n\in\mathbb{N}}$ à valeurs dans E. On dit que u est **convergente** (dans $(E, \|\cdot\|)$) s'il existe un élément $\ell \in E$ tel que la suite à valeurs réelles $(\|u_n - \ell\|)_{n\in\mathbb{N}}$ tend vers 0 quand n tend vers $+\infty$.
Dans ce cas on dit que la suite u **converge vers** ℓ.

Si une suite ne converge pas, on dit qu'elle est **divergente**.

Remarque 10.

Autrement dit, pour d la distance associée à $\|\cdot\|$, une suite (u_n) à valeurs dans E converge vers ℓ si la suite réelle $(d(u_n, \ell))_{n\in\mathbb{N}}$ tend vers 0.

Proposition 12.

Soit u une suite à valeurs dans E. Si u converge, alors il existe un *unique* ℓ tel que u converge vers ℓ.

Démonstration.

Soit $u = (u_n)$ une suite à valeurs dans E convergente. Soit $\ell_1, \ell_2 \in E$ tels que u converge à la fois vers ℓ_1 et ℓ_2. Alors on a, pour tout $n \in \mathbb{N}$:

$$\|\ell_1 - \ell_2\| = \|(\ell_1 - u_n) + (u_n - \ell_2)\| \leq \|\ell_1 - u_n\| + \|u_n - \ell_2\|.$$

On passe à la limite dans l'égalité précédente quand n tend vers $+\infty$ et on obtient $\|\ell_1 - \ell_2\| = 0$, donc $\ell_1 = \ell_2$. □

Notation 3.

Pour une suite $u = (u_n)$ à valeurs dans E qui converge vers $\ell \in E$, on notera indifféremment :

$$\lim_{n\to+\infty} u_n = \ell \text{ ou } \lim u = \ell \text{ ou } u_n \xrightarrow[n\to+\infty]{} \ell.$$

Remarque 11.

Même si les notations précédentes ne le laissent pas entendre, la notion de convergence et de limite de suite dans un espace vectoriel normé dépend très fortement de la norme sous-jacente.

Exercice 17.

On considère l'espace $C([0,1],\mathbb{R})$ et la suite $(f_n)_{n\in\mathbb{N}}$ telle que $f_n : x \mapsto x^n$.

1. Montrer que dans $(C([0,1],\mathbb{R}), \|\cdot\|_1)$, la suite (f_n) converge et déterminer sa limite.
2. Calculer la norme de (f_n) dans $(C([0,1],\mathbb{R}), \|\cdot\|_\infty)$. Qu'en conclut-on ?

Correction.

1. On a :

$$\|f_n - \mathbf{0}\|_1 = \int_0^1 |f_n(t) - 0| dt = \int_0^1 t^n dt = \left[\frac{t^{n+1}}{n+1}\right]_0^1 = \frac{1}{n+1} \xrightarrow[n\to\infty]{} 0.$$

Par suite, (f_n) converge vers la fonction nulle dans $(C([0,1],\mathbb{R}), \|\cdot\|_1)$.

2. On a, pour tout $n \in \mathbb{N}$:

$$\|f_n - \mathbf{0}\|_\infty = \sup_{t\in[0,1]} |f_n(t) - 0| = \sup_{t\in[0,1]} (t^n) = 1.$$

Donc (f_n) ne converge pas vers $\mathbf{0}$ dans $(C([0,1],\mathbb{R}), \|\cdot\|_\infty)$.

Pire ! (f_n) diverge dans $(C([0,1],\mathbb{R}), \|\cdot\|_\infty)$!

En effet, supposons par l'absurde que (f_n) converge dans $(C([0,1],\mathbb{R}), \|\cdot\|_\infty)$ vers une fonction $f \in C([0,1],\mathbb{R})$. Alors pour tout $t \in [0,1]$:

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

Comme $f_n(t) \to 0$ si $t \in [0,1[$ et $f_n(1) \to 1$, on en déduit, pour $t \in [0,1]$:

$$f(t) = \begin{cases} 0 & \text{si } 0 \leq t < 1 \\ 1 & \text{si } t = 1. \end{cases}$$

La fonction f n'est pas continue, contradiction !

Notation 4.

On note :
— $c(E)$ l'ensemble des suites à valeurs dans E convergentes ;
— $c_0(E)$ l'ensemble des suites à valeurs dans E qui convergent vers 0 ;
— $\ell^\infty(\mathbb{N}, E) = \mathcal{F}_b(\mathbb{N}, E)$ l'ensemble des suites à valeurs dans E bornées.

Remarque 12.

Comme on l'a vu dans la partie précédente, $\ell^\infty(E) = \mathcal{F}_b(\mathbb{N}, E)$ muni de la norme $\|\cdot\|_\infty$ est un espace vectoriel normé. On rappelle que pour $u = (u_n)_{n\in\mathbb{N}} \in \ell^\infty(E)$,

$$\|u\|_\infty = \sup_{n\in\mathbb{N}} \|u_n\|.$$

2. Opérations algébriques sur les suites convergentes

a. Combinaisons linéaires

Proposition 13. *Combinaison linéaire*

Soit $u = (u_n)_{n\in\mathbb{N}}$, $v = (v_n)_{n\in\mathbb{N}}$ des suites à valeurs dans E.
Si u converge vers $\ell \in E$ et v converge vers $\ell' \in E$ alors pour tous $\lambda, \mu \in \mathbb{K}$, la suite $\lambda u + \mu v$ converge vers $\lambda\ell + \mu\ell'$.

Démonstration.

On a $\lambda u + \mu v = (\lambda u_n + \mu v_n)_{n\in\mathbb{N}}$ et, d'après l'inégalité triangulaire et l'homogénéité de $\|\cdot\|$:

$$\begin{aligned}
\|(\lambda u_n + \mu v_n) - (\lambda\ell + \mu\ell')\| &= \|(\lambda u_n - \lambda\ell) + (\mu v_n - \mu\ell')\| \\
&\leq \|\lambda u_n - \lambda\ell\| + \|\mu v_n - \mu\ell'\| \\
&\leq |\lambda| \underbrace{\|u_n - \ell\|}_{\xrightarrow[n\to\infty]{} 0} + |\mu| \underbrace{\|v_n - \ell'\|}_{\xrightarrow[n\to\infty]{} 0}
\end{aligned}$$

$\square$

Corollaire 1.

Une combinaison linéaire de suites à valeurs dans E convergentes est convergente.

Démonstration.

On raisonne par récurrence en utilisant la proposition précédente. $\square$

Corollaire 2.

Les ensembles de suites $c(E)$ et $c_0(E)$ sont des espaces vectoriels sur $\mathbb{K}$ (munis de l'addition canonique des suites et de la multiplication canonique par les scalaires).

Démonstration.

La suite constante en 0 appartient clairement à $c(E)$ et d'après la proposition précédente, toute combinaison linéaire de deux éléments de $c(E)$ appartient à $c(E)$. Par suite, $c(E)$ est un sous-espace vectoriel de $E^{\mathbb{N}}$.
Concernant $c_0(E)$, il est clair que c'est un sous-espace vectoriel de $c(E)$. $\square$

Proposition 14.

Soit $u = (u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. Si u converge vers ℓ alors la suite à valeurs réelles $(\|u_n\|)_{n\in\mathbb{N}}$ converge vers $\|\ell\|$.

Démonstration.

On a, d'après le seconde inégalité triangulaire :

$$0 \leq \big|\|u_n\| - \|\ell\|\big| \leq \|u_n - \ell\| \xrightarrow[n\to\infty]{} 0.$$

$\square$

Corollaire 3.

Toute suite convergente est bornée. Autrement dit, $c(E) \subset \ell^\infty(E)$.

Démonstration.

Soit $u = (u_n) \in c(E)$. D'après la proposition précédente, la suite réelle $(\|u_n\|)$ converge et donc est bornée (exercice : montrer que toute suite à valeurs réelles convergente est bornée). Ainsi, il existe $R \geq 0$ tel que $\|(\|u_n\|)_{n\in\mathbb{N}}\|_\infty \leq R$. Par suite, $(\|u_n\|)$ appartient à $\ell^\infty(\mathbb{R})$. On a alors :

$$\underbrace{\|u\|_\infty}_{\text{norme sur } \ell^\infty(E)} = \sup_{n\in\mathbb{N}} \|u_n\| = \underbrace{\|(\|u_n\|)_{n\in\mathbb{N}}\|_\infty}_{\text{norme sur } \ell^\infty(\mathbb{R})} \leq R.$$

Donc $(u_n) \in \ell^\infty(E)$. Il en résulte que $c(E) \subset \ell^\infty(E)$. $\square$

Exercice 18.

Quelle norme peut-on considérer sur $c(E)$ et sur $c_0(E)$ pour les munir de structures d'espaces vectoriels normés ?

Correction.

$c(E)$ et $c_0(E)$ étant des sous-espaces vectoriels de $\ell^\infty(E)$, on peut induire la norme infinie sur ces deux sous-espaces.

Exercice 19.

Soit $(\lambda_n)_{n\in\mathbb{N}}$ une suite à valeurs scalaires et $(u_n)_{n\in\mathbb{N}}$ un suite à valeurs dans E. Montrer que si $(\lambda_n)_{n\in\mathbb{N}}$ et $(u_n)_{n\in\mathbb{N}}$ sont convergentes, alors la suite $(\lambda_n u_n)_{n\in\mathbb{N}}$ à valeurs dans E est convergente.

Correction.

On conjecture que la limite potentielle de $(\lambda_n u_n)_{n\in\mathbb{N}}$ est $\lambda\ell$ où $\lambda = \lim \lambda_n$ et $\ell = \lim_{n\to+\infty} u_n$. Montrons le. En utilisant l'inégalité triangulaire et l'homogénéité de $\|\cdot\|$, on a :

$$\|\lambda_n u_n - \lambda\ell\| = \|\lambda_n u_n - \lambda u_n + \lambda u_n - \lambda\ell\| \leq \underbrace{|\lambda_n - \lambda|}_{\xrightarrow[n\to\infty]{}0} \underbrace{\|u_n\|}_{\xrightarrow[n\to\infty]{}\|\ell\|} + |\lambda| \underbrace{\|u_n - \ell\|}_{\xrightarrow[n\to\infty]{}0} \xrightarrow[n\to\infty]{} 0.$$

b. Suites à valeurs dans un produit fini d'espaces vectoriels normés

Comme on l'a vu dans la partie précédente, on peut munir un produit fini d'espaces vectoriels normés de la norme produit. C'est-à-dire, pour $k \in \mathbb{N}^*$ et $(E_1, \|\cdot\|_{(1)}), \dots, (E_k, \|\cdot\|_{(k)})$ des espaces vectoriels normés, on considère la norme $\|\cdot\|$ définie, pour $x = (x_1, \dots, x_k) \in \prod_{i=1}^{k} E_i$, par :

$$\|x\| = \max_{1\leq i\leq k} \|x_i\|_{(i)}.$$

Proposition 15.

Soit $k \in \mathbb{N}^*$ et $(E_1, \|\cdot\|_{(1)}), \dots, (E_k, \|\cdot\|_{(k)})$ des espaces vectoriels normés. On munit $E = \prod_{i=1}^{k} E_i$ de la norme produit notée $\|\cdot\|$.

Soit $u = \left((u_n^{(1)})_{n\in\mathbb{N}}, \dots, (u_n^{(k)})_{n\in\mathbb{N}}\right)$ une suite à valeurs dans E. La suite u converge vers $(\ell^{(1)}, \dots, \ell^{(k)})$ dans E si, et seulement si, pour tout $i \in \{1, \dots, k\}$, $(u_n^{(i)})_{n\in\mathbb{N}}$ converge vers $\ell^{(i)}$ dans E_i.

Démonstration.

Soit $u = (u_n)_{n\in\mathbb{N}} = \left((u_n^{(1)})_{n\in\mathbb{N}}, \dots, (u_n^{(k)})_{n\in\mathbb{N}}\right)$ une suite à valeurs dans E.

- $(\Rightarrow)$. On suppose que u converge vers $\ell = (\ell^{(1)}, \dots, \ell^{(k)})$ dans E.
  Alors, pour tout $i \in \{1, \dots, k\}$, et tout $n \in \mathbb{N}$, on a :

  $$\|u_n^{(i)} - \ell^{(i)}\|_{(i)} \leq \max(\|u_n^{(1)} - \ell^{(1)}\|_{(1)}, \dots, \|u_n^{(k)} - \ell^{(k)}\|_{(k)}) = \|u_n - \ell\| \xrightarrow[n\to\infty]{} 0.$$

  Donc pour tout $i \in \{1, \dots, k\}$, $(u_n^{(i)})_{n\in\mathbb{N}}$ converge vers $\ell^{(i)}$ dans E_i.

- $(\Leftarrow)$. On suppose que pour tout $i \in \{1, \dots, k\}$, $(u_n^{(i)})_{n\in\mathbb{N}}$ converge vers $\ell^{(i)}$ dans E_i. On pose $\ell = (\ell^{(1)}, \dots, \ell^{(k)})$.
  Alors on a, pour tout $n \in \mathbb{N}$:

$$\begin{aligned}\|u_n - \ell\| &= \max(\|u_n^{(1)} - \ell^{(1)}\|_{(1)}, ..., \|u_n^{(k)} - \ell^{(k)}\|_{(k)}) \\ &\leq \underbrace{\|u_n^{(1)} - \ell^{(1)}\|_{(1)}}_{\xrightarrow[n\to\infty]{} 0} + ... + \underbrace{\|u_n^{(k)} - \ell^{(k)}\|_{(k)}}_{\xrightarrow[n\to\infty]{} 0} \xrightarrow[n\to\infty]{} 0.\end{aligned}$$

Donc u converge vers ℓ dans E.

□

Exemple 4.

Une suite à valeurs dans $\mathbb{K}^n$ converge (pour la norme infinie) si, et seulement si, chacune des n suites de ses coordonnées converge.

c. Relations de comparaisons

Dans ce paragraphe, on étend, pour des suites à valeurs vectorielles, les notations de Landau, "o" et "O" vues dans le cadre de l'analyse réelle (et complexe).

Définition 18. *Domination et négligeabilité*

Soit $u = (u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E et $a = (a_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles positives.

1. On dit que u est **dominée** par a en $+\infty$ s'il existe $M \in \mathbb{R}_+$ et il existe $N \in \mathbb{N}$ tels que pour tout $n \geq N$:
$$\|u_n\| \leq M a_n.$$
On note alors $u = O(a)$ ou $u_n = O(a_n)$.

2. On dit que u est **négligeable** devant a en $+\infty$ si pour tout $\varepsilon \in \mathbb{R}_+$ et il existe $N \in \mathbb{N}$ tels que pour tout $n \geq N$:
$$\|u_n\| \leq \varepsilon a_n.$$
On note alors $u = o(a)$ ou $u_n = o(a_n)$.

L'exercice suivant donne une vision plus intuitive des relations de domination et de négligeabilité.

Exercice 20.

Soit $u = (u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E et $a = (a_n)_{n\in\mathbb{N}}$ une suite à valeurs réelles *strictement* positives.

1. $u_n = O(a_n)$ si, et seulement si, la suite $\left(\dfrac{1}{a_n}u_n\right)_{n\in\mathbb{N}}$ à valeurs dans E est bornée.

2. $u_n = o(a_n)$ si, et seulement si, $\lim\limits_{n\to\infty} \dfrac{1}{a_n}u_n = 0_E$.

Correction.

1. En remarquant qu'une suite est bornée si, et seulement si, elle est bornée à partir d'un certain rang, on a :
$$u_n = O(a_n)$$
si, et seulement si,
$$\exists M \in \mathbb{R}_+,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geq N \Rightarrow \|u_n\| \leq M a_n,$$
si, et seulement si,
$$\exists M \in \mathbb{R}_+,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geq N \Rightarrow \|\tfrac{1}{a_n} u_n\| \leq M,$$
si, et seulement si,
$$\left(\frac{1}{a_n} u_n\right)_{n\in\mathbb{N}} \quad \text{est bornée.}$$

2. On a :
$$u_n = o(a_n)$$
si, et seulement si,
$$\forall \varepsilon \in \mathbb{R}_+^*,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geq N \Rightarrow \|u_n\| \leq \varepsilon a_n,$$
si, et seulement si,
$$\forall \varepsilon \in \mathbb{R}_+^*,\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ n \geq N \Rightarrow \|\tfrac{1}{a_n} u_n - 0_E\| \leq \varepsilon,$$
si, et seulement si,
$$\lim_{n\to\infty} \frac{1}{a_n} u_n = 0_E.$$

On peut alors comparer les suites à valeurs dans E entre elles :

Définition 19. *Domination, négligeabilité et équivalence*

Soit $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ des suites à valeurs dans E.

1. On dit que u est **dominée** par v en $+\infty$ et on note $u = O(v)$ ou $u_n = O(v_n)$ si :
$$u_n = O(\|v_n\|).$$
2. On dit que u est **négligeable** devant v en $+\infty$ et on note $u = o(v)$ ou $u_n = o(v_n)$ si :
$$u_n = o(\|v_n\|).$$
3. On dit que u **équivaut** à v en $+\infty$ et on note $u \sim v$ ou $u_n \sim v_n$ si :
$$u_n - v_n = o(\|v_n\|).$$

Remarque 13.

La relation $\sim$ est une relation d'équivalence entre les suites à valeurs dans E. On pourra donc employer la terminologie : "u et v sont équivalentes en $+\infty$" si la suite u équivaut à la suite v en $+\infty$.

Exercice 21.

Soit $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ des suites à valeurs dans E. On suppose que pour tout $n \in \mathbb{N}$, $v_n \neq 0_E$.
Montrer que si $u_n \sim v_n$ alors

$$\lim_{n\to\infty} \frac{\|u_n\|}{\|v_n\|} = 1.$$

Correction.

On a suppose $u_n \sim v_n$. Alors $u_n - v_n = o(\|v_n\|)$.
Soit $\varepsilon \in \mathbb{R}_+^*$. Alors il existe un rang $N \in \mathbb{N}$ tel que pour tout $n \geq N$,

$$\|u_n - v_n\| \leq \varepsilon \|v_n\|.$$

Soit $n \geq N$. Alors,

$$|\frac{\|u_n\|}{\|v_n\|} - 1| = \frac{1}{\|v_n\|}|\|u_n\| - \|v_n\|| \leq \frac{1}{\|v_n\|}\|u_n - v_n\| \leq \varepsilon.$$

Ceci étant vrai pour tout $\varepsilon > 0$, il en résulte que $\lim_{n\to\infty} \frac{\|u_n\|}{\|v_n\|} = 1.$

3. Suites extraites et valeurs d'adhérence

On emploie ici la même terminologie que pour les suites à valeurs réelles ou complexes :

Définition 20. *Suite extraite / sous-suite*

Soit $u = (u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E et $\varphi : \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante.
On appelle **suite extraite** ou **sous-suite** de la suite u, la suite

$$(u_{\varphi(n)})_{n\in\mathbb{N}}.$$

Remarque 14.

La fonction φ - *la fonction extractrice* - est souvent notée de la façon suivante : pour $k \in \mathbb{N}$, $n_k := \varphi(k)$. Ainsi, on désignera souvent par $(u_{n_k})_{k\in\mathbb{N}}$ une sous-suite de u.

Exercice 22.

1. Montrer que si $\varphi : \mathbb{N} \to \mathbb{N}$ est une fonction strictement croissante, alors pour tout $n \in \mathbb{N}$, $\varphi(n) \geq n$.
2. Soit u une suite à valeurs dans E et v une sous-suite de u. Montrer que si w est une sous-suite de v, alors w est une sous-suite de u.

Correction.

1. On remarque que pour tout $n \in \mathbb{N}$, $\varphi(n+1) \geq \varphi(n)+1$ car φ est strictement croissante et à valeurs entières. On prouve alors le résultat en raisonnant par récurrence sur $\mathbb{N}$ et en utilisant cette remarque.
2. Il suffit de remarquer que la composition de deux fonctions strictement croissantes est strictement croissante.

Proposition 16.

Soit u une suite à valeurs dans E et $\ell \in E$. La suite u converge vers ℓ si, et seulement si, **toute** sous-suite de u converge vers ℓ.

Démonstration.

Comme u est en particulier une sous-suite de u (la fonction extractrice est donnée par l'identité sur $\mathbb{N}$), si toute sous-suite de u converge vers ℓ, alors u converge ℓ. Montrons la réciproque.

On suppose que $u = (u_n)_{n\in\mathbb{N}}$ converge vers $\ell \in E$. Soit $v = (u_{\varphi(n)})_{n\in\mathbb{N}}$ une sous-suite de u. Soit $\varepsilon > 0$. Comme u converge, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $\|u_n - \ell\| \leq \varepsilon$. Or pour tout $n \in \mathbb{N}$, $\varphi(n) \geq n$, donc pour tout $n \geq N$, $\varphi(n) \geq N$ d'où :

$$\|u_{\varphi(n)} - \ell\| \leq \varepsilon.$$

Il en résulte que v converge vers ℓ. $\square$

Définition 21. *Valeur d'adhérence*

Soit u une suite à valeurs dans E et $x \in E$. On dit que x est une **valeur d'adhérence** de u s'il existe une sous-suite de u qui converge vers x.

Exemple 5.

— Les nombres 1 et -1 sont des valeurs d'adhérence de la suite $u = ((1-\frac{1}{n})(-1)^n)_{n\in\mathbb{N}^*}$.
— le nombre complexe i est une valeur d'adhérence de la suite $v = (e^{\frac{in\pi}{4}})_{n\in\mathbb{N}}$.

Exercice 23.

1. Pour les exemples précédents, déterminer des sous-suites qui convergent vers les valeurs d'adhérence.
2. Quelles sont les autres valeurs d'adhérence de la suite v de l'exemple précédent.

Correction.

— (u_{2n}) converge vers 1 et (u_{2n+1}) converge vers -1. (v_{2+8n}) converge vers i (elle est même constante en i).

— Les valeurs d'adhérence de v sont exactement les valeurs de v_n pour $n = 0, 1, ..., 7$.

Proposition 17.

Soit u une suite à valeurs dans E. Si u converge alors u possède une unique valeur d'adhérence.

Démonstration.

D'après la proposition 16, si u converge vers ℓ, toutes les sous-suites de u convergent également vers ℓ. Donc ℓ est une valeur d'adhérence de u et c'est la seule par unicité de la limite. □

Remarque 15.

- La contraposée de la proposition précédente nous donne une manière de prouver qu'une suite est divergente, c'est-à-dire :
  Si u possède au moins deux valeurs d'adhérence distinctes, alors u diverge.
- Attention, la réciproque de la proposition précédente est fausse ! Il existe des suites divergentes qui ne possèdent qu'une seule valeur d'adhérence, comme le montre l'exercice suivant.

Exercice 24.

On considère la suite $u = (u_n)_{n\in\mathbb{N}} = (n(1 + (-1)^n))_{n\in\mathbb{N}}$. Montrer que u possède une unique valeur d'adhérence. En déduire que la réciproque de la proposition 17 est fausse.

Correction.

La sous-suite (u_{2n+1}) est constante en 0 et donc a pour limite 0. Par suite, 0 est une valeur d'adhérence de u. Montrons que c'est la seule.

Soit $v = u_{\varphi(n)}$ une sous-suite de u. Si φ prend une infinité de valeurs impaires alors 0 est une valeur d'adhérence de v, donc quitte à passer à une sous-suite de v, on peut supposer que $\varphi(\mathbb{N}) \cap 2\mathbb{N}+1$ est fini. Par suite, il existe $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $\varphi(n)$ est pair ; donc

$$v_n = u_{\varphi(n)} = \varphi(n) \xrightarrow[n\to\infty]{} +\infty.$$

Donc 0 est bien la seule valeur d'adhérence de u.

De plus, il est clair que u ne converge pas : en effet, (u_{2n}) diverge. Par suite, la réciproque de la proposition 17 est fausse.

Exercice 25. *Caractérisation des valeurs d'adhérence*

Soit $u = (u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E et $x \in E$. Alors x est une valeur d'adhérence de u

si, et seulement si,
$$\forall \varepsilon,\ \forall N \in \mathbb{N}, \exists n \geq N,\ \|u_n - x\| \leq \varepsilon.$$

Correction.

- $(\Rightarrow)$. On suppose que x est une valeur d'adhérence de u. Alors il existe une sous-suite $(u_{\varphi(n)})$ de u qui converge vers x, i.e.
$$\forall \varepsilon,\ \exists n_0 \in \mathbb{N}, \forall n \geq n_0,\ \|u_{\varphi(n)} - x\| \leq \varepsilon.$$
Soit $\varepsilon > 0$ et $N \in \mathbb{N}$. Alors il existe un entier n_0 tel que pour tout $k \geq n_0$, $\|u_{\varphi(k)} - x\| \leq \varepsilon$.

  Considérons l'entier $n = \varphi(\max(N, n_0))$. Alors $n \geq \varphi(N) \geq N$ et $n \geq \varphi(n_0) \geq n_0$. Par suite, $n \geq N$ et $n \geq n_0$ donc :
$$\|u_n - x\| = \|u_{\varphi(\max(N,n_0))} - x\| \leq \varepsilon.$$

- $(\Leftarrow)$. On suppose
$$\forall \varepsilon,\ \forall N \in \mathbb{N}, \exists n \geq N,\ \|u_n - x\| \leq \varepsilon. \quad (*)$$
On construit, par récurrence sur $\mathbb{N}$, une suite strictement croissante d'entiers $(n_k)_{k\in\mathbb{N}}$ telle que pour tout $k \in \mathbb{N}$,
$$\|u_{n_k} - x\| \leq \frac{1}{2^k}.$$
Initialisation : D'après $(*)$, pour $\varepsilon = 1 = \frac{1}{2^0}$, il existe $n_0 \in \mathbb{N}$ tel que $\|u_n - x\| \leq \frac{1}{2^0}$.

  Hérédité : Soit $k \in \mathbb{N}$. On suppose qu'il existe $n_0 < ... < n_k$ des entiers tels que $\|u_{n_k} - x\| \leq \frac{1}{2^k}$.
  D'après $(*)$, pour $\varepsilon = \frac{1}{2^{k+1}}$ et $N = n_k$, il existe $n_{k+1} > n_k$ tel que
$$\|u_{n_{k+1}} - x\| \leq \frac{1}{2^{k+1}}.$$
Il en résulte que la sous-suite $(u_{n_k})_{k\in\mathbb{N}}$ converge vers x.

Partie C
Comparaison de normes

Dans toute cette partie, E est un espace vectoriel sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

On a discuté tout au long des parties précédentes, de propriétés d'ensembles ou de convergence de suites, qui ne sont pas en général invariantes par changement de norme dans notre espace E. Par exemple, on a vu dans les exercices précédents, que la suite (f_n) de $C([0,1],\mathbb{R})$ donnée, pour $n \in \mathbb{N}$, par $f_n : t \mapsto t^n$, est convergente pour la norme de la convergence en moyenne mais pas pour la norme de la convergence uniforme.

On peut alors légitimement se demander quelles conditions sur deux normes N_1 et N_2 sur E données pourraient nous permettre d'assurer que si une certaine propriété est vraie pour N_1, alors elle l'est également sur N_2.

Voici tout l'enjeu du chapitre "Comparaison de normes".

1. Domination de normes

Définition 22.

Soit N_1, N_2 des normes sur E. On dit que N_1 est **dominée** par N_2 s'il existe un réel $C > 0$ tel que $N_1 \leq CN_2$ i.e. pour tout $x \in E$:

$$N_1(x) \leq CN_2(x).$$

Dans ce cas, on dira également que la norme N_2 est **plus fine** que N_1.

Proposition 18.

Soit N_1, N_2 des normes sur E. N_1 est dominée par N_2, si, et seulement si, il existe $C > 0$ tel que pour tout $x \in E$, $N_2(x) \leq 1$ implique $N_1(x) \leq C$.

Autrement dit, N_1 est dominée par N_2, si, et seulement si, il existe $C > 0$ tel que la boule unité fermée pour la norme N_2 est incluse dans la boule fermée pour la norme N_1 de centre 0_E et de rayon C.

Démonstration.

- $(\Rightarrow)$. Immédiat.
- $(\Leftarrow)$. On suppose que pour tout $x \in E$, $N_2(x) \leq 1$ implique $N_1(x) \leq C$.
  Soit $x \in E$. Si $x{=}0_E$, alors $N_1(x) = 0 \leq 0 = CN_2(x)$. On suppose $x \neq 0_E$.
  Alors $u = \frac{1}{N_2(x)} x$ vérifie $N_2(u) = 1 \leq 1$ donc, par hypothèse,
  $$\frac{N_1(x)}{N_2(x)} = N_1(u) \leq C.$$
  D'où le résultat.

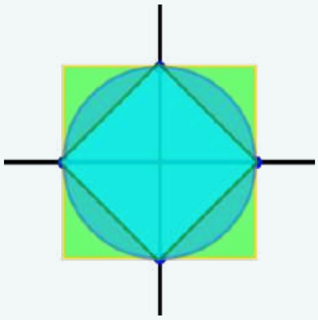
$\square$

Exemple 6.

Sur $\mathbb{R}^2$, en examinant les boules unités des normes un, deux et infini, on peut, dans un premier temps, remarquer les relations suivantes :

la norme un est plus fine que la norme deux qui est plus fine que la norme infinie, i.e.

$$\|\cdot\|_\infty \leq \|\cdot\|_2 \leq \|\cdot\|_1.$$



Exercice 26.

Soit $a, b \in \mathbb{R}$ avec $a < b$. Montrer que dans $C([a,b],\mathbb{K})$, les normes de la convergence en moyenne et de la convergence en moyenne quadratique sont dominées par la norme de la convergence uniforme.

Correction.

Soit $f \in C([a,b],\mathbb{K})$. Alors

$$\|f\|_1 = \int_a^b |f(t)|dt \leq \int_a^b \|f\|_\infty dt = (b-a)\|f\|_\infty,$$

et de même,

$$\|f\|_2 = \sqrt{\int_a^b |f(t)|^2 dt} \leq \sqrt{\int_a^b \|f\|_\infty^2 dt} = \sqrt{b-a}\|f\|_\infty.$$

Pour comprendre la terminologie de "finesse" précédente qui pourrait sembler paradoxale au premier abord, il faut étudier les conséquences de la domination en termes de boules pour chaque norme :

Proposition 19.

Soit N_1,N_2 des normes sur E. On suppose que N_1 est dominée par N_2. Alors toute boule fermée pour N_1 contient une boule fermée pour N_2.

Démonstration.

On suppose qu'il existe $C > 0$ tel que $N_1 \leq CN_2$. Alors, pour tous $x \in E$ et $r > 0$,

$$B_f^{N_2}(x, \frac{r}{C}) \subset B_f^{N_2}(x, \frac{r}{C}).$$

En effet, si $y \in B_f^{N_2}(x, \frac{r}{C})$, alors

$$N_1(y) \leq CN_2(y) \leq C\frac{r}{C} = r.$$

□

Remarque 16.

On peut remplacer les mentions "boule fermée" dans la proposition précédente par "boule ouverte" et même seulement l'une des deux.

En ce qui concerne les notions d'ensemble ou d'application bornée et de convergence de suites, la norme dominante les transmet à la norme dominée :

Proposition 20.

Soit N_1,N_2 des normes sur E. On suppose que N_1 est dominée par N_2. Alors une partie bornée pour N_2 est une partie bornée pour N_1.

Démonstration.

On suppose qu'il existe $C > 0$ tel que $N_1 \leq CN_2$. Soit B une partie bornée pour N_2. Alors il existe $R \geq 0$ tel que pour tout $x \in B$, $N_2(x) \leq R$. Ainsi, pour tout $x \in B$,

$$N_1(x) \leq CN_2(x) \leq CR.$$

Donc B est bornée. □

Corollaire 4.

Soit N_1,N_2 des normes sur E. On suppose que N_1 est dominée par N_2.

— Toute application d'un ensemble X dans E bornée pour la norme N_2, est bornée pour la norme N_1.

— Toute suite à valeurs dans E bornée pour la norme N_2 est bornée pour la norme N_1.

Démonstration.

On suppose N_1 dominée par N_2.

— Soit $f : X \to E$. Une application bornée pour N_2. Alors $f(X)$ est une partie de E bornée pour N_2. Par suite, d'après la proposition précédente, $f(X)$ est bornée pour N_1. Il en résulte que f est une application bornée pour N_1.

— On applique le point précédent en voyant une suite à valeurs dans E comme une application de $\mathbb{N}$ dans E.

□

Proposition 21.

Soit N_1,N_2 des normes sur E. On suppose que N_1 est dominée par N_2. Si une suite à valeurs dans E converge vers $\ell \in E$ pour la norme N_2, alors elle converge vers ℓ pour la norme N_1.

Démonstration.

On suppose qu'il existe $C > 0$ tel que $N_1 \leq CN_2$. Soit (u_n) une suite à valeurs dans E. On suppose que (u_n) converge vers $\ell \in E$ pour la norme N_2. Par suite,

$$N_1(u_n - \ell) \leq CN_2(u_n - \ell) \xrightarrow[n\to\infty]{} 0.$$

Par suite, (u_n) converge vers ℓ pour N_1. $\square$

Remarque 17.

Ainsi, pour montrer qu'une norme N_1 n'est pas dominée par N_2, il suffit d'exhiber une suite qui converge pour N_2 mais pas pour N_1.

Exercice 27.

Montrer que dans $C([0,1],\mathbb{R})$, la norme de la convergence uniforme n'est pas dominée par la norme de la convergence en moyenne.

Correction.

On reprend la suite de terme général $f_n : t \mapsto t^n$ qui converge vers $\mathbf{0}$ pour la norme de la convergence en moyenne, et qui ne converge pas vers $\mathbf{0}$ pour la norme de la convergence uniforme (c'est même plus fort : elle diverge pour la norme de la convergence uniforme). Par suite, d'après la contraposée de la proposition précédente, la norme de la convergence uniforme n'est pas dominée par la norme de la convergence en moyenne.

2. Normes équivalentes

a. Définition

Définition 23. *Normes équivalentes*

Soit N_1, N_2 des normes sur E. On dit que N_1 et N_2 sont équivalentes s'il existe des constantes $c > 0$ et $C > 0$ telles que :

$$cN_1(x) \leq N_2(x) \leq CN_1(x).$$

Proposition 22.

Soit N_1, N_2 des normes sur E. N_1 et N_2 sont équivalentes si, et seulement si, N_1 est dominée par N_2 et N_2 est dominée par N_1.

Démonstration.

$$N_1 \text{ est dominée par } N_2 \text{ et } N_2 \text{ est dominée par } N_1\,;$$

si, et seulement si

$$\exists K > 0,\ \forall x \in E, \quad N_1(x) \leq K N_2(x) \quad \text{et} \quad \exists C > 0,\ \forall x \in E, \quad N_2(x) \leq C N_1(x);$$

si, et seulement si

$$\exists K, C > 0,\ \forall x \in E, \quad \frac{1}{K} N_1(x) \leq N_2(x) \leq C N_1(x);$$

si, et seulement si (ici $c = 1/K$)

$$\exists c, C > 0,\ \forall x \in E, \quad c N_1(x) \leq N_2(x) \leq C N_1(x).$$

□

Exercice 28.

Soit $n \in \mathbb{N}^*$. On considère l'espace vectoriel $\mathbb{K}^n$. Montrer que les normes un, deux et infini sont équivalentes.

Correction.

On a, pour $x = (x_1, ..., x_n) \in \mathbb{K}^n$,

$$\|x\|_\infty^2 = (\max(|x_1|, ..., |x_n|))^2 \leq \max(|x_1|^2, ..., |x_n|^2) \leq \sum_{i=1}^n |x_i|^2 = \|x\|_2^2;$$

$$\|x\|_2^2 = \sum_{i=1}^n |x_i|^2 \leq \left(\sum_{i=1}^n |x_i|\right)^2 = \|x\|_1^2;$$

$$\|x_i\|_1 = \sum_{i=1}^n \underbrace{|x_i|}_{\leq \|x\|_\infty} \leq n\|x\|_\infty.$$

Donc on obtient, en utilisant ces inégalités et celles obtenues précédemment :

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\|x\|_\infty$$

D'où l'équivalence des normes un, deux et infini.

b. Propriétés invariantes par passage à une norme équivalente

Théorème 1. *Conservation du caractère borné d'une partie*

Soit N_1,N_2 deux normes sur E. On suppose que N_1 et N_2 sont équivalentes.
Une partie est bornée pour N_1 si, et seulement si, elle est bornée pour N_2.

Démonstration.

On applique la proposition 20 pour N_1 dominée par N_2 puis pour N_2 dominée par N_1. □

Corollaire 5.

Soit N_1,N_2 deux normes sur E. On suppose que N_1 et N_2 sont équivalentes.

- Une application d'un ensemble X dans E est bornée pour la norme N_1 si, et seulement si, est bornée pour la norme N_2.
- Une suite à valeurs dans E bornée pour la norme N_1 si, et seulement si, elle est bornée pour la norme N_2.

Théorème 2. *Conservation de la convergence d'une suite*

Soit N_1,N_2 des normes sur E. On suppose que N_1 et N_2 sont équivalentes. Une suite à valeurs dans E converge vers $\ell \in E$ pour la norme N_1 si, et seulement si, elle converge vers ℓ pour la norme N_2.

Démonstration.

On applique la proposition 21 pour N_1 dominée par N_2 puis pour N_2 dominée par N_1. □

c. Exercices types

Exercice 29.

On considère l'espace vectoriel $E = C^1([0,1],\mathbb{R})$ des fonctions de classe C^1 de $[0,1]$ dans $\mathbb{R}$ et la norme $\|\cdot\|_\infty$ sur E.

1. Pour $f \in E$, on pose :
$$N(f) = |f(0)| + \|f'\|_\infty$$
Montrer que N est une norme sur E.
2. Pour $f \in E$, on pose :
$$N'(f) = \|f\|_\infty + \|f'\|_\infty$$
Montrer que N' est une norme sur E.
3. Montrer que N et N' sont équivalentes.
4. Les normes N et N' sont-elles équivalentes à $\|\cdot\|_\infty$?

Correction.

1. Soit $f, g \in E$ et $\lambda \in \mathbb{R}$.
   — (Positivité) $N(f)$ positif comme somme de quantités positives.
   — (Séparation) On suppose $N(f) = 0$. Alors $|f(0)| = 0 = \|f'\|_\infty$. Par suite, $f(0) = 0$ et pour tout $t \in [0,1]$, $f'(t) = 0$. Donc f est constante en 0 i.e. est égale à la fonction nulle.
   — (Homogénéité) Par la linéarité de la dérivée et l'homogénéité de $\|\cdot\|_\infty$, on a :
   $$N(\lambda f) = |(\lambda f)(0)| + \|(\lambda f)'\|_\infty = |\lambda||f(0)| + |\lambda|\|f'\|_\infty = |\lambda| N(f).$$
   — (Inégalité triangulaire) Par la linéarité de la dérivée et l'inégalité triangulaire de $\|\cdot\|_\infty$, on a :
   $$N(f+g) = |(f+g)(0)| + \|(f+g)'\|_\infty \leq |f(0)| + |g(0)| + \|f'\|_\infty + \|g'\|_\infty = N(f) + N(g).$$
2. Soit $f, g \in E$ et $\lambda \in \mathbb{R}$.
   — (Positivité) $N'(f)$ positif comme somme de quantités positives.
   — (Séparation) On suppose $N'(f) = 0$. Alors, en particulier, $\|f\|_\infty = 0$. Par suite, d'après l'axiome de séparation pour $\|\cdot\|_\infty$, f est égale à la fonction nulle.
   — (Homogénéité) Par la linéarité de la dérivée et l'homogénéité de $\|\cdot\|_\infty$, on a :
   $$N'(\lambda f) = \|(\lambda f)\|_\infty + \|(\lambda f)'\|_\infty = |\lambda|\|f\|_\infty + |\lambda|\|f'\|_\infty = |\lambda| N'(f).$$
   — (Inégalité triangulaire) Par la linéarité de la dérivée et l'inégalité triangulaire de $\|\cdot\|_\infty$, on a :
   $$N'(f+g) = \|(f+g)\|_\infty + \|(f+g)'\|_\infty \leq \|f\|_\infty + \|g\|_\infty + \|f'\|_\infty + \|g'\|_\infty = N'(f) + N'(g).$$
3. On a tout d'abord,
   $$N(f) = \underbrace{|f(0)|}_{\leq \|f\|_\infty} + \|f'\|_\infty \leq N'(f).$$
   Puis, pour tout $x \in [0,1]$,
   $$|f(x)| = |f(0) + \int_0^x f'(t)dt| \leq |f(0)| + \int_0^x \|f'\|_\infty dt = |f(0)| + x\|f'\|_\infty \leq N(f).$$
   Donc, $\|f\|_\infty \leq N(f)$. Par suite,
   $$N'(f) = \|f\|_\infty + \|f'\|_\infty \leq \underbrace{\|f\|_\infty}_{\leq N(f)} + \underbrace{|f(0)| + \|f'\|_\infty}_{=N(f)} \leq 2N(f).$$
   Conclusion :
   $$N \leq N' \leq 2N,$$
   donc N et N' sont équivalentes.
4. Prenons la suite de terme général $f_n : t \mapsto t^n$ ($n \in \mathbb{N}^*$). Alors :
   $$\|f_n\|_\infty = 1,$$
   et
   $$N(f_n) = n \xrightarrow[n\to\infty]{} \infty.$$
   Donc (f_n) est une suite bornée pour $\|\cdot\|_\infty$ mais pas pour N. donc ces normes ne sont pas équivalentes. Et comme N' est équivalente à N, N' ne peut-être équivalente à $\|\cdot\|_\infty$.

Exercice 30.

Soit $a \geq 0$. Pour $P \in \mathbb{R}[X]$, on définit

$$N_a(P) = |P(a)| + \int_0^1 |P'(t)|dt.$$

1. Démontrer que N_a est une norme sur $\mathbb{R}[X]$.
2. Soit $a, b \geq 0$ avec $a < b$ et $b > 1$. Démontrer que N_a et N_b ne sont pas équivalentes.
3. Démontrer que si $(a, b) \in [0, 1]^2$, alors N_a et N_b sont équivalentes.

Correction.

1. Le point "délicat" est la séparation : si $N_a(P) = 0$, alors $|P(a)| = 0$ et $\int_0^1 |P'(t)|dt = 0$. Or, comme $|P'|$ est une fonction continue, positive et d'intégrale nulle sur $[0, 1]$, alors $P'(x) = 0$ pour tout $x \in [0, 1]$. Alors P' est un polynôme qui possède une infinité de racine ; par suite $P' = 0$ et donc P est un polynôme constant. Or $P(a) = 0$, donc P est le polynôme nul.
2. On suppose que N_a et N_b sont équivalentes. Alors, il existe deux constantes $C_1 > 0$ et $C_2 > 0$ tels que, pour tout $P \in \mathbb{R}[X]$, on a :
$$C_1 N_a(P) \leq N_b(P) \leq C_2 N_a(P).$$
Pour $n \geq 0$, soit $P(X) = X^n$. On a
$$N_a(P) = a^n + n \int_0^1 t^{n-1}dt = a^n + 1 \text{ et } N_b(P) = b^n + 1.$$
On en déduit alors que, pour tout $n \geq 0$,
$$b^n + 1 \leq C_2(a^n + 1) \Leftrightarrow 1 + \frac{1}{b^n} \leq C_2 \left(\frac{a}{b}\right)^n + \frac{C_2}{b^n}.$$
Or, le membre de droite tend vers 1 et le membre de gauche vers 0. On obtient en passant à la limite $1 \leq 0$, ce qui est absurde. L'hypothèse de départ est donc fausse, et N_a et N_b ne sont pas équivalentes.
3. Supposons par exemple $a \leq b$. Alors
$$P(b) - P(a) = \int_a^b P'(t)dt \leq \int_0^1 |P'(t)|dt.$$
Ainsi,
$$|P(b)| \leq |P(a)| + \int_0^1 |P'(t)|dt \leq N_a(P).$$
Il vient
$$N_b(P) \leq N_a(P) + \int_0^1 |P'(t)|dt \leq 2N_a(P).$$
On a de la même façon
$$|P(a)| \leq |P(b)| + \int_0^1 |P'(t)|dt \leq N_b(P)$$
et donc
$$N_a(P) \leq 2N_b(P).$$
Les deux normes sont bien équivalentes.

Partie D

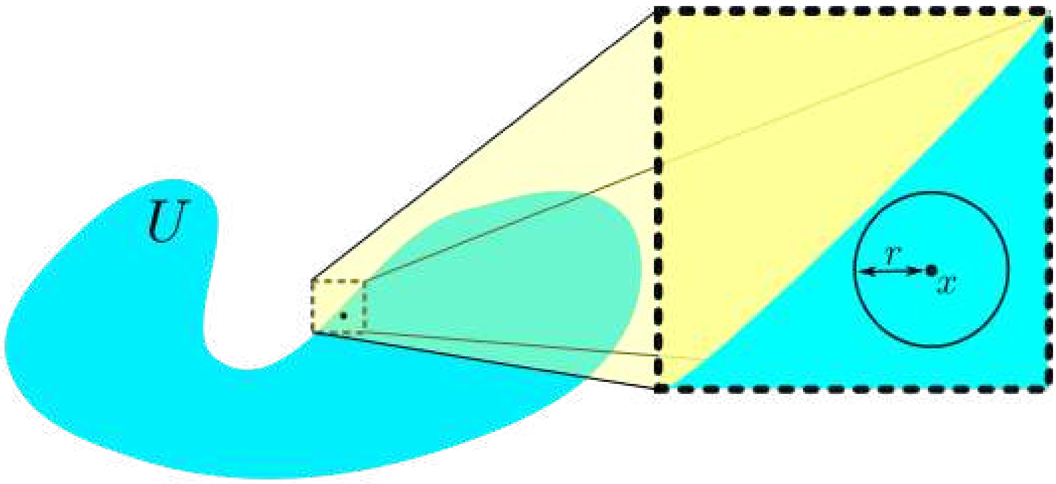
Topologie d'un espace vectoriel normé

Dans cette partie, $(E, \|\cdot\|)$ désigne un espace vectoriel normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

1. Ouverts

Définition 24. *Partie ouverte*

Soit U une partie de E. On dit que U est un **ouvert** ou une **partie ouverte** de $(E, \|\cdot\|)$ si pour tout $x \in U$, il existe $r > 0$ tel que $B_f(x, r) \subset U$.



Remarque 18.

On peut remplacer la boule fermée $B_f(x, r)$ par la boule ouverte $B(x, r)$ dans la définition.

Exercice 31.

1. Montrer l'équivalence entre la définition utilisant des boules fermées et celle utilisant des boules ouvertes.
2. Écrire la définition de partie ouverte avec des quantificateurs et en traduisant l'inclusion de $B_f(x, r)$ dans U.

Correction.

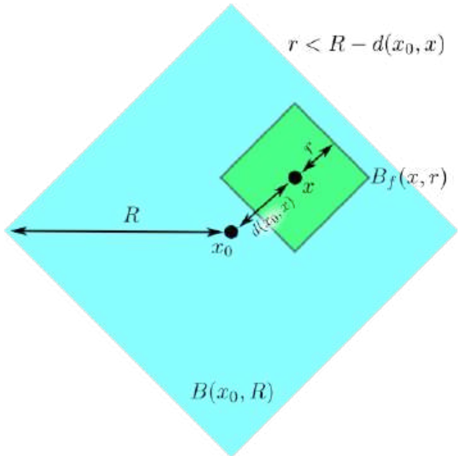
1. Il suffit de remarquer que pour tout $r > 0$, $B(x,r) \subset B_f(x,r)$ et $B_f(x, \frac{r}{2}) \subset B(x,r)$.
2. U est ouvert si, et seulement si :

$$\forall x \in U,\ \exists r > 0,\ \forall y \in E,\ \|x - y\| \leq r \Rightarrow y \in U.$$

On justifie ici la terminologie de boule *ouverte* employée dans ce chapitre.

Proposition 23.

Une boule ouverte est un ouvert de E.



Démonstration.

On considère une boule ouverte $B := B(x_0, R)$ de centre $x_0 \in E$ et de rayon $R > 0$. Soit $x \in B$. Alors $r = \frac{1}{2}(R - d(x_0, x)) > 0$ et on a, pour tout $y \in B_f(x,r)$,

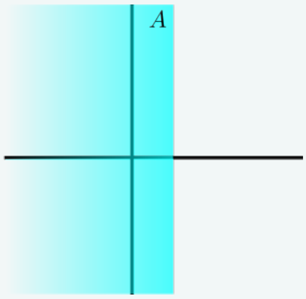
$$d(y, x_0) \leq d(y, x) + d(x, x_0) \leq r + d(x_0, x) < R.$$

Par suite $B_f(x,r) \subset B$. Donc B est un ouvert de E. $\square$

Exemple 7.

— L'espace E et l'ensemble vide $\emptyset$ sont des ouverts de E.

— Le sous-ensemble $A = \{(a,b) \mid a < \frac{1}{3}\}$ de $\mathbb{R}^2$ muni de la norme deux est un ouvert.



Exercice 32.

1. Montrer l'affirmation de l'exemple précédent.
2. Soit $a, b \in \mathbb{R}$. Que dire de l'intervalle $]a, b[$ dans $\mathbb{R}$ muni de la valeur absolue ? dans $\mathbb{C}$ muni du module ?
3. Montrer que dans $C([0,1], \mathbb{R})$ muni de la norme infinie (i.e. la norme de la convergence uniforme), le sous-ensemble des fonctions strictement positives est un ouvert de E.
4. Montrer que dans $\mathcal{F}_b(\mathbb{R}, \mathbb{R})$ muni de la norme infinie, le sous-ensemble des fonctions strictement positives n'est pas ouvert.

Correction.

1. Soit $M = (x, y) \in A$. Alors $x < y$. On note $r = \frac{1}{2}|\frac{1}{3} - x|$.
   Soit $M' = (u, v) \in B_f(M, r)$. Alors on a :
   $$r \geq d(M, M') = \sqrt{(x-u)^2 + (y-v)^2} \geq |x - u| \geq |u| - |x|.$$
   Or $r = \frac{1}{2}|\frac{1}{3} - x| \leq \frac{1}{2}(|\frac{1}{3}| + |x|) \leq \frac{1}{2}(|\frac{1}{3}|) + |x|$,
   par suite,
   $$|u| \leq \frac{1}{2}(|\frac{1}{3}|) \leq \frac{1}{3}.$$
   Donc $B_f(M, r) \subset A$. D'où A est ouvert.
2. On note U le sous-ensemble en question. Soit $f \in U$. f est continue sur $[0,1]$ donc elle est bornée et atteint ses bornes. Notons m son minimum sur $[0,1]$. Alors $B_f(f, \frac{m}{2}) \subset U$. En effet, si $g \in B_f(f, \frac{m}{2})$, alors, pour tout $x \in [0,1]$,
   $$|g(x) - f(x)| \leq \|f - g\|_\infty \leq \frac{m}{2}.$$
   Par suite, $f(x) - g(x) \leq \frac{m}{2}$ et donc :
   $$g(x) \geq f(x) - \frac{m}{2} > 0.$$
   Donc g est strictement positive.
3. Notons X le sous-ensemble en question et considérons $f : t \mapsto e^{-t^2}$. Comme f tend vers 0 en ∞, pour tout $r > 0$, on pourra trouver une fonction dans $B_f(f, r)$ dont le graphe passe en dessous de l'axe des abscisses pour t assez grand ; X ne peut donc pas être ouvert.

Plus précisément, étant donné $r > 0$, exhibons une fonction $g \in B_f(f, r)$ qui n'est pas dans X.

On note $M = \begin{cases} \sqrt{-\ln(r)} & \text{si } r \leq 1 \\ 0 & \text{si } r > 1 \end{cases}$. Pour $|t| \geq M$, $e^{-t^2} - r \leq 0$. Donc la fonction $g : \mathbb{R} \to \mathbb{R}$, définie, pour $t \in \mathbb{R}$ par $g(t) = f(t) - r$ appartient à la boule $B_f(f, r)$ (car $\|f - g\|_\infty = r$) et n'est pas strictement positive, et donc n'appartient pas à X.

Proposition 24.

— La réunion d'une famille **quelconque** d'ouverts de E est un ouvert de E.
— L'intersection d'une famille **finie** d'ouverts est un ouvert.

Démonstration.

— Soit $(U_i)_{i \in I}$ une famille quelconque d'ouverts de E. Notons $U = \bigcup_{i \in I} U_i$. Pour tout $x \in U$, il existe $i \in I$ tel que $x \in U_i$. Par suite, comme U_i est ouvert, il existe $r > 0$ tel que $B_f(x, r) \subset U_i \subset U$. Par suite, U est ouvert.
— Soit $U_1, \ldots, U_n$ des ouverts de E. On note $U = \bigcap_{i=1}^n U_i$. Soit $x \in U$. Alors, pour tout $1 \leq i \leq n$, $x \in U_i$ ouvert, donc il existe $r_i > 0$ tel que $B_f(x, r_i) \subset U_i$.
Soit $r = \min(r_1, \ldots, r_n)$. Alors $r > 0$ et pour tout $i \in \{1, \ldots, n\}$,

$$B_f(x, r) \subset B_f(x, r_i) \subset U_i.$$

Il en résulte que $B_f(x, r) \subset U$. Donc U est ouvert.

□

Exercice 33.

Montrer qu'en général, une intersection quelconque d'ouverts n'est pas un ouvert.

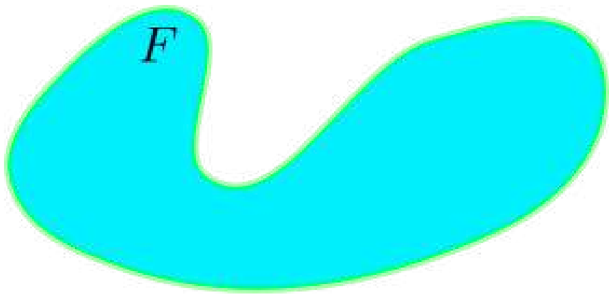
Correction.

L'intersection de la famille $\left(\left] -\frac{1}{n}, \frac{1}{n} \right[\right)_{n \in \mathbb{N}^*}$ est égale au singleton $\{0\}$ (à démontrer correctement) qui n'est pas un ouvert de $\mathbb{R}$.

2. Fermés

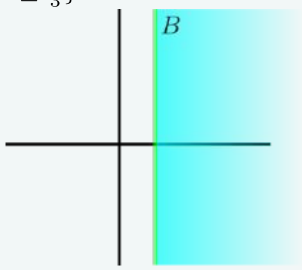
Définition 25. *Fermé*

Soit F une partie de E. On dit que F est un **fermé** ou une **partie fermée** de $(E, \|\cdot\|)$ si son complémentaire F^c est un ouvert de E.



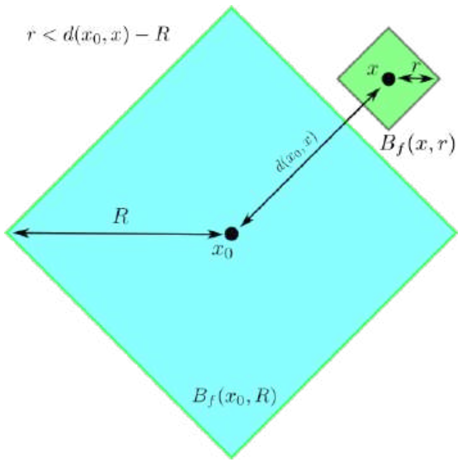
Exemple 8.

- L'espace E et l'ensemble vide $\emptyset$ sont des fermés de E.
- Le sous-ensemble $B = \{(a, b) \mid a \geq \frac{1}{3}\}$ de $\mathbb{R}^2$ muni de la norme deux est un fermé.



Proposition 25.

Une boule fermée est un fermé de E.



Démonstration.

On considère une boule fermée $F := B_f(x_0, R)$ de centre $x_0 \in E$ et de rayon $R > 0$. Soit $x \notin F$. Alors $r = \frac{1}{2}(d(x_0, x) - R) > 0$ et on a, pour tout $y \in B_f(x, r)$,

$$d(x, x_0) \leq d(x, y) + d(y, x_0) \leq r + d(y, x_0) < d(x_0, x) - R + d(y, x_0).$$

Alors $d(y, x_0) > R$. Par suite $y \notin F$. Donc F^c est un ouvert de E.
Il en résulte que F est un fermé de E. □

Proposition 26.

— L'intersection d'une famille **quelconque** de fermés de E est un fermé de E.
— La réunion d'une famille **finie** de fermés de E est un fermé de E.

Démonstration.

On utilise les propriétés du passage au complémentaire :

— On remarque que pour des ensembles A et B,

$$(A \cap B)^c = A^c \cup B^c.$$

Soit $(F_i)_{i \in I}$ une famille quelconque de fermés et $F = \bigcap_{i \in I} F_i$ l'intersection de cette famille. Alors on a :

$$F^c = \left(\bigcap_{i \in I} F_i\right)^c = \bigcup_{i \in I} F_i^c.$$

Or chaque F_i^c est ouvert car F_i est fermé, donc d'après la proposition 24, F^c est ouvert comme réunion quelconque d'ouverts.
Par suite, F est fermé.

— On raisonne de la même façon en remarquant que pour des ensembles A et B,

$$(A \cup B)^c = A^c \cap B^c.$$

□

Remarque 19.

Comme pour les intersections d'ouverts, une réunion quelconque de fermés n'est pas fermée en général : on peut considérer par exemple la famille de fermés $\left(\left[\frac{1}{n}, 1 - \frac{1}{n}\right]\right)_{n \in \mathbb{N}^*}$.

Exercice 34.

Montrer que les ensembles suivants sont des fermés :

1. Un singleton.
2. Une sphère.
3. $\mathbb{N}$ (dans $\mathbb{R}$).
4. $\mathbb{R}$ (dans $\mathbb{C}$).

Correction.

— On considère le singleton $\{x_0\}$ dans E. Soit $x \in E \setminus \{x_0\}$, et $r = d(x, x_0) > 0$. Alors pour tout $y \in B_f(x, \frac{r}{2})$, $d(x_0, y) \geq d(x_0, x) - d(x, y) \geq \frac{r}{2} > 0$. Donc $y \in E \setminus \{x_0\}$. Par suite $\{x_0\}$ est fermé.

— Soit $S(x_0, r)$ une sphère de E. Alors on a :

$$S(x_0, r) = B_f(x_0, r) \setminus B(x_0, r) = B_f(x_0, r) \cap (B(x_0, r))^c .$$

Par suite, $S(x_0, r)$ est fermé comme intersection de fermés.

— Soit $x \notin \mathbb{N}$. On note $r = \min(x - E(x), E(x) + 1 - x) > 0$. Alors $B_f(x, \frac{r}{2}) \subset \mathbb{R}$. Donc $\mathbb{N}^c$ est un ouvert de $\mathbb{R}$. Par suite, $\mathbb{N}$ est un fermé de $\mathbb{R}$.

— Soit $z \notin \mathbb{R}$. On note $r = |\mathrm{Im}(z)| > 0$. Alors $B_f(z, \frac{r}{2}) \subset \mathbb{C}$. Donc $\mathbb{R}^c$ est un ouvert de $\mathbb{C}$. Par suite, $\mathbb{R}$ est un fermé de $\mathbb{C}$.

Exercice 35.

Soit F une partie fermée de E et $x \in E$. Montrer que si $x \notin F$, alors $d(x, F) > 0$.

Correction.

L'élément x appartient à F^c qui est ouvert, donc il existe $r > 0$ tel que $B_f(x, r) \subset F^c$. Par suite, pour tout $y \in F$, $y \notin B_f(x, r)$ d'où $d(x, y) \geq r$. Par suite,

$$d(x, F) \geq r > 0.$$

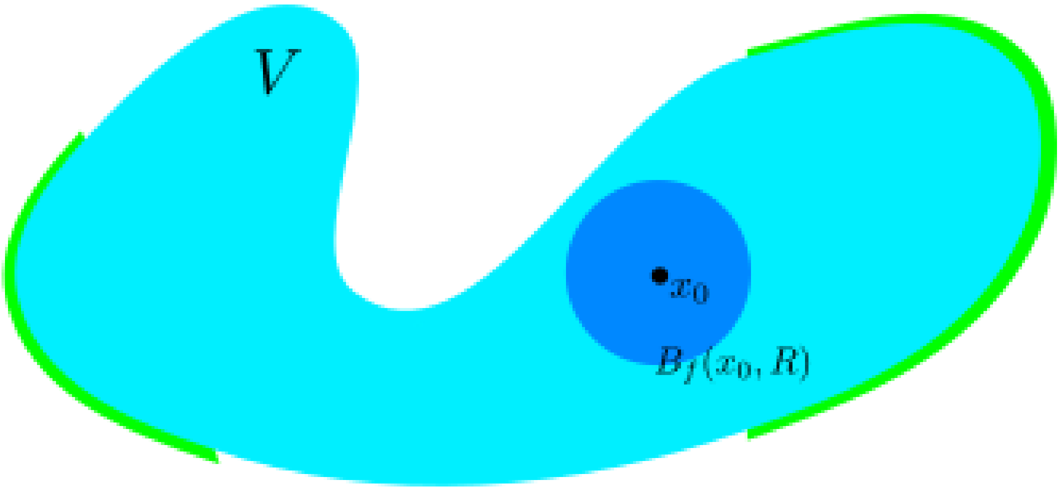
3. Voisinages

Définition 26. *Voisinage*

Soit $x_0 \in E$ et V une partie de E. On dit que V est un **voisinage de** x_0 s'il existe $r > 0$ tel que

$$B(x_0, r) \subset V.$$

On note $\mathcal{V}(x_0)$ l'ensemble des voisinages de x_0.



Remarque 20.

On peut remplacer la boule ouverte $B(x_0, r)$ par la boule fermée $B_f(x_0, r)$ dans la définition.

Exercice 36.

1. Montrer l'équivalence entre la définition utilisant des boules ouvertes et celle utilisant des boules fermées.
2. Écrire la définition de voisinage de x_0 avec des quantificateurs et en traduisant l'inclusion de $B(x_0, r)$ dans V.

Correction.

1. Il suffit de remarquer que pour tout $r > 0$, $B(x_0, r) \subset B_f(x_0, r)$ et $B_f(x_0, \frac{r}{2}) \subset B(x_0, r)$.
2. V est un voisinage de x_0 si, et seulement si :

$$\exists r > 0,\ \forall x \in E,\ \|x - x_0\| \leq r \Rightarrow x \in V.$$

Proposition 27.

Soit U une partie de E. Alors U est ouvert si, et seulement si, U est un voisinage de chacun de ses points.

Démonstration.

- ($\Rightarrow$). On suppose U ouvert. Soit $x \in U$. Comme U est ouvert, il existe $r > 0$ tel que $B_f(x, r) \subset U$. Donc a fortiori, $B(x, r) \subset U$. Donc U est un voisinage de x.
- ($\Leftarrow$). On suppose que U est un voisinage de chacun de ses points. Soit $x \in U$. Comme U

est un voisinage de x, alors il existe $r > 0$ tel que $B(x, r) \subset U$. Par suite, $B_f(x, \frac{r}{2}) \subset U$. Donc U est ouvert.

□

Exercice 37.

Soit $V \subset E$ et $x_0 \in E$. Montrer que V est un voisinage de x_0 si, et seulement si, il existe U un ouvert tel que $x_0 \in U$ et $U \subset V$.

Correction.

- ($\Rightarrow$). On suppose $V \in \mathcal{V}(x_0)$. Alors il existe $r > 0$ tel que l'ouvert $B(x_0, r)$ est inclus dans V.
- ($\Leftarrow$). On suppose qu'il $U \subset V$ avec $x_0 \in U$. Alors il existe $r > 0$ tel que :

$$B(x_0, r) \subset B_f(x_0, r) \subset U \subset V,$$

donc V est un voisinage de x_0.

Proposition 28.

Soit $x_0 \in E$.
- L'intersection d'une famille **finie** de voisinages de x_0 est un voisinage de x_0.
- Tout ensemble contenant un voisinage de x_0 est un voisinage de x_0.
- La réunion d'une famille **quelconque** de voisinages de x_0 est un voisinage de x_0.

Exercice 38.

Montrer, en s'inspirant de la proposition 24, la première assertion de la proposition précédente.

Démonstration.

- Arguments similaires à ceux de la démonstration de la proposition 24.
- Immédiat
- On utilise le point précédent : une réunion de voisinages de x_0 contient chacun des voisinages en question, c'est donc un voisinage de x_0.

□

4. Topologie d'un espace produit

Proposition 29.

Soit $k \in \mathbb{N}^*$, $(E_1, N_1), ..., (E_k, N_k)$ des espaces vectoriels normés. On considère $E = \prod_{i=1}^k E_i$ muni de la norme produit $\|\cdot\|$. Soit $U_1 \subset E_1$,...,$U_k \subset E_k$.
Si, pour tout $1 \leq i \leq k$, U_i est un ouvert de E_i, alors $U_1 \times ... \times U_k$ est un ouvert de E.

Démonstration.

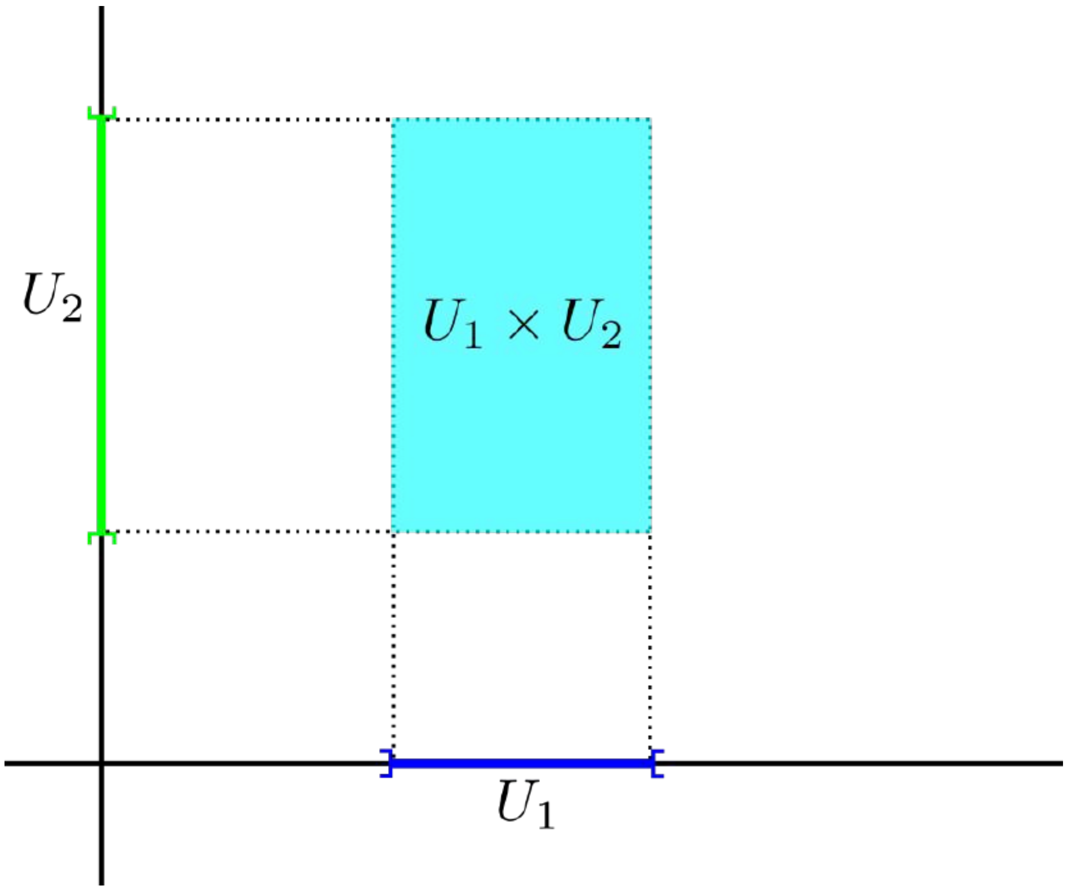
On suppose que pour tout $1 \leq i \leq k$, U_i est un ouvert de E_i. Soit $x = (x_1, ..., x_k) \in E$. Alors pour tout $1 \leq i \leq k$, il existe $r_i > 0$ tel que $B_f^{N_i}(x_i, r_i) \subset U_i$ car U_i est ouvert.

On note $r = \min(r_1, ..., r_k)$. Soit $y = (y_1, ..., y_k) \in B_f^{\|\cdot\|}(x, r)$. Alors on a :

$$r \geq \|y - x\| = \max_{1 \leq i \leq k}(N_i(y_i - x_i)).$$

Par suite, pour tout $1 \leq i \leq k$, $N_i(y_i - x_i) \leq r \leq r_i$ et donc $y_i \in B_f^{N_i}(x_i, r_i) \subset U_i$.

Il en résulte que $y \in U_1 \times ... \times U_k$. D'où $B_f^{\|\cdot\|}(x, r) \subset \prod_{i=1}^k U_i$.
Donc $U_1 \times ... \times U_k$ est un ouvert de E muni de la norme produit. □



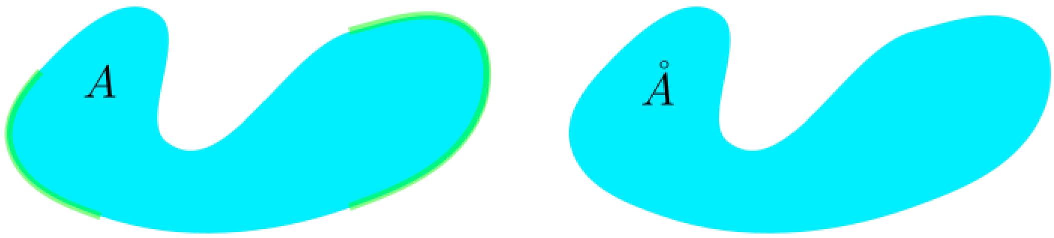
5. Intérieur

Définition 27. *Point intérieur et intérieur d'une partie*

Soit $A \subset E$.

Soit $x \in E$. On dit que x est un **point intérieur** à A si A est un voisinage de x i.e. s'il existe $r > 0$ telle que $B(x, r) \subset A$.

On appelle **intérieur** de A l'ensemble noté $\mathring{A}$ des points intérieurs à A.



Remarque 21.

- Si x est un point intérieur à A, alors $x \in A$ d'après la définition. Donc $\mathring{A} \subset A$.
- Si A, B sont des parties de E avec $A \subset B$, alors $\mathring{A} \subset \mathring{B}$. La réciproque est fausse.

Exemple 9.

- Soit $a, b \in \mathbb{R}$. Alors $\overset{\circ}{\widehat{[a, b]}} =]a, b[$.
- Soit $A = \{(a, b) \mid a \geq \frac{1}{3}\} \subset \mathbb{R}^2$. Alors $\mathring{A} = \{(a, b) \mid a > \frac{1}{3}\}$.

Proposition 30.

Soit $A \subset E$.

- $\mathring{A}$ est le plus grand ouvert inclus dans A.
- A est ouvert si, et seulement si, $A = \mathring{A}$.

Démonstration.

- Soit U un ouvert inclus dans A. Alors tout point de U est intérieur à A car pour tout $x \in U$, il existe $r > 0$ tel que $x \in B(x, r) \subset B_f(x, r) \subset U \subset A$. Donc $U \subset \mathring{A}$. Il reste à prouver que $\mathring{A}$ est ouvert.
  Soit $x \in \mathring{A}$. Alors il existe $r > 0$ tel que $B(x, r) \subset A$. Or $B(x, r)$ est un ouvert, donc, d'après ce qui précède, $B(x, r) \subset \mathring{A}$; par suite $\mathring{A}$ est ouvert.
  Il en résulte que $\mathring{A}$ est le plus grand ouvert contenu dans A.
- Si A est ouvert, alors $A \subset \mathring{A}$ d'après ce qui précède. Réciproquement, toujours d'après ce qui précède, $\mathring{A}$ est ouvert, donc A est ouvert.

□

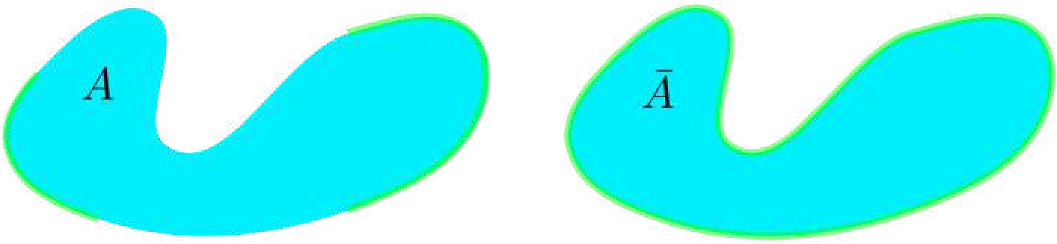
6. Adhérence

Définition 28. *Point adhérent et adhérence d'une partie*

Soit $A \subset E$.
Soit $x \in E$. On dit que x est un **point adhérent** à A si, pour tout $r > 0$,

$$B(x, r) \cap A \neq \emptyset.$$

On appelle **adhérence** de A l'ensemble noté $\overline{A}$ des points adhérents à A.



Remarque 22.

- Pour tout $A \subset E$, $A \subset \overline{A}$.
- Si A, B sont des parties de E avec $A \subset B$, alors $\overline{A} \subset \overline{B}$. La réciproque est fausse.

Exemple 10.

- Soit $a, b \in \mathbb{R}$. Alors $\overline{]a, b[} = [a, b]$.
- Soit $A = \{(a, b) \mid a > \frac{1}{3}\} \subset \mathbb{R}^2$. Alors $\overline{A} = \{(a, b) \mid a \geq \frac{1}{3}\}$.

Proposition 31.

Soit $A \subset E$. Alors :

$$\left(\overline{A}\right)^c = \overset{\circ}{\widehat{A^c}}, \quad \overline{A} = \left(\overset{\circ}{\widehat{A^c}}\right)^c \quad \text{et} \quad \mathring{A} = \left(\overline{(A^c)}\right)^c$$

Démonstration.

Les deux dernières relations découlent clairement de la première. Prouvons $\left(\overline{A}\right)^c = \overset{\circ}{\widehat{A^c}}$.

Si $x \notin \overline{A}$, alors il existe $r > 0$ tel que $B(x, r) \cap A = \emptyset$ donc $B(x, r) \subset A^c$. Par suite $x \in \overset{\circ}{\widehat{A^c}}$. Donc $\left(\overline{A}\right)^c \subset \overset{\circ}{\widehat{A^c}}$.
Réciproquement, les mêmes arguments employés dans l'autre "sens" donnent l'autre inclusion. $\square$

Exercice 39.

Soit $A \subset E$ une partie non vide et $x \in E$. Montrer que :

$$x \in \overline{A} \text{ si, et seulement si, } d(x, A) = 0.$$

Correction.

- $(\Rightarrow)$. On suppose $x \in \overline{A}$.
  Alors pour tout $r > 0$, $B(x, r) \cap A \neq \emptyset$. Par suite, pour tout $r > 0$, il existe $a \in A$ tel que $d(x, A) \leq d(x, a) < r$. Donc $d(x, A) = 0$.
- $(\Leftarrow)$. On suppose $d(x, A) = 0$.
  Alors pour tout $r > 0$, il existe $a \in A$ tel que $d(x, a) < r$. Donc $a \in B(x, r) \cap A$. D'où $B(x, r) \cap A$ non vide. Par suite $x \in \overline{A}$.

Proposition 32.

Soit $A \subset E$.
- $\overline{A}$ est le plus petit fermé contenant A.
- A est fermé si, et seulement si, $A = \overline{A}$.

Démonstration.

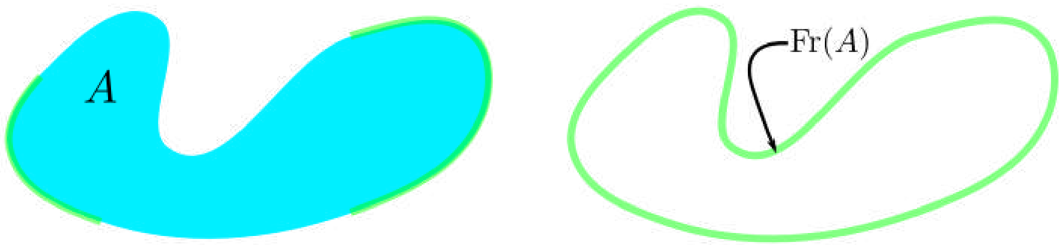
On utilise les résultats de la proposition 31 pour se ramener au cas de l'intérieur de A et on applique la proposition 30. □

7. Frontière

Définition 29. *Frontière d'une partie*

Soit $A \subset E$.
On appelle **frontière** de A l'ensemble noté $\mathrm{Fr}(A)$ (ou ∂A) défini par :

$$\mathrm{Fr}(A) = \overline{A} \setminus \mathring{A}.$$



Exemple 11.

- La frontière de $[a, b[$ est la paire $\{a, b\}$.
- La frontière de $A = \{(a, b) \mid a < \frac{1}{3}\}$ dans $\mathbb{R}^2$ muni de la norme euclidienne, est l'ensemble $\mathrm{Fr}(A) = \{(a, b) \mid a = \frac{1}{3}\}$.

Exercice 40.

Soit $A \subset E$.

- Montrer que $\mathrm{Fr}(A) = \overline{A} \cap \overline{A^c}$.
- Montrer que $\mathring{A} \cup \mathrm{Fr}(A) \cup (\overline{A})^c = E$ et que les intersections deux à deux de ces trois ensembles sont vides. (On dira que ces trois ensembles forment une partition de E).

Correction.

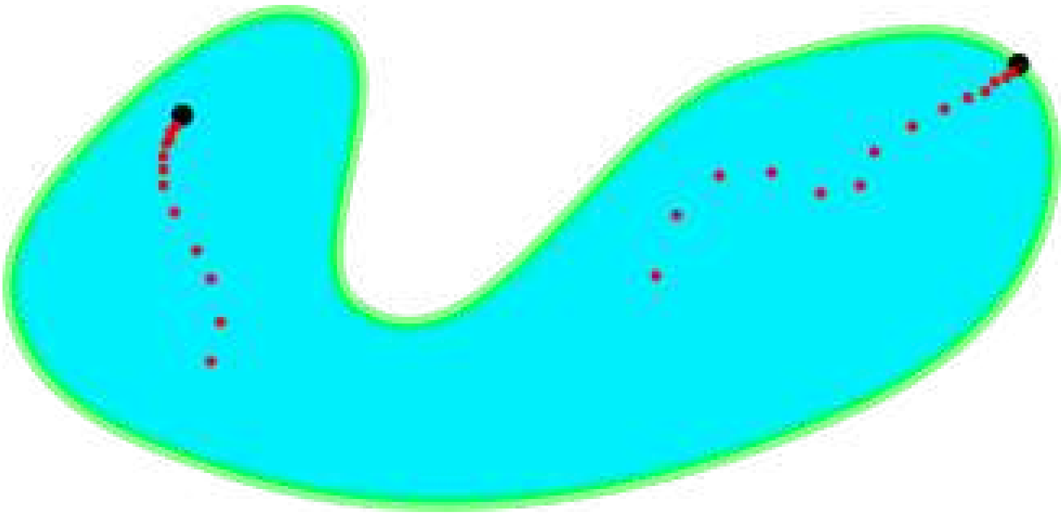
- On a $\overline{A^c} = \left(\mathring{A}\right)^c$. Or
$$\mathrm{Fr}(A) = \overline{A} \cap (\mathring{A})^c,$$
d'où le résultat.
- Comme $\mathring{A} \subset \overline{A}$, la relation $\mathrm{Fr}(A) = \overline{A} \setminus \mathring{A}$ assure que $\mathring{A}$ et $\mathrm{Fr}(A)$ forment une partition de $\overline{A}$. Le résultat découle immédiatement de cette remarque.

8. Caractérisation séquentielle des fermés

a. Caractérisation séquentielle des points adhérents.

Proposition 33. *Caractérisation séquentielle des points adhérents*

Soit $A \subset E$ et $x \in E$.
L'élément x appartient à $\overline{A}$ si, et seulement si, il existe une suite à valeurs dans A qui converge vers x.



Démonstration.

- $(\Rightarrow)$. On suppose $x \in \overline{A}$. Alors, pour tout $n \in \mathbb{N}$, il existe

$$a_n \in B(x, \frac{1}{2^n}) \cap A.$$

On a donc construit une suite (a_n) à valeurs dans A qui converge vers x ; en effet :

$$\|a_n - x\| \leq \frac{1}{2^n} \xrightarrow[n\to\infty]{} 0.$$

- $(\Leftarrow)$. On suppose qu'il existe une suite (a_n) à valeurs dans A qui converge vers x. Soit $r > 0$. Montrons que $B(x,r) \cap A \neq \emptyset$. Comme (a_n) tend vers x, il existe $N \in \mathbb{N}$ tel que

$$\|a_N - x\| < r.$$

Par suite, $a_N \in B(x,r) \cap A$.

□

Exercice 41.

Montrer que l'adhérence d'un sous-espace vectoriel est aussi un sous-espace vectoriel.

Correction.

Soit F un sous-espace vectoriel de E. Montrons que $\overline{F}$ est une sous-espace vectoriel de E.

— F est non vide. En effet $0_E \in F \subset \overline{F}$.

— Soit $x, y \in \overline{F}$ et $\lambda, \mu \in \mathbb{K}$. On cherche à montrer que $\lambda x + \mu y \in \overline{F}$.
Comme x, y sont des points adhérents à F, d'après la proposition 33, il existe des suites $(x_n), (y_n)$ à valeurs dans F tels que $x_n \xrightarrow[n\to\infty]{} x$ et $y_n \xrightarrow[n\to\infty]{} y$.

Par suite, on a, pour tout $n \in \mathbb{N}$, $\lambda x_n + \mu y_n \in F$, et

$$\|\lambda x_n + \mu y_n - (\lambda x + \mu y)\| = \|\lambda(x_n - x) + \mu(y_n - y)\| \leq |\lambda| \underbrace{\|x_n - x\|}_{\xrightarrow[n\to\infty]{} 0} + |\mu| \underbrace{\|y_n - y\|}_{\xrightarrow[n\to\infty]{} 0} \xrightarrow[n\to\infty]{} 0.$$

Donc la suite $(\lambda x_n + \mu y_n)$ à valeurs dans F converge vers $\lambda x + \mu y$. D'après la proposition 33, $\lambda x + \mu y \in \overline{F}$. Il en résulte que $\overline{F}$ est un sous-espace vectoriel de E.

b. Caractérisation séquentielle des fermés.

Théorème 3. *Caractérisation séquentielle des fermés*

Soit $A \subset E$
La partie A est fermée si, et seulement si, la limite de toute suite convergente à valeurs dans A appartient à A.

Démonstration.

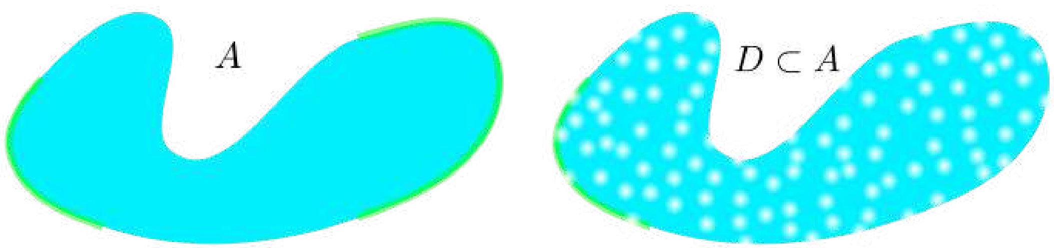
- $(\Rightarrow)$. On suppose A fermée. Soit (a_n) une suite à valeurs dans A qui converge vers $l \in E$. D'après la caractérisation des points adhérents, l appartient à $\overline{A}$. Or A est fermée donc $l \in \overline{A} = A$.
- $(\Leftarrow)$. On suppose que la limite de toute suite convergente à valeurs dans A appartient à A. Montrons que $\overline{A} \subset A$.
Soit $x \in \overline{A}$. D'après la caractérisation séquentielle des points adhérents, il existe une suite à valeurs dans A qui converge vers x. Par hypothèse $x = \lim a_n \in A$. Par suite, $\overline{A} \subset A$. L'inclusion réciproque étant triviale, on a $\overline{A} = A$. Par suite A est fermée.

□

9. Densité

Définition 30. *Densité*

Soit A une partie de E. On dit qu'une partie D de A est **dense** dans A si $A \subset \overline{D}$.



Proposition 34.

Soit $A \subset E$ et $D \subset A$. On a équivalence entre :

i) D est dense dans A ;

ii) pour tout $a \in A$ et pour tout $r > 0$, il existe $x \in D$ tel que $\|x - a\| \leq r$;

iii) pour tout $a \in A$, il existe une suite d'éléments de D qui converge vers a.

Démonstration.

les propriétés ii) et iii) sont simplement des reformulations de la densité de D dans A. $\square$

Exemple 12.

— L'ensemble $\mathbb{R}^*$ est dense dans $\mathbb{R}$.
— L'intervalle $]a, b[$ est dense dans $[a, b]$.

Exercice 42.

Montrer que :

1. Les ensembles $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans $\mathbb{R}$.
2. Soit $n \in \mathbb{N}$. L'ensemble $GL_n(\mathbb{K})$ des matrices inversibles est dense dans $\mathcal{M}_n(\mathbb{K})$ pour la norme infinie. (On verra dans la suite que ceci est vrai pour toutes les normes sur $\mathcal{M}_n(\mathbb{K})$).

   Indication : On admettra que, pour $M \in \mathcal{M}_n(\mathbb{R})$, l'application $\chi_M : t \mapsto \det(M - tI_n)$ est une fonction polynomiale de degré exactement n. (Nous montrerons ce résultat au chapitre IV sur la réduction de matrices)

Correction.

— Premièrement, montrons la densité de $\mathbb{Q}$ dans $\mathbb{R}$. Soit $x \in \mathbb{R}$. Exhibons une suite $(u_n)_{n\in\mathbb{N}^*}$ de rationnels qui converge vers x.
Voici un exemple classique d'une telle suite : pour $n \in \mathbb{N}^*$, on pose

$$u_n = \frac{2}{n^2} \sum_{i=1}^{n} E(ix).$$

Par les propriétés de la partie entière, on a, pour $i = 1, ..., n$:

$$ix - 1 < E(ix) \leq ix,$$

d'où en sommant ces inégalités de $i = 1$ à n, puis en multipliant par $\frac{2}{n^2}$, on obtient :

$$\frac{n+1}{n}x - \frac{1}{n} < u_n \leq \frac{n+1}{n}x.$$

(On rappelle que $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$).
Or $\frac{n+1}{n}x - \frac{1}{n} \xrightarrow[n\to\infty]{} x$ et $\frac{n+1}{n}x \xrightarrow[n\to\infty]{} x$; donc d'après le théorème des gendarmes, (u_n) converge vers x.

De plus, il est clair que pour tout $n \in \mathbb{N}^*$, $u_n \in \mathbb{Q}$. Il en résulte que $\mathbb{Q}$ est dense dans $\mathbb{R}$.

Nous allons utiliser le résultat précédent pour montrer la densité de $\mathbb{R} \setminus \mathbb{Q}$ dans $\mathbb{R}$.
Soit $x \in \mathbb{R}$. Montrons qu'il existe une suite d'irrationnels qui converge vers x.
Première remarque : si $y \in \mathbb{R}$ est irrationnel, pour tous rationnels $r, s \neq 0$, $r + sy$ aussi. Prenons un irrationnel au hasard, disons $\sqrt{2}$ (exercice : montrer que $\sqrt{2}$ est irrationnel !). Alors d'après la densité de $\mathbb{Q}$ dans $\mathbb{R}$, il existe une suite (u_n) de rationnels qui converge vers x. Posons, pour $n \in \mathbb{N}^*$,

$$v_n = u_n + \frac{\sqrt{2}}{n}.$$

Alors d'après la remarque initial, la suite (v_n) est à valeurs dans $\mathbb{R} \setminus \mathbb{Q}$, et on a :

$$\lim_{n \to \infty} v_n = x.$$

Par suite, $\mathbb{R} \setminus \mathbb{Q}$ est dense dans $\mathbb{R}$.

— Soit $n \in \mathbb{N}^*$. Soit $M \in \mathcal{M}_n(\mathbb{K})$. Montrons qu'il existe une suite à valeurs dans $GL_n(\mathbb{K})$ qui converge vers M.
On définit la suite $(M_k)_{k \in \mathbb{N}}$ à valeurs dans $\mathcal{M}_n(\mathbb{R})$ telle que, pour $k \in \mathbb{N}$:

$$M_k = M - \frac{1}{2^k} I_n.$$

Comme l'application χ_M est une fonction polynomiale de degré n, alors elle s'annule au plus n fois sur $\mathbb{R}$. Par suite, à partir d'un certain rang $N \in \mathbb{N}$, on a, pour tout $k \geq N$, $\det(M_k) = \chi_M(2^{-k}) \neq 0$.
On pose alors, pour $k \in \mathbb{N}$,

$$M'_k = M_{k+N}$$

Alors $(M'_k)_{k \in \mathbb{N}}$ est une suite à valeurs dans $GL_n(\mathbb{K})$ et on a, pour tout $k \in \mathbb{N}$:

$$\|M - M'_k\|_\infty = 2^{-k-N} \|I_n\|_\infty = 2^{-k-N} \xrightarrow[k \to \infty]{} 0.$$

Par suite, (M'_k) converge vers M. Il en résulte que $GL_n(\mathbb{K})$ est dense dans $\mathcal{M}_n(\mathbb{K})$.

10. Ouverts, fermés et voisinages relatifs

Définition 31. *Ouvert relatif, fermé relatif et voisinage relatif*

Soit $A \subset E$.

— Soit $U \subset E$. On dit que U est un **ouvert relatif** de A s'il existe un ouvert U' de E tel que $U = U' \cap A$.

— Soit $F \subset E$. On dit que F est un **fermé relatif** de A s'il existe un fermé F' de E tel que $F = F' \cap A$.

— Soit $x_0 \in E$ et $V \subset E$. On dit que V est un **voisinage de x_0 relatif** de A s'il existe un voisinage V' de x_0 dans E tel que $V = V' \cap A$.

Remarque 23.

Si $A = E$, les notions *relatives* coïncident avec les définitions déjà vues d'ouverts, de fermés et de voisinages.

Mais si A est strictement inclus dans E, les ouverts, fermés et voisinages relatifs ne sont, en général, pas des ouverts, des fermés et des voisinages de E respectivement.

Exercice 43.

1. — Dans $\mathbb{R}$, montrer que $]0,1]$ est un ouvert relatif de $[-1,1]$. Que remarque-t-on ?
   — Donner un exemple de fermé relatif qui n'est pas fermé dans $\mathbb{R}$ et un exemple de voisinage relatif qui n'est pas un voisinage dans $\mathbb{C}$.
2. Soit $A \subset E$. Montrer que A est un ouvert relatif de A et un fermé relatif de A.

Correction.

1. — On a $]0,1] =]0,+\infty[\cap [-1,1]$. Donc $]0,1]$ est un ouvert relatif de $[-1,1]$. On a donc un exemple d'ouvert relatif qui n'est pas un ouvert de l'espace sous-jacent.
   — On peut prendre par exemple, dans $\mathbb{R}$, $[0,1[$ qui n'est pas fermé mais qui est un fermé relatif de $]-1,1[$.
   Pour le deuxième exemple, dans $\mathbb{C}$, on peut considérer $[-1,1]$ qui n'est pas un voisinage de 0 dans $\mathbb{C}$ mais qui est un voisinage de 0 relatif de $\mathbb{R}$ (en effet, $[-1,1] = \mathbb{D} \cap \mathbb{R}$ où $\mathbb{D} = \{z \in \mathbb{C} \mid |z| \leq 1\}$ qui est bien un voisinage de 0 dans $\mathbb{C}$).
2. On a $A = E \cap A$ et E est à la fois un ouvert de E et un fermé de E. D'où le résultat.

11. Topologie et comparaison de normes

a. Conservation des ouverts, fermés et voisinages

Proposition 35.

Soit N_1,N_2 des normes sur E. On suppose que N_1 est dominée par N_2. Alors un ouvert (resp. un fermé, resp. un voisinage d'un point $x_0 \in E$) pour la norme N_1 est un ouvert pour la norme N_2 (resp. un fermé, resp. un voisinage de x_0).

Démonstration.

On suppose que N_1 est dominée par N_2. Soit U un ouvert de E pour N_1. Alors pour tout $x \in U$, il existe $r > 0$ tel que $B_f^{N_1}(x,r) \subset U$. Or d'après la proposition précédente, toute boule de N_1 contient une boule de N_2 ; donc pour tout $x \in U$, il existe $r, R > 0$, tels que

$$B_f^{N_2}(x,R) \subset B_f^{N_1}(x,r) \subset U.$$

Par suite U est ouvert pour la norme N_2. En utilisant les caractérisations de fermé et de voisinage en terme d'ouvert, on obtient le résultat grâce au raisonnement précédent. $\square$

Théorème 4. *Conservation des ouverts, fermés et voisinages*

Soit N_1,N_2 des normes sur E. On suppose que N_1 et N_2 sont équivalentes.
Une partie est un ouvert (resp. un fermé, resp. un voisinage d'un point $x_0 \in E$) pour la norme N_1 si, et seulement si, elle est un ouvert pour la norme N_2 (resp. un fermé, resp. un voisinage de x_0).

Démonstration.

On applique la proposition 35 pour N_1 dominée par N_2 puis pour N_2 dominée par N_1. □

b. Exercice

Exercice 44. *Banque CCP*

On note E l'espace vectoriel des applications continues de $[0;1]$ dans $\mathbb{R}$.
On pose : $\forall f \in E$, $N_\infty(f) = \sup\limits_{x\in[0;1]} |f(x)|$ et $N_1(f) = \displaystyle\int_0^1 |f(t)|\mathrm{d}t$.

1. (a) Démontrer que N_∞ et N_1 sont deux normes sur E.
   (b) Démontrer qu'il existe $k > 0$ tel que, pour tout f de E, $N_1(f) \leq kN_\infty(f)$.
   (c) Démontrer que tout ouvert pour la norme N_1 est un ouvert pour la norme N_∞.
2. Démontrer que les normes N_1 et N_∞ ne sont pas équivalentes.

Correction.

1.(a) Voire le paragraphe "normes sur des espaces de fonctions".

(b) $k = 1$ convient car, $\forall f \in E$, $\displaystyle\int_0^1 |f(t)|\,\mathrm{d}t \leqslant \int_0^1 N_\infty(f)\mathrm{d}t = N_\infty(f)$.

(c) L'application identité de E, muni de la norme N_∞, vers E, muni de la norme N_1, est continue car linéaire et vérifiant $\forall f \in E$, $N_1(f) \leqslant kN_\infty(f)$.
L'image réciproque d'un ouvert par une application continue étant un ouvert, on en déduit que : un ouvert pour la norme N_1 est un ouvert pour la norme N_∞.
On peut aussi raisonner de façon plus élémentaire par inclusion de boules et retour à la définition d'un ouvert.

2. Pour $f_n(t) = t^n$, on a $N_1(f_n) = \dfrac{1}{n+1}$ et $N_\infty(f_n) = 1$, donc $\lim\limits_{n\to+\infty} \dfrac{N_\infty(f_n)}{N_1(f_n)} = +\infty$.
Donc ces deux normes ne sont donc pas équivalentes.

Partie E

Limites et continuité

Dans toute cette partie, $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ désignent deux espaces vectoriels normés sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On suppose que E est de dimension non nulle.
Quand on considérera une partie A incluse dans E, on supposera toujours - sans le dire - que A est non vide.

1. Limite d'une application

a. Généralités

Définition 32. *Limite*

Soit $A \subset E$, $f : A \to F$ une application, $a \in \overline{A}$ et $\ell \in F$.
On dit que f **admet pour** ℓ **limite en** a ou **tend vers** ℓ **en** a si :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in A, \quad \|x - a\|_E \leq \delta \ \Rightarrow\ \|f(x) - \ell\|_F \leq \varepsilon.$$

Proposition 36.

Soit $A \subset E$, $f : A \to F$ une application, $a \in \overline{A}$ et $\ell \in F$. Alors les assertions suivantes sont équivalentes :

i) f tend vers ℓ en a ;

ii) $\forall \varepsilon > 0,\ \exists \delta > 0,\ f(B_f(a, \delta) \cap A) \subset B_f(\ell, \varepsilon)$;

iii) pour tout voisinage V de ℓ dans F, il existe un voisinage W de a dans E tel que $f(W \cap A) \subset V$.

Démonstration.

Il s'agit simplement de reformulations de la définition. □

Proposition 37.

Soit $A \subset E$, $f : A \to F$ et $a \in \overline{A}$.
Si f admet une limite en a, alors cette limite est unique.

Démonstration.

On suppose que f tend vers l et ℓ' en a. Alors, pour tout $x \in A$, on a :

$$\|\ell - \ell'\|_F = \|\ell - f(x) + f(x) - \ell\|_F \leq \|f(x) - \ell\|_F + \|f(x) - \ell'\|_F.$$

Soit $\varepsilon > 0$. Alors il existe $\delta, \delta' > 0$ tels que pour tout $x \in A$:

$$\|x-a\|_E \leq \delta \Rightarrow \|f(x)-\ell\|_F \leq \frac{\varepsilon}{2}$$

et

$$\|x-a\|_E \leq \delta' \Rightarrow \|f(x)-\ell'\|_F \leq \frac{\varepsilon}{2}.$$

On note $\eta = \min(\delta, \delta')$. Par suite, pour tout $x \in A$ tel que $\|x-a\|_E \leq \eta$:

$$\|\ell-\ell'\|_F = \leq \underbrace{\|f(x)-\ell\|_F}_{\leq \varepsilon/2} + \underbrace{\|f(x)-\ell'\|_F}_{\leq \varepsilon/2} \leq 2\frac{\varepsilon}{2} = \varepsilon.$$

Ceci étant vrai pour tout $\varepsilon > 0$, on en déduite que $\|\ell-\ell'\|_E = 0$, et donc, d'après l'axiome de séparation de $\|\cdot\|_E$, $\ell = \ell'$. $\square$

Notation 5.

Soit $A \subset E$, $f : A \to F$, $a \in \overline{A}$ et $\ell \in F$. Si f admet ℓ pour limite en a, on note :

$$\lim_{x \to a} f(x) = \ell.$$

b. Limites et infini

Dans la définition suivante, on se placera dans le cas où $F = \mathbb{R}$ pour considérer des limites infinies :

Définition. Limite infinie

Soit $A \subset E$, $f : A \to \mathbb{R}$ et $a \in \overline{A}$.

— On dit que f tend vers $+\infty$ en a, et on note $\lim_{x \to a} f(x) = +\infty$, si :

$$\forall R \geq 0,\ \exists \delta > 0,\ \forall x \in A, \quad \|x-a\|_E \leq \delta \ \Rightarrow\ f(x) \geq R.$$

— On dit que f tend vers $-\infty$ en a, et on note $\lim_{x \to a} f(x) = -\infty$, si :

$$\forall R \geq 0,\ \exists \delta > 0,\ \forall x \in A, \quad \|x-a\|_E \leq \delta \ \Rightarrow\ f(x) \leq -R.$$

Pour cette définition, on se place dans le cas où $A = E = \mathbb{R}$ pour considérer des limites en $\pm\infty$:

Définition. Limite en $\pm\infty$

Soit $f : \mathbb{R} \to F$ et $\ell \in F$

— On dit que f tend vers ℓ en $+\infty$, et on note $\lim_{x \to +\infty} f(x) = \ell$, si :

$$\forall \varepsilon \geq 0,\ \exists R > 0,\ \forall x \in \mathbb{R}, \quad x \geq R \ \Rightarrow\ \|f(x)-\ell\|_F \leq \varepsilon.$$

— On dit que f tend vers ℓ en $-\infty$, et on note $\lim_{x \to -\infty} f(x) = \ell$, si :

$$\forall \varepsilon \geq 0,\ \exists R > 0,\ \forall x \in \mathbb{R}, \quad x \leq -R \ \Rightarrow\ \|f(x)-\ell\|_F \leq \varepsilon.$$

On considère désormais le cas où A n'est pas bornée et $\|x\| \to +\infty$:

Définition.

Soit $A \subset E$ une partie non bornée, $f : A \to F$ et $\ell \in F$.
On dit que f **tend vers** ℓ **quand** $\|x\|$ **tend vers** $+\infty$, et on note $\lim\limits_{\|x\| \to +\infty} f(x) = \ell$ si :

$$\forall \varepsilon \geq 0,\ \exists R > 0,\ \forall x \in A, \quad \|x\|_E \geq R \ \Rightarrow\ \|f(x) - \ell\|_F \leq \varepsilon.$$

c. Caractérisation séquentielle de la limite

Théorème 5. *Caractérisation séquentielle de la limite*

Soit $A \subset E$, $f : A \to F$, $a \in \overline{A}$ et $\ell \in F$.
Alors f tend vers ℓ en a si, et seulement si, pour tout suite $(u_n)_{n\in\mathbb{N}}$ à valeurs dans A telle que $u_n \xrightarrow[n\to\infty]{} a$,

$$f(u_n) \xrightarrow[n\to\infty]{} \ell.$$

Autrement dit, f tend vers ℓ en a si, et seulement si, pour tout suite $(u_n)_{n\in\mathbb{N}}$ à valeurs dans A qui converge vers a, la suite $(f(u_n))_{n\in\mathbb{N}}$ à valeurs dans F converge vers ℓ.

Démonstration.

— $(\Rightarrow)$. On suppose $\lim_{x\to a} f(x) = \ell$. Soit (u_n) une suite à valeurs dans A tel que $u_n \to a$. Soit $\varepsilon > 0$. Alors il existe $\delta > 0$ tel que pour tout $x \in A$ tel que $\|x-a\|_E \leq \delta$, $\|f(x)-\ell\|_F \leq \varepsilon$.
La suite (u_n) converge vers a donc il existe un rang $N \in \mathbb{N}$, tel que $n \geq N$,

$$\|u_n - a\| \leq \delta.$$

Par suite, $\|f(u_n) - \ell\|_F \leq \varepsilon$.
Ceci étant vrai pour tout $\varepsilon > 0$, il en résulte que la suite $(f(u_n))$ converge vers ℓ.

— $(\Leftarrow)$. On procède par contraposition. On suppose que f ne tend pas vers ℓ en a.
Exhibons une suite (x_n) à valeurs dans A qui tend vers a et telle que la suite $(f(u_n))$ ne tende pas vers ℓ.
Comme f ne tend pas vers ℓ en a, alors :

$$\exists \varepsilon > 0,\ \forall \delta > 0\ \exists x \in A ; \|x - a\|_E \leq \delta \text{ et } \|f(x) - \ell\|_F > \varepsilon.$$

Considérons ce $\varepsilon > 0$ et soit $n \in \mathbb{N}$. Alors pour $\delta = \frac{1}{2^n}$, il existe $u_n \in A$ avec $\|u_n - a\|_E \leq \delta$ et tel que $\|f(u_n) - \ell\|_F > \varepsilon$.
Ainsi la suite (u_n) est à valeurs dans A, converge vers a mais pour tout $n \in \mathbb{N}$, $\|f(u_n) - \ell\|_F > \varepsilon$, donc $(f(u_n))$ ne converge pas vers ℓ.

□

2. Propriétés des limites

a. Opérations algébriques

Proposition 38. *Espace vectoriel des fonctions convergentes*

Soit $A \subset E$ et $a \in \overline{A}$. L'ensemble des fonctions de A dans F qui admettent une limite en a est un espace vectoriel sur $\mathbb{K}$, et l'application qui à f associe sa limite en a est linéaire.

Démonstration.

On peut utiliser les résultats établis dans la partie sur les suites à valeurs dans un espace vectoriel normé en appliquant la caractérisation de la limite par les suites.

Voici une preuve directe :
On note $\mathcal{F}$ l'ensemble des fonctions de A dans F qui admettent une limite en a. Montrons que c'est un sous-espace vectoriel de l'espace vectoriel $\mathcal{F}(A, F)$ de toutes les fonctions de A dans F.

— L'application nulle $\mathbf{0} : x \mapsto 0_F$ de A dans F admet pour limite 0_F en a. Donc $\mathbf{0}$ appartient à $\mathcal{F}$.
— Soit $\lambda, \mu \in \mathbb{K}$ et $f, f' \in \mathcal{F}$. On note $\ell, \ell' \in F$ leurs limites respectives en a. Alors on a, pour tout $x \in A$:
$$\begin{aligned} \|(\lambda f + \mu f')(x)) - (\lambda\ell + \mu\ell')\|_F &= \|\lambda(f(x) - \ell) + \mu(f'(x) - \ell')\|_F \\ &\leq |\lambda| \underbrace{\|f(x) - \ell\|_F}_{\xrightarrow[x\to a]{} 0} + |\mu| \underbrace{\|f'(x) - \ell'\|_F}_{\xrightarrow[x\to a]{} 0} \xrightarrow[x\to a]{} 0. \end{aligned}$$
Donc $\lambda f + \mu f'$ admet une limite en a égale $\lambda\ell + \mu\ell' \in F$ (car F est un espace vectoriel. D'où $\lambda f + \mu f'$ appartient à $\mathcal{F}$.

Soit $L_a : \mathcal{F} \to \mathbb{K}$ l'application $L_a : f \mapsto \lim_{x\to a} f(x)$. En reprenant les arguments précédents, on obtient que pour $f, f' \in \mathcal{F}$ de limites respectives $\ell, \ell' \in F$ en a :
$$L_a(\lambda f + \mu f') = \lambda\ell + \mu\ell' = \lambda L_a(f) + \mu L_a(f').$$
Donc L_a est une application linéaire de $\mathcal{F}$ dans $\mathbb{K}$. $\square$

Proposition 39.

Soit $A \subset E$, $a \in \overline{A}$, $f : A \to F$, $\varphi : A \to \mathbb{K}$, $\ell \in F$ et $\lambda \in \mathbb{K}$.
Si $\lim_{x\to a} f(x) = \ell$ et $\lim_{x\to a} \varphi(x) = \lambda$, alors l'application $\varphi.f : A \to F$ définie par :
$$(\varphi.f) : x \mapsto \varphi(x)f(x)$$
admet pour limite $\lambda\ell$ en a.

Démonstration.

On peut utiliser un raisonnement direct ou se servir des résultats sur la convergence des suites et la caractérisations de la limite par les suites.
Utilisons les suites :
On suppose $\lim_{x\to a} f(x) = \ell \in F$ et $\lim_{x\to a} \varphi(x) = \lambda \in \mathbb{K}$. Soit (x_n) une suite à valeurs dans A

qui converge vers a. D'après le théorème 5, on a $f(x_n) \xrightarrow[n\to\infty]{} l$ et $\varphi(x_n) \xrightarrow[n\to\infty]{} \lambda$. Par suite (on reconnaît ici la preuve de l'exercice 19), pour tout $n \in \mathbb{N}$:

$$\|\varphi(x_n)f(x_n) - \lambda l\|_F \leq \underbrace{|\varphi(x_n) - \lambda|}_{\xrightarrow[n\to\infty]{}0} \underbrace{\|f(x_n)\|_F}_{\xrightarrow[n\to\infty]{}\|l\|_F} + |\lambda| \underbrace{\|f(x_n) - l\|_F}_{\xrightarrow[n\to\infty]{}0} \xrightarrow[n\to\infty]{} 0.$$

Ceci étant vrai pour toute suite (x_n) à valeurs dans A qui converge vers a, d'après le théorème 5, $\varphi.f$ admet une limite en a égale à λl. □

b. Espace produit

On considère dans ce paragraphe F_1 et F_2, deux espaces vectoriels normés sur $\mathbb{K}$ et l'espace vectoriel $F = F_1 \times F_2$ muni de la norme produit.

Proposition 40.

Soit $A \subset E$, $a \in \overline{A}$, $\ell = (\ell_1, \ell_2) \in F_1 \times F_2$ et $f = (f_1, f_2) : A \to F_1 \times F_2$.
Alors $\lim_{x\to a} f(x) = \ell$ si, et seulement si, $\lim_{x\to a} f_1(x) = \ell_1$ et $\lim_{x\to a} f_2(x) = \ell_2$.

Démonstration.

D'après la proposition 15, une suite $u = ((u_n^1), (u_n^2))$ à valeurs dans F converge vers ℓ si, et seulement si, (u_n^1) converge vers ℓ_1 et (u_n^2) converge vers ℓ_2.
Ainsi,

$$\lim_{x\to a} f(x) = \ell\,;$$

si, et seulement si,

pour toute suite (x_n) à valeurs dans A qui converge vers a, la suite $u = ((f_1(x_n)), (f_2(x_n)))$ à valeurs dans F converge vers $\ell = (\ell_1, \ell_2)$;

si, et seulement si,

pour toute suite (x_n) à valeurs dans A qui converge vers a, $(f_1(x_n))$ converge vers ℓ_1 et $(f_2(x_n))$ converge vers ℓ_2 ;

si, et seulement si,

$$\lim_{x\to a} f_1(x) = \ell_1 \text{ et } \lim_{x\to a} f_2(x) = \ell_2.$$

□

Exercice 45.

Montrer que cette proposition est vraie dans le cas où l'on considère un produit fini $F_1 \times ... \times F_n$ de n espaces vectoriels normés ($n \in \mathbb{N}^*$).

Correction.

On raisonne par récurrence en utilisant la proposition précédente.

c. Limite d'une composée

Proposition 41. *Limite d'une application composée*

Soit $(E, \|\cdot\|_E), (F, \|\cdot\|_F), (G, \|\cdot\|_G)$ des espaces vectoriels normés sur K ; $A \subset E$, $B \subset F$; $a \in \overline{A}$, $b \in \overline{B}$ et $\ell \in G$.
Soit $f : A \to F$ et $g : B \to G$. Si $f(A) \subset B$, $\lim_{x \to a} f(x) = b$ et $\lim_{y \to b} g(y) = \ell$, alors

$$\lim_{x \to a} g \circ f(x) = \ell.$$

Démonstration.

On suppose $f(A) \subset B$, $\lim_{x \to a} f(x) = b$ et $\lim_{y \to} g(y) = \ell$.
Soit (x_n) une suite à valeurs dans A qui converge vers a. Comme $\lim_{x \to a} f(x) = b$ et $\lim_{y \to b} g(y) = \ell$, d'après le théorème 5, la suite $(f(x_n)$ à valeurs dans B converge vers b et donc la suite $(g \circ f(x_n))$ à valeurs dans G converge vers ℓ.
Ceci étant vrai pour toute suite (x_n) à valeurs dans A qui converge vers a, d'après le théorème 5, $\lim_{x \to a} g \circ f(x) = \ell$. □

3. Applications continues

a. Continuité locale

Définition 33. *Application continue en un point*

Soit $A \subset E$, $a \in A$ et $f : A \to F$.
On dit que f est **continue en** a si f admet une limite en a et $\lim_{x \to a} f(x) = f(a)$.

Théorème 6. *Caractérisation séquentielle de la continuité*

Soit $A \subset E$, $a \in A$ et $f : A \to F$.
L'application f est continue en a si, et seulement si, pour toute suite (x_n) à valeurs dans A qui converge vers a,

$$f(x_n) \xrightarrow[n \to \infty]{} f(a).$$

Démonstration.

On applique le théorème 5 : $\lim_{x \to a} f(x) = f(a)$, si, et seulement si, pour toute suite (x_n) à valeurs dans A qui converge vers a, $f(x_n) \xrightarrow[n \to \infty]{} f(a)$. □

b. Continuité globale

Définition 34. *Application continue*

Soit $A \subset E$ et $f : A \to F$.
On dit que f est **continue sur** A si f est continue en tout point de A.

Exercice 46.

1. Traduire cette définition avec des quantificateurs.
2. Montrer que l'application de E dans $\mathbb{R}$ donnée par $x \mapsto \|x\|_E$ est continue sur E.

Correction.

1. f est continue sur A si, et seulement si,
$$\forall a \in A,\ \forall \varepsilon > 0,\ \exists \delta > 0\ \forall x \in A, \quad \|x-a\|_E \ \Rightarrow\ \|f(x)-f(a)\|_F \leq \varepsilon.$$
2. Soit $a \in E$. On a, pour toute suite (x_n) à valeurs dans A qui converge vers a, $\lim_{n\to\infty} \|x_n\|_E = \|a\|_E$; donc $\lim_{x\to a} \|x\|_E = \|a\|_E$. Ainsi l'application $x \mapsto \|x\|_E$ est continue en tout point a de E donc elle est continue sur E.

Notation 6.

Soit $A \subset E$. On note $C(A, F)$ l'ensemble des applications continues de A dans F.

c. Propriétés des applications continues

Les propositions suivantes découlent des propriétés des limites (paragraphe 2) établies précédemment.

Proposition 42. *Combinaisons linéaires*

Soit $A \subset E$. L'ensemble $C(A, F)$ est un espace vectoriel sur $\mathbb{K}$.

Proposition 43. *Continuité d'une application dans un produit*

Soit $A \subset E$, F_1, F_2 des espaces vectoriels normés et on considère $F = F_1 \times F_2$ muni de la norme produit.
Soit $f = (f_1, f_2)$ une application de A dans F. Alors f est continue sur A si, et seulement si, f_1 et f_2 sont continues sur A.

Démonstration (sans utiliser la caractérisation séquentielle).

— ($\Rightarrow$) On suppose f continue sur A. Montrons que les f_i sont continues sur A.
Comme f est continue sur A, pour tout $a \in A$ et pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $x \in A$,

$$\|x - a\|_E \leq \delta \;\Rightarrow\; \|f(x) - f(a)\| = \max(\|f_1(x) - f_1(a)\|_1, \|f_2(x) - f_2(a)\|_2) \leq \varepsilon.$$

Soit $i \in \{1, 2\}$ et $a \in A$. Montrons que f_i est continue en a. Soit $\varepsilon > 0$. Soit $x \in A$. Si $\|x - a\|_E \leq \delta$, on a :

$$\|f_i(x) - f_i(a)\|_i \leq \max(\|f_1(x) - f_1(a)\|_1, \|f_2(x) - f_2(a)\|_2) = \|f(x) - f(a)\| \leq \varepsilon$$

Donc f_i est continue en a. Ceci étant vrai pour tout $a \in A$, f_i est continue sur A.

— ($\Leftarrow$) On suppose f_1, f_2 continues sur A. Montrons que f est continue sur A.
Pour $i \in \{1, 2\}$, comme f_i est continue en A, pour tout $a \in A$ et pour tout $\varepsilon > 0$, il existe $\delta_i > 0$ tel que pour tout $x \in A$,

$$\|x - a\|_E \leq \delta_i \;\Rightarrow\; \|f_i(x) - f_i(a)\|_i) \leq \varepsilon.$$

Soit $a \in A$. Montrons que f est continue en a. Soit $\varepsilon > 0$. On pose $\delta = \max(\delta_1, \delta_2)$. Soit $x \in A$. Si $\|x - a\|_E \leq \delta$, on a :

$$\|f(x) - f(a)\| = \max(\underbrace{\|f_1(x) - f_1(a)\|_1}_{\leq \varepsilon}, \underbrace{\|f_2(x) - f_2(a)\|_2}_{\leq \varepsilon}) \leq \varepsilon$$

Donc f est continue en a. Ceci étant vrai pour tout $a \in A$, f est continue sur A.

$\square$

Remarque 24.

En raisonnant par récurrence, on montre que le résultat précédent est toujours valable pour un produit fini $F_1 \times ... \times F_n$ de n espaces vectoriels normés ($n \in \mathbb{N}^*$).

Proposition 44. *Composition d'applications continues*

La composée de deux applications continues est continue.

Exercice 47.

1. Écrire la proposition suivante en décrivant précisément les hypothèses.
2. Soit $A \subset E$ et $f : A \to F$ une application continue. Montrer que $x \mapsto \|f(x)\|_F$ est une application continue sur A

Correction.

1. Soit $(E, \|\cdot\|_E), (F, \|\cdot\|_F), (G, \|\cdot\|_G)$ des espaces vectoriels normés sur K ; $A \subset E$, $B \subset F$. Soit $f : A \to F$ et $g : B \to G$. Si $f(A) \subset B$, f est continue sur A et g est continue sur B alors $g \circ f$ est continue sur A.
2. $f : A \to F$ est continue sur $A \subset F$ et $\left|\begin{array}{l} F \to \mathbb{R} \\ y \mapsto \|y\|_F \end{array}\right.$ est continue sur F (voire exercice 46 2)) donc par composition, $x \mapsto \|f(x)\|_F$ est continue sur A.

Théorème 7. *Continuité et densité*

Soit $A \subset E$, D une partie dense dans A et $f, g : A \to F$ des applications continues sur A. Si pour tout $x \in D$, $f(x) = g(x)$ (i.e. $f_{|D} = g_{|D}$) alors $f = g$.

Démonstration.

On suppose que $f_{|D} = g_{|D}$. Soit $x \in A$. Comme D est dense dans A, $xA \subset\in \overline{D}$, donc il existe (x_n) à valeurs dans D tel que $x_n \to x$. Par suite, on a :

$$0 \leq \|f(x) - g(x)\|_F \leq \underbrace{\|f(x) - f(x_n)\|_F}_{\to 0} + \underbrace{\|f(x_n) - g(x_n)\|_F}_{=0} + \underbrace{\|g(x_n) - g(x)\|_F}_{\to 0} \to 0.$$

Donc, par l'axiome de séparation, $f(x) = g(x)$. □

4. Applications lipschitziennes

Définition 35.

Soit $k \in \mathbb{R}_+$, $A \subset E$ et $f : A \to F$.
On dit que f est k**-lipschitzienne** ou **lipschitzienne de rapport** k si, pour tout $x, y \in A$:

$$\|f(x) - f(y)\|_F \leq k\|x - y\|_E.$$

On dit que f est **lipschitzienne** s'il existe $k \in \mathbb{R}_+$ tel que f est lipschitzienne de rapport k.

Exemple 13.

— Une norme est une application 1-lipschitzienne.

> En effet, pour tous $x, y \in E$, $\|x - y\|_E \leq \|x - y\|_E$ donc $\|\cdot\|_E$ est 1-lipschitzienne.

— Soit $((E_1, \|\cdot\|_1), ..., (E_n, \|\cdot\|_n))$ des espaces vectoriels normés et $E = E_1 \times ... \times E_n$ muni de la norme produit. Les applications coordonnées $(i = 1, ..., n)$:

$$\varphi_i : \left|\begin{array}{ccc} E & \to & E_i \\ x = (x_1, ..., x_n) & \mapsto & x_i \end{array}\right.$$

sont 1-lipschitziennes.

> En effet, pour tous $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in E$,
> $$\|\varphi_i(x) - \varphi_i(y)\|_i = \|x_i - y_i\|_i \leq \max_{1\leq j\leq n} \|x_j - y_j\|_j = \|x - y\|_E.$$
> donc φ_i est 1-lipschitzienne.

— Soit $A \subset E$. L'application $\begin{vmatrix} E & \to & \mathbb{R} \\ x & \mapsto & d(x, A) \end{matrix}$ est 1-lipschitzienne.

> En effet, pour tous $x, y \in E$, (voire exercice 7)
> $$|d(x, A) - d(y, A)| \leq d(x, y) = \|x - y\|_E$$
> donc $x \mapsto d(x, A)$ est 1-lipschitzienne.

Exercice 48.

1. Montrer que $x \mapsto \cos(x)$ est lipschitzienne.
2. Montrer que $\begin{vmatrix} E \times E & \to & \mathbb{R} \\ (x, y) & \mapsto & d(x, y) \end{matrix}$ est lipschitzienne où $E \times E$ est muni de la norme produit

Correction.

1. $x \mapsto \cos(x)$ est dérivable sur $\mathbb{R}$ donc d'après le théorème des accroissements finis, pour tous $x < y$, il existe $c \in]x, y[$ tel que
$$\cos(x) - \cos(y) = -\sin(c)(x - y);$$
or $|\sin(c)| \leq 1$ donc :
$$|\cos(x) - \cos(y)| \leq |x - y|.$$
Cette inégalité étant également valable pour $x = y$, on obtient que cos est 1-lipschitzienne.
2. On a, pour tous $(x, y), (u, v) \in E^2$:
$$\begin{aligned} |d(x, y) - d(u, v)| &\leq |d(x, y) - d(y, u)| + |d(y, u) - d(u, v)| \\ &\leq d(x, u) + d(y, v) \\ &\leq 2\max(d(x, u), d(y, v)) = 2\|(x, y) - (u, v)\|_\infty. \end{aligned}$$
Donc cette application est 2-lipschitzienne.

Remarque 25.

La notion de fonction lipschitzienne est invariante par passage à des normes équivalentes (sur l'espace de départ ou celui d'arrivée). Par contre, la constante peut tout de même changer.

Proposition 45.

Toute application lipschitzienne est continue.

Démonstration.

Soit $A \subset E$ et $f : A \to \mathbb{R}$. On suppose que f est k-lipschitzienne. Si $k = 0$, f est la fonction nulle et donc continue. Supposons $k \neq 0$.
Soit $x \in E$ et $\varepsilon > 0$. On pose $\delta = \frac{\varepsilon}{k} > 0$. Alors pour tout $y \in E$, tel que $\|x - y\|_E \leq \delta$, on a :

$$\|f(x) - f(y)\|_F \leq k\|x - y\| \leq k\delta = \varepsilon.$$

Ceci étant vrai pour tout $x \in A$ et tour $\varepsilon > 0$, f est continue sur A. □

5. Continuité et topologie

Théorème 8.

Soit $A \subset E$ et $f : A \to F$. On suppose f continue sur A.

- Soit U un ouvert de F. L'image réciproque $f^{-1}(U)$ de U par f est un ouvert relatif de A.
- Soit C un fermé de F. L'image réciproque $f^{-1}(C)$ de C par f est un fermé relatif de A.

Démonstration.

On suppose f continue sur A. Soit U un ouvert de F et $a \in f^{-1}(U)$. Comme $f(a) \in U$ et U est un ouvert, il existe $r > 0$ tel que $B_f(f(a), r) \subset U$. Par suite, f étant continue en a, il existe $\delta > 0$ tel que $f(B_f(a, \delta) \cap A) \subset B_f(f(a), r) \subset U$. Donc

$$B_f(a, \delta) \cap A \subset f^{-1}(f(B_f(a, \delta) \cap A)) \subset f^{-1}(U).$$

D'où $f^{-1}(U)$ est un voisinage de a relatif de A. Il en résulte que $f^{-1}(U)$ est un voisinage relatif de A de chacun des points de A, donc $f^{-1}(U)$ est un ouvert relatif de A. Pour le cas fermé, on remarque que $f^{-1}(X^c) = (f^{-1}(X))^c$ et utilise le fait qu'un fermé est le complémentaire d'un ouvert. □

Proposition 46. *Réciproque du théorème précédent*

Soit $A \subset E$ et $f : A \to F$. Si l'image réciproque par f de tout ouvert (resp. fermé) de F est un ouvert (resp. fermé) relatif de A, alors f est continue.

Démonstration.

On suppose que l'image réciproque par f de tout ouvert de F est un ouvert relatif de A. Montrons que f est continue.
Soit $a \in A$ et V un voisinage de $f(a)$ dans F. Alors il existe un ouvert $U \subset V$ tel que $f(a) \in U$.
Par suite, $W = f^{-1}(U)$ est un voisinage de a dans E. En effet, W est un ouvert car c'est l'image réciproque d'un ouvert par f et $a \in W$ car $f(a) \in U$.
De plus, on a $f(W \cap A) \subset V$; donc ceci étant vrai pour tout voisinage V de $f(a)$ et tout $a \in A$, f est continue sur A. $\square$

Exemple 14.

— Une boule ouverte (resp. fermée) est un ouvert (resp. fermé) de E. On l'avait déjà démontré directement dans la partie B.

> En effet, soit $x_0 \in E$ et $r > 0$.
> Notons $N : E \to \mathbb{R}$ la norme sur E. L'intervalle $]-\infty, r[$ (resp. $]-\infty, r]$) est un ouvert (resp. fermé) de $\mathbb{R}$ et, de plus, l'application $f : E \to \mathbb{R}$, définie, pour tout $x \in E$, par :
> $$f(x) = N(x - x_0),$$
> est une application continue comme composée d'applications continues. Or, on a :
> $$B(x_0, r) = f^{-1}(]-\infty, r[) \quad \text{et} \quad B_f(x_0, r) = f^{-1}(]-\infty, r]),$$
> donc $B(x_0, r)$ est un ouvert de E et $B_f(x_0, r)$ est un fermé de E.

— L'ensemble $\{(x, y) \mid x + y > 0\}$ est un ouvert de $\mathbb{R}^2$ et $\{(x, y) \mid x + y = 2\}$ est un fermé de $\mathbb{R}^2$.

> En effet, l'application $f : \mathbb{R}^2 \to \mathbb{R}$ telle que $f : (x, y) \mapsto x + y$ est une application continue (pour n'importe quelle norme sur $\mathbb{R}^2$) - *À démontrer pour votre norme préférée*. Or, on a :
> $$\{(x, y) \mid x + y > 0\} = f^{-1}(\mathbb{R}_+^*) \quad \text{et} \quad \{(x, y) \mid x + y = 2\} = f^{-1}(\{2\}).$$
> $\mathbb{R}_+^*$ étant un ouvert de $\mathbb{R}$ et $\{2\}$ étant un fermé de $\mathbb{R}$, on obtient le résultat.

Exercice 49.

Soit $n \in \mathbb{N}^*$. Montrer que l'ensemble $GL_n(\mathbb{K})$ est un ouvert de $\mathcal{M}_n(\mathbb{K})$ (muni de la norme infinie par exemple) et que $SL_n(\mathbb{K})$ est un fermé de $\mathcal{M}_n(\mathbb{K})$.

Correction.

Soit $n \in \mathbb{N}^*$. On a :
$$GL_n(\mathbb{K}) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid \det(M) \neq 0\} = \det{}^{-1}(\mathbb{K}^*), \text{ et}$$

$$SL_n(\mathbb{K}) = \{M \in \mathcal{M}_n(\mathbb{K}) \mid \det(M) = \pm 1\} = \det^{-1}(\{-1, 1\}).$$

Comme $\mathbb{K}^*$ est un ouvert de $\mathbb{K}$ (réunion d'ouverts) et $\{-1, 1\}$ est un fermé de $\mathbb{K}$ (réunion finie de fermés) ; il suffit de montrer que $\det : \mathcal{M}_n(\mathbb{K}) \to \mathbb{K}$ est continue pour obtenir le résultat.

Le déterminant est une fonction polynomiale de degré n à n^2 variables (les coefficients matriciels) : en effet, on peut, soit le montrer, sans expliciter de formule, par récurrence (en utilisant le développement par les lignes par exemple) ; soit montrer la formule suivante : pour $M = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$,

$$\det(M) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(j),j},$$

où $\mathcal{S}_n$ est l'ensemble des permutations de l'ensemble $\{1, ..., n\}$ et $\varepsilon : \mathcal{S}_n \to \{-1, 1\}$ est l'application signature.

Une fonction polynomiale sur $\mathcal{M}_n(\mathbb{K})$ (que l'on peut voir pour simplifier comme $\mathbb{K}^{n^2}$) est continue comme somme de produits des applications "coordonnées" ($\varphi_{i,j}(M) = a_{i,j}$) qui sont continues (voire l'exemple 13 + la proposition 45).
Par suite det est continue sur $\mathcal{M}_n(\mathbb{K})$.

6. Continuité uniforme

Définition 36. *Continuité uniforme*

Soit $A \subset E$ et $f : A \to F$. On dit que f est **uniformément continue** si :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x, y \in A \quad \|x - y\|_E \leq \delta \ \Rightarrow\ \|f(x) - f(y)\|_F \leq \varepsilon.$$

Proposition 47.

— Toute application uniformément continue est continue.
— Toute application lipschitzienne est uniformément continue.

Démonstration.

Soit $A \subset E$ et $f : A \to F$ une application.

— On suppose f uniformément continue sur A. Alors pour tout $a \in A$,

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in A \quad \|x - y\|_E \leq \delta \ \Rightarrow\ \|f(x) - f(y)\|_F \leq \varepsilon.$$

Donc f est continue sur A.

— On suppose f k-lipschitzienne. Si $k = 0$, alors la fonction f est nulle, donc trivialement, f est uniformément continue. On suppose $k \neq 0$.
Soit $\varepsilon > 0$. Alors pour $\delta = \frac{\varepsilon}{k}$, et $x, y \in E$ tels que $\|x - y\|_E \leq \delta$:

$$\|f(x) - f(y)\|_F \leq k\|x - y\|_E \leq k\delta = \varepsilon.$$

Par suite, f est uniformément continue.

□

Remarque 26.

Attention, les réciproques des implications précédentes sont fausses :

— $f : x \mapsto e^x$ est continue mais pas uniformément continue sur $\mathbb{R}$.

> En effet, supposons par l'absurde que f est uniformément continue sur $\mathbb{R}$. Soit $\varepsilon > 0$. Alors il existe $\delta > 0$ tel que pour tous $x, y \in \mathbb{R}$, si $|x-y| \leq \delta$ alors $|e^x - e^y| \leq \varepsilon$.
> Pour $x \in \mathbb{R}$ et $y = x + \delta$, on a $|x-y| \leq \delta$ et
> $$|e^x - e^y| = e^x|e^{y-x}| = e^x e^\delta \xrightarrow[x \to +\infty]{} +\infty.$$

— $f : x \mapsto \sqrt{x}$ est uniformément continue mais pas lipschitzienne sur $\mathbb{R}$.

> Montrons premièrement que f est uniformément continue. Pour $x, y \in \mathbb{R}_+$ avec $x \leq y$, on a $\sqrt{xy} \geq x$, et donc :
> $$(\sqrt{y} - \sqrt{x})^2 = y - 2\sqrt{xy} + x \leq y - x,$$
> d'où $\sqrt{y} - \sqrt{x} \leq \sqrt{y-x}$ et de même, pour $y \leq x$, $\sqrt{x} - \sqrt{y} \leq \sqrt{x-y}$.
> Par suite, pour tous $x, y \in \mathbb{R}_+$:
> $$|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x-y|}.$$
> Soit $\varepsilon > 0$. Alors pour $\delta = \varepsilon^2$ et $x, y \in \mathbb{R}$ avec $|x-y| \leq \delta$,
> $$|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x-y|} \leq \sqrt{\delta} = \varepsilon.$$
> Il en résulte que f est uniformément continue. Montrons désormais que f n'est pas lipschitzienne sur $\mathbb{R}$. Le "problème" se situe au voisinage de 0. Pour $x > 0$, on a :
> $$\frac{|f(x) - f(0)|}{|x-0|} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}} \xrightarrow[x \to 0^+]{} +\infty.$$
> f n'est donc pas lipschitzienne sur $\mathbb{R}$.

7. Continuité, applications linéaires et multilinéaires

a. Continuité et applications linéaires

On rappelle la notation $\mathcal{L}(E, F)$ pour désigner l'ensemble des applications linéaires de E dans F.

Théorème 9.

Soit $f \in \mathcal{L}(E, F)$. Alors les assertions suivantes sont équivalentes :

i) f est continue ;

ii) f est continue en 0_E ;

iii) il existe $k \geq 0$ tel que, pour tout $x \in E$,

$$\|f(x)\|_F \leq k\|x\|_E.$$

iv) f est lipschitzienne.

Démonstration.

Montrons i)⇒ii)⇒iii)⇒iv)⇒i)

- i)⇒ii). Évident.
- ii)⇒iii). On suppose f est continue en 0_E :

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in E \quad \|x - 0_E\|_E \leq \delta \ \Rightarrow\ \|f(x) - f(0_E)\|_F \leq \varepsilon.$$

  Considérons le cas $\varepsilon = 1$. Soit $x \in E$ avec $x \neq 0_E$. On pose $y = \frac{\delta}{\|x\|_E}x$. Par construction, on a $\|y\|_E = \delta \leq \delta$, donc

$$\frac{\delta}{\|x\|_E}\|f(x)\|_F = \|f(y)\|_F \leq \varepsilon = 1.$$

  Par suite, on a

$$\|f(x)\|_F \leq \frac{1}{\delta}\|x\|_E.$$

  Cette égalité étant vraie pour $x = 0_E$, on a pour $k = \frac{1}{\delta}$ et pour tout $x \in E$:

$$\|f(x)\|_F \leq k\|x\|_E.$$

- iii)⇒iv). On suppose qu'il existe $k \geq 0$ tel que, pour tout $x \in E$,

$$\|f(x)\|_F \leq k\|x\|_E.$$

  Soit $x, y \in E$. Comme f est linéaire, on a :

$$\|f(x) - f(y)\|_F = \|f(x - y)\|_F \leq k\|x - y\|_E.$$

- iv)⇒i). On applique la proposition 45.

□

Notation 7.

On note $\mathcal{L}_c(E, F)$ l'ensemble des applications linéaires continues de $(E, \|\cdot\|_E)$ dans $(F, \|\cdot\|_F)$.

Proposition 48.

L'ensemble $\mathcal{L}_c(E, F)$ est un espace vectoriel sur $\mathbb{K}$.

Démonstration.

On a $\mathcal{L}_c(E,F) = C(E,F) \cap \mathcal{L}(E,F)$ donc $\mathcal{L}_c(E,F)$ est un sous-espace vectoriel de $\mathcal{F}(E,F)$ (espace vectoriel des fonctions de E dans F) comme intersection de sous-espaces vectoriels de $\mathcal{F}(E,F)$. □

Proposition 49.

Soit $f \in \mathcal{L}(E,F)$. Alors f est continue si, et seulement si, l'application de $E \setminus \{0_E\}$ dans $\mathbb{R}$:

$$x \mapsto \frac{\|f(x)\|_F}{\|x\|_E}$$

est bornée sur $E \setminus \{0_E\}$.

Démonstration.

$$f \text{ est continue}$$

si, et seulement si,

$$\text{il existe } k \geq 0 \text{ tel que, pour tout } x \in E,\ \|f(x)\|_F \leq k\|x\|_E.$$

si, et seulement si,

$$\text{il existe } k \geq 0 \text{ tel que, pour tout } x \in E \setminus \{0_E\},\ \frac{\|f(x)\|_F}{\|x\|_E} \leq k.$$

si, et seulement si,

$$x \mapsto \frac{\|f(x)\|_F}{\|x\|_E} \text{ est bornée sur } E \setminus \{0_E\}.$$

□

Exercice 50.

Soit $E = C([0,1],\mathbb{R})$. On considère l'application $\varphi : f \mapsto f(1)$.

— Montrer que φ est linéaire.

— Montrer que φ est continue pour $\|\cdot\|_\infty$ mais pas pour $\|\cdot\|_1$.

Correction.

— Soit $\lambda, \mu \in \mathbb{R}$ et $f, g \in E$. On a :

$$\varphi(\lambda f + \mu g) = (\lambda f + \mu g)(1) = \lambda f(1) + \mu g(1) = \lambda\varphi(f) + \mu\varphi(g);$$

donc φ est linéaire.

— On a, pour $f \in E$,

$$|\varphi(f)| = |f(1)| \leq \sup_{x\in[0,1]} |f(x)| = \|f\|_\infty.$$

Donc φ est continue pour $\|\cdot\|_\infty$.

On pose, pour $n \in \mathbb{N}$, $f_n : x \mapsto x^n$. On a :

$$\frac{|\varphi(f_n)|}{\|f_n\|_1} = n + 1.$$

Par suite, $f \mapsto \frac{|\varphi(f)|}{\|f\|_1}$ n'est pas bornée. Donc φ n'est pas continue pour $\|f\|_1$.

Proposition 50.

Soit N_1, N_2 des normes sur E. Alors N_1, N_2 sont équivalentes si, et seulement si,

$$Id_E : (E, N_1) \to (E, N_2)$$

et

$$Id_E : (E, N_2) \to (E, N_1).$$

sont continues.

Démonstration.

$Id_E : (E, N_1) \to (E, N_2)$ est continue si, et seulement si, il existe $k \geq 0$ tel que pour tout $x \in E$, $N_2(x) \leq kN_1(x)$ i.e. N_2 est dominée par N_1.
Et de même, $Id_E : (E, N_2) \to (E, N_1)$ est continue si, et seulement si, N_1 est dominée par N_2.
On en déduit le résultat. $\square$

b. Normes d'opérateur (ou normes subordonnées)

Lemme 1.

Soit $f \in \mathcal{L}_c(E, F)$. L'ensemble $\{k \in \mathbb{R}_+ \mid \forall\, x \in E,\ \|f(x)\|_F \leq k\|x\|_E\}$ est un partie non vide et minorée de $\mathbb{R}$.

Correction.

On note $K = \{k \in \mathbb{R}_+ \mid \forall\, x \in E,\ \|f(x)\|_F \leq k\|x\|_E\}$. Comme f est une application linéaire continue, d'après le théorème 9, il existe $k_0 \geq 0$ tel que pour tout $x \in E$, $\|f(x)\|_F \leq k\|x\|_E$. Ainsi, k_0 appartient à K donc ce dernier est non vide. De plus, il est minoré par 0, d'où le résultat.

Le lemme précédent permet de justifier l'existence de la norme d'opérateur définie ci-après :

Définition 37.

Soit $f \in \mathcal{L}_c(E, F)$. On appelle **norme d'opérateur** de f associée aux normes $\|\cdot\|_E$ et $\|\cdot\|_F$ ou encore **norme subordonnée** de f aux normes $\|\cdot\|_E$ et $\|\cdot\|_F$ et on note $|||f|||$, ou encore $\|f\|_{\mathrm{op}}$), la quantité :

$$|||f||| = \inf\{k \in \mathbb{R}_+ \mid \forall\, x \in E,\ \|f(x)\|_F \leq k\|x\|_E\}.$$

Proposition 51.

Soit $f \in \mathcal{L}_c(E,F)$. Alors on a :

i) $\displaystyle |||f||| = \sup_{\|x\|_E \neq 0} \frac{\|f(x)\|_F}{\|x\|_E}$;

ii) $\displaystyle |||f||| = \sup_{\|x\|_E \leq 1} \frac{\|f(x)\|_F}{\|x\|_E}$;

iii) $\displaystyle |||f||| = \sup_{\|x\|_E = 1} \|f(x)\|_F$.

Démonstration.

On note $A = \{k \in \mathbb{R}_+ \mid \forall\, x \in E,\ \|f(x)\|_F \leq k\|x\|_E\}$.

i) On a note $B = \left\{\dfrac{\|f(x)\|_F}{\|x\|_E} \mid x \in E \setminus \{0_E\}\right\}$. Montrons tout d'abord que B admet une borne supérieure. Comme E n'est pas réduit à 0_E, il existe $x_0 \in E \setminus \{0_E\}$, ainsi B est non vide car $\frac{\|f(x_0)\|_F}{\|x_0\|_E}$ appartient à B.
De plus, comme f est continue sur E, il existe $k \geq 0$ tel que pour tout $x \in E$, $\|f(x)\|_F \leq k\|x\|_E$. Ainsi, en particulier, pour tout $x \in E \setminus \{0_E\}$,

$$\frac{\|f(x)\|_F}{\|x\|_E} \leq k$$

Par suite, B est une partie non vide et majorée (par k) de $\mathbb{R}$ donc elle admet une borne supérieure. Notons b cette borne supérieure.

Montrons que $b = |||f|||$. Commençons par montrer que $b \geq |||f|||$. Comme b est un majorant de B, on a, pour tout $x \in E \setminus \{0_E\}$, $\frac{\|f(x)\|_F}{\|x\|_E} \leq b$ et donc pour tout $x \in E$,

$$\|f(x)\|_F \leq b\|x\|_E$$

(cette inégalité étant bien vérifiée pour 0_E également). Ainsi, b appartient à A. Or $|||f|||$ étant la borne inférieure de A et donc en particulier un minorant de A, on a $b \geq |||f|||$.
De plus, si $k \in A$, on a, pour tout $x \in E \setminus \{0_E\}$:

$$\frac{\|f(x)\|_F}{\|x\|_E} \leq k$$

donc k est un majorant de B, d'où, b étant son plus petit majorant, $k \geq b$. Ceci étant vrai pour tout $k \in A$, b est un minorant de A. Or $|||f|||$ étant son plus grand minorant, on obtient $|||f||| \geq b$.
Il en résulte que $b = |||f|||$.

ii) On note $C = \left\{\dfrac{\|f(x)\|_F}{\|x\|_E} \mid \|x\|_E \leq 1\right\}$. Alors $C \subset B \cup \{0\}$ (le 0 provient du fait que 0_E fait partie des vecteurs de normes plus petites que 1 mais pas de $E \setminus \{0_E\}$). De plus, on peut remarquer que B et $B \cup \{0\}$ ont la même borne supérieure car 0 est un minorant de B. Donc, par inclusion, C possède une borne supérieure c et $c \leq b$. Montrons l'inégalité dans l'autre sens.
Soit $x \in E \setminus \{0_E\}$. On pose $y = \frac{x}{\|x\|_E}$. Alors on a $\|y\|_E = 1$ et par linéarité de f et par

homogénéité de la norme $\|\cdot\|_F$:

$$\frac{\|f(x)\|_F}{\|x\|_E} = \left\| f\left(\frac{x}{\|x\|_E}\right)\right\|_F = \|f(y)\|_F = \frac{\|f(y)\|_F}{\|y\|_E}.$$

Or $\|y\|_E = 1 \leq 1$ donc $\frac{\|f(x)\|_F}{\|x\|_E}$ appartient à C. D'où $B \subset C$ et ainsi, $b \leq c$.
Ainsi, $b = c$.

iii) Par un raisonnement similaire à celui du point précédent, on peut prouver que $\{\|f(x)\|_F \mid \|x\|_E = 1\} = B$ et donc que leurs bornes supérieures sont égales.

□

La proposition suivante justifie la terminilogie de "norme" pour la norme d'opérateur :

Proposition 52.

L'application $|||\cdot||| : \mathcal{L}_c(E,F) \to \mathbb{K}$ est une norme sur $\mathcal{L}_c(E,F)$.

Démonstration.

D'après le lemme 1, l'application $|||\cdot|||$ est bien définie sur $\mathcal{L}_c(E,F)$.
Soit $f, g \in \mathcal{L}_c(E,F)$ et $\lambda \in \mathbb{K}$.

— Positivité : on a $|||f||| = \sup_{\|x\|_E=1} \underbrace{\|f(x)\|_F}_{\geq 0} \geq 0$.

— Séparation : on suppose $|||f||| = 0$. Alors pour tout $x \in E$,

$$\|f(x)\|_F \leq 0\|x\|_E = 0$$

Ainsi, par positivité et séparation de la norme $\|\cdot\|_F$, pour tout $x \in E$, $f(x) = 0_F$. Par suite, $f = \mathbf{0}$.

— Homogénéité : on a, par homogénéité de $\|\cdot\|_F$ et les propriétés usuelles de la borne supérieure :

$$|||\lambda f||| = \sup_{\|x\|_E=1} \underbrace{\|\lambda f(x)\|_F}_{=|\lambda|.\|f(x)\|_F} = |\lambda|.|||f|||.$$

— Inégalité Triangulaire : on a, par inégalité triangulaire de $\|\cdot\|_F$ et les propriétés usuelles de la borne supérieure :

$$|||f+g||| = \sup_{\|x\|_E=1} \underbrace{\|f(x)+g(x)\|_F}_{\leq \|f(x)\|_F + \|g(x)\|_F} = |||f||| + |||g|||.$$

□

Proposition 53.

Soit $f \in \mathcal{L}_c(E,F)$ et $g \in \mathcal{L}_c(F,G)$. Alors $g \circ f \in \mathcal{L}_c(E,G)$ et on a :

$$|||g \circ f||| \leq |||g|||.|||f|||.$$

Démonstration.

Soit $f \in \mathcal{L}_c(E,F)$ et $g \in \mathcal{L}_c(F,G)$. Alors $g \circ f$ est linéaire continue comme composée d'applications linéaires et continues. On remarque que pour tout $x \in E$ et tout $y \in F$, on a :

$$\|f(x)\|_F \leq |||f|||.\|x\|_E \text{ et } \|g(y)\|_G \leq |||g|||.\|y\|_F$$

Donc, pour tout $x \in E$,

$$\|g \circ f(x)\|_G = \|g(f(x))\|_G \leq |||g|||.\|f(x)\|_E \leq |||g|||.|||f|||.\|x\|_E$$

D'où $|||g \circ f||| \leq |||g|||.|||f|||$. □

Méthodes pour calculer une norme subordonnée :
On considère une application linéaire $f : E \to F$ où E et F sont munis de $\|\cdot\|_E$ et $\|\cdot\|_F$ respectivement.

On cherche le "meilleur" $k \geq 0$ possible tel que, pour tout $x \in E$, $\|f(x)\|_F \leq k\|x\|_E$.
Une fois un tel k trouvé, alors f est continue et $|||f||| \leq k$.

> En effet, f est continue en vertu du Théorème 9 et comme $|||f|||$ est la borne inférieure de $A = \{k' \in \mathbb{R}_+ \mid \forall\, x \in E,\ \|f(x)\|_F \leq k'\|x\|_E\}$ et que k appartient à A, $|||f||| \leq k$.

Ensuite, on démontre que notre k est bien le meilleur d'une des façons suivantes :

— on cherche $x_0 \in E \setminus \{0_E\}$ tel que $\|f(x_0)\|_F = k\|x_0\|_E$ ou même $x_0 \in E$ avec $\|x_0\|_E = 1$ tel que $\|f(x_0)\|_F = k$. Dans ce cas, on conclut que $|||f||| = k$.

> En effet, on a $\frac{\|f(x_0)\|_F}{\|x_0\|_E} = k$; or k étant un majorant de $B = \left\{\frac{\|f(x)\|_F}{\|x\|_E} \mid x \in E \setminus \{0_E\}\right\}$ car $|||f||| \leq k$; on a $k = \max B = \sup B = |||f|||$.

— on cherche une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans $E \setminus \{0_E\}$ tel que $\frac{\|f(x_n)\|_F}{\|x_0\|_E} \xrightarrow[n\to+\infty]{} k$. Dans ce cas, on conclut que $|||f||| = k$.

> En effet, k est un majorant de $B = \left\{\frac{\|f(x)\|_F}{\|x\|_E} \mid x \in E \setminus \{0_E\}\right\}$ car $|||f||| \leq k$, et on a trouvé une suite à valeurs dans B qui converge vers k. Ainsi, par caractérisation séquentielle de la borne supérieure $k = \sup B = |||f|||$.

Exemple 15.

On considère l'application linéaire $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ telle que pour $(x,y) \in \mathbb{R}^2$, $f(x,y) = (x+y, x-y)$. Alors la norme d'opérateur de f associée aux normes $\|\cdot\|_2$ (pour l'espace de départ) et $\|\cdot\|_2$ (pour l'espace d'arrivée) est $|||f||| = \sqrt{2}$.

En effet, on a :

$$\|f(x,y)\|_2 = \sqrt{(x+y)^2+(x-y)^2} = \sqrt{2}\sqrt{x^2+y^2} = \sqrt{2}\|(x,y)\|_2$$

donc en particulier, $\|f(x,y)\|_2 \leq \sqrt{2}\|(x,y)\|_2$ d'où f est continue et $|||f||| \leq \sqrt{2}$ et comme on a égalité pour un certain $(x_0,y_0) \in \mathbb{R}^2$ (en fait pour tous, d'après la première égalité !), on obtient $|||f||| = \sqrt{2}$

Exercice 51.

On considère $E = \mathbb{R}[X]$. Dans la suite, l'espace d'arrivée des fonctions étant $\mathbb{R}$, on le munit de sa norme canonique même si on ne le précise pas !

1. On munit E de la norme $\|P\| = \sup\limits_{t\in[-1,1]} |P(t)|$ et on considère la forme linéaire $f : P \mapsto P(0)$.
   Montrer que f est continue et déterminer la norme subordonnée de f.
2. On munit E de la norme $\|P\| = \sup\limits_{n\in\mathbb{N}} \dfrac{|P^{(n)}(0)|}{n!}$ et on considère la forme linéaire $g : P \mapsto P(\frac{1}{2})$
   Après avoir prouver que $\|\cdot\|$ est bien une norme, montrer que g est continue et déterminer la norme subordonnée de g.

Correction.

1. Pour tout $P \in E$, on a :
   $$|f(P)| = |P(0)| \leq \|P\|$$
   donc f est continue sur E muni de $\|\cdot\|$ et $|||f||| \leq 1$. De plus, on a, pour P le polynôme constant en 1 : $|f(P)| = 1 = \|P\|$ donc $|||f||| = 1$.

2. On rappelle ici le formule de Taylor pour les polynômes qui nous servira ici : pour $P \in E$ et $a \in \mathbb{R}$, on a :
   $$P = \sum_{n=0}^{+\infty} \frac{P^{(n)}(a)}{n!}(X-a)^n \quad \text{(somme finie)}$$
   Pour montrer que $\|\cdot\|$, seul l'axiome de séparation peut poser vraiment problème ; montrons seulement celui-ci : soit $P \in E$ tel que $\|P\| = 0$. Alors, pour tout $n \in \mathbb{N}$, $P^{(n)}(0) = 0$. Ainsi, d'après la formule de Taylor pour les polynômes appliqué en $a = 0$, on obtient :
   $$P = \sum_{n=0}^{+\infty} \frac{P^{(n)}(0)}{n!}X^n = \sum_{n=0}^{+\infty} 0.X^n = 0$$
   D'où l'axiome de séparation.

   Passons à la norme subordonnée. Soit $P \in E$. D'après la formule de Taylor pour les polynômes appliquée en 0, on a, en evaluant en $\frac{1}{2}$:

$$P\left(\frac{1}{2}\right)=\sum_{n=0}^{+\infty}\frac{P^{(n)}(0)}{n!}\left(\frac{1}{2}\right)^n$$

Donc

$$|g(P)|=\left|P\left(\frac{1}{2}\right)\right|\leq\sum_{n=0}^{+\infty}\underbrace{\frac{|P^{(n)}(0)|}{n!}}_{\leq\|P\|}\left(\frac{1}{2}\right)^n\leq\left(\sum_{n=0}^{+\infty}\left(\frac{1}{2}\right)^n\right)\|P\|=2\|P\|.$$

Par suite, $|||g|||\leq 2$.
De plus, considérons la suite $(P_n)_{n\in\mathbb{N}}$ de polynômes non nuls où pour tout $n\in\mathbb{N}$, $P_n=\sum_{k=0}^n X^k$. Alors, comme $\|P_n\|=1$ car $P^{(k)}(0)=k!$ pour tout $k\leq n$, on a :

$$\frac{|g(P_n)|}{\|P_n\|}=\frac{\sum_{k=0}^n\left(\frac{1}{2}\right)^k}{1}=2-\frac{1}{2^n}\xrightarrow[n\to+\infty]{}2$$

Ainsi, $|||g|||=2$.

Exercice 52.

On considère $E=\mathbb{R}^n$ et les deux normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sur E. Montrer que l'application identité Id_E est continue de E dans E dans toutes les façons de munir E (au départ et à l'arrivée) et calculer sa norme d'opérateur dans chacun de ces cas.

c. Continuité et applications multilinéaires

Théorème 10.

Soit E,F,G des espaces vectoriels normés et $f:E\times F\to G$ une application bilinéaire où $E\times F$ est muni de la norme produit. On a équivalence entre les assertions suivantes :

i) f est continue sur $E\times F$;

ii) f est continue en $(0_E,0_F)$;

iii) f vérifie : il existe $k\geq 0$ telle que pour tout $(x,y)\in E\times F$:

$$\|f(x,y)\|_G\leq k\|x\|_E.\|y\|_F.$$

Démonstration.

On montre i)⇒ii)⇒iii)⇒i)

- i)⇒ii) Si f est continue sur $E\times F$ alors f est continue en $(0_E,0_F)\in E\times F$.
- ii)⇒iii) On suppose f continue en $(0_E,0_F)$.
  Alors pour tout $\varepsilon>0$, il existe $\delta>0$ tel que pour tout $(x,y)\in E\times F$,

$$\|(x,y)\|\leq\delta\ \Rightarrow\ \|f(x,y)\|_G\leq\varepsilon.$$

Pour $\varepsilon = 1$, il existe $\delta > 0$ vérifiant la proposition ci-dessus. Notons $k = \frac{1}{\delta^2} > 0$. Soit $(x, y) \in E \times F$ avec $(x, y) \neq (0_E, 0_F)$. On pose :

$$x' = \frac{\delta}{\|x\|_E} x \quad \text{et} \quad y' = \frac{\delta}{\|y\|_E} y.$$

Alors $\|(x, y)\| = \max(\|x'\|_E, \|y'\|_F) = \delta$, donc

$$\|f(x', y')\|_G \leq \varepsilon = 1.$$

De plus, par bilinéarité, $\|f(x', y')\|_G = \frac{\delta^2}{\|x\|_E\|y\|_F}\|f(x, y)\|_G$, donc

$$\|f(x, y)\|_G \leq k\|x\|_E\|y\|_F.$$

Et on remarque que cette inégalité est également vraie pour $(x, y) = (0_E, 0_F)$.

- iii)$\Rightarrow$i) On suppose ii). Soit $(x_0, y_0) \in E \times F$. Soit $\varepsilon > 0$ et on pose $\delta = \max\left(\sqrt{\frac{\varepsilon}{2k}}, \frac{\varepsilon}{2k(\|x_0\|_E + \|y_0\|_F)}\right)$,
  Alors on a :
  $$\begin{aligned}
  \|f(x, y) - f(x_0, y_0)\|_G &\leq \|f(x, y) - f(x, y_0)\|_G + \|f(x, y_0) - f(x_0, y_0)\|_G \\
  &\leq \|f(x, y - y_0)\|_G + \|f(x - x_0, y_0)\|_G \\
  &\leq k(\|x\|_E.\|y - y_0\|_F + \|x - x_0\|_E.\|y_0\|_F) \\
  &\leq k(\underbrace{\|x\|_E}_{\leq \|x_0\| + \delta} + \|y_0\|_F)\|(x - x_0, y - y_0)\|. \\
  &\leq k\delta^2 + k\delta(\|x_0\|_E + \|y_0\|_F) \leq \varepsilon
  \end{aligned}$$
  Donc f est continue.

□

Exercice 53.

Soit E un espace préhilbertien, $(\cdot|\cdot)$ un produit scalaire sur E et $\|\cdot\|$ la norme associée à ce produit scalaire. Montrer que $(\cdot|\cdot) : E \times E \to \mathbb{K}$ est continue.

Correction.

Soit $x, y \in E$. D'après le théorème de Cauchy-Schwarz, on a :

$$|(x|y)| \leq \|x\|.\|y\|.$$

Par suite, $(\cdot|\cdot)$ est continue.

Théorème 11.

Soit $E_1, ..., E_m, G$ des espaces vectoriels normés sur $\mathbb{K}$ et $f : E \to G$ une application m-linéaire où $E = E_1 \times ... \times E_m$ est muni de la norme produit. On a équivalence entre les assertions suivantes :

i) f est continue sur E ;

ii) il existe $k \geq 0$ telle que pour tout $(x_1, ..., x_m) \in E$:

$$\|f(x_1, ..., x_m)\|_G \leq k\|x_1\|_{E_1} ... \|x_m\|_{E_m}.$$

Démonstration.

On raisonne de la même manière que dans la démonstration du théorème 10. □

Partie F
Compacité

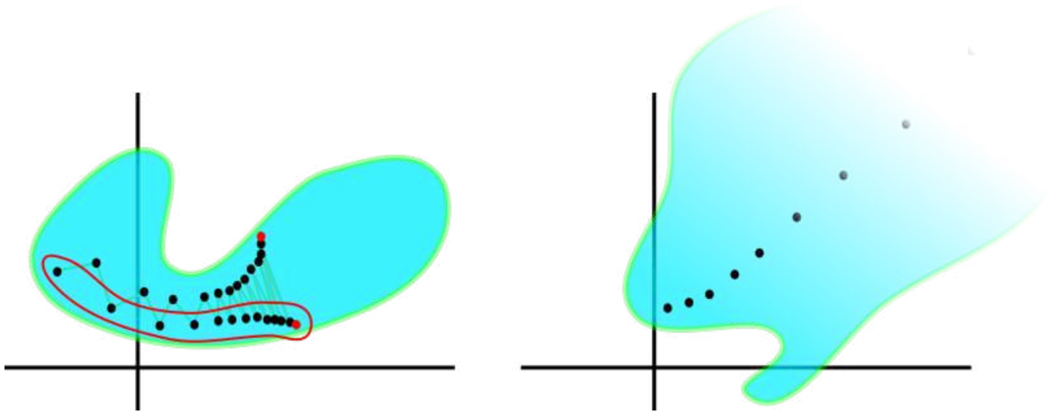
Dans toute cette partie, E désigne un espace vectoriel normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ muni d'une norme $\|\cdot\|$.

1. Définition

Définition 38. *Partie compacte*

Soit $A \subset E$. On dit que A est une partie **compacte** de E si toute suite à valeurs dans A possède au moins une valeur d'adhérence dans A.

Autrement dit, A est compacte si, de toute suite à valeurs dans A, on peut extraire une sous-suite qui converge dans A.



Exemple 16.

— L'ensemble vide est compact ; toute partie finie de E est compacte.

Montrons que toute partie finie A de E est compacte. Notons $N = \#A$. Alors il existe $x_1, \dots, x_N$ deux à deux distincts tels que $A = \{x_1, \dots, x_N\}$.
Soit $u = (u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans A. Intuitivement, la partie A possédant un nombre fini d'éléments, la suite va devoir "passer" une infinité de fois par au moins un des x_i, disons x_{i_0} et donc la sous-suite des termes valant x_{i_0} étant constante, elle converge vers $x_{i_0} \in A$. Voyons comment construire une telle sous-suite rigoureusement :
Pour $i \in [\![1, N]\!]$, on note $U_i = \{n \in \mathbb{N} \mid u_n = x_i\}$. Alors la famille $(U_i)_{i \in [\![1,N]\!]}$ vérifie :

- pour tous $i, j \in [\![1, N]\!]$ avec $i \neq j$, $U_i \cap U_j = \emptyset$. En effet, si $n \in U_i$, alors $u_n = x_i \neq x_j$ donc $n \notin U_j$.
- $\displaystyle\bigcup_{i=1}^{N} U_i = \mathbb{N}$. En effet

Remarque : certains U_i pouvant potentiellement être vides, il ne s'agit pas d'une partition de $\mathbb{N}$ à proprement parler.

La réunion étant finie et valant $\mathbb{N}$ qui est infini, il existe $i_0 \in [\![1, N]\!]$ tel que U_{i_0} est infini (par l'absurde, si tous les U_i étaient finis, leur réunion finie le serait aussi, ce qui est absurde car elle est égale à $\mathbb{N}$).
On construit alors la sous-suite de u dont les indices sont les éléments de U_{i_0} dans l'ordre. Pour ce faire, on utilise une méthode que l'on reverra quand nous étudierons la dénombrabilité, la construction d'une bijection de $\mathbb{N}$ dans un sous-ensemble infini de $\mathbb{N}$: si B est un sous-ensemble infini de $\mathbb{N}$, la fonction $\varphi : \mathbb{N} \to \mathbb{N}$ telle que :

$$\begin{cases} \varphi(0) = \min(B) \\ \varphi(n+1) = \min(B \setminus \{\varphi(0), \dots, \varphi(n)\}) \text{ pour } n \in \mathbb{N} \end{cases}$$

est bien définie, strictement croissante et d'image B (on laisse le soin au lecteur de prouver ces affirmations). Ainsi, en appliquant cette construction à $B = U_{i_0}$, on obtient une sous-suite $(u_{\varphi(n)})_{n \in \mathbb{N}}$ de u qui vérifie :

$$\text{pour tout } n \in \mathbb{N},\ u_{\varphi(n)} = x_{i_0} \text{ car } \varphi(n) \in U_{i_0}.$$

Par suite, $(u_{\varphi(n)})_{n \in \mathbb{N}}$ est constante en x_{i_0} et donc converge vers $x_{i_0} \in A$.
Il en résulte que A est compacte.

- Dans $\mathbb{R}$, les parties fermées bornées (les segments par exemple) sont compacts : il s'agit (pour "compact" au sens de la définition 38) du théorème de Bolzano-Weierstrass dans $\mathbb{R}$.

2. Propriétés

Proposition 54.

Soit $A \subset E$. Si A est compacte, alors A est fermée et bornée dans E.

Démonstration.

On suppose A compacte.

— Montrons que A est fermée. Soit (x_n) une suite à valeurs dans A qui converge vers $l \in E$. La suite (x_n) est à valeurs dans A qui est compact, donc on peut en extraire une sous-suite qui converge dans A. Or, toute sous-suite d'une suite convergente converge vers la même limite, donc $l \in A$. Ceci étant vrai pour toute suite (x_n) à valeurs dans A convergente, d'après la caractérisation séquentielle des fermés, A est fermé.

Montrons par l'absurde que A est bornée. Supposons que A n'est pas bornée. Alors pour tout $n \in \mathbb{N}$, il existe $x_n \in A$ tel que $\|x_n\| \geq n$. Alors chaque sous-suite $(x_{\varphi(n)})$ de (x_n) n'est pas bornée (car $x_{\varphi(n)} \geq \varphi(n)$) donc (x_n) ne possède pas de valeur d'adhérence. Contradiction.
Par suite, A est bornée.

□

Remarque 27.

Attention, la réciproque est fausse en général ! Voire l'exercice suivant

Exercice 54.

On considère $E = \mathcal{F}_b(\mathbb{R}, \mathbb{R})$ muni de sa norme canonique.

— Montrer que $B_f(0_E, 1)$ n'est pas compact.

— Que dire de la réciproque de la proposition précédente ?

Indication.

Si une suite (u_n) vérifie qu'il existe $\eta > 0$ tel que pour tous $n, m \in \mathbb{N}$ avec $n \neq m$:

$$\|u_n - u_m\| \geq \eta,$$

alors (u_n) ne possède aucune valeur d'adhérence.

Ainsi, pour montrer qu'une partie A n'est pas compacte, il suffit d'exhiber une suite vérifiant la propriété ci-dessus.

Correction.

— Considérons la suite (f_n) à valeurs dans E de terme général

$$f_n : x \mapsto f_n(x) = \begin{cases} 1 & \text{si } x \in [-n, n] \\ 0 & \text{sinon} \end{cases}$$

.

Alors clairement, (f_n) est à valeurs dans $B_f(0_E, 1)$ et on a, pour tous n, m avec $n \neq m$:

$$\|f_n - f_m\|_F = 1.$$

Par suite, (f_n) ne possède pas de valeur d'adhérence, donc $B_f(0_E, 1)$ n'est pas compact.

— L'ensemble $B_f(0_E, 1)$ est fermé et borné mais n'est pas compact. La réciproque de la proposition en question est donc fausse.

Proposition 55.

Soit $A, B \subset E$. On suppose $B \subset A$. Si A est compacte et B est fermée, alors B est compacte.

Démonstration.

On suppose $B \subset A$, A compacte et B fermée. Soit (x_n) à valeurs dans B. Comme $B \subset A$, (x_n) est à valeurs dans A compacte donc il existe une sous-suite (u_n) de (x_n) qui converge dans A. Or B est fermée donc toute suite à valeurs dans B converge dans B. Il en résulte que (u_n) converge dans B. Ceci étant vrai pour toute suite (x_n) à valeurs dans B, b est compacte. □

Théorème 12.

Une suite à valeurs dans un compact est convergente, si, et seulement si, elle possède une unique valeur d'adhérence.

Démonstration.

- $(\Rightarrow)$. Une suite convergente ne possède qu'une seule valeur d'adhérence : sa limite.
- $(\Leftarrow)$. Soit $A \subset E$ un compact et (u_n) une suite à valeurs dans A.
  On suppose que (u_n) possède une unique valeur d'adhérence a. Supposons par l'absurde que (u_n) ne converge pas. Alors, en particulier, (u_n) ne converge pas vers a.
  Par suite, il existe $\varepsilon > 0$ tel que pour tout $n \in \mathbb{N}$, il existe $k \geq n$ tel que $\|u_k - a\| \geq \varepsilon$.
  Notons, $\varphi(0) = \min\{k \in \mathbb{N} \mid \|u_k - a\| \geq \varepsilon\}$ et pour $n \in \mathbb{N}^*$,
  $$\varphi(n) = \min\{k > \varphi(n-1) \mid \|u_k - a\| \geq \varepsilon\}.$$
  Alors $\varphi : \mathbb{N} \to \mathbb{N}$ est strictement croissante, donc $(u_{\varphi(n)})$ est une sous-suite de (u_n) qui ne possède pas a pour valeur d'adhérence.
  Or $(u_{\varphi(n)})$ est une suite à valeur dans A compact et donc possède une valeur d'adhérence. Cela implique que (u_n) possède au moins deux valeurs d'adhérence. Contradiction.

□

Proposition 56.

Soit E, F des espaces vectoriels normés et $A \subset E$, $B \subset F$. Si A est un compact de E et B un compact de F, alors $A \times B$ est un compact de $E \times F$ muni de la norme produit.

Démonstration.

On suppose A est un compact de E et B un compact de F. Soit $((x_n, y_n))_{n\in\mathbb{N}}$ une suite à valeurs dans $A \times B$. Alors (x_n) est à valeurs dans A compact. Par suite, il existe une sous-suite $(x_{\varphi(n)})$ de (x_n) qui converge vers un certain $x \in A$. La suite $(y_{\varphi(n)})$ est à valeurs dans B compact, donc on peut en extraire une sous-suite $(y_{\varphi\circ\psi(n)})$ qui converge vers un certain $y \in B$.
Alors la sous-suite $((x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)}))$ de $((x_n, y_n))_{n\in\mathbb{N}}$ converge vers $(x, y) \in A \times B$. Donc $A \times B$ est compact. □

Corollaire 6.

Un produit d'une famille finie de compacts est un compact de l'espace produit muni de la norme produit.

Démonstration.

On raisonne par récurrence en appliquant la proposition précédente. □

Exercice 55.

Soit $A \subset \mathbb{R}^n$ où $\mathbb{R}^n$ est muni de la norme infini. Montrer que A est compact si, et seulement si, A est fermée bornée.

Correction.

Si A est compact alors A est fermée bornée. Montrons la réciproque pour $A \subset \mathbb{R}^n$ muni de la norme produit $\|\cdot\|$.
On suppose A fermée bornée. Alors il existe $R \geq 0$ tel que $A \subset B_f(0_E, R) = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$. Or, pour la norme produit sur $\mathbb{R}^n$, on a :

$$B_f(0_E, R) = [-R, R]^n.$$

Donc $B_f(0_E, R)$ est compact comme produit d'une famille finie de compacts. Or A est fermé dans le compact $B_f(0_E, R)$, donc A est compact.

3. Applications continues sur un compact

Dans ce paragraphe, F est un espace vectoriel normé.

Proposition 57.

Soit $A \subset E$ et $f : A \to F$ une application continue sur A. Si A est un compact de E, alors $f(A)$ est un compact de F.
Autrement dit, l'image direct d'un compact par une application continue est un compact.

Démonstration.

On suppose que A est compact. Soit $(f(x_n))$ une suite à valeurs dans $f(A)$. Alors (x_n) est une suite à valeurs dans A compact donc il existe une sous-suite $(x_{\varphi(n)})$ de (x_n) qui converge vers un certain $x \in A$. Par la caractérisation séquentielle de la continuité,

$$f(x_{\varphi(n)}) \xrightarrow[n\to\infty]{} f(x).$$

Donc la sous-suite $(f(x_{\varphi(n)}))$ de $(f(x_n))$ converge vers $f(x) \in f(A)$. Il en résulte que $f(A)$ est compact. □

Corollaire 7. *Conséquences pour $F = \mathbb{R}$*

Soit A une partie non vide de $\mathbb{R}$ et $f : A \to \mathbb{R}$ une application continue. Si A est compacte, alors f est bornée et atteint ses bornes.

Démonstration.

On suppose A compacte. Alors d'après la proposition 57, $f(A)$ est un compact de $\mathbb{R}$. C'est donc un fermé borné, donc f est borné et comme $f(A)$ est fermé, par la caractérisation séquentielle des bornes supérieures et inférieures, les bornes sont atteintes. □

Exercice 56.

Soit A, B des parties compactes non vide de E.

1. Montrer que pour tout $x \in E$, il existe $a \in A$ tel que $d(x, a) = d(x, A)$.
2. Montrer qu'il existe $a \in A$ et $b \in B$ tels que $d(a, b) = d(A, B)$.

Correction.

1. Soit $x \in E$. L'application $f : y \mapsto d(x, y)$ est continue sur E dans $\mathbb{R}$ et donc sa restriction à A est continue de A dans $\mathbb{R}$. Comme A est compacte, alors f est bornée et atteint sa borne inférieure sur A. Donc il existe $a \in A$ tel que :
$$d(x, a) = \inf_{y \in A} d(x, y) = d(x, A).$$
2. $A \times B$ est compact comme produit de compacts et $(x, y) \mapsto d(x, y)$ est continue de $E \times E$ dans $\mathbb{R}$ donc sa restriction à $A \times B$ est bornée et atteint ses bornes. Par suite il existe $(a, b) \in A \times B$ tel que :
$$d(a, b) = \inf_{(x,y) \in A \times B} d(x, y) = d(A, B)$$

Théorème 13. *Théorème de Heine*

Soit $A \subset E$ et $f : A \to F$ une application continue. Si A est compact, alors f est uniformément continue sur A.

Démonstration.

On suppose que A est compact. Montrons par l'absurde que f est uniformément continue.
On suppose que f n'est pas uniformément continue. Alors,
$$\exists \varepsilon > 0,\ \forall \delta > 0,\ \exists x, y \in A, \quad \|x - y\|_E \leq \delta \ \text{ et } \ \|f(x) - f(y)\|_F > \varepsilon.$$
Considérons, pour $n \in \mathbb{N}^*$, $\delta = \frac{1}{n}$. Alors il existe $x_n, y_n \in A$ tels que $\|x_n - y_n\|_E \leq \frac{1}{n}$ et $\|f(x) - f(y)\|_F > \varepsilon$.

Alors (x_n) (et (y_n) également) est une suite à valeurs dans A compact, donc on peut extraire une sous-suite $(x_{\varphi(n)})$ qui converge vers $a \in A$. On a alors, pour tout $n \in \mathbb{N}^*$,

$$\|y_{\varphi(n)} - a\|_E \leq \|y_{\varphi(n)} - x_{\varphi(n)}\|_E + \|x_{\varphi(n)} - a\|_E \leq \frac{1}{\varphi(n)} + \|x_{\varphi(n)} - a\|_E \xrightarrow[n\to\infty]{} 0,$$

donc $(y_{\varphi(n)})$ converge vers a.
Par suite, par continuité de f, on a

$$f(y_{\varphi(n)}) - f(x_{\varphi(n)}) \xrightarrow[n\to\infty]{} f(a) - f(a) = 0.$$

Or, pour tout $n \in \mathbb{N}^*$, $\|f(y_{\varphi(n)}) - f(x_{\varphi(n)})\|_F > \varepsilon$. Contradiction. $\square$

Partie G

Connexité par arcs

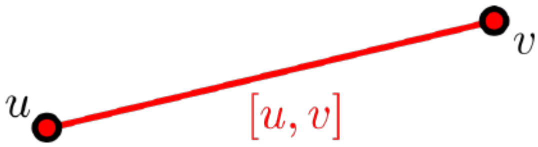
Dans toute cette partie, E désigne un espace vectoriel normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ muni d'une norme $\|\cdot\|$.

1. Parties convexes d'un espace vectoriel

Définition 39. *Segment*

Soit u, v des points de E. Le **segment** entre u et v, noté $[u, v]$ est l'ensemble défini par :

$$[u, v] := \{tu + (1-t)v \mid t \in [0, 1]\}$$



Proposition 58.

Soit $x, y, z \in \mathbb{R}$ avec $x \leq z$. On a $x \leq y \leq z$ si et seulement si, il existe $t \in [0, 1]$ tel que $y = tx + (1-t)z$.
Dans ce cas, et si $z \neq x$, on a $t = \dfrac{z-y}{z-x}$.

Démonstration.

($\Rightarrow$) Tout d'abord, si $x = z$, alors $x = y = z$ et tout $t \in [0, 1]$ convient.
On suppose $x \neq z$. On peut remarquer que y est le barycentre de la famille $((x, z-y), (z, y-x))$. En effet, on a :

$$(z-y)x + (y-x)z = (z-x)y.$$

Ainsi, en posant $t = \dfrac{z-y}{z-x}$, on obtient $1 - t = \dfrac{y-x}{z-x}$, d'où, finalement :

$$y = tx + (1-t)z.$$

($\Leftarrow$) On suppose qu'il existe $t \in [0, 1]$ tel que $y = tx + (1-t)z$. Comme $x \leq z$, $t \geq 0$ et $1 - t \geq 0$, on a :

$$x = tx + (1-t)x \geq y = tx + (1-t)z \leq tz + (1-t)z = z$$

et de plus, si $z \neq x$, comme on vient de montrer que $x \leq y \leq z$ d'après le point précédent, $y = t'x + (1-t')z$ avec $t' = \frac{z-y}{z-x}$.

Ainsi, on a :

$$\begin{aligned} 0 &= y - y \\ &= tx + (1-t)z - (t'x + (1-t')z) \\ 0 &= (t-t')(x-z) \end{aligned}$$

Or comme $x \neq z$, et que $\mathbb{R}$ est un anneau intègre, on obtient $t = t' = \frac{z-y}{z-x}$.

□

Exercice 57.

Soit $x, z \in \mathbb{R}$ avec $x \leq z$. Montrer que la notation $[x, z]$ en tant qu'intervalle de $\mathbb{R}$ coïncide bien avec la définition précédente de segment.

Correction.

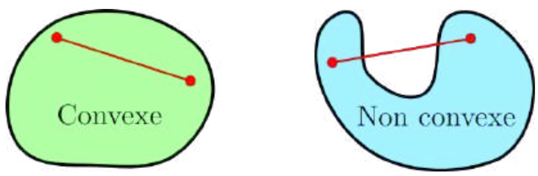
On note $[x, z]_1 = \{y \in \mathbb{R} \mid x \leq y \leq z\}$ (notation en tant qu'intervalle de $\mathbb{R}$) ;
et $[x, z]_2 = \{tx + (1-t)z \mid t \in [0, 1]\}$ (notation définie précédemment).

D'après la proposition précédente, on a $y \in [x, z]_1$ i.e. $x \leq y \leq z$ si, et seulement si, il existe $t \in [0, 1]$ tel que $y = tx + (1-t)z$ i.e. $y \in [x, z]_2$.
D'où $[x, z]_1 = [x, z]_2$. Les notations coïncident donc bien !

Définition 40. *Partie convexe*

Soit A une partie de E.
On dit que A est **convexe** si, pour tous $u, v \in A$, le segment $[u, v]$ est inclus dans A ; i.e.

$$\forall u, v \in A,\ \forall t \in [0, 1], \quad tu + (1-t)v \in A.$$



Exemple 17.

Les segments, les sous-espaces vectoriels sont des parties convexes.

— **Segments**. Soit $x, y \in E$. Soit $u, v \in [x, y]$. Alors il existe $\alpha, \beta \in [0,1]$ tels que $u = \alpha x + (1-\alpha)y$ et $v = \beta x + (1-\beta)y$.
Par suite, on a, pour tout $t \in [0,1]$:

$$tu + (1-t)v = t\,(\alpha x + (1-\alpha)y) + (1-t)\,(\beta x + (1-\beta)y) = sx + (1-s)y,$$

où $s = t\alpha + (1-t)\beta$.
Or s est compris entre 0 et 1 car $s \in [\alpha, \beta] \subset [0,1]$.
Il en résulte que $tu + (1-t)v \in [x, y]$; d'où $[x, y]$ est une partie convexe de E.

— **Sous-espace vectoriel**. Il suffit de remarquer qu'un sous-espace vectoriel est stable par combinaison linéaire.

Exercice 58.

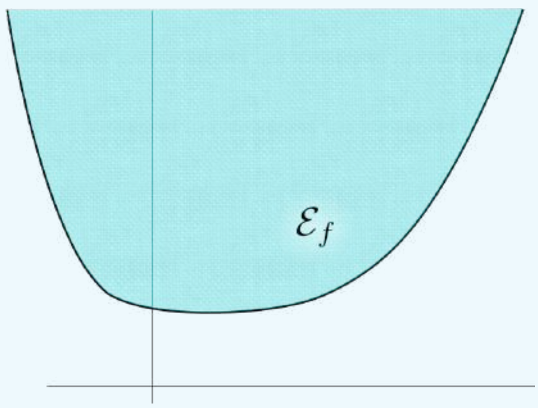
Soit f une fonction d'un intervalle I dans $\mathbb{R}$. On définit l'épigraphe $\mathcal{E}_f$ de f par :

$$\mathcal{E}_f = \{(x, y) \in I \times \mathbb{R} \mid y \geq f(x)\}.$$

1. Dessiner le graphe d'une fonction f (quelconque) puis représenter l'épigraphe $\mathcal{E}_f$ de f sur ce dessin.
2. Montrer que : f est une fonction convexe sur I si, et seulement si, $\mathcal{E}_f$ est une partie convexe de $\mathbb{R}^2$.

Correction.

1.



2. On procède par double implication.
   - ($\Rightarrow$). On suppose f convexe sur I. Montrons que $\mathcal{E}_f$ est convexe.
     Soit $M_1 = (x_1, y_1)$, $M_2 = (x_2, y_2)$ des points de $\mathcal{E}_f$ et $M = (x, y) \in [M_1, M_2]$. Alors il existe $t \in [0,1]$ tel que $M = tM_1 + (1-t)M_2$, d'où :

$$x = tx_1 + (1-t)x_2 \text{ et } y = ty_1 + (1-t)y_2.$$

La fonction f étant convexe, on a $f(x) \leq tf(x_1)+(1-t)f(x_2)$. Or, M_1, M_2 appartiennent à l'épigraphe donc $f(x_1) \leq y_1$ et $f(x_2) \leq y_2$. Par suite, comme t et $(1-t)$ sont positifs,

$$f(x) \leq ty_1 + (1-t)y_2 = y.$$

Donc M appartient à $\mathcal{E}_f$.

Il en résulte que tout segment de $\mathcal{E}_f$ est inclus dans $\mathcal{E}_f$ donc $\mathcal{E}_f$ est convexe.

- ($\Leftarrow$). On suppose $\mathcal{E}_f$ convexe. Montrons que f est convexe sur I.

  Soit $x_1, x_2 \in I$. Alors $M_1 = (x_1, f(x_1))$ et $M_2 = (x_2, f(x_2))$ appartiennent à $\mathcal{E}_f$ (car $f(x_i) \leq f(x_i)$, $i = 1, 2$). Par convexité de $\mathcal{E}_f$, le segment $[M_1, M_2]$ est inclus dans $\mathcal{E}_f$. Autrement dit, pour tout $t \in [0,1]$, $(x, y) := t(x_1, f(x_1)) + (1-t)(x_2, f(x_2))$ vérifie $f(x) \leq y$. Donc, pour tout $t \in [0,1]$:

  $$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

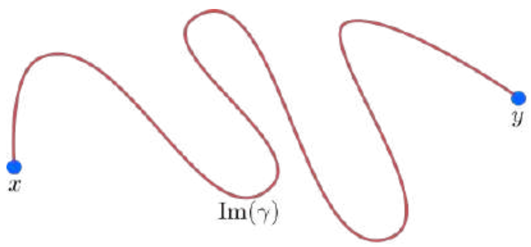
  Ceci étant vrai pour tous x_1, x_2 dans I, il en résulte que f est convexe sur I.

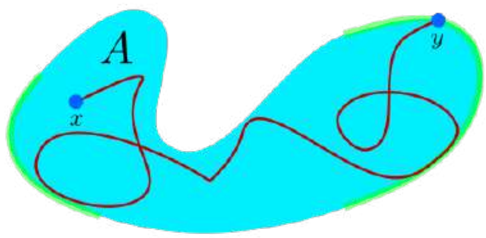
2. Chemins

Définition 41. *Chemin*

Soit $x, y \in E$. On appelle **chemin d'extrémités** x **et** y ou **chemin joignant** x **à** y, toute application continue $\gamma : [0,1] \to E$ telle que $\gamma(0) = x$ et $\gamma(1) = y$.

Soit $A \subset E$ et γ un chemin d'extrémités x et y. On dit que γ est un chemin d'extrémités x et y **dans A** si $\mathrm{Im}(\gamma) \subset A$.





Exemple 18.

Pour $x, y \in E$, l'application γ de $[0,1]$ dans E telle que $\gamma : t \mapsto (1-t)x + ty$ est un chemin d'extrémités x et y. On remarque que pour ce chemin, $\mathrm{Im}(\gamma) = [x, y]$ où $[x, y]$ est le segment d'extrémités x, y.

En effet, on a $\mathrm{Im}(\gamma) = \{\gamma(t) \mid t \in [0,1]\} = \{(1-t)x + ty \mid t \in [0,1]\} = [y,x] = [x,y]$.
De plus, $\gamma(0) = x$ et $\gamma(1) = y$; il nous reste à montrer que γ est continue sur $[0,1]$: pour tout $t, s \in [0,1]$, on a :

$$\|\gamma(t) - \gamma(s)\| = \|(1-t)x + ty - ((1-s)x + sy)\| = |t-s|.\|x-y\|$$

donc γ est $\|x-y\|$-lipschitzienne sur $[0,1]$ et donc continue sur $[0,1]$.

Remarque 28. *Re-paramétrisation*

Soit $x, y \in E$.
- Soit $a, b \in \mathbb{R}$, $a < b$. Si $g : [a,b] \to E$ est une application continue telle que $g(a) = x$ et $g(b) = y$ alors il existe un chemin joignant x à y.

  En effet, $\gamma : t \mapsto g((1-t)a + tb)$ est une application continue telle que $\gamma(0) = g(a) = x$ et $\gamma(1) = g(b) = y$.
- Si γ est un chemin joignant x à y alors $\tilde{\gamma} : t \mapsto \gamma(1-t)$ est un chemin joignant y à x.

3. Connexité par arcs et composantes connexes par arcs

Définition 42.

Soit $A \subset E$. On définit la relation binaire $\mathcal{R}$ suivante sur A : pour $x, y \in A$,

$$x \,\mathcal{R}\, y \text{ si, et seulement si, il existe un chemin d'extrémités } x \text{ et } y \text{ dans } A.$$

Proposition 59.

Soit $A \subset E$. La relation binaire $\mathcal{R}$ est une relation d'équivalence sur A.

Démonstration.

Montrons que la relation binaire $\mathcal{R}$ est une relation d'équivalence sur A, c'est-à-dire $\mathcal{R}$ est réflexive, symétrique et transitive.
- *Réflexive.* Pour tout $x \in A$, l'application de $[0,1]$ dans E constante en x est un chemin d'extrémités x et x ; donc $x \,\mathcal{R}\, x$.
- *Symétrique.* Soit $x, y \in A$ tels que $x \,\mathcal{R}\, y$. Alors il existe un chemin γ joignant x à y dans A. Le chemin $\tilde{\gamma} : t \mapsto \gamma(1-t)$ est un chemin joignant y à x et $\tilde{\gamma}$ est dans A, en effet, $\mathrm{Im}(\tilde{\gamma}) = \mathrm{Im}(\gamma) \subset A$. Donc $y \,\mathcal{R}\, x$.
- *Transitive.* Soit $x, y, z \in A$ tels que $x \,\mathcal{R}\, y$ et $y \,\mathcal{R}\, z$. Soit γ_{xy} un chemin joignant x à y et

γ_{yz} un chemin joignant y à z. Alors l'application $g : [0,2] \to E$ telle que, pour $t \in [0,2]$:

$$g(t) = \begin{cases} \gamma_{xy}(t) & \text{si } t \in [0,1], \\ \gamma_{yz}(t-1) & \text{si } t \in]1,2]. \end{cases}$$

est une application continue de $[0,2]$ dans A telle que $g(0) = \gamma_{xy}(0) = x$ et $g(2) = \gamma_{yz}(1) = z$.
En effet, g est continue sur $[0,1]$ et sur $]1,2]$ (comme composée d'applications continues sur cet intervalle) et de plus on a

$$\lim_{t \to 1^+} g(t) = \lim_{t \to 1^+} \gamma_{yz}(t-1) = \gamma_{yz}(0) = y = \gamma_{xy}(1) = g(1);$$

donc g est continue sur $[0,2]$. On a également $\mathrm{Im}(g) = \mathrm{Im}(\gamma_{xy}) \cup \mathrm{Im}(\gamma_{yz}) \subset A$.
Par suite, la re-paramétrisation de g donnée par $\gamma_{xz} : t \mapsto g(2t)$ est un chemin joignant x à z de A (car $\mathrm{Im}(\gamma_{xz}) = \mathrm{Im}(g) \subset A$).

Donc $\mathcal{R}$ est une relation d'équivalence sur A.

□

Définition 43. *Composantes connexes par arcs*

Soit $A \subset E$. On appelle **composantes connexes par arcs de** A les classes d'équivalence de la relation $\mathcal{R}$.

Définition 44. *Connexité par arcs*

Soit $A \subset E$. On dit que A est **connexe par arcs** si A possède une unique composante connexe par arcs.

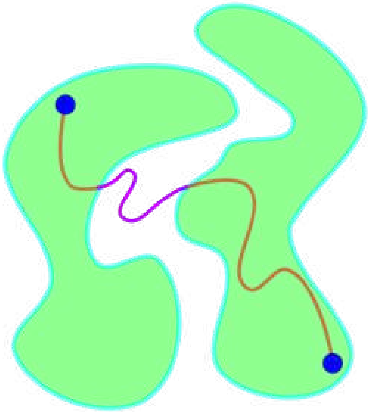
Proposition 60. *Caractérisation À CONNAÎTRE !*

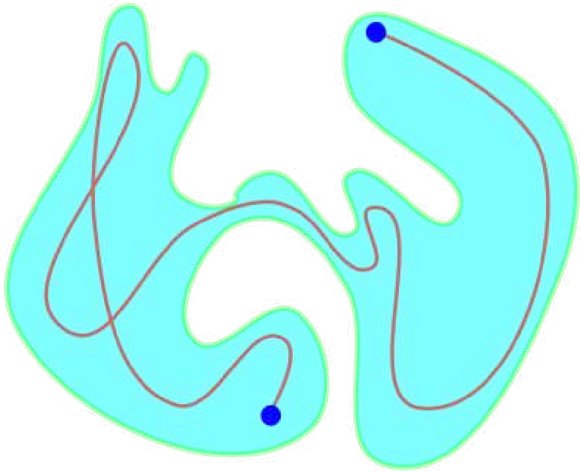
Soit $A \subset E$. A est connexe par arcs si pour tout couple (x,y) de points de A, il existe un chemin joignant x à y dans A.

Démonstration.

- $(\Rightarrow)$. On suppose A connexe par arcs. Soit $x, y \in A$. Comme la relation $\mathcal{R}$ n'a qu'une seule classe d'équivalence, on a $x \,\mathcal{R}\, y$ i.e. il existe un chemin joignant x et y dans A.
- $(\Leftarrow)$. On suppose que pour tous $x, y \in A$, il existe un chemin joignant x et y dans A ; alors $x \,\mathcal{R}\, y$. Donc A possède une unique composante connexe par arcs, d'où A est connexe par arcs.

□





Exemple 19.

— Soit $A \subset E$. Si C est une composante connexe par arcs de A, alors C est connexe par arcs.

> Soit $x, y \in C$. Alors x, y appartiennent à une même classe d'équivalence pour $\mathcal{R}$ donc $x \mathcal{R} y$ i.e. il existe un chemin joignant x à y. Donc C est connexe par arcs.

— Soit $A \subset E$. Si A est convexe, alors A est connexe par arcs.

> On suppose A convexe. Soit $x, y \in A$. Alors $[x, y] \subset A$ et donc, d'après l'exemple 18, $\gamma : t \mapsto (1-t)x + ty$ est un chemin joignant x à y dans A.
> Donc A est connexe par arcs.

— Soit $A = \mathbb{R}^* \subset \mathbb{R}$. Alors A n'est par connexe par arcs et ses composantes connexes sont $\mathbb{R}_-^*$ et $\mathbb{R}_+^*$.

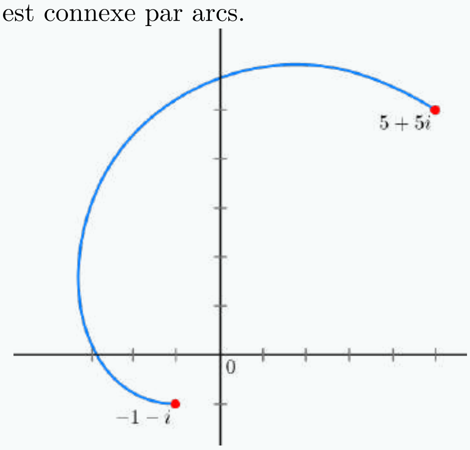
> Montrons que $\mathbb{R}^*$ n'est pas connexe par arcs. Soit $\gamma : [0,1] \to \mathbb{R}$ un chemin entre -1 et 1. En particulier, γ est continue et $\gamma(0) < 0$, $\gamma(1) > 0$ donc d'après le théorème des valeurs intermédiaires, il existe $t \in]0,1[$ tel que $\gamma(t) = 0$. Par suite, $\mathrm{Im}(\gamma) \not\subseteq \mathbb{R}^*$. Ainsi, il n'existe aucun chemin entre -1 et 1 dans $\mathbb{R}^*$: ce dernier n'est donc pas connexe par arcs.
> De plus, $\mathbb{R}_-^*$ et $\mathbb{R}_+^*$ sont convexes (exercice) donc sont connexes par arcs et $\mathbb{R}^* = \mathbb{R}_-^* \sqcup \mathbb{R}_+^*$. Ainsi, les composantes connexes de $\mathbb{R}^*$ sont $\mathbb{R}_-^*$ et $\mathbb{R}_+^*$.

— Soit $A = \mathbb{C}^* \subset \mathbb{C}$. Alors A est connexe par arcs.

Soit $z, z' \in \mathbb{C}^*$. Alors il existe $r, r' > 0$ et $\theta, \theta' \in \mathbb{R}$ tels que $z = re^{i\theta}$ et $z' = r'e^{i\theta'}$.
On pose $\gamma : [0,1] \to \mathbb{C}$ telle que :

$$\gamma(t) = ((1-t)r + tr')\, e^{i\left((1-t)\theta + t\theta'\right)}$$

Alors γ est continue sur $[0,1]$ comme composée et produit de fonctions continues (fonctions affines et $x \mapsto e^{ix}$).
On a $\gamma(0) = re^{i\theta} = z$, $\gamma(1) = r'e^{i\theta'} = z'$ et pour tout $t \in [0,1]$, $\gamma(t) \neq 0$ car $|\gamma(t)| = (1-t)r + tr' \geq \min(r, r') > 0$.
Par suite, γ est un chemin dans $\mathbb{C}^*$ joignant z à z'.
Il en résulte que $\mathbb{C}^*$ est connexe par arcs.



Exemple de l'image d'un tel chemin entre $-1-i$ et $5+5i$.

Exercice 59.

— Déterminer si les ensembles suivants sont connexes par arcs et si ce n'est pas le cas, en déterminer les composantes connexes par arcs.

1. $B_f(0_E, 1)$.
2. $A = \mathbb{C} \setminus \mathbb{R}^*$.
3. $A \cup B$ avec $A \cap B \neq \emptyset$.
4. $S(0_E, 1)$ (discuter selon la dimension et le corps des scalaires).

— Que dire de l'intersection de deux parties de E connexes par arcs ?

Correction.

— 1. $B_f(0_E, 1)$ est convexe et donc connexe par arcs.

2. $A = \mathbb{C} \setminus \mathbb{R}^*$. Pour $z, z' \in \mathbb{C} \setminus \mathbb{R}^*$, l'application

$$\gamma : t \mapsto \begin{cases} (1-2t)z & \text{si } t \in [0, \frac{1}{2}] \\ (2t-1)z' & \text{si } t \in]\frac{1}{2}, 1] \end{cases}$$

3. Soit $x, y \in A \cup B$. Le seul cas "difficile" est $x \in A$ et $y \in B$. Soit $z \in A \cap B$. Comme $z \in A$ et A est connexe par arcs, il existe un chemin joignant x à z dans $A \subset A \cup B$ et comme $z \in B$ et B est connexe par arcs, il existe un chemin joignant z à y dans $B \subset A \cup B$. La re-paramétrisation de la concaténation de ces deux chemins est un chemin joignant x à y dans $A \cup B$.

4. Soit E un espace vectoriel normé sur $\mathbb{K}$.
Si E est de dimension 1 sur $\mathbb{K} = \mathbb{R}$, pour tout $u, v \in E$, u, v sont colinéaires, et donc si u, v sont unitaires, $v = \pm u$ (car dans $\mathbb{R}$ seuls -1 et 1 sont de valeur absolue 1). Par suite, pour u un vecteur unitaire, $S(0_E, 1) = \{u, -u\}$. Donc $S(0_E, 1)$ n'est pas connexe par arcs (à justifier soigneusement en utilisant le théorème des valeurs intermédiares) dans ce cas.

   On suppose que $\dim(E) \geq 2$ ou $\mathbb{K} = \mathbb{C}$. Remarquons tout d'abord que pour tout $x \in S(0_E, 1)$, il existe $z \in S(0_E, 1)$ tel que $z \neq \pm x$. En effet :
1er cas : $\mathbb{K} = \mathbb{C}$. Soit $x \in S(0_E, 1)$. Alors $z = ix \in S(0_E, 1)$ et $z \neq \pm x$.
2eme cas : $\dim(E) \geq 2$. Soit $x \in S(0_E, 1)$ et $y \in E$ tel que $\{x, y\}$ est une famille libre (possible car $\dim(E) \geq 2$). Alors $z = \frac{y}{\|y\|} \in S(0_E, 1)$ et $z \neq \pm x$ car x et z ne sont pas colinéaires.

   Pour $x, y \in E$, on note $S_{xy} : \left| \begin{array}{ccc} [0,1] & \to & E \\ t & \mapsto & (1-t)x + ty \end{array} \right.$ et $\psi : \left| \begin{array}{ccc} E \setminus \{0_E\} & \to & E \\ u & \mapsto & \frac{u}{\|u\|} \end{array} \right.$.
Alors S_{xy} est continue sur $[0, 1]$ et ψ est continue sur $E \setminus \{0_E\}$.

   Démontrons maintenant que $S(0_E, 1)$ est connexe par arcs. Soit $x, y \in S(0_E, 1)$.
1er cas : $y \neq -x$. Alors, comme x, y sont de norme 1, $0_E \notin [x, y]$. Par suite, $\mathrm{Im}(S_{xy}) \subset E \setminus \{0_E\}$ donc $\gamma_{xy} = \psi \circ S_{x,y} : [0, 1] \to E$ est continue (comme composée d'applications continues), $\gamma_{xy}(0) = x$, $\gamma_{xy}(1) = y$ et il est clair que $\mathrm{Im}(\gamma_{xy}) \subset S(0_E, 1)$. Il en résulte que γ_{xy} est un chemin joignant x à y.

   2eme cas : $y = -x$. Alors, d'après la remarque précédente, il existe $z \in S(0_E, 1)$ tel que $z \neq x$ et $z \neq y$. En reprenant le raisonnement et les notations du premier cas, alors γ_{xz} et γ_{zy} sont des chemins dans $S(0_E, 1)$ joignant respectivement x à z et z à y. Donc la re-paramétrisation de la concaténation de ces deux chemins est un chemin joignant x à y dans $S(0_E, 1)$.

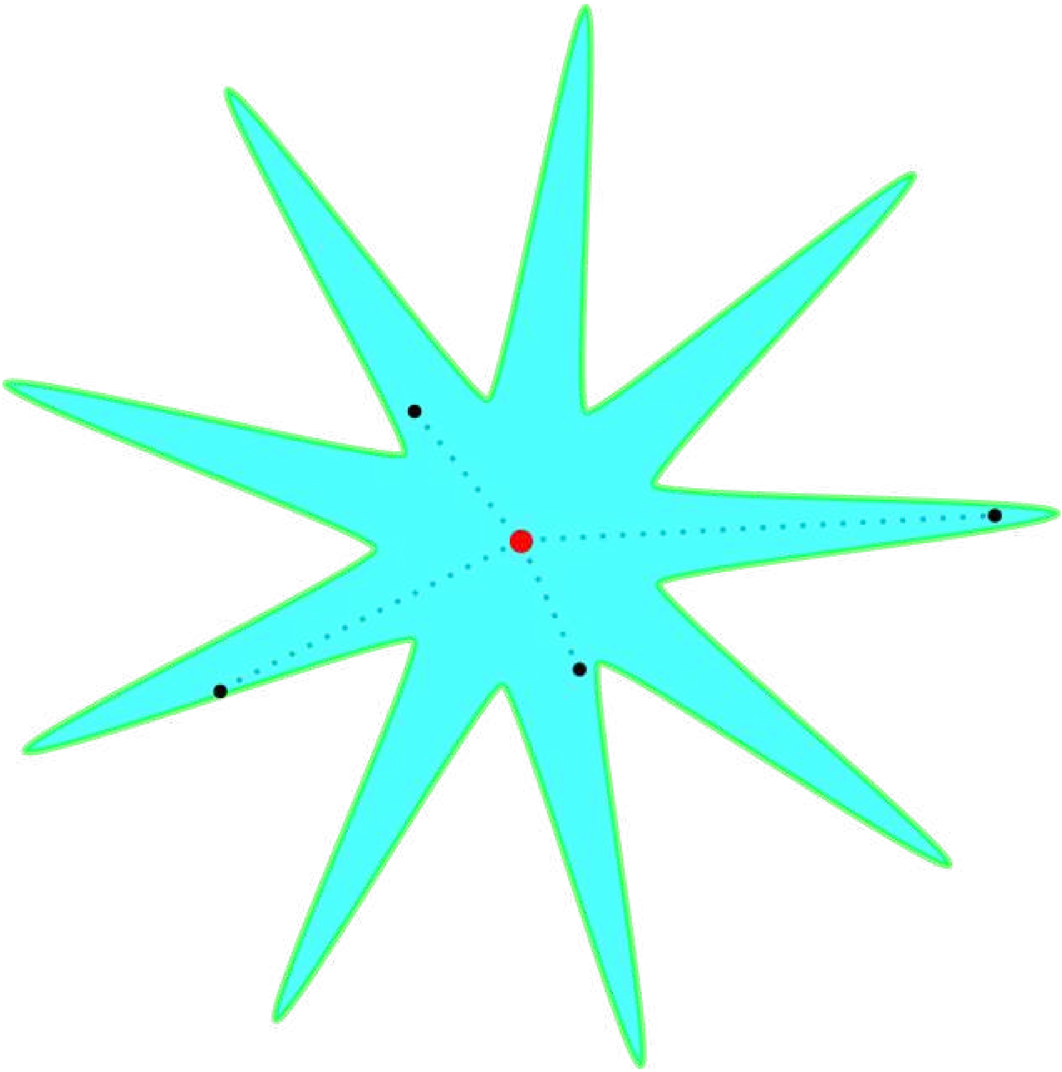
   Il en résulte que $S(0_E, 1)$ est connexe par arcs.

— En général, une intersection de parties connexes par arcs n'est pas connexe par arcs : par exemple, pour $A = \mathbb{C} \setminus \mathbb{R}_+$ et $B = \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0\}$ qui sont connexes par arcs, $A \cap B$ ne l'est pas !

4. Parties étoilées

Définition 45. *Partie étoilée*

Soit $A \subset E$.
Soit $a \in A$. On dit que A est **étoilée en** a si, pour tout $x \in A$, $[a, x] \in A$.
On dit que A est **étoilée** s'il existe $a \in A$ tel que A est étoilée en a.



Proposition 61.

— Une partie convexe de E est étoilée en chacun de ses points.
— Une partie étoilée de E est connexe par arcs.

Démonstration.

Soit $A \subset E$.

— On suppose A convexe. Soit $a \in A$. Comme A est convexe, alors pour tout $x \in A$, $[a, x] \subset A$. Donc a est étoilée en a.

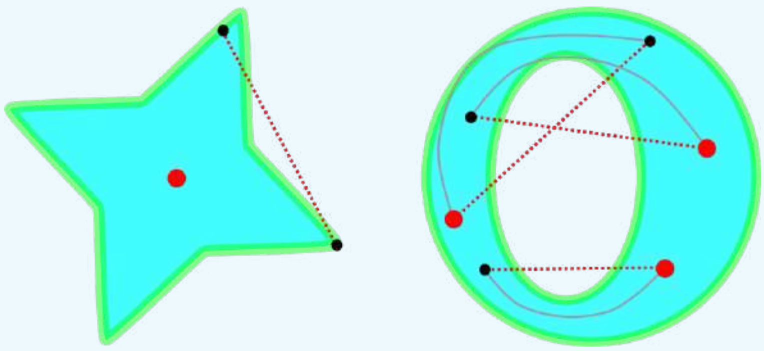
— On suppose qu'il existe $a \in A$ tel que A est étoilée en a. Pour tous $x, y \in A$, la re-paramétrisation de la concaténation des chemins dont les images sont les segments $[x, a]$ et $[a, y]$ est une chemin de x à y dans A. Donc A est connexe par arcs.

□

Exercice 60.

1. Prouver grâce à des dessins convaincants que la réciproque de chacune des propositions précédentes est fausse.
2. Montrer que $A = \mathbb{R} \cup i\mathbb{R}$ est une partie étoilée de $\mathbb{C}$.

Correction.



1.

2. Pour tout $x \in \mathbb{R}$, $[0, x] \subset \mathbb{R} \subset A$ et $[0, ix] \subset i\mathbb{R} \subset A$, donc A est étoilée en 0.

5. Parties connexes par arcs de $\mathbb{R}$

On rappelle qu'un intervalle de $\mathbb{R}$ est une partie de $\mathbb{R}$ de la forme, avec $a, b \in \mathbb{R}$, $a \leq b$ ou, dans les cas "ouverts" $a = -\infty$ ou $b = +\infty$:

$$[a, b] := \{x \in a \leq x \leq b\}; \quad [a, b[:= \{x \in a \leq x < b\};$$

$$]a, b] := \{x \in a < x \leq b\}; \text{ ou }]a, b[:= \{x \in a < x < b\}.$$

Proposition 62.

Les parties convexes de $\mathbb{R}$ sont les intervalles.

Démonstration.

Dans cette démonstration, on notera (pour] ou [;) pour] ou [et $\lhd$ pour $<$ ou $\leq$. Soit $I = (a, b)$ un intervalle de $\mathbb{R}$. Alors pour tous $x, y \in I$ et tout $z \in [x, y]$,

$$a \lhd x \leq z \leq y \lhd b,$$

donc $[x, y] \subset I$. Par suite, I est convexe.

Réciproquement, si I est une partie convexe (non vide) de $\mathbb{R}$. Alors on note $a = \inf(I)$ (possiblement $-\infty$) et $b = \sup(I)$ (possiblement $+\infty$). Alors $a \leq b$ et $I \subset [a, b]$ (ou ouvert en a et/ou en b dans les cas infini). On a alors les quatre cas suivants :

- *1er cas :* $a, b \in I$. Comme I est convexe, $[a, b] \subset I$. Donc $I = [a, b]$.
- *2eme cas :* $a \in I, b \notin I$. Soit $x \in I$. Alors $a \leq x < b$ donc $x \in [a, b[$. Réciproquement, si $x \in [a, b[$, on a $x < b$, donc x n'est pas un majorant de I. Par suite, il existe $y \in I$ tel que $x \leq y < b$. Comme I est convexe, $x \in [a, y] \subset I$. Par suite, $I = [a, b[$.
- *3eme cas :* $a \notin I, b \in I$. Alors, par un raisonnement similaire au 2eme cas, on obtient $I =]a, b]$.
- *4eme cas :* $a \notin I, b \notin I$. Alors, par un raisonnement similaire au 2eme cas, on obtient $I =]a, b[$.

Dans tous les cas, la partie convexe I de $\mathbb{R}$ est un intervalle. □

Proposition 63.

Les parties connexes par arcs de $\mathbb{R}$ sont les intervalles.

Démonstration.

Montrons que pour toute partie I de $\mathbb{R}$, I est connexe par arcs, si, et seulement si, I est convexe. D'après l'exemple 19 et la proposition 5, si I est un intervalle alors I est convexe, donc I est connexe par arcs.

Réciproquement, si I est connexe par arcs, alors pour tous $x, y \in I$, il existe une fonction continue γ_{xy} de $[0, 1]$ dans $\mathbb{R}$ telle que $\gamma(0) = x$, $\gamma(1) = y$ et $\mathrm{Im}(\gamma) \subset I$.
Soit $x, y \in I$ et $z \in [x, y]$ avec $x < z < y$. On note $\Gamma_z = \{t \in [0, 1] \mid \gamma(t) \leq z\}$. Alors, comme $\gamma(0) = x \leq z$, $0 \in \Gamma_z$, donc Γ_z est non vide. De plus, Γ_z est majoré par 1, donc Γ_z possède une borne supérieure $M := \sup(\Gamma_z)$.
Par la caractérisation séquentielle de la borne supérieure, il existe (u_n) à valeurs dans Γ_z telle que $u_n \to M$. Or γ_{xy} est continue donc :

$$\gamma_{xy}(u_n) \xrightarrow[n\to\infty]{} \gamma_{xy}(M).$$

Par suite, $\gamma_{xy}(M) \leq z$. Comme $M = \sup(\Gamma_z)$, M appartient à la frontière $\mathrm{Fr}(\Gamma_z)$ de Γ_z. Par suite M est un point adhérent à $\Gamma_z^c = \{t \in [0, 1] \mid \gamma_{xy}(t) > z\}$ (qui est non vide : 1 y appartient), donc, comme précédemment, en considérant une suite à valeurs dans Γ_z^c qui converge vers M et la continuité de γ_{xz}, on obtient : $\gamma_{xy}(M) \geq z$.

Il en résulte que $[x, y] \subset \gamma_{xy}([0, 1]) \subset I$. Donc I est convexe. □

6. Image continue d'une partie connexe par arcs

Théorème 14.

Soit F un espace vectoriel normé, $A \subset E$ et $f : A \to F$ une application continue.
Si A est une partie connexe par arcs de E, alors $f(A)$ est une partie connexe par arcs de F.

Autrement dit : l'image directe d'un connexe par arcs par une application continue est connexe par arcs.

Démonstration.

On suppose A connexe par arcs. Soit $f(x), f(y) \in f(A)$ (où $x, y \in A$). Comme $x, y \in A$ connexe par arcs, il existe γ un chemin joignant x à y dans A. Alors l'application composée $f \circ \gamma$ est continue sur $[0,1]$, $f \circ \gamma(0) = f(x)$, $f \circ \gamma(1) = f(y)$ et $\mathrm{Im}(f \circ \gamma) \subset f(A)$. Donc $f \circ \gamma$ est un chemin joignant $f(x)$ à $f(y)$ dans $f(A)$. Il en résulte que $f(A)$ est connexe par arcs. $\square$

Corollaire 8. *Théorème des valeurs intermédiaires*

Soit $A \subset E$ et $f : A \to \mathbb{R}$ une application continue.
Si A est une partie connexe par arcs de E, alors $f(A)$ est un intervalle ; autrement dit, pour tous $x, y \in A$,

$$[f(x), f(y)] \subset f(A).$$

Démonstration.

On applique le théorème 6 au cas $F = \mathbb{R}$ en utilisant la proposition 5 qui assure qu'une partie connexe par arcs de $\mathbb{R}$ est un intervalle. $\square$

Exercice 61.

Montrer qu'il n'existe pas de bijection continue de $\mathbb{C}$ dans $\mathbb{R}$.

Correction.

Si f est une bijection continue de $\mathbb{C}$ dans $\mathbb{R}$, alors la restriction g de f sur $\mathbb{C} \setminus \{f^{-1}(0)\}$ est bijective et continue sur $\mathbb{C} \setminus \{f^{-1}(0)\}$ dans $\mathbb{R}^*$. Or $\mathbb{C} \setminus \{f^{-1}(0)\}$ est connexe par arcs, et $\mathbb{R}^*$ ne l'est pas. Contradiction car l'image d'un connexe par arcs par une application continue est connexe par arcs.

Partie H

Espaces vectoriels normés de dimension finie

Dans le cas particulier des espaces vectoriels de dimension finie, nous allons voir que la plupart des notions étudiées depuis le début du chapitre sont invariantes par changement de norme. Ainsi, en dimension finie, la topologie s'en trouve grandement simplifiée : l'étude des ouverts et fermés, de la convergence de suite, de limites ou de la continuité des fonctions,... ne dépendent pas du choix des normes !

Dans toute cette partie, n désigne un entier naturel non nul et $\mathbb{K}$ un corps égal à $\mathbb{R}$ ou $\mathbb{C}$.

1. Équivalence des normes en dimension finie

Lemme 2.

Soit $A \subset \mathbb{K}^n$ où $\mathbb{K}^n$ est muni de la norme infini. A est compact si, et seulement si, A est fermé borné.

Démonstration.

Voir la correction de l'exercice 55. □

Théorème 15.

Dans $\mathbb{K}^n$, toutes les normes sont équivalentes.

Démonstration Non exigible.

Soit $\|\cdot\|$ une norme sur $\mathbb{K}^n$. Montrons que $\|\cdot\|$ et la norme infini $\|\cdot\|_\infty$ sont équivalentes.
On note $(e_1, ..., e_n)$ la base canonique de $\mathbb{K}^n$.

- Montrons que $\|\cdot\|_\infty$ est plus fine que $\|\cdot\|$. Pour $x = (x_1, ..., x_n) \in \mathbb{K}^n$, on a :

$$\|x\| \leq \sum_{i=1}^{n} |x_i| \|e_i\| \leq C\|x\|_\infty$$

  où $C = \sum_{i=1}^n \|e_i\|$.
  Donc il existe $C > 0$ tel que $\|\cdot\| \leq C\|\cdot\|_\infty$.
- Montrons que $\|\cdot\|$ est plus fine que $\|\cdot\|_\infty$.
  1. Montrons tout d'abord que $\|\cdot\|$ est continue sur $\mathbb{K}^n$ muni de la norme $\|\cdot\|_\infty$. Pour tous $x, y \in \mathbb{K}^n$,
  $$|\|x\| - \|y\|| \leq \|x - y\| \leq C\|x - y\|_\infty;$$
  donc $\|\cdot\|$ est continue sur $\mathbb{K}^n$ car lipschitzienne sur $\mathbb{K}^n$.
  2. Montrons ensuite que la sphère unité S_∞ de la norme $\|\cdot\|_\infty$ est compacte dans $\mathbb{K}^n$ muni de $\|\cdot\|_\infty$. Toute sphère dans un espace vectoriel normé est fermée et bornée pour

sa propre norme, donc S_∞ est fermée bornée pour la norme $\|\cdot\|_\infty$. Donc, d'après le lemme 2, S_∞ est compacte dans $\mathbb{K}^n$ muni de la norme $\|\cdot\|_\infty$.

Utilisons désormais ces deux points précédents : Pour la norme $\|\cdot\|_\infty$, S_∞ est compact et $\|\cdot\|$ est continue de $\mathbb{K}^n$ dans $\mathbb{R}$, donc d'après le corollaire 3, $\|\cdot\|$ est bornée sur S_∞ et atteint ses bornes. En particulier, elle atteint sa borne inférieure m en $x_0 \in S_\infty$ et on remarque que $m > 0$ car $x_0 \neq (0, ..., 0)$.
Par suite, pour tout $x \in \mathbb{K}^n$ avec $x \neq (0, ..., 0)$, $\frac{x}{\|x\|_\infty} \in S_\infty$ et on a :

$$m \leq \|\frac{x}{\|x\|_\infty}\| = \frac{1}{\|x\|_\infty}\|x\|.$$

Ce qui implique, l'égalité suivante, vraie également pour $x = (0, ..., 0)$:

$$\|x\|_\infty \leq \frac{1}{m}\|x\|.$$

Donc $\|\cdot\|_\infty$ est dominée par $\|\cdot\|$.
Il en résulte que $\|\cdot\|$ et $\|\cdot\|_\infty$ sont équivalentes.
Ainsi, par transitivité, toutes les normes sont équivalentes sur $\mathbb{K}^n$. □

Théorème 16.

Soit E un espace vectoriel de dimension finie. Alors toutes les normes sur E sont équivalentes.

Démonstration.

Soit E un espace vectoriel de dimension n. Alors il existe un isomorphisme f de E dans $\mathbb{K}^n$.
Soit $\|\cdot\|_1$ et $\|\cdot\|_2$ des normes sur E. Alors d'après la proposition 11, $\|f(\cdot)\|_1$ et $\|f(\cdot)\|_2$ sont des normes sur $\mathbb{K}^n$. Par suite, d'après le théorème précédent, il existe $c, C > 0$ tels que :

$$c\|f(\cdot)\|_1 \leq \|f(\cdot)\|_2 \leq C\|f(\cdot)\|_1.$$

Or f est bijective donc en composant par f^{-1}, on obtient :

$$c\|\cdot\|_1 \leq \|\cdot\|_2 \leq C\|\cdot\|_1.$$

Il en résulte que $\|\cdot\|_2$ et $\|\cdot\|_1$ sont équivalentes. □

2. Conséquences topologiques

Remarque 29.

On vient de voir qu'en dimension finie, toutes les normes sont équivalentes. D'après les résultats obtenus dans le chapitre relatif à la comparaison de normes, on remarque que de nombreuses propriétés topologiques sont indépendantes de la norme choisie sur un espace vectoriel de dimension finie. Ainsi, les notions suivantes ne dépendent pas de la norme choisie :

— les ouverts, les fermés, les voisinages ;

— l'intérieur, l'adhérence, la frontière des parties ;

— les limites des suites/fonctions, la continuité des fonctions ;

— les parties bornées, compactes, connexes par arcs.

Il sera donc légitime que lorsqu'on parlera dans la suite d'un espace vectoriel normé, on ne précisera plus de quelle norme on munit cet espace et, de plus, on utilisera la norme la plus adaptée dans les démonstrations.

Proposition 64. *convergence et limites en dimension finie*

Soit E un espace vectoriel de dimension finie n sur $\mathbb{K}$, $\mathcal{B} = (e_1, ..., e_n)$ une base de E, $\ell = \sum_{i=1}^n \ell_i e_i \in E$, A une partie d'un espace vectoriel normé quelconque et $a \in A$.

- Soit $(u_k)_{k\in\mathbb{N}}$ une suite à valeurs dans E. On note $(u_k^{(1)}), ..., (u_k^{(n)})$ les suites des coordonnées dans $\mathcal{B}$, i.e. pour tout $k \in \mathbb{N}$:

$$u_k = \sum_{i=1}^n u_k^{(i)} e_i.$$

Alors la suite $(u_k)_{k\in\mathbb{N}}$ à valeurs dans E converge vers ℓ si, et seulement si, pour tout $i \in [\![1, n]\!]$, la suite $(u_k^{(i)})_{k\in\mathbb{N}}$ à valeurs dans $\mathbb{K}$ converge vers ℓ_i.

- Soit $f : A \to E$. On note $(f_1, ..., f_n)$ les applications coordonnées dans $\mathcal{B}$, i.e. pour tout $x \in A$:

$$f(x) = \sum_{i=1}^n f_i(x) e_i.$$

Alors, l'application f admet ℓ comme limite en a si, et seulement si, pour tout $i \in [\![1, n]\!]$, l'application f_i admet ℓ_i comme limite en a.

Démonstration.

On muni E de la norme $\|\cdot\|$ suivante : pour $x = \sum_{i=1}^n x_i e_i \in E$,

$$\|x\| = \max(|x_1|, ..., |x_n|).$$

L'application $f : x = \sum_{i=1}^n x_i e_i \mapsto (x_1, ..., x_n)$ est un isomorphisme de E dans $\mathbb{K}^n$ et on a pour tout $x \in E$, $\|x\| = \|f(x)\|_\infty$. En appliquant les résultats de convergence des suites dans un espace produit muni de la norme produit, on obtient :

$$(u_k)_{k\in\mathbb{N}} \text{ à valeurs dans } E \text{ converge vers } \ell\,;$$

si, et seulement si

$$(f(u_k))_{k\in\mathbb{N}} \text{ à valeurs dans } \mathbb{K}^n \text{ converge vers } f(\ell)\,;$$

si, et seulement si

$$((x_1, ..., x_n))_{k\in\mathbb{N}} \text{ à valeurs dans } \mathbb{K}^n \text{ converge vers } (\ell_1, ..., \ell_n)\,;$$

si, et seulement si

$$\text{pour tout } i \in [\![1,k]\!], \text{ la suite } (u_k^{(i)})_{k\in\mathbb{N}} \text{ à valeurs dans } \mathbb{K} \text{ converge vers } \ell_i.$$

Pour la seconde partie de la proposition, on utilise la caractérisation séquentielle de la limite et la première partie de la proposition. □

3. Compacité en dimension finie

Dans cette partie, on suppose connu le théorème de Bolzano-Weierstrass dans $\mathbb{R}$ (vu en 1ere année).

Lemme 3.

Dans $\mathbb{C}$ muni du module, toute boule fermée est compacte.

Démonstration.

Soit $R \geq 0$ et $B_f(R) = \{z \in \mathbb{C} \mid |z| \leq R\}$. Soit (z_n) une suite à valeurs dans $B_f(R)$, avec pour tout $n \in \mathbb{N}$, $z_n = x_n + iy_n$ (où $x_n, y_n \in \mathbb{R}$).
Alors pour tout $n \in \mathbb{N}$, $|x_n| \leq |z_n| \leq R$ et $|y_n| \leq |z_n| \leq R$, donc $(x_n), (y_n)$ sont deux suites à valeurs dans le compact $[-R, R]$.
Par suite, la suite $((x_n, y_n))_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{R}^2$, est à valeurs dans le compact $[-R, R]^2$, et donc possède une sous-suite $((x_{\varphi(n)}, y_{\varphi(n)}))_{n\in\mathbb{N}}$ qui converge vers $(x, y) \in [-R, R]^2$. De plus, on a pour tout $n \in \mathbb{N}$,

$$\sqrt{x_n^2 + y_n^2} = |z| \leq R,$$

donc, comme $(u, v) \mapsto \sqrt{u^2 + v^2}$ est continue, en passant à la limite dans cette inégalité, on obtient :

$$\sqrt{x^2 + y^2} = |z| \leq R.$$

Alors $z = x + iy$ appartient à $B_f(R)$ et de plus, la sous-suite $(x_{\varphi(n)} + iy_{\varphi(n)})$ de (z_n) converge vers z.
Il en résulte que $B_f(R)$ est compact. □

Théorème 17. *Théorème de Bolzano-Weierstrass*

Soit E un espace vectoriel de dimension finie. De toute suite à valeurs dans E bornée on peut extraire une sous-suite convergente.

Autrement dit : toute suite à valeurs dans E bornée possède au moins une valeur d'adhérence.

Démonstration.

Soit E un espace de dimension finie n et $\mathcal{B} = (e_1, ..., e_n)$ une base de E. On munit E de la norme $\|\cdot\|$ définie, pour $x = \sum_{i=1}^n x_i e_i \in E$, par $\|x\| = \max(|x_1|, ..., |x_n|)$.

Soit (u_k) une suite à valeurs dans E bornée et $(u_k^{(i)})_{k\in\mathbb{N}}$ ses suites coordonnées dans la base $\mathcal{B}$. Comme (u_n) est bornée, il existe $\mathbb{R} \geq 0$ tel que pour tout $k \in \mathbb{N}$, $\|u_k\| \leq R$, et donc, pour tout

$i \in [\![1, k]\!]$,

$$|u_k^{(i)}| \leq R.$$

Par suite, pour tout $i \in [\![1, k]\!]$, $(u_k^{(i)})_{k\in\mathbb{N}}$ est à valeurs dans $B_f(R)$ qui est compact (où $B_f(R)$ désigne $[-R, R]$ si $\mathbb{K} = \mathbb{R}$ ou $\{z \in \mathbb{C} \mid |z| \leq R\}$ si $\mathbb{K} = \mathbb{C}$).
Par suite, chaque $(u_k^{(i)})_{k\in\mathbb{N}}$ possède au moins une valeur d'adhérence.
D'après la proposition sur le convergence des suites en dimension finie, il en résulte que (u_k) possède au moins une valeur d'adhérence (attention, cette implication n'est pas triviale. exercice : prouver cette implication). □

Théorème 18.

Soit E un espace vectoriel de dimension finie. Les parties compacts de E sont les parties fermées bornées.

Démonstration.

Un compact est fermé borné.
Soit A une partie fermée bornée de E. Soit (u_n) une suite à valeurs dans A. Alors (u_n) est bornée donc, d'après le théorème précédent, il existe une sous-suite $(u_{\varphi(n)})$ qui converge. Or $(u_{\varphi(n)})$ est une suite à valeurs dans A fermée, donc elle converge dans A. Par suite, A est compact. □

Proposition 65.

Soit E un espace vectoriel de dimension finie. Tout sous-espace vectoriel de E est fermé.

Démonstration.

Soit F un sous-espace vectoriel de E. Soit (u_n) une suite à valeurs dans F qui converge vers $l \in E$.
Comme (u_n) converge, alors elle est bornée. Or F est de dimension finie donc d'après le théorème de Bolzano-Weierstrass appliqué dans l'espace vectoriel F, il existe une sous-suite de (u_n) qui converge dans F.
Or toute valeur d'adhérence d'une suite convergente est égale à sa limite, donc $l \in F$. Donc F est fermé dans E. □

Exercice 62.

Soit $E = \mathcal{F}_b(\mathbb{R}, \mathbb{R})$ muni de sa norme canonique.

1. Montrer que $F = \{f \in E \mid \exists R \geq 0,\ \forall x \notin [-R, R],\ f(x) = 0\}$ est un sous-espace vectoriel de E.
2. Montrer que F n'est pas fermé.

Correction.

1. La fonction nulle appartient à F et pour $\lambda, \mu \in \mathbb{R}$, et pour $f \in F$ nulle en dehors de $[-R_1, R_1]$, $g \in F$ nulle en dehors de $[-R_2, R_2]$, $\lambda f + \mu g$ est nulle en dehors de $[-\max(R_1, R_2), \max(R_1, R_2)]$. Donc F est un sous-espace vectoriel de E.
2. Soit $f : x \mapsto \frac{1}{1+x^2}$. Alors $f \in E$ et $f \notin F$. Considérons la suite de terme général

$$f_n : x \mapsto \begin{cases} f(x) & \text{si } x \in [-n, n] \\ 0 & \text{sinon.} \end{cases}$$

Alors,

$$\|f_n - f\|_\infty = \frac{1}{1+n^2} \xrightarrow[n\to\infty]{} 0.$$

Donc la suite (f_n) à valeurs dans F converge vers $f \notin F$. Donc F n'est pas fermé.

4. Continuité des applications linéaires, multilinéaires et polynomiales

a. Continuité des applications linéaires

Théorème 19.

Soit E, F des espaces vectoriels normés sur $\mathbb{K}$. Si E est de dimension finie, toute application linéaire de E dans F est continue.

Autrement dit, si E est de dimension finie, $\mathcal{L}_c(E, F) = \mathcal{L}(E, F)$.

Démonstration.

On suppose que E est de dimension finie n. Soit $f : E \to F$ une application linéaire et $\mathcal{B} = (e_1, ..., e_n)$ une base de E. On munit E de la norme $\|\cdot\|$ définie, pour $x = \sum_{i=1}^n x_i e_i \in E$, par $\|x\| = \max(|x_1|, ..., |x_n|)$. On note $K = \sum_{i=1}^n \|f(e_i)\|_F$.

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

$$\|f(x)\|_F = \|\sum_{i=1}^n x_i f(e_i)\|_F \leq \sum_{i=1}^n \underbrace{|x_i|}_{\leq \|x\|_F} \|f(e_i)\|_F \leq K\|x\|.$$

Par suite, f est continue. □

b. Continuité des applications multilinéaires

Théorème 20.

Soit E, F, G des espaces vectoriels sur $\mathbb{K}$. Si E et F sont de dimension finie, toute application bilinéaire de $E \times F$ dans G est continue.

Démonstration.

On suppose E de dimension n et F de dimension m. Soit $\mathcal{B}_E = (e_1, ..., e_n)$ une base de E et $\mathcal{B}_F = (\varepsilon_1, ..., \varepsilon_m)$ une base de F. On munit E et F des normes $\|\cdot\|_E$ et $\|\cdot\|_F$ suivante : pour $x = \sum_{i=1}^n x_i e_i \in E$ et $y = \sum_{j=1}^m y_j \varepsilon_j \in F$,

$$\|x\|_E = \max(|x_1|, ..., |x_n|) \quad \text{et} \quad \|y\|_E = \max(|y_1|, ..., |y_m|).$$

Soit $f : E \times F \to G$ une application bilinéaire. Pour $x = \sum_{i=1}^n x_i e_i \in E$ et $y = \sum_{j=1}^m y_j \varepsilon_j \in F$, on a :

$$f(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j f(e_i, \varepsilon_j),$$

d'où, pour $k = \sum_{i=1}^n \sum_{j=1}^m x_i y_j \|f(e_i, \varepsilon_j)\|_G$:

$$\|f(x, y)\|_G \leq k \|x\|_E \|y\|_F.$$

Par suite, f est continue sur $E \times F$. □

Exemple 20.

Soit E un espace euclidien, $(\cdot|\cdot)$ un produit scalaire sur E. Alors $(\cdot|\cdot) : E \times E \to \mathbb{R}$ est continue.

Théorème 21.

Soit $E_1, ..., E_k$ et F des espaces vectoriels sur $\mathbb{K}$. Si $E_1, ..., E_k$ sont de dimension finie, toute application multilinéaire de $E_1 \times ... \times E_k$ dans F est continue.

Démonstration.

On raisonne de la même manière que dans la démonstration du théorème 20. □

Exemple 21.

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension n et $\mathcal{B}$ une base de E. Alors l'application

$$\det_{\mathcal{B}} : E^n \to \mathbb{K}$$

est continue.
En effet, $\det_{\mathcal{B}} : (x_1, ..., x_n) \mapsto \det_{\mathcal{B}}(x_1, ..., x_n)$ est une application n-linéaire de E^n dans $\mathbb{K}$.

c. Continuité des fonctions polynomiales

Définition 46. *Fonction polynomiale*

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie n, $\mathcal{B} = (e_1, ..., e_n)$ une base de E et $f : E \to \mathbb{K}$ une application.

On dit que f est une **fonction polynomiale** s'il existe une famille $(\lambda_{k_1,\dots,k_n})_{(k_1,\dots,k_n)\in\mathbb{N}^n}$ de scalaires presque tous nuls telle que pour tout $x = \sum_{i=1}^n x_i e_i$,

$$f(x) = \sum_{(k_1,\dots,k_n)\in\mathbb{N}^n} \lambda_{k_1,\dots,k_n} x_1^{k_1} \dots x_n^{k_n}.$$

Dans ce cas, on dit que le **degré** de f est l'entier :

$$\deg(f) = \max\{k_1 + \dots + k_n \mid \lambda_{k_1,\dots,k_n} \neq 0\}.$$

Exemple 22.

Sur $\mathbb{R}^3$ muni de sa base canonique,

$$f : (x, y, z) \mapsto x^4 - y^2 z - 2z^3 + 5x^2y^2z^2$$

est une fonction polynomiale de degré 6.

Remarque 30.

Soit E un espace vectoriel de dimension n et $\mathcal{B}$ une base de E. On peut remarquer les faits suivants :

- Soit u l'isomorphisme canonique de $\mathbb{K}^n$ dans E muni de $\mathcal{B}$. Une application f est polynomiale sur E muni de $\mathcal{B}$, si et seulement si, $f \circ u$ est polynomiale sur $\mathbb{K}^n$ muni de sa base canonique.
- La définition de fonction polynomiale ne dépend pas de la base choisie.

Théorème 22.

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie n, $\mathcal{B}$ une base de E et $f : E \to \mathbb{K}$ une application.

Si f est polynomiale, alors f est continue sur E.

Démonstration.

D'après la remarque précédente, il suffit de montrer le résultat pour f une fonction polynomiale sur $\mathbb{K}^n$.

Une fonction polynomiale f sur $\mathbb{K}^n$ est une combinaison linéaire de produits d'applications coordonnées $(x_1, \dots, x_n) \mapsto x_i$ qui sont continues sur $\mathbb{K}^n$. Par suite, f est continue sur $\mathbb{K}^n$. $\square$

Exemple 23.

L'application

$$\det : \left| \begin{array}{ccc} M_n(\mathbb{K}) & \to & \mathbb{K} \\ M & \mapsto & \det(M) \end{array} \right.$$

est continue.

En effet, $\det((a_{i,j})_{1\leq i,j\leq n}) = \sum_{\sigma\in\mathcal{S}_n} \varepsilon(\sigma)\prod_{i=1}^{n} a_{\sigma(j),j}$, donc det est une application polynomiale. (On peut également, sans se soucier des coefficients polynomiaux de la fonction, raisonner par récurrence en utilisant le développement par rapport aux lignes (ou colonnes))

Exercice 63.

Soit $n \in \mathbb{N}^*$. Montrer que $GL_n(\mathbb{R})$ n'est pas connexe par arcs.

Correction.

L'application det : $GL_n(\mathbb{R}) \to \mathbb{R}$ est continue et d'image $\mathbb{R}^*$. Comme $\mathbb{R}^*$ n'est pas connexe par arcs, alors $GL_n(\mathbb{R})$ n'est pas connexe par arcs. En effet, l'image d'une partie connexe par arcs par une application continue est connexe par arcs.

Chapitre IV

Réduction des endomorphismes et des matrices carrées

Table des matières

Dans ce chapitre, n désigne un entier naturel non nul et E désigne un espace vectoriel sur $\mathbb{K}$ où $\mathbb{K}$ est un sous-corps de $\mathbb{C}$. On se limitera dans les manipulations au cas $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$.

Partie A
Rappels et compléments d'algèbre linéaire

1. Somme finie de sous-espaces vectoriels

Dans ce paragraphe, on généralise la notion de somme de deux sous-espaces vectoriels vue en Sup' au cas d'un nombre fini de sous-espaces.

Définition 1. *Somme finie de sous-espaces vectoriels*

Soit $m \in \mathbb{N}^*$ et $F_1, \dots, F_m$ des sous-espaces vectoriels de E. On appelle **somme de** $F_1, \dots, F_m$ et on note $F_1 + \dots + F_m$ ou encore $\sum_{i=1}^{m} F_i$, le sous-ensemble de E :

$$\sum_{i=1}^{m} F_i = \left\{ \sum_{i=1}^{m} x_i \mid \forall\, i \in [\![1, m]\!],\ x_i \in F_i \right\}.$$

Proposition 1.

Soit $m \in \mathbb{N}^*$ et $F_1, \dots, F_m$ des sous-espaces vectoriels de E. Alors

$$\sum_{i=1}^{m} F_i = \mathrm{Vect}\left(\bigcup_{i=1}^{m} F_i \right).$$

En particulier, $\sum_{i=1}^{m} F_i$ est un sous-espace vectoriel de E.

Démonstration.

On pose $F = \sum_{i=1}^{m} F_i$ et $U = \bigcup_{i=1}^{m} F_i$.
Montrons tout d'abord que F est un sous-espace vectoriel de E :
On a $F \subset E$ car pour tout $i \in [\![1, m]\!]$ et pour tout $x_i \in F_i \subset E$, E étant stable par combinaisons linéaires, $x_1 + \dots + x_m \in E$.

- On a $0_E = \sum_{i=1}^{m} \underbrace{0_E}_{\in F_i} \in F$ car pour chaque $i \in [\![1, m]\!]$, F_i est un sous-espace vectoriel de E et donc contient 0_E.
- Soit $\lambda, \mu \in \mathbb{K}$ et $x = x_1 + \dots + x_n, y = y_1 + \dots + y_n \in F$ avec, pour tout $i \in [\![1, m]\!]$,

$x_i, y_i \in F_i$. On a :

$$\lambda x + \mu y = \sum_{i=1}^{m} \underbrace{(\lambda x_i + \mu y_i)}_{\in F_i} \in F$$

car pour chaque $i \in [\![1, m]\!]$, F_i est un sous-espace vectoriel de E et donc est stable par combinaisons linéaires.

Par suite, F est un sous-espace vectoriel de E. *On pouvait également montrer que F est un sous-espace de E comme l'image directe du sous-espace vectoriel $F_1 \times ... \times F_m$ de E^m par l'application linéaire $f : (x_1, ..., x_m) \mapsto x_1 + ... + x_m$ de E^m dans E.*

Montrons que $F = \mathrm{Vect}(U)$. Soit $x \in U$. Alors il existe $j \in [\![1, m]\!]$ tel que $x \in F_j$. Pour chaque $i \in [\![1, m]\!]$, on pose $y_i = \begin{cases} 0_E & \text{si } i \neq j \\ x & \text{si } i = j \end{cases}$. Alors, pour tout $i \in [\![1, m]\!]$, $y_i \in F_i$ et donc

$$x = \sum_{i=1}^{m} y_i \in F$$

Par suite, $U \subset F$.

Maintenant, soit G un sous-espace vectoriel de E contenant U. Soit $x = x_1 + ... + x_m \in F$ avec, pour tout $i \in [\![1, m]\!]$, $x_i \in F_i$. Comme, pour tout $i \in [\![1, m]\!]$, $x_i \in U \subset G$ et G est stable par combinaisons linéaires, $x \in G$. Par suite, $F \subset G$.

Ainsi, F est le plus petit sous-espace vectoriel de E contenant U i.e. $F = \mathrm{Vect}(U)$. □

Corollaire 1.

Soit $m \in \mathbb{N}^*$ et $F_1, ..., F_m$ des sous-espaces vectoriels de E. Si pour chaque $i \in [\![1, m$, $\mathcal{F}_i$ est une famille génératrice de F_i, alors la famille $\mathcal{F}$ obtenue en concaténant, i.e. en mettant bout à bout, les familles $\mathcal{F}_1, ..., \mathcal{F}_m$ est une famille génératrice de la somme $F_1 + ... + F_m$.

Démonstration.

Comme tout élément de $\mathcal{F}$ appartient à $U = \bigcup_{i=1}^m F_i$, on a $\mathrm{Vect}(\mathcal{F}) \subset \mathrm{Vect}(U) = F_1 + ... + F_m$.

Pour l'inclusion réciproque, on remarque que pour tout $i \in [\![1, m$, tout élément de F_i est combinaison linéaire d'éléments de la famille $\mathcal{F}_i$ et donc de la famille $\mathcal{F}$. Par suite, $U \subset \mathrm{Vect}(\mathcal{F})$ qui est un sous-espace vectoriel de E. Or $\mathrm{Vect}(U)$ est le plus petit sous-espace vectoriel contenant U donc $F_1 + ... + F_m = \mathrm{Vect}(U) \subset \mathrm{Vect}(\mathcal{F})$.

Par suite, $\mathcal{F}$ est une famille génératrice de la somme. □

Définition 2. *Sous-espaces vectoriels en somme directe*

Soit $m \in \mathbb{N}^*$ et $F_1, ..., F_m$ des sous-espaces vectoriels de E.

On dit que $F_1, ..., F_m$ sont en **somme directe** si, pour tout $y \in \sum_{i=1}^m F_i$:

il existe un **unique** m-uplet $(x_1, ..., x_m) \in F_1 \times ... \times F_m$ tel que $y = x_1 + ... + x_m$;

autrement dit, y se décompose de manière unique sous la forme $y = x_1 + ... + x_m$ avec, pour tout $i \in [\![1, m]\!]$, $x_i \in F_i$.

Dans ce cas, la somme $\sum_{i=1}^m F_i$ est appelée **somme directe** de $F_1, ..., F_m$ et on note $\bigoplus_{i=1}^m F_i = \sum_{i=1}^m F_i$ ou encore $F_1 \oplus ... \oplus F_m = \sum_{i=1}^m F_i$.

Proposition 2.

Soit $m \in \mathbb{N}^*$ et $F_1, ..., F_m$ des sous-espaces vectoriels de E. Les assertions suivantes sont équivalentes :

i) $F_1, ..., F_m$ sont en somme directe ;

ii) pour tout $y = x_1 + ... + x_m \in \sum_{i=1}^m F_i$ avec, pour tout $i \in [\![1, m]\!]$, $x_i \in F_i$:
$$y = 0_E \text{ implique, pour tout } i \in [\![1, m]\!],\ x_i = 0_E.$$

iii) pour tout $i \in [\![1, m]\!]$, $\left(\sum_{\substack{j=1 \\ j \neq i}}^m F_j\right) \cap F_i = \{0_E\}$.

iv) pour tout $i \in [\![2, m]\!]$, $(F_1 + ... + F_{i-1}) \cap F_i = \{0_E\}$.

Démonstration.

i) $\Rightarrow$ ii) On suppose i). Soit $y = x_1 + ... + x_m \in \sum_{i=1}^m F_i$ avec, pour tout $i \in [\![1, m]\!]$, $x_i \in F_i$. On suppose $y = 0_E$. Comme les F_i sont des sous-espaces vectoriels de E et donc contiennent 0_E, celui-ci admet la décomposition $y = 0_E = \sum_{i=1}^m \underbrace{0_E}_{\in F_i}$ dans la somme $F_1 + ... + F_m$. La somme étant directe, par unicité de la décomposition de y dans la somme, on a, pour tout $i \in [\![1, m]\!]$, $x_i = 0_E$.

ii) $\Rightarrow$ iii) On suppose ii). Soit $i \in [\![1, m]\!]$. On pose $F = \sum_{\substack{j=1 \\ j \neq i}}^m F_j$. Soit $x \in F \cap F_i$. Comme $x \in F$, pour tout $j \in [\![1, m]\!]$ avec $j \neq i$, il existe $x_j \in F_j$ tels que $x = \sum_{\substack{j=1 \\ j \neq i}}^m x_j$. On pose $x_i = -x \in F_i$. Alors on a :
$$\sum_{j=1}^m x_j = x_i + \sum_{\substack{j=1 \\ j \neq i}}^m x_j = -x + x = 0_E$$
Ainsi, par hypothèse, pour tout $j \in [\![1, m]\!]$, $x_j = 0_E$. En particulier, on a $x = -x_i = -0_E = 0_E$.
Par suite, $F \cap F_j = \{0_E\}$.

iii) $\Rightarrow$ iv) On suppose iii). Soit $i \in [\![2, m]\!]$. Alors $F_1 + ... + F_{i-1} \subset F = \sum_{\substack{j=1 \\ j \neq i}}^m F_j$ et donc $(F_1 + ... + F_{i-1}) \cap F_i \subset F \cap F_i = \{0_E\}$. Par suite, $(F_1 + ... + F_{i-1}) \cap F_i = \{0_E\}$.

iv) $\Rightarrow$ i) On suppose iv). Soit $y \in \sum_{i=1}^{m} F_i$ et $y = x_1 + ... + x_m$, $y = x'_1 + ... + x'_m$ des décompositions de z dans la somme $\sum_{i=1}^{m} F_i$ où, pour tout $i \in [\![1, m]\!]$, $x_i, x'_i \in F_i$.
On a $0_E = y - y = \sum_{i=1}^{m}(x_i - x'_i)$ donc, les F_i étant des sous-espaces vectoriels de E et ainsi étant stables par combinaisons linéaires, $x'_m - x_m \in F_m$ et

$$x'_m - x_m = \sum_{i=1}^{m-1} \underbrace{(x_i - x'_i)}_{F_i} \in F_1 + ... + F_{m-1}$$

Donc $x'_m - x_m \in (F_1 + ... + F_{m-1}) \cap F_m = \{0_E\}$ par hypothèse. Par suite $x'_m = x_m$.
On réitère le même raisonnement de proche en proche pour $i = m-1, ..., 2$ pour obtenir $x'_i = x_i$. Puis pour $i = 1$, on arrive alors à $x_1 - x'_1 = 0_E$ d'où $x'_1 = x_1$.
Il en résulte que la décomposition de y dans la somme $F_1 + ... + F_m$ est unique.

□

Exemple 1.

— Dans $\mathbb{R}[X]$, les sous-espaces $F_i = \mathrm{Vect}(X^i)$ pour $i = 0, ..., m \in \mathbb{N}^*$ sont en somme directe et

$$\bigoplus_{i=0}^{m} F_i = \mathbb{R}_m[X].$$

— On considère $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ et pour $I \subset \mathbb{R}$, on note $F_I = \{f \in E \mid \forall\, x \notin I,\ f(x) = 0\}$ exercice : montrer que F_I est un sous-espace vectoriel de E. Alors les sous-espaces vectoriels $F_{]-\infty,0]}$, $F_{[1,8]}$, et $F_{[33,100]}$ sont en somme directe et

$$F_{]-\infty,0]} \oplus F_{[1,8]}) \oplus F_{[33,100]} = F_{]-\infty,0]\cup[1,8]\cup[33,100]}.$$

> On a $F_{]-\infty,0]} \cap F_{[1,8]} = \{\mathbf{0}\}$ car si une fonction f appartient à cette intersection, elle est nulle en dehors de $]-\infty, 0]$ et en dehors de $[1, 8]$ qui sont des intervalles disjoints.
>
> De plus, on a $F_{]-\infty,0]} + F_{[1,8]} = F_{]-\infty,0]\cup[1,8]}$.
> En effet, si $f = f_1 + f_2 \in F_{]-\infty,0]} + F_{[1,8]}$, alors, pour tout $x \notin]-\infty, 0] \cup [1, 8]$, $f(x) = 0$ car $x \notin]-\infty, 0]$, d'où $f_1(x) = 0$ et $x \notin [1, 8]$, d'où $f_2(x) = 0$. Ainsi, $F_{]-\infty,0]} + F_{[1,8]} \subset F_{]-\infty,0]\cup[1,8]}$.
> Et si $f \in F_{]-\infty,0]\cup[1,8]}$, on pose :
>
> $$f_1 : x \mapsto \begin{cases} f(x) & \text{si } x \in]-\infty, 0] \\ 0 & \text{sinon} \end{cases} \quad \text{et } f_2 : x \mapsto \begin{cases} f(x) & \text{si } x \in [1, 8] \\ 0 & \text{sinon} \end{cases}$$
>
> Alors $f_1 \in F_{]-\infty,0]}$ et $f_2 \in F_{[1,8]}$ et $]-\infty, 0]$ et $[1, 8]$ étant disjoints, $f = f_1 + f_2 \in F_{]-\infty,0]} + F_{[1,8]}$. Ainsi $F_{]-\infty,0]\cup[1,8]} \subset F_{]-\infty,0]} + F_{[1,8]}$.
> Maintenant, comme $]-\infty, 0] \cup [1, 8]$ et $[33, 100]$ sont disjoints, comme précédemment, on a :
>
> $$(F_{]-\infty,0]} + F_{[1,8]}) \cap F_{[33,100]} = F_{]-\infty,0]\cup[1,8]} \cap F_{[33,100]} = \{\mathbf{0}\}$$
>
> Il en résulte que $F_{]-\infty,0]}$, $F_{[1,8]}$, et $F_{[33,100]}$ sont en somme directe.

Exercice 1.

On note $S_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) \mid {}^tM = M\}$ (ensemble des matrices symétriques) et $A_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) \mid {}^tM = -M\}$ (ensemble des matrices antisymétriques).
Montrer que $M_n(\mathbb{R}) = S_n(\mathbb{R}) \oplus A_n(\mathbb{R})$.

Exercice 2.

On considère la matrice $M = \begin{pmatrix} 0 & 2 & -1 \\ 3 & 1 & -3 \\ 2 & 2 & -3 \end{pmatrix}$ et $f \in \mathcal{L}(\mathbb{R}^3)$ l'endomorphisme canoniquement associé à M.
Montrer que les sous-espaces vectoriels $\mathrm{Ker}(f - \mathrm{Id}_E)$, $\mathrm{Ker}(f + \mathrm{Id}_E)$, $\mathrm{Ker}(f + 2\mathrm{Id}_E)$ sont en somme directe et déterminer leur somme.

Correction.

On note $\mathcal{B}$ la base canonique de $\mathbb{R}^3$. Alors $M = \mathrm{Mat}_{\mathcal{B}}(f)$. On pose $F_1 = \mathrm{Ker}(f - \mathrm{Id}_E)$, $F_2 = \mathrm{Ker}(f + \mathrm{Id}_E)$, $F_3 = \mathrm{Ker}(f + 2\mathrm{Id}_E)$.
Soit $x = x_1 + x_2 + x_3 \in F_1 + F_2 + F_3$ avec $x_i \in F_i$. Alors on a :

$$f(x_1) = x_1; \qquad f(x_2) = -2x_2 \quad \text{et} \quad f(x_3) = -x_3.$$

et donc,

$$f^2(x_1) = x_1; \qquad f^2(x_2) = 4x_2 \quad \text{et} \quad f^2(x_3) = x_3.$$

Supposons $x = 0_{\mathbb{R}^3}$. Alors, par linéarité de f, $f(x) = 0_E = f^2(x)$. Par suite, on a le système :

$$\begin{cases} x_1 & + & x_2 & + & x_3 & = & 0_E \\ x_1 & + & -2x_2 & + & -x_3 & = & 0_E \\ x_1 & + & 4x_2 & + & x_3 & = & 0_E \end{cases}$$

On résout pour trouver $x_1 = x_2 = x_3 = 0_E$. Par suite, les sous-espaces sont en somme directe.
On peut remarquer qu'on a jamais utilisé la matrice M... nous y verrons plus clair dans la suite du chapitre !

Remarque 1.

Attention ! On peut montrer que si $F_1, ..., F_m$ sont en somme directe, alors pour tous $i, j \in [\![1, m]\!]$ avec $i \neq j$, $F_i \cap F_j = \{0_E\}$. Mais **la réciproque est fausse** comme on peut s'en apercevoir dans l'exercice suivant :

Exercice 3.

On considère l'espace vectoriel $\mathbb{R}^2$. Soit $F = \mathrm{Vect}((1, 0))$, $G = \mathrm{Vect}((0, 1))$ et $H = \mathrm{Vect}((1, 1))$.

1. Déterminer $F + G + H$.
2. (a) Déterminer les intersections deux à deux entre F, G et H.
   (b) La somme $F + G + H$ est-elle directe ?

(c) Que dire de la dernière affirmation de la remarque précédente.

3. Montrer la première affirmation de la remarque i.e. si $F_1, ..., F_m$ sont en somme directe, alors pour tous $i, j \in [\![1, m]\!]$ avec $i \neq j$, $F_i \cap F_j = \{0_E\}$.

Correction.

1. Soit $(x, y) \in \mathbb{R}^2$. On a :
$$(x, y) = x(1, 0) + y(0, 1) + 0(1, 1) \in F + G + H$$
donc $\mathbb{R}^2 \subset F + G + H$. L'inclusion réciproque est vraie car $F + G + H$ est un sous-espace vectoriel de $\mathbb{R}^2$.
Par suite, $F + G + H = \mathbb{R}^2$.
2. (a) Si $u \in F \cap G$, alors il existe $\lambda, v \in \mathbb{R}$ tels que $u = (\lambda, 0)$ et $u = (0, v)$ d'où $\lambda = 0$ et $v = 0$. Par suite $u = (0, 0)$. Ainsi, $F \cap G = \{(0, 0)\}$.
Par des raisonnements similaires, on trouve $G \cap H = \{(0, 0)\} = H \cap F$.
(b) La somme $F + G + H$ n'est pas directe car $(1, 1)$ admet dans $F + G + H$ les décompositions $(1, 1) = 0(1, 0) + 0(0, 1) + 1(1, 1)$ et $(1, 1) = 1(1, 0) + 1(0, 1) + 0(1, 1)$ qui sont différentes.
(c) Les intersections deux à deux des facteurs de la somme sont toutes réduites à 0_E mais la somme n'est pas directe. La réciproque de l'implication énoncée dans la remarque précédente est donc fausse, comme annoncé !
3. On suppose $F_1, ..., F_m$ en somme directe. Soit $i, j \in [\![1, m]\!]$ avec $i \neq j$. Quitte à les échanger, on peut supposer $j < i$. D'après le iv) de la proposition 2, on a $(F_1 + ... + F_{i-1}) \cap F_i = \{0_E\}$ Or comme $j < i$ on a $F_j \subset F_1 + ... + F_{i-1}$ donc $F_j \cap F_i \subset (F_1 + ... + F_{i-1}) \cap F_i = \{0_E\}$ d'où $F_j \cap F_i = \{0_E\}$.

Grâce à l'unicité de la décomposition dans une somme directe, on peut bien définir les applications qui à un vecteur de la somme associent chaque composante de sa décomposition :

Définition 3. *Projecteurs associés à une somme directe*

Soit $m \in \mathbb{N}^*$ et $F_1, ..., F_m$ des sous-espaces vectoriels de E en somme directe tels que $E = \bigoplus_{i=1}^{m} F_i$.

On appelle **projecteurs associés à la décomposition en somme directe** $E = \bigoplus_{i=1}^{m} F_i$ les applications $p_1, ..., p_m$ où, pour tout $i \in [\![1, m]\!]$,

$$\text{pour } x = x_1 + ... + x_m \in E \text{ avec, pour tout } j \in [\![1, m]\!],\ x_j \in F_i,$$
$$p_i(x) = x_i.$$

Proposition 3.

Soit $m \in \mathbb{N}^*$, $F_1, ..., F_m$ des sous-espaces vectoriels de E en somme directe, tels que $E = \bigoplus_{i=1}^{m} F_i$ et $p_1, ..., p_m$ les projecteurs associés. Alors :

- pour tout $i \in [\![1, m]\!]$, p_i est un projecteur ; plus précisément, p_i est la projection sur F_i parallèlement à $\bigoplus_{\substack{j=1 \\ j \neq i}}^{m} F_j$.
- $p_1 + ... + p_n = \mathrm{Id}_E$ et pour tout $i, j \in [\![1, m]\!]$ avec $i \neq j$, $p_i \circ p_j = \mathbf{0}$.

Proposition 4.

Soit $m \in \mathbb{N}^*$, $F_1, ..., F_m$ des sous-espaces vectoriels de E. On a :

$$\dim\left(\sum_{i=1}^{m} F_i\right) \leq \sum_{i=1}^{m} \dim(F_i).$$

Et il y a égalité si, et seulement si, la somme est directe.

Démonstration.

On considère l'application linéaire f de $F_1 \times ... \times F_m$ dans $F_1 + ... + F_m$ tel que $f : (x_1, ..., x_m) \mapsto x_1 + ... + x_m$.
Par définition de la somme, f est surjective, donc :

$$\dim\left(\sum_{i=1}^{m} F_i\right) \leq \dim(F_1 \times ... \times F_m) = \sum_{i=1}^{m} \dim(F_i).$$

De plus, comme f est surjective, il y a égalité si, et seulement si, f est injective d'après le théorème du rang (on est bien en dimension finie ici, car cet énoncé n'a aucun intérêt en dimension infinie !).
Or, f étant linéaire, f est injective si, et seulement si, pour tout $x = (x_1, ..., x_m) \in F_1 \times ... \times F_m$, $x_1 + ... + x_m = f(x) = 0_E$ implique $x = (0_E, ...0_E)$ i.e. pour tout $i \in [\![1, m]\!]$, $x_i = 0_E$; d'après la proposition 2, ceci est équivalent à $F_1, ..., F_m$ sont en somme directe.

$\square$

Définition-Proposition 4. *Base adaptée à une somme directe*

Soit $m \in \mathbb{N}^*$, $F_1, ..., F_m$ des sous-espaces vectoriels de E en somme directe. Si, pour chaque $i \in [\![1, m]\!]$, $\mathcal{B}_i$ une base de F_i, alors la famille $\mathcal{B}$ obtenue en concaténant, i.e. en mettant bout à bout, les bases $\mathcal{B}_1, ..., \mathcal{B}_m$ est une base $F_1 \oplus ... \oplus F_m$.
Une telle base $\mathcal{B}$ est appelée **base adpatée à la somme directe** $F_1 \oplus ... \oplus F_m$.

Démonstration.

$\square$

2. Matrices semblables

a. Matrices équivalentes

Définition 5. *Matrices équivalentes*

Soit $A, B \in M_n(\mathbb{K})$. On dit que A et B sont **équivalentes** s'il existe $P, Q \in GL_n(\mathbb{K})$ tels que

$$B = Q^{-1}AP.$$

Exercice 4.

Montrer que la relation "*être équivalentes*" est une relation d'équivalence sur $M_n(\mathbb{K})$.

Proposition 5.

Soit $A \in M_n(\mathbb{K})$. Alors $\mathrm{rg}(A) = r \in [\![0, n]\!]$ si, et seulement si, A est équivalente à la matrice

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

b. Matrices semblables

Définition 6. *Matrices semblables*

Soit $A, B \in M_n(\mathbb{K})$. On dit que A et B sont **semblables** s'il existe $P \in GL_n(\mathbb{K})$ tels que

$$B = P^{-1}AP.$$

Exemple 2.

Soit $\mathcal{B}$ et $\mathcal{B}'$ des bases de E et $u \in \mathcal{L}(E)$. On note $M = \mathrm{Mat}_{\mathcal{B}}(u)$ la matrice de u dans la base $\mathcal{B}$ et $M' = \mathrm{Mat}_{\mathcal{B}'}(u)$ la matrice de u dans la base $\mathcal{B}'$. Alors M et M' sont semblables, en effet, on a

$$M' = P^{-1}MP$$

où P est la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$.

Remarque 2.

- Deux matrices A et B qui sont semblables ont le même déterminant.
- Deux matrices A et B qui sont semblables ont la même trace. En vertu de cette remarque et de l'exemple ci-dessus, cela permet de définir la trace d'un endomorphisme : la trace d'un endomorphisme est la trace d'une matrice de cet endomorphisme dans une base quelconque.

3. Sous-espaces stables et endomorphismes induits

Définition 7. *Sous-espace stable*

Soit F un sous-espace vectoriel de E et $u \in \mathcal{L}(E)$. On dit que F est **stable** par u si $u(F) \subset F$, i.e. pour tout $x \in F$, $u(x) \in F$.

Exemple 3.

— Les sous-espaces vectoriels $\{0\}$ et E sont stables par tout endomorphisme de E.
— Une homothétie (i.e. $\lambda \mathrm{Id}_E$ où $\lambda \in \mathbb{K}$) stabilise tous les sous-espaces vectoriels de E.
— Une intersection ou une somme de sous-espaces stables par un endomorphisme u est un sous-espace stable par u.

Exercice 5.

1. On suppose que pour tout $x \in E$, la famille $(x, u(x))$ est liée. Montrer que u est une homothétie.
2. En déduire que les seuls endomorphismes qui stabilisent tous les sous-espaces vectoriels de E sont les homothéties.

Correction.

1. On suppose que pour tout $x \in E$, $(x, u(x))$ est liée. Alors pour tout $x \neq 0$, il existe un unique $\lambda_x \in \mathbb{K}$ tel que $u(x) = \lambda_x x$. Montrons que pour tous $x, y \in E$ non nuls, $\lambda_x = \lambda_y$.
   — *1er cas :* x et y sont colinéaires. Alors il existe $\mu \in \mathbb{K}$ tel que $x = \mu y$, d'où :
   $$\lambda_x x = u(x) = \mu u(y) = \mu \lambda_y y = \lambda_y x;$$
   donc $\lambda_x = \lambda_y$.
   — *2nd cas :* (x, y) est libre. Alors on a
   $$\lambda_{x+y}(x+y) = u(x+y) = u(x) + u(y) = \lambda_x x + \lambda_y y.$$
   Par suite,
   $$(\lambda_{x+y} - \lambda_x)x + (\lambda_{x+y} - \lambda_y)y = 0_E,$$
   or (x, y) est libre donc $\lambda_{x+y} - \lambda_x = 0$ et $\lambda_{x+y} - \lambda_y = 0$. Et donc $\lambda_x = \lambda x + y = \lambda_y$.
   Il en résulte qu'il existe $\lambda \in E$ tel que $u(x) = x$ (cette égalité étant trivialement vraie pour $x = 0_E$).
2. Une homothéties stabilise tous les sous-espaces vectoriels. Réciproquement, si u est un endomorphisme qui stabilise tous les sous-espaces vectoriels, alors, pour tout $x \in E$, u stabilise $\mathbb{K}x = \mathrm{Vect}(x)$. Ainsi, pour tout $x \in E$, $u(x) \in \mathbb{K}x$ i.e. $(x, u(x))$ est liée. Par suite, d'après la question précédente, u est une homothétie.

Proposition 6.

Soit F un sous-espace vectoriel de E, $(e_i)_{i\in I}$ une famille génératrice de F et $u \in \mathcal{L}(E)$. Alors F est stable par u si, et seulement si, pour tout $i \in I$, $u(e_i) \in F$.

Démonstration.

- $(\Rightarrow)$. On suppose F stable par u. Alors pour tout $x \in F$, $u(x) \in F$, donc en particulier, comme chaque $e_i \in F$ pour $i \in I$, $u(e_i) \in F$.
- $(\Leftarrow)$. On suppose que pour tout $i \in I$, $u(e_i) \in F$. Soit $x \in F$. Comme $(e_i)_{i\in I}$ est génératrice, alors il existe une famille $(\lambda_i)_{i\in I})$ presque tous nuls telle que $x = \sum_{i\in I} \lambda_i e_i$. Par suite, on a :

$$u(x) = \sum_{i\in I} \lambda_i \underbrace{u(e_i)}_{\in F},$$

donc, comme F est un sous-espace vectoriel, $u(x) \in F$. Il en résulte que F est stable par u.

□

Remarque 3.

Pour $x \in E$, $\mathbb{K}x$ est un sous-espace vectoriel de E. D'après la proposition précédente, ce sous-espace est stable par u si, et seulement si, il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$.

Proposition 7.

Soit $u, v \in \mathcal{L}(E)$. Si u et v commutent i.e. $u \circ v = v \circ u$, alors $\mathrm{Ker}(v)$ et $\mathrm{Im}(v)$ sont stables par u.

Démonstration.

On suppose que u et v commutent.

- Soit $x \in \mathrm{Ker}(v)$. Montrons que $u(x) \in \mathrm{Ker}(v)$. On a :

$$v(u(x)) = v \circ u(x) = u \circ v(x) = u(v(x)) = u(0_E) = 0_E$$

car $v \circ u = u \circ v$ et u est linéaire. Par suite, $\mathrm{Ker}(v)$ est stable par u.
- Soit $v(x) \in \mathrm{Im}(v)$ où $x \in E$. Montrons que $u(v(x)) \in \mathrm{Im}(v)$. On a :

$$u(v(x)) = u \circ v(x) = v \circ u(x) = v(u(x)) \in \mathrm{Im}(v)$$

car $u \circ v = v \circ u$ et $u(x) \in E$. Par suite $\mathrm{Im}(v)$ est stable par u.

□

Définition 8. *Endomorphisme induit*

Soit F un sous-espace vectoriel de E stable par un endomorphisme u de E. On appelle **endomorphisme induit par u sur F** et l'endomorphisme $u_F \in \mathcal{L}(F)$ défini par $u_F = u_{|F}$ i.e. pour tout $x \in F$

$$u_F(x) = u(x)$$

Proposition 8.

On suppose E de dimension finie n. Soit u un endomorphisme de F un sous-espace vectoriel de E de dimension p et $\mathcal{B} = (e_1, ..., e_n)$ une **base adaptée** à F i.e. $\mathcal{B}' = (e_1, ..., e_p)$ est une base de F.
Alors F est stable par u si, et seulement si, la matrice $M = \mathrm{Mat}_{\mathcal{B}}(u)$ dans la base $\mathcal{B}$ est triangulaire supérieure par bloc, i.e.

$$M = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$$

avec $A \in M_p(\mathbb{K})$.
Dans ce cas, A est la matrice $\mathrm{Mat}_{\mathcal{B}'}(u_F)$ de l'endomorphisme induit u_F par u sur F.

Démonstration.

On note $M = (m_{ij})_{1 \leq i,j \leq n} = \mathrm{Mat}_{\mathcal{B}}(u)$. Alors on a, pour $i \in [\![1, n]\!]$, $u(e_i) = \sum_{j=1}^n m_{ij} e_i$. On note : $A = (m_{ij})_{1 \leq i,j \leq p}$, $B = (m_{ij})_{\substack{p+1 \leq j \leq n \\ 1 \leq i \leq p}}$, $C = (m_{ij})_{p+1 \leq i,j \leq n}$ et $D = (m_{ij})_{\substack{1 \leq j \leq p \\ p+1 \leq i \leq n}}$.
On remarque que $(e_1, ..., e_p)$ est en particulier une famille génératrice de F.
Ainsi,

F est stable par u

si, et seulement si,

pour tout $j \in [\![1, p]\!]$, $u(e_j) = \sum_{j=1}^n m_{ij} e_i \in F$

si, et seulement si,

pour tout $j \in [\![1, p]\!]$ et $i \in [\![p+1, n]\!]$, $m_{ij} = 0$

si, et seulement si,

$D = (0)_{\substack{1 \leq j \leq p \\ p+1 \leq i \leq n}}$

si, et seulement si,

$$M = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$$

$\square$

Partie B

Éléments propres

1. Éléments propres d'un endomorphisme

a. Définitions

Définition 9. *Valeur/vecteur propre d'un endomorphisme*

Soit $u \in \mathcal{L}(E)$.

— On dit que $\lambda \in \mathbb{K}$ est une **valeur propre** de u s'il existe $x \in E$ **non nul** tel que

$$u(x) = \lambda x.$$

— Soit $\lambda \in \mathbb{K}$ une valeur propre de u. On dit que $x \in E$ est un **vecteur propre de** u **associé à** λ si :

$$x \neq 0_E \quad \text{et} \quad u(x) = \lambda x.$$

Proposition 9.

Soit $u \in \mathcal{L}(E)$, $\lambda \in \mathbb{K}$ et $x \in E \setminus \{0_E\}$.

— Le scalaire λ est une valeur propre de u si, et seulement si, $\mathrm{Ker}(u - \lambda \mathrm{Id}_E) \neq \{0_E\}$ - autrement dit, si, et seulement si, $u - \lambda \mathrm{Id}_E$ n'est pas injectif.

— Le vecteur x est un vecteur propre de u si, et seulement si, $u(x)$ est colinéaire à x.

Démonstration.

•

$$\lambda \text{ est une valeur propre de } u$$

si, et seulement si,

$$\text{il existe } x \in E \setminus \{0_E\} \text{ tel que } u(x) = \lambda x$$

si, et seulement si,

$$\text{il existe } x \in E \setminus \{0_E\} \text{ tel que } u - \lambda \mathrm{Id}_E(x) = 0_E$$

si, et seulement si,

$$\text{il existe } x \in E \setminus \{0_E\} \text{ tel que } x \in \mathrm{Ker}(u - \lambda \mathrm{Id}_E)(x)$$

si, et seulement si,

$$\mathrm{Ker}(u - \lambda \mathrm{Id}_E)(x) \neq \{0_E\}.$$

•

$$x \text{ est un vecteur propre de } u$$

si, et seulement si,

$$\text{il existe } \lambda \in \mathbb{K} \text{ tel que } u(x) = \lambda x$$

si, et seulement si,

$$u(x) \text{ et } x \text{ sont colinéaires.}$$

$\square$

Exercice 6.

Soit $f \in \mathcal{L}(\mathbb{R}^3)$ tel que $f : (x, y, z) \to (2y, 2x, 2z)$. Montrer que $(1, 1, 0)$, $(0, 0, 1)$ et $(1, -1, 0)$ sont des vecteurs propres de f. À quelle valeur propre chacun d'entre eux est-il associé ?

Démonstration.

On a :

$$f(1, 1, 0) = (2, 2, 0) = 2(1, 1, 0)$$
$$f(0, 0, 1) = (0, 0, 1) = 2(0, 0, 1)$$
$$f(1, -1, 0) = (-2, 2, 0) = (-2)(1, -1, 0)$$

Donc $(1, 1, 0)$ et $(0, 0, 1)$ sont des vecteurs propres de f associés à la valeur propre 2 et $(1, -1, 0)$ est un vecteur propre de f associé à la valeur propre -2. $\square$

Définition 10. *Sous-espace propre d'un endomorphisme*

Soit $u \in \mathcal{L}(E)$ et λ une valeur propre de u. On appelle **sous-espace propre de u associé à la valeur propre** λ le sous-espace vectoriel de E noté $E_\lambda(u)$ et défini par :

$$E_\lambda(u) = \mathrm{Ker}(u - \lambda \mathrm{Id}_E) = \{x \in E \mid u(x) = \lambda x\}.$$

Autrement dit, $E_\lambda(u)$ est l'ensemble contenant 0_E et l'ensemble des vecteurs propres associés à la valeur propre λ.

Définition 11. *Spectre d'un endomorphisme*

On suppose que E est de **dimension finie**. Soit $u \in \mathcal{L}(E)$. Le **spectre** de u, noté $\mathrm{Sp}(u)$, est l'ensemble des valeurs propres de u i.e.

$$\mathrm{Sp}(u) = \{\lambda \in \mathbb{K} \mid \exists x \in E \setminus \{0_E\},\ u(x) = \lambda x\}.$$

Remarque 4.

— le vecteur nul 0_E n'est JAMAIS un vecteur propre ! Par contre, il appartient à tout sous-espace propre.

— 0 est valeur propre de u si, et seulement si, u n'est pas injectif. Dans ce cas $E_0(u) = \mathrm{Ker}(u)$.

— Tout sous-espace propre $E_\lambda(u)$ est de dimension supérieure ou égale à 1.

Exercice 7.

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$ une valeur propre de u.

1. On suppose que $\lambda \neq 0$. Montrer que $E_\lambda(u) \subset \mathrm{Im}(u)$.
2. On suppose u inversible. Montrer que $\lambda \neq 0$ et que $\frac{1}{\lambda}$ est une valeur propre de u^{-1}. Que dire de $E_{\frac{1}{\lambda}}(u^{-1})$?

Correction.

1. Soit $x \in E_\lambda(u)$. Alors $u(x) = \lambda x$ et donc, par linéarité de u,
$$x = u(\frac{1}{\lambda}x) \in \mathrm{Im}(f).$$
Par suite, $E_\lambda(u) \subset \mathrm{Im}(u)$.
2. Comme $\mathrm{Ker}(u) = \{0_E\}$, 0 n'est pas valeur propre de u. Par suite, $\lambda \neq 0$.
Soit $x \in E_\lambda(u)$. Alors $u(x) = \lambda x$ et donc, par linéarité de u^{-1}, $x = u^{-1}(u(x)) = \lambda u^{-1}(x)$. Par suite, on a :
$$u^{-1}(x) = \frac{1}{\lambda}x,$$
et donc $\frac{1}{\lambda}$ est une valeur propre de u^{-1} et x est un vecteur propre de u^{-1} associé à $\frac{1}{\lambda}$. Ainsi, $E_\lambda(u) \subset E_{\frac{1}{\lambda}}(u^{-1})$. Et réciproquement, si $x \in E_{\frac{1}{\lambda}}(u^{-1})$, par un raisonnement similaire, on obtient $u(x) = \lambda x$. Il en résulte que
$$E_{\frac{1}{\lambda}}(u^{-1}) = E_\lambda(u).$$

b. Exemples

On applique les transformations suivantes à la première image. Déterminons les valeurs propres et leurs directions propres associées pour chacune des transformations. Une direction propre correspond à une direction qui reste inchangée après transformation et une valeur propre correspond à l'échelle de la modification (en tenant compte du changement de sens grâce au signe) après transformation dans la direction propre qui lui est associée.

Symétrie Dilatation

Exemple 4.

— Soit $\lambda \in \mathbb{K}$. Alors l'homothétie $\lambda \mathrm{Id}_E$ admet λ pour unique valeur propre et $E_\lambda(\lambda \mathrm{Id}_E) = E$.

— Une rotation non triviale (i.e. d'angle différent d'un multiple de π) dans le plan euclidien n'admet pas de valeur propre.

— Soit $p \in \mathcal{L}(E)$ un projecteur non trivial de E i.e. $p^2 = p$ et $p \neq 0, \mathrm{Id}_E$. Alors p admet pour valeurs propres 0 et 1 et on a :

$$E_0(p) = \mathrm{Ker}(p) \quad \text{et} \quad E_1(p) = \mathrm{Im}(p)$$

> Si λ est une valeur propre de p, alors pour x un vecteur propre de p associé à λ, on a :
> $$\lambda^2 x = p^2(x) = p(x) = \lambda x.$$
> Comme $x \neq 0_E$, on en déduit que $\lambda = 0$ ou $\lambda = 1$.
>
> Montrons que 0 et 1 sont bien valeurs propres de p :
>
> — Comme $p \neq \mathbf{0}$, il existe $x \in E$ tel que $p(x) \neq 0_E$. Ainsi, pour $y = p(x) \neq 0_E$, on a :
> $$p(y) = p(p(x)) = p(x) = y = 1.y \quad \text{car } p^2 = p$$
> Donc, y étant un vecteur non nul, 1 est bien valeur propre de p.
>
> — Comme $p \neq \mathrm{Id}_E$, il existe $x \in E$ tel que $p(x) \neq x$. Ainsi, pour $y = p(x) - x \neq 0_E$, on a, par linéarité de p :
> $$p(y) = p(p(x)) - p(x) = 0_E = 0.y \quad \text{car } p^2 = p$$
> Donc, y étant un vecteur non nul, 0 est bien valeur propre de p.
>
> Déterminons désormais les sous-espaces propres associés à 0 et 1 :
>
> - $\lambda = 0$. Pour tout endomorphisme qui admet 0 pour valeur propre, le sous-espace propre associé à 0 est égal à son noyau, donc $E_0(p) = \mathrm{Ker}(u)$.
> - $\lambda = 1$. Pour tout endomorphisme u qui admet $\lambda \neq 0$ pour valeur propre, $E_\lambda(u) \subset \mathrm{Im}(u)$. Par suite, $E_1(p) \subset \mathrm{Im}(p)$.
>
>   Réciproquement, pour $y = p(x) \in \mathrm{Im}(p)$ avec $x \in E$, on a :
> $$p(y) = p(p(x)) = p^2(x) = p(x) = y.$$
>   d'où $y \in E_1(p)$.
>   Par suite, $\mathrm{Im}(p) \subset E_1(p)$.
>   Il en résulte que $E_1(p) = \mathrm{Im}(p)$.

— Soit F, G deux sous-espaces supplémentaires non triviaux. La symétrie s par rapport à F parallèlement à G admet pour valeur propre 1 et -1 et et on a :

$$E_1(s) = F \quad \text{et} \quad E_{-1}(s) = G$$

> Soit p le projecteur sur F parallèlement à G. Alors on a $s = 2p - \mathrm{Id}_E$, donc, pour $\lambda \in \mathbb{K}$ et $x \in E$,
> $$s(x) = \lambda x \Leftrightarrow p(x) = \frac{1+\lambda}{2}x.$$
> Par suite, comme p est non trivial, d'après l'exemple précédent, s admet 1 et -1 pour valeurs propres et
> $$E_1(s) = E_1(p) = \mathrm{Ker}(p) = F$$
> et
> $$E_{-1}(s) = E_0(p) = \mathrm{Im}(p) = G.$$

— Soit $f \in \mathcal{L}(\mathbb{R}^2)$ tel que $f : (x, y) \mapsto (2x, x + y)$. Alors $\mathrm{Sp}(f) = \{2, 1\}$; et $E_2(f) = \mathrm{Vect}((1, 1))$ et $E_1(f) = \mathrm{Vect}((0, 1))$.

> On a, pour $\lambda \in \mathbb{K}$:
> $$(*)\ f(x,y) = \lambda(x,y) \Leftrightarrow \left\{ \begin{array}{ccccc} 2x & & & = & \lambda x \\ x & + & y & = & \lambda y \end{array} \right. \Leftrightarrow \left\{ \begin{array}{ccccc} (2-\lambda)x & & & = & 0 \\ x & + & (1-\lambda)y & = & 0 \end{array} \right.$$
> *1er cas :* $\lambda \neq 2$ et $\lambda \neq 1$. Alors $(*)$ est équivalent à
> $$\left\{ \begin{array}{ccc} x & = & 0 \\ y & = & 0 \end{array} \right.$$
> Donc, dans ce cas, $(x, y) = (0, 0)$ est la seule solution de $f(x, y) = \lambda(x, y)$ donc λ n'est pas valeur propre de f.
> *2eme cas :* $\lambda = 2$. Alors
> $$(*) \Leftrightarrow \left\{ \begin{array}{ccccc} x & - & y & = & 0 \end{array} \right.$$
> Donc, dans ce cas, l'ensemble des solutions de $f(x, y) = 2(x, y)$ est $\{(x, y) \mid x - y = 0\} = \mathrm{Vect}((1, 1)) \neq \{0_E\}$ donc $\lambda = 2$ est valeur propre de f et $E_2(f) = \mathrm{Vect}((1, 1))$.
>
> *2eme cas :* $\lambda = 2$. Alors
> $$(*) \Leftrightarrow \left\{ \begin{array}{ccc} x & = & 0 \end{array} \right.$$
> Donc, dans ce cas, l'ensemble des solutions de $f(x, y) = (x, y)$ est $\{(x, y) \mid x = 0\} = \mathrm{Vect}((0, 1)) \neq \{0_E\}$ donc $\lambda = 1$ est valeur propre de f et $E_1(f) = \mathrm{Vect}((0, 1))$.

— Soit $E = C^\infty(\mathbb{R})$ et $D \in \mathcal{L}(E)$ tel que $D : f \mapsto f'$. Alors pour tout $\lambda \in \mathbb{R}$, λ est une valeur propre et $x \mapsto e^{\lambda x}$ est un vecteur propre associé à λ.

Soit $\lambda \in \mathbb{R}$. Pour $f \in E$, on a $f \in \operatorname{Ker}(u - \lambda \mathrm{Id}_E)$ si, et seulement si, $f' - \lambda f = 0$, c'est à dire, f est solution de l'équation différentielle homogène $y' - \lambda y = 0$. Cette équation à pour ensemble de solution $\{x \mapsto C.e^{\lambda x} \mid C \in \mathbb{R}\} \neq \{\mathbf{0}\}$; donc λ est une valeur propre de D et on a :

$$E_\lambda(D) = \{x \mapsto C.e^{\lambda x} \mid C \in \mathbb{R}\} = \mathbb{R}.f$$

où f est le vecteur propre de D associé à λ défini par $f : x \mapsto e^{\lambda x}$.

Exercice 8.

1. Que dire de l'endomorphisme nul 0 ? de l'identité Id_E ? et, en général, d'une homothétie ?
2. Que dire d'une rotation dans $\mathbb{R}^3$?
3. Que dire de l'application $\Delta : P \to P'$ de $\mathbb{R}[X]$ dans lui-même ?

Correction.

1. Pour tout $x \in E \setminus \{0_E\}$, on a $0_E = 0(x) = \lambda x$ si, et seulement si, $\lambda = 0$, donc λ est la seule valeur propre de 0 et $E_0(0) = E$.
   Pour tout $x \in E \setminus \{0_E\}$, on a $x\mathrm{Id}_E = \lambda x$ si, et seulement si, $\lambda = 1$, donc λ est la seule valeur propre de Id_E et $E_1(\mathrm{Id}_E) = E$.
   En raisonnant de la même façon, on obtient que pour une homothétie $\alpha\mathrm{Id}_E$ avec $\alpha \in \mathbb{K}$, α est sa seul valeur propre et $E_\alpha(\alpha\mathrm{Id}_E) = E$.
2. Une rotation de $\mathbb{R}^3$ (d'angle différent d'un multiple de π) n'admet qu'une seule valeur propre. Il s'agit de la valeur propre 1 dont le sous-espace propre associé est l'axe de la rotation.
3. Pour $\lambda \in \mathbb{K}^*$, $P' = \lambda P$ implique $\deg(P) = \deg(P') = \deg(P) - 1$. Ainsi $P = 0$ est la seule solution de $P = \lambda P'$, donc si $\lambda \neq 0$, λ n'est pas une valeur propre de Δ.
   Pour $\lambda = 0$, $P' = 0$ a pour solutions les polynômes constants. Ainsi, 0 est la seule valeur propre de Δ et $E_0(\Delta) = \operatorname{Ker}(\Delta) = P = a_0 \mid a_0 \in \mathbb{K}$.

2. Propriétés des sous-espaces propres

Proposition 10.

Soit $u, v \in \mathcal{L}(E)$. Si u et v commutent, i.e. $u \circ v = v \circ u$ alors les sous-espaces propres de u sont stables par v et les sous-espaces propres de v sont stables par u.

Démonstration.

On suppose que u et v commutent. Comme u commute avec Id_E, alors, pour tout $\lambda \in \mathbb{K}$, u commute avec $v - \mathrm{Id}_E$. Ainsi, d'après la proposition 7, $\mathrm{Ker}(v - \lambda \mathrm{Id}_E)$ est stable par u. Par suite, si $\lambda \in \mathbb{K}$ est une valeur propre de u, $E_\lambda(v) = \mathrm{Ker}(v - \lambda \mathrm{Id}_E)$ est stable par u.

On raisonne de même pour la stabilité par v des sous-espaces propres de u. $\square$

Proposition 11.

Soit $u \in \mathcal{L}(E)$ et $\lambda, \mu \in \mathbb{K}$ des valeurs propres de u. Si $\lambda \neq \mu$, alors $E_\lambda(u)$ et $E_\mu(u)$ sont en somme directe i.e.

$$E_\lambda(u) \cap E_\mu(u) = \{0_E\}.$$

Démonstration.

Raisonnons par contraposée. On suppose que $E_\lambda(u) \cap E_\mu(u) \neq \{0_E\}$. Alors il $x \neq 0_E$ tel que $x \in E_\lambda(u)$ et $x \in E_\mu(u)$. Par suite, x est un vecteur propre de u associé à λ et à μ donc

$$\lambda x = u(x) = \mu x.$$

Or $x \neq 0_E$, donc $\lambda = \mu$. $\square$

Corollaire 2.

Soit $u \in \mathcal{L}(E)$ et $\lambda, \mu \in \mathbb{K}$ des valeurs propres de u. Si $\lambda \neq \mu$, alors, pour tous vecteurs propres x et y associés à λ et μ respectivement, la famille (x, y) est libre.

Démonstration.

On suppose $\lambda \neq \mu$. Soit $x \in E_\lambda(u) \setminus \{0_E\}$ et $y \in E_\lambda(u) \setminus \{0_E\}$. D'après la proposition précédente, $E_\lambda(u)$ et $E_\mu(u)$ sont en somme directe, donc (x, y) est libre.

Exercice : Soit F, G des sous-espaces vectoriels de E tels que F et G sont en somme directe. Montrer que toute famille (x, y) avec $x \in F \setminus \{0_E\}$ et $y \in G \setminus \{0_E\}$ est libre. $\square$

Proposition 12.

Soit $u \in \mathcal{L}(E)$, $k \in \mathbb{N}^*$ et $\lambda_1, ..., \lambda_k \in \mathbb{K}$ des valeurs propres de u. Si $\lambda_1, ..., \lambda_k$ sont toutes distinctes, alors les sous-espaces propres associés $E_{\lambda_1}(u), ..., E_{\lambda_k}(u)$ sont en somme directe.

Démonstration.

Montrons, par récurrence sur $\mathbb{N} \setminus \{0, 1\}$, que pour tout $k \in \mathbb{N} \setminus \{0, 1\}$, la propriété $\mathcal{P}_k =$"pour k-uplets $(\lambda_1, ..., \lambda_k)$ de valeurs propres distinctes de u, les sous-espaces propres associés sont en somme directe".
L'initialisation $k = 2$ est donnée par la proposition préccédente.

Hérédité : Soit k un entier plus grand que 2. On suppose $\mathcal{P}_k$ vraie.
Soit $\lambda_1, ..., \lambda_{k+1}$ des valeurs propres distinctes de u. Soit $x = x_1 + ... + x_{k+1} \in \sum_{i=1}^{k+1} E_{\lambda_i}(u)$ où, pour tout $i \in [\![1, k+1]\!]$, $x_i \in E_{\lambda_i}(u)$. On suppose $x = 0_E$. Alors, par linéarité de u, on a d'une part $u(x) = 0_E$ et d'autre part :

$$u(x) = \sum_{i=1}^{k+1} u(x_i) = \sum_{i=1}^{k+1} \lambda_i x_i.$$

Par suite, on a :

$$0_E = u(x) - \lambda_{k+1} x = \sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) x_i.$$

Or, pour tout $i \in [\![1, k]\!]$, $E_{\lambda_i}(u)$ étant un sous-espace vectoriel, $(\lambda_i - \lambda_{k+1}) x_i \in E_{\lambda_i}(u)$.
Ainsi, $\sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) x_i$ appartient à la somme $\bigoplus_{i=1}^{k} E_{\lambda_i}(u)$ qui est bien directe par hypothèse de récurrence. Ainsi, comme $\sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) x_i = 0_E$, on a pour tout $i \in [\![1, k]\!]$, $(\lambda_i - \lambda_{k+1}) x_i = 0_E$, d'où $x_i = 0_E$ car $\lambda_i - \lambda_{k+1} \neq 0$.
Et de plus, on a alors, $x_{k+1} = x = 0_E$, d'où, pour tout $i \in [\![1, k+1]\!]$, $x_i = 0_E$.
Il en résulte que la somme $\sum_{i=1}^{k+1} E_{\lambda_i}(u)$ est directe.
Ce qui achève le raisonnement par récurrence. Ainsi, pour tout entier $k \geq 2$, $\mathcal{P}_k$ est vraie. □

Corollaire 3.

Soit $u \in \mathcal{L}(E)$, $k \in \mathbb{N}$ et $\lambda_1, ..., \lambda_k \in \mathbb{K}$ des valeurs propres de u. Si $\lambda_1, ..., \lambda_k$ sont deux à deux distinctes, alors

$$\sum_{i=1}^{k} \dim(E_{\lambda_i}(u)) \leq \dim(E).$$

Démonstration.

On suppose $\lambda_1, ..., \lambda_k$ deux à deux distinctes. Alors les sous-espaces propres $E_{\lambda_i}(u)$ sont en somme directe et on a :

$$\sum_{i=1}^{k} \dim(E_{\lambda_i}(u)) = \dim\left(\bigoplus_{i=1}^{k} E_{\lambda_i}(u)\right) \leq \dim(E).$$

□

Théorème 1.

On suppose E de dimension finie n. Tout endomorphisme u de E admet au plus n valeurs propres distinctes ; autrement dit :

$$\#\mathrm{Sp}(u) \leq n.$$

Démonstration.

Soit $u \in \mathcal{L}(E)$. On suppose par l'absurde que $\#\mathrm{Sp}(u) > n$. Alors il existe $n+1$ valeurs propres de u deux à deux distinctes $\lambda_1, ..., \lambda_{n+1}$. Pour chaque sous-espace propre $E_{\lambda_i}(u)$, on a $\dim(E_{\lambda_i}(u)) \geq 1$, donc :

$$n+1 \leq \sum_{i=1}^{k} \dim(E_{\lambda_i}(u)) \leq n.$$

Contradiction. Par suite $\#\mathrm{Sp}(u) \leq n$. □

Remarque 5.

Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E stable par u. Les valeurs propres de l'endomorphisme $u_F \in \mathcal{L}(F)$ induit par u sur F sont les valeurs propres λ de u telles que $E_\lambda(u) \cap F \neq \{0\}$. Dans ce cas,

$$E_\lambda(u_F) = E_\lambda(u) \cap F.$$

3. Éléments propres d'une matrice carrée

a. Définitions

Définition 12. *Éléments propres d'une matrice*

Soit $A \in M_n(\mathbb{K})$.

— On dit que $\lambda \in \mathbb{K}$ est une **valeur propre** de A s'il existe $X \in M_{n,1}(\mathbb{K})$ **non nulle** telle que

$$AX = \lambda X.$$

— Soit $\lambda \in \mathbb{K}$ une valeur propre de A. On dit que $X \in M_{n,1}(\mathbb{K})$ est un **vecteur propre de** A **associé à** λ si :

$$X \neq 0_{n,1} \quad \text{et} \quad AX = \lambda X.$$

— Soit $\lambda \in \mathbb{K}$ une valeur propre de A. On appelle **sous-espace propre associé de** A **à** λ le sous-espace vectoriel noté $E_\lambda(A)$ de $M_{n,1}(\mathbb{K})$ défini par :

$$E_\lambda(A) = \mathrm{Ker}(A - \lambda I_n) = \{X \in M_{n,1}(\mathbb{K}) \mid AX = \lambda X\}.$$

— On appelle **spectre** de A et on note $\mathrm{Sp}(A)$, l'ensemble des valeurs propres de A.

Remarque 6.

Soit $A \in M_n(\mathbb{K})$. On remarque que $\lambda \in \mathrm{Sp}(A)$ si, et seulement si, $A - \lambda I_n \notin GL_n(\mathbb{K})$.

Exercice 9.

Déterminer les valeurs, vecteurs et sous-espaces propres de $A = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in M_n(\mathbb{K})$.

Correction.

$\lambda \in \mathbb{K}$ est valeur propre de A si, et seulement si, pour tout $j \in [\![1, n]\!]$, il existe $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \neq 0_{n,1}$ tel que :

$$\sum_{i=1}^{n} x_i = \lambda x_j.$$

Pour $\lambda = 0$, ces équations deviennent : $\sum_{i=1}^{n} x_i = 0$; dont il existe des solutions non nulles. Par suite, 0 est une valeur propre de A et on a

$$E_0(A) = \operatorname{Ker}(A) = \{X \mid \sum_{i=1}^{n} x_i = 0\} = \{X = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ -\sum_{i=1}^{n-1} x_i \end{pmatrix}\}$$

Pour $\lambda \neq 0$, on obtient $x_i = x_j$ $\forall i, j \in [\![1, n]\!]$. Ainsi, ce système d'équations admet des solutions non nulle dans le seul cas $\lambda = n$. Par suite, la deuxième et dernière valeur propre de A est n et on a

$$E_n(A) = \operatorname{Ker}(A - nI_n) = \{\begin{pmatrix} \alpha \\ \vdots \\ \alpha \end{pmatrix} \mid \alpha \in \mathbb{K}\} = \mathbb{K}.\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

On remarque également que $\dim(E_0(A)) = n - 1$ et que $\dim(E_n(A)) = 1$ donc on a :

$$M_{n,1}(\mathbb{K}) = E_0(A) \oplus E_n(A).$$

b. Propriétés du spectre d'une matrice

Proposition 13.

On suppose E de dimension finie n. Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de E, $u \in \mathcal{L}(E)$ et $A = \operatorname{Mat}_{\mathcal{B}}(u) \in M_n(\mathbb{K})$. Alors on a $\operatorname{Sp}(A) = \operatorname{Sp}(u)$.
De plus, pour tout $\lambda \in \operatorname{Sp}(u)$,

$$x = \sum_{i=1}^{n} x_i e_i \in E_\lambda(u) \text{ si, et seulement si, } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in E_\lambda(A).$$

Démonstration.

L'application $\varphi_{\mathcal{B}} : E \to M_{n,1}(\mathbb{K})$ tel que

$$\varphi_{\mathcal{B}} : x = \sum_{i=1}^{n} x_i e_i \mapsto X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

est un isomorphisme.
Ainsi, l'équation $MX = \lambda X$ est équivalente à l'équation $u(x) = \lambda x$ d'où le résultat. $\square$

Proposition 14.

Soit $A, B \in M_n(\mathbb{K})$. Si A et B sont semblables, alors $\mathrm{Sp}(A) = \mathrm{Sp}(B)$ et pour tout $\lambda \in \mathrm{Sp}(A)$, $E_\lambda(A) = PE_\lambda(B)$ où $P \in GL_n(\mathbb{K})$ vérifie $B = P^{-1}AP$.

Démonstration.

On peut voir deux matrices semblables comme les matrices d'un même endomorphisme dans deux bases différentes. On obtient alors le résultat souhaité en appliquant la proposition précédente. $\square$

Proposition 15.

Soit $\mathbb{K}'$ un sous-corps de $\mathbb{K}$ et $A \in M_n(\mathbb{K}')$. Alors $\mathrm{Sp}_{\mathbb{K}'}(A) \subset \mathrm{Sp}_{\mathbb{K}}(A)$.

Démonstration.

Soit $\lambda \in \mathbb{K}'$ une valeur propre de $A \in M_n(\mathbb{K}') \subset M_n(\mathbb{K})$ et $X \in \mathbb{M}_{n,1}(\mathbb{K}')$ un vecteur propre de A. Comme $\mathbb{M}_{n,1}(\mathbb{K}') \subset \mathbb{M}_{n,1}(\mathbb{K})$, alors X vu comme matrice à coefficients dans $\mathbb{K}$ vérifie l'équation $AX = \lambda X$. Donc $\lambda \in \mathrm{Sp}_{\mathbb{K}}(A)$. $\square$

Exercice 10.

Illustrer le résultat précédent en déterminant les spectres dans $\mathbb{R}$ puis dans $\mathbb{C}$ de $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

Démonstration.

Soit $\lambda \in \mathbb{K}$. On a, pour $\big(x, y\big) \in \mathbb{M}_{n,1}(\mathbb{K})$,

$$M\,\big(x, y\big) = \lambda X \;\Leftrightarrow\; \begin{cases} (1+\lambda^2)x = 0 \\ (1+\lambda^2)y = 0 \end{cases}$$

Par suite, si $\mathbb{K} = \mathbb{R}$, $1+\lambda^2 > 0$, l'unique solution de ce système est $\big(0, 0\big)$ (et ce, pour toute valeur de λ). Donc $\mathrm{Sp}_{\mathbb{R}}(M) = \emptyset$.

Dans le cas où $\mathbb{K} = \mathbb{C}$, on a $1 + \lambda^2 = 0$ si, seulement si $\lambda = \pm i$. Ainsi, $M\left(x, y\right) = \lambda X$ possède des solutions non nulles si, et seulement si, $\lambda = \pm i$. Les valeurs propres de M dans $\mathbb{C}$ sont donc i et $-i$, d'où $\mathrm{Sp}_{\mathbb{C}}(M) = \{i, -i\}$.

$\square$

Partie C
Polynôme caractéristique

Dans cette partie, l'espace vectoriel E est supposé de dimension finie n.

1. Polynôme caractéristique

a. Polynôme caractéristique d'une matrice carrée

L'application $M \mapsto \det(M)$ est une fonction polynomiale en les coefficients de M. Ainsi, pour $A \in M_n(\mathbb{K})$ fixée, l'application $\lambda \mapsto \det(\lambda I_n - A)$ est une fonction polynomiale de la variable λ ; ce qui justifie la définition suivante :

Définition 13. *Polynôme caractéristique d'une matrice carrée*

Soit $A \in M_n(\mathbb{K})$. On appelle **polynôme caractéristique de** A et on note $\chi_A(X)$ l'unique polynôme de $\mathbb{K}[X]$ tel que, pour tout $\lambda \in \mathbb{K}$:

$$\chi_A(\lambda) = \det(\lambda I_n - A).$$

Remarque 7.

On notera directement $\chi_A(X) = \det(XI_n - A)$. Pour justifier cette notation, il faudrait pouvoir définir le déterminant d'une matrice à coefficients polynomiaux. Et c'est possible : au lieu d'utiliser le corps de base $\mathbb{K}$ pour les coefficients, on utilise le corps $\mathbb{K}(X)$ des fractions rationnelles. La théorie reste la même.

Proposition 16.

Soit $A \in M_n(\mathbb{K})$. Le polynôme caractéristique χ_A est un polynôme unitaire de degré n et on a :

$$\chi_A(X) = X^n - \mathrm{Tr}(A)X^{n-1} + ... + (-1)^n \det(A).$$

Démonstration.

On a, pour $\lambda \in \mathbb{K}$,

$$\chi_A(\lambda) = \det(\lambda I_n - A) = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & & \vdots \\ \vdots & & \ddots & \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{pmatrix}$$

En utilisant la base canonique $\mathcal{B} = (e_1, ..., e_n)$ de $M_{n,1}(\mathbb{K})$, on a, pour $C_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \in M_{n,1}(\mathbb{K})$:

$$\chi_A(\lambda) = \det_{\mathcal{B}}(\lambda e_1 - C_1, ..., \lambda e_n - C_n).$$

L'application $\det_{\mathcal{B}}$ est multilinéaire, donc en développant l'expression précédente on remarque que l'on obtient un polynôme de degré au plus n et on a, pour $0 \leq k \leq n$ où c_{n-k} est le coefficient de $\chi_A(\lambda)$ correspondant à λ^{n-k} :

$$\begin{aligned} c_{n-k}\lambda^{n-k} &= \sum_{i_1 < \ldots < i_k \in [\![1,n]\!]} \det_{\mathcal{B}}(\lambda e_1, ..., -C_{i_1}, ..., \lambda e_j, ..., -C_{i_k}, ..., \lambda e_n) \\ &= (-1)^k \lambda^{n-k} \sum_{i_1 < \ldots < i_k \in [\![1,n]\!]} \det_{\mathcal{B}}(e_1, ..., C_{i_1}, ..., e_j, ..., C_{i_k}, ..., e_n). \end{aligned}$$

Par suite, on obtient le résultat en évaluant, c_{n-k} pour $k = 0, 1$ et n :

- $c_n = \det_{\mathcal{B}}(e_1, ..., e_n) = 1$
- $c_{n-1} = -\sum_{i=1}^{n} \det_{\mathcal{B}}(e_1, ..., C_i, ..., e_n) = -\sum_{i=1}^{n} a_{ii} = -\mathrm{Tr}(A)$.
- $c_0 = (-1)^n \det_{\mathcal{B}}(C_1, ..., C_n) = \det(A)$.

$\square$

Exercice 11.

Soit $A = (a_{ij}) \in M_3(\mathbb{K})$. Exprimer le coefficient c_1 du monôme de degré 1 dans $\chi_A(X)$ en fonction des a_{ij}.

Correction.

On utilise les notations de la démonstrations précédente :

$$\begin{aligned} c_1 &= \det_{\mathcal{B}}(C_1, C_2, e_3) + \det_{\mathcal{B}}(C_1, e_2, C_3) + \det_{\mathcal{B}}(e_1, C_2, C_3) \\ &= \begin{vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & a_{13} \\ a_{21} & 1 & a_{23} \\ a_{31} & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}. \end{aligned}$$

Théorème 2.

Soit $A \in M_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. Le scalaire λ est une valeur propre de A si, et seulement si, λ est une racine du polynôme caractéristique de A. Autrement dit :

$$\lambda \in \mathrm{Sp}(A) \iff \chi_A(\lambda) = 0.$$

Démonstration.

$\lambda \in \mathrm{Sp}(A)$

si, et seulement si,

$\mathrm{Ker}(A - \lambda I_n) \neq \{0\}$

si, et seulement si,

$A - \lambda I_n \notin GL_n(\mathbb{K})$

si, et seulement si,

$\lambda I_n - A \notin GL_n(\mathbb{K})$

si, et seulement si,

$\det(\lambda I_n - A) = 0$

si, et seulement si,

$\chi_A(\lambda) = 0$.

□

Corollaire 4.

Soit $A \in M_n(\mathbb{K})$. *On rappelle que A possède au plus n valeurs propres.*

— Si $\mathbb{K} = \mathbb{C}$, alors A a au moins une valeur propre.

— Si $\mathbb{K} = \mathbb{R}$ et n est impair, alors A a au moins une valeur propre.

Démonstration.

On note χ_A le polynôme caractéristique de A.

— Si $\mathbb{K} = \mathbb{C}$, d'après le théorème de D'Alembert-Gauss, χ_A possède au moins une racine, donc d'après le théorème 2, A possède au moins une valeur propre.

— Si $\mathbb{K} = \mathbb{R}$ et n est impair, on a $\deg(\chi_A) = n$. Par suite, en appliquant le théorème des valeurs intermédiaires ou en raisonnant en terme de facteurs irréductibles, on peut montrer que χ_A possède au moins une racine, donc d'après le théorème 2, A possède au moins une valeur propre.

□

Méthode : Calcul des éléments propres d'une matrice $A \in M_n(\mathbb{K})$ dans le corps $\mathbb{K}$.

- On calcule le polynôme caractéristique χ_A de A.
- On factorise dans $\mathbb{K}$ le polynôme caractéristique χ_A de A et on détermine toutes ses racines.

- Chaque racine $\lambda \in \mathbb{K}$ de χ_A étant une valeur propre de χ_A, on résout le système

$$MX = \lambda X,$$

qui, NÉCESSAIREMENT, admet une infinité de solution (car λ est une valeur propre de A).

- Pour chaque racine λ de χ_A, le sous-espace propre associé à λ est égal à l'ensemble des solutions du système précédent :

$$E_\lambda(A) = \{X \in M_{n,1}(\mathbb{K}) \mid MX = \lambda X\}.$$

En pratique, on cherchera une base $(X_1, ..., X_k)$ de l'ensemble des solutions de $MX = \lambda X$ i.e. une famille libre maximale de vecteurs propres associés à λ, afin d'écrire :

$$E_\lambda(A) = \text{Vect}(X_1, ..., X_n).$$

Exercice 12.

Calculer les valeurs propres et les sous-espaces propres des matrices suivantes dans $\mathbb{R}$ puis dans $\mathbb{C}$:

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 1 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 6 & 0 \\ -2 & 0 & 7 \end{pmatrix} \quad D = \begin{pmatrix} 0 & 2 & -2 \\ \frac{3}{2} & 2 & -4 \\ 2 & 2 & -4 \end{pmatrix}$$
$$E = \begin{pmatrix} 1 & 2 & 2 \\ 4 & 2 & 8 \\ -2 & 2 & -4 \end{pmatrix}$$

Correction.

1. $\chi_A = X^2 - 3X$, d'où $\text{Sp}(A) = \{0, 3\}$ et on a :

$$E_0(A) = \text{Vect}(\begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix}) \text{ et } E_3(A) = \text{Vect}(\begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

2. $\chi_B = X^3 - X^2 - 3X - 1$, d'où $\text{Sp}(B) = \{-1, 1-\sqrt{2}, 1+\sqrt{2}\}$ et on a :

$$E_1(B) = \text{Vect}(\begin{pmatrix} 1 \\ \frac{1}{2} \\ -\frac{5}{2} \end{pmatrix}), \ E_{1-\sqrt{2}}(B) = \text{Vect}(\begin{pmatrix} 0 \\ 1 \\ \sqrt{2}-1 \end{pmatrix}) \text{ et } E_{1+\sqrt{2}}(B) = \text{Vect}(\begin{pmatrix} 0 \\ 1 \\ -\sqrt{2}-1 \end{pmatrix})$$

3. $\chi_C = X^3 - 15X^2 + 72X - 108$, d'où $\text{Sp}(B) = \{3, 6\}$ et on a :

$$E_3(C) = \text{Vect}(\begin{pmatrix} 1 \\ 0 \\ \frac{1}{2} \end{pmatrix}) \text{ et } E_6(C) = \text{Vect}(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix})$$

4. $\chi_D = X^3 + 2X^2 + X + 2$, d'où $\text{Sp}_{\mathbb{R}}(D) = \{-2\}$ et $\text{Sp}(D) = \{\pm i\}$. On a :

$$E_{-2}(D) = \text{Vect}(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix})$$

et dans le cas de $\mathbb{C}$, on a de plus :

$$E_i(D) = \mathrm{Vect}(\begin{pmatrix} 1 \\ \frac{1}{2}(i+2) \\ 1 \end{pmatrix}) \text{ et } E_{-i}(D) = \mathrm{Vect}(\begin{pmatrix} 1 \\ -\frac{1}{2}(i+2) \\ 1 \end{pmatrix})$$

5. $\chi_E = X^3 + X^2 - 30X$, d'où $\mathrm{Sp}(E) = \{-6, 0, 5\}$ et on a :

$$E_{-6}(E) = \mathrm{Vect}(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}),\ E_0(E) = \mathrm{Vect}(\begin{pmatrix} 1 \\ 0 \\ -\frac{1}{2} \end{pmatrix}) \text{ et } E_5(E) = \mathrm{Vect}(\begin{pmatrix} 1 \\ \frac{20}{11} \\ \frac{2}{11} \end{pmatrix})$$

Exercice 13. *Matrice compagnon*

Soit $n \in \mathbb{N}^*$, $a_0, ..., a_{n-1} \in \mathbb{K}$ et

$$A = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & & \vdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

Montrer que $\chi_A = X^n + a_{n-1}X^{n-1} + ... + a_1 X + a_0$.
En déduire que pour tout polynôme unitaire $P \in \mathbb{K}[X]$, il existe une matrice $A \in M_n(\mathbb{K})$ telle que $P = \chi_A$.

Correction.

Voici deux méthodes pour obtenir le résultat (on explicite ici seulement la deuxième) :
1) On développe le déterminant $\det(\lambda I_n - A)$ par rapport à la dernière colonne.

2) On a $\det(\lambda I_n - A) = \begin{vmatrix} X & 0 & \cdots & \cdots & 0 & a_0 \\ -1 & X & \cdots & \cdots & 0 & a_1 \\ 0 & -1 & \ddots & & \vdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \cdots \\ 0 & 0 & \cdots & -1 & X & a_{n-2} \\ 0 & 0 & \cdots & 0 & -1 & X + a_{n-1} \end{vmatrix}$

En faisant l'opération : $L_0 \leftarrow \sum_{i=0}^{n-1} X^i L_i$, on obtient :

$$\det(\lambda I_n - A) = \begin{vmatrix} 0 & 0 & \cdots & \cdots & 0 & P(X) \\ -1 & X & \cdots & \cdots & 0 & a_1 \\ 0 & -1 & \ddots & & \vdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \cdots \\ 0 & 0 & \cdots & -1 & X & a_{n-2} \\ 0 & 0 & \cdots & 0 & -1 & X + a_{n-1} \end{vmatrix}$$

où $P(X) = X^n + a_{n-1}X^{p-1} + ... + a_1X + a_0$.
On obtient alors le résultat en développant par rapport à la 1ere ligne.
Pour $P = X^n + a_{n-1}X^{p-1} + ... + a_1X + a_0$ un polynôme unitaire de $\mathbb{K}[X]$, la matrice compagnon A de la question précédente a pour polynôme caractéristique le polynôme P.

Proposition 17.

Soit $A \in M_n(\mathbb{K})$. Si A est triangulaire (supérieure ou inférieure), alors $\chi_A = \prod_{i=1}^{n}(X - \alpha_i)$ où $\alpha_1, ..., \alpha_n$ sont les coefficients diagonaux de A.

Démonstration.

On a, pour $\lambda \in K$,

$$\chi_A(\lambda) = \det(\lambda I_n - A) = \begin{vmatrix} \lambda - \alpha_1 & * & \dots & * \\ 0 & \lambda - \alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda - \alpha_n \end{vmatrix}$$

D'où $\chi_A(\lambda) = \prod_{i=1}^{n}(\lambda - \alpha_i)$. □

b. Polynôme caractéristique d'un endomorphisme

Lemme 1.

Soit $A, B \in M_n(\mathbb{K})$. Si A et B sont semblables, alors $\chi_A = \chi_B$.

Démonstration.

On suppose A et B semblables. Alors il existe $P \in \mathbb{K}[X]$ tel que $B = PAP^{-1}$. Par suite, on a :

$$\chi_B = \det(XI_n - PAP^{-1}) = \det(P(XI_n - A)P^{-1}) = \frac{\det(P)}{\det(P)}\det(\lambda I_n - A) = \chi_A.$$

□

Définition 14.

Soit $u \in \mathcal{L}(E)$. On appelle **polynôme caractéristique** de u et on note $\chi_u(X)$ le polynôme caractéristique de toute matrice représentant u, i.e. si $\mathcal{B}$ est une base de E et si $A = \mathrm{Mat}_{\mathcal{B}}(u)$,

$$\chi_u := \chi_A$$

Remarque 8.

Le lemme précédent nous permet d'affirmer que le polynôme caractéristique d'un endomorphisme est bien défini : en effet, si A et B sont des matrices représentant u, elles sont semblables et donc ont même polynôme caractéristique.

Proposition 18.

Soit $u \in \mathcal{L}(E)$. On a, pour tout $\lambda \in \mathbb{K}$:

$$\chi_u(\lambda) = \det(\lambda \mathrm{Id}_E - u) = \lambda^n - \mathrm{Tr}(u)\lambda^{n-1} + \ldots + (-1)^n \det(u).$$

Démonstration.

Il suffit d'écrire $\chi_u = \chi_A$ avec A une matrice représentant u. On a alors

$$\chi_u = \chi_A = \lambda^n - \mathrm{Tr}(A)\lambda^{n-1} + \ldots + (-1)^n \det(A) = \lambda^n - \mathrm{Tr}(u)\lambda^{n-1} + \ldots + (-1)^n \det(u);$$

et de plus, la matrice $\lambda I_n - A$ est une matrice représentant $\lambda \mathrm{Id}_E - u$, donc

$$\chi_u = \chi_A = \det(\lambda I_n - A) = \det(\lambda \mathrm{Id}_E - u).$$

□

Théorème 3.

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Le scalaire λ est une valeur propre de u si, et seulement si, λ est une racine du polynôme caractéristique de u. Autrement dit :

$$\lambda \in \mathrm{Sp}(u) \;\Leftrightarrow\; \chi_u(\lambda) = 0.$$

Démonstration.

On écrit $\chi_u = \chi_A$ avec A une matrice représentant u et on a, pour $\lambda \in \mathbb{K}$:

$$\lambda \in \mathrm{Sp}(u) \;\Leftrightarrow\; \lambda \in \mathrm{Sp}(A) \;\Leftrightarrow\; \chi_A(\lambda) = 0 \;\Leftrightarrow\; \chi_u(\lambda) = 0.$$

□

Remarque 9.

Comme pour le cas des matrices, on en déduit que si $\mathbb{K} = \mathbb{C}$ ou si $\mathbb{K} = \mathbb{R}$ avec $\dim(E)$ impair, alors tout endomorphisme de E possède au moins une valeur propre.

c. Polynôme caractéristique d'un endomorphisme induit

Proposition 19.

Soit F un sous-espace vectoriel de E et $u \in \mathcal{L}(E)$. Si F est stable par u, alors le polynôme caractéristique χ_{u_F} de l'endomorphisme u_F induit par u sur F divise χ_u

Démonstration.

On suppose que F est stable par u. Sot $\mathcal{B} = (e_1, ..., e_n)$ une base de E adaptée à F où $\mathcal{B}' = (e_1, ..., e_p)$ forme une base de F. On pose $M = \mathrm{Mat}_{\mathcal{B}}(u)$ et $A = \mathrm{Mat}_{\mathcal{B}'}(u_F)$. Alors il existe $B \in M_{p,n-p}(\mathbb{K})$ et $C \in M_{n-p,n-p}(\mathbb{K})$ telles que :

$$M = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right)$$

Par suite, on a, en notant $Q = \det(XI_{n-p} - C) \in \mathbb{K}[X]$:

$$\begin{aligned} \chi_u &= \det(XI_n - M) \\ &= \left|\begin{array}{c|c} XI_p - A & -B \\ \hline 0 & XI_{n-p} - C \end{array}\right| \\ &= \det(XI_p - A).\det(XI_{n-p} - C) \\ &= \chi_A.Q \\ \chi_u &= \chi_{u_F}Q. \end{aligned}$$

Il en résulte que $\chi_{u_F}|\chi_u$. □

Remarque 10.

- On a alors $\mathrm{Sp}(u_F) \subset \mathrm{Sp}(u)$;
- Si χ_u est scindé (resp. scindé à racines simples) alors χ_{u_F} l'est aussi ;
- Par une récurrence finie, on obtient que si $E = \bigoplus_{i=1}^{k} F_i$ et chaque F_i est stable par u, alors

$$\chi_u = \prod_{i=1}^{k} \chi_{u_{F_i}} = \chi_{u_{F_1}} \cdots \chi_{u_{F_k}}.$$

2. Ordre de multiplicité d'une valeur propre

Définition 15. *Multiplicité d'une valeur propre*

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathrm{Sp}(u)$. On définit l'**ordre de multiplicité** - ou plus simplement la **multiplicité**) de la valeur propre λ de u et on note $m(\lambda)$ l'ordre de multiplicité de λ comme racine du polynôme caractéristique χ_u de u.
On définit de même la multiplicité d'une valeur propre d'une matrice $A \in M_n(\mathbb{K})$.

Remarque 11.

— Autrement dit, si $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ...\lambda_k$ deux à deux distinctes alors

$$\chi_u = P\prod_{i=1}^{k}(X-\lambda_i)^{m_i},$$

où $P \in \mathbb{K}[X]$ n'a pas de racine dans $\mathbb{K}$ et on a, pour tout $i \in [\![1, k]\!]$:

$$m(\lambda_i) = m_i.$$

— On a donc :

$$\deg(P) + m(\lambda_1) + ... + m(\lambda_k) = n.$$

— En particulier, pour λ une valeur propre, on a : $1 \leq m(\lambda) \leq n = \dim(E)$.

Proposition 20.

Soit $u \in \mathcal{L}(E)$ et λ une valeur propre de u. On a :

$$1 \leq \dim(E_\lambda(u)) \leq m(\lambda).$$

Démonstration.

On note $F = E_\lambda(u)$. Alors F est stable par u et l'endomorphisme induit $u_F \in \mathcal{L}(F)$ de u sur F est égal à l'homothétie $\lambda\mathrm{Id}_F$. Comme F est un sous-espace propre de u, on a $p = \dim(F) \geq 1$ et d'après la proposition précédente, on a :

$$(X-\lambda)^p = \chi_{u_F} | \chi_u = Q(X-\lambda)^{m(\lambda)},$$

avec $Q \in \mathbb{K}[X]$ et $(X-\lambda)$ premiers entre eux. Donc, d'après le lemme de Gauss, $(X-\lambda)^p | (X-\lambda)^{m(\lambda)}$.

Il en résulte que $1 \leq p = \dim(E_\lambda(u)) \leq m(\lambda)$. $\square$

Corollaire 5.

Soit $u \in \mathcal{L}(E)$. Si λ est une valeur propre simple de u, alors $\dim(E_\lambda(u)) = 1$.

Démonstration.

On suppose que λ est une valeur propre simple de u, d'après la proposition précédente, $1 \leq \dim(E_\lambda(u)) \leq 1$ donc $\dim(E_\lambda(u)) = 1$. $\square$

Partie D

Diagonalisation et trigonalisation

Dans cette partie, E désigne un espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

1. Endomorphismes et matrices diagonalisables

Définition 16. *Endomorphisme/matrice diagonalisable*

— Soit $u \in \mathcal{L}(E)$. On dit que u est **diagonalisable** s'il existe une base dans laquelle la matrice de u est diagonale.

— Soit $A \in M_n(\mathbb{K})$. On dit que A est **diagonalisable** si elle semblable à une matrice diagonale, i.e. s'il existe $D \in M_n(\mathbb{K})$ diagonale et $P \in GL_n(\mathbb{K})$ tels que :

$$A = PDP^{-1}.$$

Proposition 21.

Soit $u \in \mathcal{L}(E)$. Alors u est diagonalisable si, et seulement si, il existe une base de vecteurs propres de u.

Démonstration.

- $(\Rightarrow)$. Si u est diagonalisable, il existe une base $\mathcal{B} = (e_1, ..., e_n)$ de E telle que $A = \mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{diag}(\alpha_1, ..., \alpha_n)$. Par suite, on a, pour chaque $i \in [\![1,]\!]n$, par définition des coefficients de A,
$$u(e_i) = \alpha_i e_i \text{ et } e_i \neq 0_E.$$
Donc les éléments de $\mathcal{B}$ sont des vecteurs propres de u.
- $(\Leftarrow)$. Si $\mathcal{B} = (e_1, ..., e_n)$ est une base de vecteurs propres associés respectivement à $\alpha_1, ..., \alpha_n \in \mathbb{K}$, alors :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{array}{cc} \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_n \end{pmatrix} & \begin{matrix} e_1 \\ \vdots \\ \vdots \\ e_n \end{matrix} \\ \begin{matrix} u(e_1) & \dots & \dots & u(e_n) \end{matrix} & \end{array}.$$

Donc la matrice de u est diagonale dans la base $\mathcal{B}$.

$\square$

Exercice 14.

1. Soit $\mathbb{K}_n[X]$ l'espace vectoriel des polynômes de degré inférieur ou égal à n. L'endomorphisme de $\mathbb{K}_n[X]$, $\Delta_n : P \mapsto P'$ est-il diagonalisable ?
2. Montrer que que les projecteurs et les symétries de E sont diagonalisables.

Correction.

1. Δ ne possède qu'une seule valeur propre 0, et les vecteurs propres associés à 0 sont les polynômes constants (non nuls). Ainsi, on ne peut pas obtenir une base de $\mathbb{K}_n[X]$ formée de vecteurs de propres de Δ_n.
2. Soit p le projecteur sur F parallèlement à G. Alors $E = F \oplus G$. On considère alors $\mathcal{B} = (e_1, ..., e_p, e_{p+1}, ..., e_n)$ une base de E adaptée à $F \oplus G$. Par suite, on a :

$$\mathrm{Mat}_{\mathcal{B}}(p) = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right)$$

Donc la matrice de p dans $\mathcal{B}$ est diagonale, d'où p est diagonalisable.

Soit s la symétrie associée au projecteur p sur F parallèlement à G i.e. $s = 2p - \mathrm{Id}_E$. On considère de nouveau la base adaptée $\mathcal{B} = (e_1, ..., e_p, e_{p+1}, ..., e_n)$ précédente. Par suite, on a :

$$\mathrm{Mat}_{\mathcal{B}}(s) = \left(\begin{array}{c|c} 2I_p & 0 \\ \hline 0 & 0 \end{array}\right) - \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & I_{n-p} \end{array}\right) = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_{n-p} \end{array}\right)$$

Donc la matrice de s dans $\mathcal{B}$ est diagonale, d'où s est diagonalisable.

Proposition 22.

Soit $u \in \mathcal{L}(E)$ et $A \in M_n(\mathbb{K})$ une matrice représentant u dans une certaine base de E. Alors A est diagonalisable si, et seulement si, u est diagonalisable.

Démonstration.

u est diagonalisable si, et seulement si, il existe une matrice D diagonale représentant u. Or A et D représentent toutes deux u si, et seulement si, A et D sont semblables. Donc u est diagonalisable si, et seulement si, A est diagonalisable. □

Corollaire 6.

Soit $A \in M_n(\mathbb{K})$. Alors A est diagonalisable si, et seulement si, l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A est diagonalisable.

Démonstration.

On applique la proposition précédente au cas particulier : $E = \mathbb{K}^n$, $A = (a_{i,j}) \in M_n(\mathbb{K})$ et

$u \in \mathcal{L}(\mathbb{K}^n)$ tel que :

$$u : (x_1, ..., x_n) \mapsto \left(\sum_{j=1}^{n} a_{1,j} x_j, ..., \sum_{j=1}^{n} a_{n,j} x_j \right).$$

□

Proposition 23.

Soit $A \in M_n(\mathbb{K})$. Alors A est diagonalisable si, et seulement si, il existe D une matrice diagonale tel que $A = PDP^{-1}$ où $P = \left(\begin{array}{c|c|c} C_1 & \dots & C_n \end{array} \right)$ et $C_1, ..., C_n$ constituent une base de $M_{n,1}(\mathbb{K})$ formée de vecteurs propres de A.

Démonstration.

On suppose A diagonalisable. Alors l'endomorphisme u de $\mathbb{K}^n$ canoniquement associé à A est diagonalisable, donc il existe une base $\mathcal{B}' = (\varepsilon_1, ..., \varepsilon_n)$ de $\mathbb{K}^n$ formée de vecteurs propres de u. Soit P la matrice de passage de la base canonique $\mathcal{B}$ de $\mathbb{K}^n$ vers la base $\mathcal{B}'$. La formule de changement de base pour les matrices représentant un endomorphisme nous donne

$$\mathrm{Mat}_{\mathcal{B}'}(u) = P^{-1} \mathrm{Mat}_{\mathcal{B}}(u) P,$$

Or $\mathrm{Mat}_{\mathcal{B}}(u) = A$ et $\mathrm{Mat}_{\mathcal{B}'} = D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ où λ_i est la valeur propre associé à ε_i. Par suite,

$$A = PDP^{-1}.$$

□

2. Diagonalisation

Proposition 24.

Soit $u \in \mathcal{L}(E)$ avec $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distincts. Les assertions suivantes sont équivalentes :

i) u est diagonalisable ;

ii) $E = \bigoplus_{i=1}^{k} E_{\lambda_i}(u)$;

iii) $n = \dim(E) = \sum_{i=1}^{k} \dim(E_{\lambda_i}(u))$.

Démonstration.

On démontre ii)⇔iii), i)⇐ii) puis i)⇒ii).

- ii)⇔iii). Les $E_{\lambda_i}(u)$ sont en somme directe, donc on a

$$\dim(\bigoplus_{i=1}^{k} E_{\lambda_i}(u)) = \sum_{i=1}^{k} \dim(E_{\lambda_i}(u)).$$

Ainsi, $\bigoplus_{i=1}^{k} E_{\lambda_i}(u) = E$ si, et seulement si, $\sum_{i=1}^{k} \dim(E_{\lambda_i}(u)) = n$.

- i)⇐ii). On suppose u diagonalisable. Alors il existe une base de E formée de vecteurs propres de u i.e. formée d'éléments appartenant aux sous-espaces propres de u. Par suite, tout élément de E se décompose en somme d'éléments des sous-espaces propres qui sont en somme directe ; donc E est égal à la somme directe des sous-espaces propres.
- ii)⇐i). On suppose $\bigoplus_{i=1}^{k} E_{\lambda_i}(u) = E$. Si on considère une base $\mathcal{B}$ de E adapté à cette somme directe, on a :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda_1 I_{\dim(E_{\lambda_1}(u))} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_k I_{\dim(E_{\lambda_k}(u))} \end{pmatrix}$$

qui est une matrice diagonale, donc u est diagonalisable.

□

Proposition 25.

Soit $A \in M_n(\mathbb{K})$ avec $\mathrm{Sp}(A) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distincts. Les assertions suivantes sont équivalentes :

i) A est diagonalisable ;

ii) $M_{n,1}(\mathbb{K}) = \bigoplus_{i=1}^{k} E_{\lambda_i}(A)$;

iii) $n = \sum_{i=1}^{k} \dim(E_{\lambda_i}(A))$.

Démonstration.

On applique la proposition précédente à l'endomorphisme canoniquement associé à A. □

Exercice 15.

1. Soit $A \in M_n(\mathbb{K})$. On suppose que A possède une unique valeur propre λ. Montrer que A est diagonalisable si, et seulement si, $A = \lambda I_n$.
2. Soit $A \in M_n(\mathbb{K})$ une matrice nilpotente i.e. vérifiant qu'il existe $k \in \mathbb{N}^*$ tel que $A^k = 0_n$. Montrer que A est diagonalisable si, et seulement si, $A = 0$.

Correction.

1. On suppose que λ est la seule valeur propre de A. Si A est diagonalisable, alors il existe D diagonale et $P \in GL_n(\mathbb{K})$ tels que $A = PDP^{-1}$. Comme A et D sont semblables, ils ont même polynôme caractéristique et donc même spectre $\mathrm{Sp}(A) = \{\lambda\} = \mathrm{Sp}(D)$. Or D s'écrit sous la forme $\mathrm{diag}(\alpha_1, ..., \alpha_n)$ et donc son spectre vérifie :
$$\{\alpha_1, ..., \alpha_n\} = \mathrm{Sp}(D) = \{\lambda\}.$$
Par suite, $\alpha_1 = ...\alpha_n = \lambda$ et donc $D = \lambda I_n$. Il en résulte que
$$A = PDP^{-1} = P\lambda I_n P^{-1} = \lambda PP^{-1} = \lambda I_n.$$

2. Soit A une matrice nilpotente. Alors il existe $k \in \mathbb{N}^*$ tel que $A^k = 0_n$. Ainsi le polynôme X^k est un polynôme annulateur pour A. Or, si P est un polynôme annulateur de A, toute valeur propre de A est racine de P. Or le polynôme X^k n'a que 0 pour racine. Donc si A possède une valeur propre, ça ne peut être que 0.
De plus, 0 est bien valeur propre de A, car A n'est pas inversible. Ainsi, A n'a que 0 pour valeur propre, donc, d'après la question 1., A est diagonalisable si, et seulement si, $A = 0I_n = 0_n$.

Remarque 12.

Soit $u \in \mathcal{L}(E)$ et $\mathrm{sp}(u) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distinctes. Si u est diagonalisable, alors $E = \bigoplus_{i=1}^k E_{\lambda_i}(u)$ et si on note p_{λ_m} le projecteur sur $E_{\lambda_m}(u)$ parallèlement à $\bigoplus_{\substack{i=1 \\ i\neq m}}^k E_{\lambda_i}(u)$, alors
$$u = \lambda_1 p_{\lambda_1} + ... + \lambda_k p_{\lambda_k}$$

On rappelle les notions de polynômes scindés et polynômes scindés à racines simples.

Définition 17. *Polynôme scindé*

Soit $P \in \mathbb{K}[X]$. On dit que P est un **polynôme scindé** sur $\mathbb{K}$ si ses facteurs irréductibles dans $\mathbb{K}[X]$ sont tous de degré 1 i.e. il existe $\lambda \in \mathbb{K}$, $k \in \mathbb{N}^*$, $\lambda_1, ..., \lambda_k \in \mathbb{K}$ et $\alpha_1, ..., \alpha_k \in \mathbb{N}^*$ tels que :
$$P = \lambda \prod_{i=1}^{k} (X - \lambda_i)^{\alpha_i}$$
Si de plus, dans la décomposition précédente, $\alpha_1 = 1, ..., \alpha_k = 1$ i.e. les racines $\lambda_1, ..., \lambda_k$ de P sont simples, on dit que P est scindé **à racines simples** sur $\mathbb{K}$ ou encore P est **simplement** scindé sur $\mathbb{K}$.

Théorème 4. *Théorème de diagonalisation d'un endomorphisme*

Soit $u \in \mathcal{L}(E)$. Alors u est diagonalisable si, et seulement si, il vérifie les deux conditions suivantes :

i) le polynôme caractéristique χ_u de u est scindé.

ii) la multiplicité de chaque valeur propre de u est égale à la dimension de son sous-espace propre associé, i.e. pour tout $\lambda \in \mathrm{Sp}(u)$,

$$m(\lambda) = \dim(E_\lambda(u)).$$

Démonstration.

- $(\Rightarrow)$. On suppose u diagonalisable. On note $\lambda_1, ..., \lambda_k$ ses valeurs propres (deux à deux distinctes). Alors $E = \bigoplus_{i=1}^k E_{\lambda_i}(u)$ et l'endomorphisme u_i induit sur $E_{\lambda_i}(u)$ par u est égal à l'homothétie $u_i = \lambda_i \mathrm{Id}_{E_{\lambda_i}(u)}$.
  De plus, en notant $d_i = \dim(E_{\lambda_i}(u))$ on a
  $$\chi_u = \chi_{u_1} ... \chi_{u_k} = (X-\lambda_1)^{d_1} ... (X-\lambda_k)^{d_k}.$$
  Donc, χ_u est scindé et pour tout $i \in [\![1, k]\!]$, $m(\lambda_i) = d_i$.
- $(\Leftarrow)$. On suppose i) et ii). D'après i), on a $\chi_u = \prod_{i=1}^k (X-\lambda_i)^{m_i}$ où les λ_i sont deux à deux distincts. Donc $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ et on a, d'après ii) :
  $$n = \deg(\chi_u) = \sum_{i=1}^n m_i = \sum_{i=1}^n m(\lambda_i) = \sum_{i=1}^n \dim(E_{\lambda_i}(u)).$$
  Donc d'après la proposition 24, u est diagonalisable.

$\square$

Théorème 5. *Théorème de diagonalisation d'une matrice*

Soit $A \in M_n(\mathbb{K})$. Alors A est diagonalisable si, et seulement si, il vérifie les deux conditions suivantes :

i) le polynôme caractéristique χ_A de A est scindé.

ii) la multiplicité de chaque valeur propre de A est égale à la dimension de son sous-espace propre associé, i.e. pour tout $\lambda \in \mathrm{Sp}(A)$,

$$m(\lambda) = \dim(E_\lambda(A)).$$

Démonstration.

On raisonne de la même manière que pour le théorème précédent. $\square$

Corollaire 7.

Soit $u \in \mathcal{L}(E)$ et $A \in M_n(\mathbb{K})$. On rappelle que $\dim(E) = n$.

— Si le polynôme caractéristique de u est scindé à racines simples i.e. si u possède n valeurs propres distinctes, alors u est diagonalisable.

— Si le polynôme caractéristique de A est scindé à racines simples i.e. si A possède n valeurs propres distinctes, alors A est diagonalisable.

Démonstration.

Si χ_u est scindé à racines simples alors, pour tout $\lambda \in \mathrm{Sp}(u)$, on a $1 \leq \dim(E_\lambda(u)) \leq m(\lambda) = 1$, donc $\dim(E_\lambda(u)) = m(\lambda)$. On applique alors le théorème précédent. □

Proposition 26. *Forme de la matrice diagonalisée*

Soit $A \in M_n(\mathbb{K})$ où $\mathrm{Sp}(A) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distinctes. Si A est diagonalisable, alors $A = PDP^{-1}$ où :

$$D = \begin{pmatrix} \lambda_1 I_{m(\lambda_1)} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_k I_{m(\lambda_k)} \end{pmatrix}$$

et P est la matrice de passage de la base canonique $\mathcal{B}$ de $M_n(\mathbb{K})$ vers une base $\mathcal{B}' = (C_1, ..., C_n)$ adaptée à la somme directe $\bigoplus_{i=1}^{k} E_{\lambda_i}(A)$, i.e.

$$P = \left(\begin{array}{c|c|c} C_1 & ... & C_n \end{array} \right)$$

Remarque 13.

Soit $u \in \mathcal{L}(E)$ avec $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ où $\lambda_1, ..., \lambda_k$ deux à deux distinctes et A sa matrice dans une certaine base $\mathcal{B}$. Si u est diagonalisable, alors $A = PDP^{-1}$ où D à la même forme que dans la proposition précédente et P est la matrice de passage de la base $\mathcal{B}$ de $M_n(\mathbb{K})$ vers une base $\mathcal{B}' = (C_1, ..., C_n)$ adaptée à la somme directe $\bigoplus_{i=1}^{k} E_{\lambda_i}(u)$.

Méthode : Diagonaliser une matrice $A \in M_n(\mathbb{K})$ dans $\mathbb{K}$.

- On calcule le polynôme caractéristique χ_A de A. S'il est scindé dans $\mathbb{K}$, on continue ; s'il ne l'est pas, A n'est pas diagonalisable.
- On calcule les éléments propres de A et on détermine la dimension de chaque sous-espace propre de A. Si la multiplicité de **chaque** valeur propre est égale à la dimension du sous-espace associé, alors A est diagonalisable et on continue ; sinon A n'est pas diagonalisable.
- On met A sous la forme $A = PDP^{-1}$ où P est la matrice formée par les vecteurs propres de A

Exercice 16.

Diagonaliser (si c'est possible) les matrices suivantes dans $\mathbb{R}$ puis $\mathbb{C}$:

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 1 & 10 \\ 0 & 6 & 8 \\ 0 & 0 & -1 \end{pmatrix} \quad C = \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & -1 \\ 0 & 2 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} i & -1 & i \\ 0 & 1-3i & -2 \\ 0 & -4 & 1+3i \end{pmatrix}$$

Correction.

1. $\chi_A = X^3 + 3X^2 - 2 = (X-1)(X+2)^2$, d'où $\mathrm{Sp}(A) = \{-2, 1\}$ et on a :

$$E_1(A) = \mathrm{Vect}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}) \text{ et } E_3(A) = \mathrm{Vect}(\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix})$$

   d'où A est diagonalisable et $A = PDP^{-1}$ avec

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$$

2. $\chi_B = X^3 - 7X^2 + 4X + 12 = (X+1)(X-2)(X+6)$ donc B est diagonalisable (polynôme scindé à racine simples) et $\mathrm{Sp}(B) = \{-1, 2, 6\}$ et on a :

$$E_{-1}(B) = \mathrm{Vect}(\begin{pmatrix} 62 \\ 24 \\ -21 \end{pmatrix}), \quad E_2(B) = \mathrm{Vect}(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix})$$

$$E_6(B) = \mathrm{Vect}(\begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}).$$

   d'où B est diagonalisable et $B = PDP^{-1}$ avec

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix} \quad P = \begin{pmatrix} 62 & 1 & 1 \\ 24 & 0 & 4 \\ -21 & 0 & 0 \end{pmatrix}$$

3. $\chi_C = X^3 + 4X = X(X^2+4) = X(X-2i)(X+2i)$, d'où C n'est pas diagonalisable dans $\mathbb{R}$ (car son polynôme caractéristique n'est pas scindé dans $\mathbb{R}[X]$) et $\mathrm{Sp}_{\mathbb{R}}(B) = \{0\}$. Par contre, C est diagonalisable dans $\mathbb{C}$ (polynôme scindé à racines simples dans $\mathbb{C}[X]$) et on trouve :

$$E_0(C) = \mathrm{Vect}(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix});$$

$$E_{2i}(C) = \mathrm{Vect}(\begin{pmatrix} -1 \\ i \\ 1 \end{pmatrix}), \quad E_{-2i}(C) = \overline{E_{2i}(C)} = \mathrm{Vect}(\begin{pmatrix} -1 \\ -i \\ 1 \end{pmatrix}).$$

Donc $C = PDP^{-1}$ avec :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2i & 0 \\ 0 & 0 & -2i \end{pmatrix} \quad P = \begin{pmatrix} 1 & -1 & -1 \\ 0 & i & -i \\ 1 & 1 & 1 \end{pmatrix}$$

4. $\chi_D = (X-i)(X-(1-i))(X-(1+i))$, d'où D est diagonalisable (polynôme scindé à racines simples) et on trouve :

$$E_i(D) = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right);$$

$$E_{1-i}(D) = \mathrm{Vect}\left(\begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}\right), \quad E_{1+i}(D) = \mathrm{Vect}\left(\begin{pmatrix} i \\ i \\ 2 \end{pmatrix}\right).$$

Donc $D = P\mathbb{D}P^{-1}$ avec :

$$\mathbb{D} = \begin{pmatrix} i & 0 & 0 \\ 0 & 1-i & 0 \\ 0 & 0 & 1+i \end{pmatrix} \quad P = \begin{pmatrix} 1 & 0 & i \\ 0 & i & i \\ 0 & 1 & 2 \end{pmatrix}$$

Exercice 17.

Chercher, si c'est possible, une base qui diagonalise l'endomorphisme de $\mathbb{R}_2[X]$ défini par :

$$u : P = a + bX + cX^2 \mapsto u(P) = (3b+c) + bX + (-a+3b+3c)X^2$$

et, le cas échéant, donner sa matrice dans cette base.

Correction.

$$\mathcal{B} = \{1 - X - X^2, X, 1 - 3X - 2X^2\} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

3. Endomorphismes et matrices trigonalisables

Définition 18. *Endomorphisme/matrice trigonalisable*

— Soit $u \in \mathcal{L}(E)$. On dit que u est **trigonalisable** s'il existe une base dans laquelle la matrice de u est triangulaire supérieure.

— Soit $A \in M_n(\mathbb{K})$. On dit que A est **trigonalisable** si elle semblable à une matrice triangulaire supérieure, i.e. s'il existe $T \in M_n(\mathbb{K})$ triangulaire supérieure et $P \in GL_n(\mathbb{K})$ tels que :

$$A = PTP^{-1}.$$

Proposition 27. *Forme d'une matrice trigonalisée*

Soit $A \in M_n(\mathbb{K})$ et $\mathrm{Sp}(A) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distincts. Si A est trigonalisable, alors A est semblable à :

$$T = \begin{pmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & \cdots & \cdots & * \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & \lambda_1 & \ddots & & * & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & 0 & & \ddots & \lambda_k & \ddots & \vdots \\ \vdots & & & & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \lambda_k \end{pmatrix}$$

Démonstration.

A et T ont même polynôme caractéristique qui est scindé et dont les racines sont $\lambda_1, ..., \lambda_k$ de multiplicités respectives $m(\lambda_1), ..., m(\lambda_k)$. Or T étant triangulaire, les coefficients diagonaux de T sont exactement les racines de χ_T et le nombre d'apparition d'un coefficient sur la diagonale est exactement sa multiplicité dans χ_T. D'où la forme annoncée pour T. □

Proposition 28.

Soit $u \in \mathcal{L}(E)$ et $A \in M_n(\mathbb{K})$ une matrice représentant u dans une certaine base de E. Alors A est trigonalisable si, et seulement si, u est trigonalisable.

Démonstration.

u est trigonalisable si, et seulement si, il existe une matrice T triangulaire représentant u. Or A et T représentent toutes deux u si, et seulement si, A et T sont semblables. Donc u est trigonalisable si, et seulement si, A est trigonalisable. □

Corollaire 8.

Soit $A \in M_n(\mathbb{K})$. Alors A est trigonalisable si, et seulement si, l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A est trigonalisable.

Démonstration.

On applique la proposition précédente au cas particulier : $E = \mathbb{K}^n$, $A = (a_{i,j}) \in M_n(\mathbb{K})$ et

$u \in \mathcal{L}(\mathbb{K}^n)$ tel que :

$$u : (x_1, ..., x_n) \mapsto \left(\sum_{j=1}^{n} a_{1,j} x_j, ..., \sum_{j=1}^{n} a_{n,j} x_j \right).$$

□

4. Trigonalisation

Théorème 6. *Théorème de trigonalisation*

— Soit $A \in M_n(\mathbb{K})$. Alors A est trigonalisable si, et seulement si, son polynôme caractéristique χ_A est scindé.
— Soit $u \in \mathcal{L}(E)$. Alors u est trigonalisable si, et seulement si, son polynôme caractéristique χ_u est scindé.

Démonstration.

On démontre la partie concernant les matrices. Pour les endomorphismes, il suffit d'utiliser l'équivalence $u \in \mathcal{L}(E)$ est trigonalisable si, et seulement si, une matrice représentant u est trigonalisable et de remarquer que u et sa matrice ont le même polynôme caractéristique.

- $(\Rightarrow)$. Si A est trigonalisable, alors il existe une matrice triangulaire supérieure T semblable à A. Par suite on a $\chi_A = \chi_T$ et le polynôme caractéristique d'une matrice triangulaire est scindé. Donc χ_A est scindé.
- $(\Leftarrow)$. On considère la propriété

$$\mathcal{P}_n : \text{"}\forall A \in M_n(\mathbb{K}),\ \chi_A \text{ est scindé } \Rightarrow A \text{ est trigonalisable."}$$

Montrons que, pour tout $n \in \mathbb{N}^*$, $\mathcal{P}_n$ est vraie par récurrence $n \in \mathbb{N}^*$.

— *Initialisation.* Pour $n = 1$, la propriété $\mathcal{P}_0$ est triviale : toute matrice de dimension 1 est triangulaire !

— *Hérédité.* Soit $n \in \mathbb{N}^*$. On suppose la propriété $\mathcal{P}_n$ vraie.
Soit $A \in M_{n+1}(\mathbb{K})$. On suppose que on polynôme caractéristique χ_A est scindé. Par suite, χ_A admet au moins une racine λ qui est valeur propre de A. Soit $C_1 \in M_{n+1,1}(\mathbb{K})$ un vecteur propre de A associé à λ. On complète C_1 en une base $\mathcal{B} = \{C_1, C_2, ..., C_{n+1}\}$ de $M_{n+1,1}(\mathbb{K})$. Alors, en posant $Q = \left(C_1 \mid \dots \mid C_{n+1} \right)$ i.e. Q est la matrice de passage de la base canonique de $M_{n+1,1}(\mathbb{K})$ vers $\mathcal{B}$, on a

$$Q^{-1}AQ = \left(\begin{array}{c|c} \lambda & B \\ \hline 0 & C \end{array} \right).$$

où $B \in M_{1,n}(\mathbb{K})$ et $C \in M_n(\mathbb{K})$.
Alors on a :

$$\chi_A = \chi_{Q^{-1}AQ} = (X - \lambda)\chi_C$$

Or comme χ_A est scindé et $\chi_C | \chi_A$, alors χ_C est scindé et ainsi, par hypothèse de récurrence, C est trigonalisable. Par suite, il existe $T' \in M_n(\mathbb{K})$ triangulaire et $R \in$

$GL_n(\mathbb{K})$ tels que $C = RT'R^{-1}$. Alors, si on pose :

$$P' = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & R \end{array}\right) \quad \text{et} \quad P = QP'$$

on obtient :

$$P^{-1}AP = P'^{-1}Q^{-1}AQP' = P'^{-1}\left(\begin{array}{c|c} \lambda & B \\ \hline 0 & C \end{array}\right)P' = \left(\begin{array}{c|c} \lambda & BR \\ \hline 0 & T' \end{array}\right).$$

Donc $T = P^{-1}AP$ est triangulaire ; d'où A est trigonalisable. Par suite, $\mathcal{P}_{n+1}$ est vraie. Ce qui achève la récurrence.

□

Corollaire 9.

— Toute matrice de $M_n(\mathbb{C})$ est trigonalisable.

— Si $\mathbb{K} = \mathbb{C}$, tout endomorphisme de E est trigonalisable.

Démonstration.

Soit A une matrice de $\mathbb{M}_n(\mathbb{C})$. Alors son polynôme caractéristique χ_A appartient à $\mathbb{C}[X]$ donc d'après le théorème de D'Alembert-Gauss, χ_A est scindé. Il en résulte que A est trigonalisable d'après le théorème précédent.
Même raisonnement pour un endomorphisme. □

Proposition 29.

Soit $u \in \mathcal{L}(E)$ avec $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distincts. Si u est trigonalisable, alors :

$$\mathrm{Tr}(u) = \sum_{i=1}^{k} m(\lambda_i)\lambda_i \quad \text{et} \quad \det(u) = \prod_{i=1}^{k} \lambda_k^{m(\lambda_k)}.$$

Démonstration.

On suppose u trigonalisable. Alors il existe T triangulaire qui représente u et T est de la forme :

$$T = \begin{pmatrix} \lambda_1 & * & \dots & \dots & \dots & \dots & * \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \lambda_1 & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \lambda_k & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & * \\ 0 & \dots & \dots & \dots & \dots & 0 & \lambda_k \end{pmatrix}$$

d'où le résultat. □

Méthode : Trigonalisation d'une matrice.

- On calcule le polynôme caractéristique de la matrice. S'il est **scindé**, la matrice est trigonalisable, on continue.
- On détermine les sous-espaces propres ; on compare la dimension de chacun de ces sous-espaces et la multiplicité des valeurs propres correspondantes. Si chaque dimension est égale à la multiplicité correspondante, on diagonalise ; sinon, on doit trigonaliser.

  Dans le cas général, il n'y a pas de méthode à connaître ; mais nous allons voir comment trigonaliser une matrice A dans les différents cas possibles en dimension 3 sur des exemples. Dans la suite, u désignera l'endomorphisme canonique de $\mathbb{K}^3$ associé à A.

Méthode : Trigonalisation d'une matrice de $M_3(\mathbb{K})$ non diagonalisable.

1er cas : Deux valeurs propres distinctes de multiplicité 1 et 2 et chaque sous-espace propre de dimension 1.

<u>Exemple représentatif :</u>

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

On a $\chi_A = (X-1)(X-2)^2$ et les sous-espaces propres sont :

$$E_1(A) = \text{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) \text{ et } E_2(A) = \text{Vect}\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right)$$

- On forme une base $\mathcal{B} = (e_1, e_2, e_3)$ de $\mathbb{K}^3$ en prenant $e_1 = (1,1,1)$ et $e_2 = (0,1,-1)$ et en choisissant e_3 de manière à compléter en une base la famille e_1, e_2.
- On obtient alors $A = PTP^{-1}$ où :

$$T = \text{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & * \\ 0 & 2 & * \\ 0 & 0 & 2 \end{pmatrix}$$

  et P est formé des vecteurs e_1, e_2, e_3 mis en colonne.

Par exemple : On choisit $e_3 = e_1 \wedge e_2 = (-2,1,1)$ et on a :

$$u(e_3) = {}^t\left(A\begin{pmatrix}-2\\1\\1\end{pmatrix}\right) = (-8,4,-8) = -4e_1 + 6e_2 + \mathbf{2}e_3.$$

d'où, dans ce cas, $T = \begin{pmatrix}1&0&-4\\0&2&6\\0&0&2\end{pmatrix}$ et $P = \begin{pmatrix}1&0&-2\\1&1&1\\1&-1&1\end{pmatrix}$

Méthode : Trigonalisation d'une matrice de $M_3(\mathbb{K})$ non diagonalisable.

2eme cas : Une valeur propre triple et le sous-espace propre associé de dimension 2.

Exemple représentatif :

$$A = \begin{pmatrix}1&0&0\\0&0&-1\\0&1&2\end{pmatrix}$$

On a $\chi_A = (X-1)^3$ et le sous-espace propre associé à 1 est :

$$E_1(A) = \mathrm{Vect}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\-1\\1\end{pmatrix}\right)$$

- On forme une base $\mathcal{B} = (e_1, e_2, e_3)$ de $\mathbb{K}^3$ en prenant $e_1 = (1,0,0)$ et $e_2 = (0,-1,1)$ et en choisissant e_3 de manière à compléter en une base la famille e_1, e_2.
- On obtient alors $A = PTP^{-1}$ où :

$$T = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix}1&0&*\\0&1&*\\0&0&1\end{pmatrix}$$

et P est formé des vecteurs e_1, e_2, e_3 mis en colonne.

Par exemple : On choisit $e_3 = (0,1,1)$ et on a $u(e_3) = (8,4,-8) = 2e_2 + \mathbf{1}.e_3$. d'où, dans ce cas, $T = \begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}$ et $P = \begin{pmatrix}1&0&0\\0&-1&1\\0&1&1\end{pmatrix}$

Méthode : Trigonalisation d'une matrice de $M_3(\mathbb{K})$ non diagonalisable.

3eme cas : Une valeur propre triple λ et le sous-espace propre associé de dimension 1.

On utilise ici la méthode de réduction de Jordan (par souci de simplicité) :
On cherche une base $\mathcal{B} = (e_1, e_2, e_3)$ de $\mathbb{K}^3$ telle que :

- On cherche $e_3 \notin \mathrm{Ker}\left((u - \lambda \mathrm{Id}_E)^2\right)$;
- on pose $e_2 = u(e_3) - \lambda e_3$; (d'où $u(e_3) = e_2 + \lambda e_3$) ;
- on pose $e_1 = u(e_2) - \lambda e_2$; (d'où $u(e_2) = e_1 + \lambda e_2$).

Et on prouvera plus tard qu'on a nécessairement $u(e_1) = \lambda e_1$ grâce au théorème de Cayley-

Hamilton. Ainsi, on obtient $A = PTP^{-1}$ où :

$$T = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

et P est formé des vecteurs e_1, e_2, e_3 mis en colonne.

Exercice 18.

Trigonaliser les matrices suivantes :

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 2 & -4 & 1 \\ 2 & -1 & -2 \end{pmatrix} \quad B = \begin{pmatrix} \frac{1}{2}+i & \frac{1}{2} & -i \\ -\frac{1}{2} & -\frac{1}{2}+i & -i \\ 0 & 0 & i \end{pmatrix} C = \begin{pmatrix} 1 & -1 & 0 \\ 2 & -2 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

Correction.

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1}$$

$$B = \begin{pmatrix} -i & -i & 0 \\ i & -i & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} i & 1 & 0 \\ 0 & i & 1 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} -i & -i & 0 \\ i & -i & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -3 & 1 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1}$$

5. Endomorphismes nilpotents et matrices nilpotentes

On rappelle ici la notion de nilpotence évoquée dans le chapitre Structures algébriques usuelles :

Définition 19. *Endomorphisme nilpotent/Matrice nilpotente*

— Soit $u \in \mathcal{L}(E)$. On dit que u est **nilpotent** s'il existe $k \in \mathbb{N}^*$ tel que $u^k = 0$. On appelle alors **indice de nilpotence** le plus petit entier $p \in \mathbb{N}^*$ tel que $u^p = 0$.

— Soit $A \in M_n(\mathbb{K})$. On dit que A est **nilpotente** s'il existe $k \in \mathbb{N}^*$ tel que $A^k = 0_n$. On appelle alors **indice de nilpotence** le plus petit entier $p \in \mathbb{N}^*$ tel que $A^p = 0_n$.

Exemple 5.

Une matrice triangulaire dont la diagonale est composée de 0 - on appelle ce type de matrices des matrices triangulaires **strictes** - est nilpotente.

Proposition 30.

Soit $u \in \mathcal{L}(E)$. On a équivalence entre les assertions :

i) u est nilpotent ;

ii) $\chi_u = X^n$ (où $n = \dim(E)$) ;

On a le même résultat pour $A \in M_n(\mathbb{K})$.

Démonstration.

- i)$\Rightarrow$ ii). On suppose u nilpotent d'indice p. Soit $A \in M_n(\mathbb{K})$ une matrice représentant u dans une certaine base $\mathcal{B}$ de E. La matrice A est trigonalisable dans $M_n(\mathbb{C})$. Soit $\lambda \in \mathbb{C}$ une valeur propre de A et $X \in M_{n,1}(\mathbb{C})$ un vecteur propre associé. Alors on a, pour tout $k \in \mathbb{N}$:
$$A^k X = \lambda^k X.$$
Or $A^p = 0$ donc $\lambda^p X = 0$ avec $X \neq 0$, d'où $\lambda^p = 0$. Ainsi, $\lambda = 0$, donc 0 est la seule valeur propre de A. A étant trigonalisable, son polynôme caractéristique est donc $\chi_A = X^n$. Par suite, $\chi_u = \chi_A = X^n$.
- ii)$\Rightarrow$ i). On suppose $\chi_u = X^n$. Comme χ_u est scindé, u est trigonalisable et donc il existe une base $\mathcal{B}$ dans laquelle la matrice T de u est triangulaire stricte car 0 est la seule valeur propre de u. Or T est nilpotente car triangulaire stricte, donc il existe $k \geq 1$ tel que $\text{Mat}_{\mathcal{B}}(u^k) = T^k = 0_n$. Ainsi, $u^k = 0$.

□

Corollaire 10.

— Soit $u \in \mathcal{L}(E)$. Alors u est nilpotent si, et seulement si, u est trigonalisable et $\text{Sp}(u) = \{0\}$.

— Soit $A \in M_n(\mathbb{K})$. Alors A est nilpotente si, et seulement si, A est trigonalisable et $\text{Sp}(A) = \{0\}$.

Démonstration.

On a u est nilpotent si, et seulement si, $\chi_u = X^n$ si, et seulement si, u est trigonalisable et son unique valeur propre est 0. □

Proposition 31.

Soit $u \in \mathcal{L}(E)$. Si u est nilpotent d'indice p, alors :

— pour tout $x \in E$ tel que $u^{p-1}(x) \neq 0_E$, la famille $(x, u(x), ..., u^{p-1}(x))$ est libre ;

— $p \leq n = \dim(E)$.

Démonstration.

— Soit $a_0, ..., a_{p-1} \in \mathbb{K}$ tels que $\sum_{i=0}^{p-1} a_i u^i(x) = 0_E$. On a :

$$0_E = u^{p-1}(\sum_{i=0}^{p-1} a_i u^i(x)) = a_0 u^{p-1}(x).$$

donc $a_0 = 0$ car $u^{p-1}(x) \neq 0_E$, puis, on a :

$$0_E = u^{p-2}(\sum_{i=1}^{p-1} a_i u^i(x)) = a_1 u^{p-1}(x).$$

d'où $a_1 = 0$. On continue ainsi de proche en proche pour trouver finalement :

$$a_0 = a_1 = ... = a_{p-1} = 0.$$

Donc $(x, u(x), ..., u^{p-1}(x))$ est une famille libre.

— Comme p est le plus petit entier de $\mathbb{N}^*$ tel que $u^p = 0$, alors $u^{p-1} \neq 0$. Par suite, il existe $x \neq 0_E$ tel que $u^{p-1}(x) \neq 0_E$. Ainsi, en utilisant le point précédent, la famille $(x, u(x), ..., u^{p-1}(x))$ est une famille libre de E de p vecteurs, par suite, $p \leq \dim(E)$.

□

Partie E

Polynômes annulateurs et réduction

Dans cette partie, l'espace vectoriel E est supposé de dimension finie n.

1. Rappels et compléments sur les polynômes annulateurs

a. Rappels

Soit $P = \sum_{i=0}^{k} a_i X^i \in \mathbb{K}[X]$. On rappelle que les polynômes P en $u \in \mathcal{L}(E)$ et en $A \in M_n(\mathbb{K})$ sont définis par :

$$P(u) = \sum_{i=0}^{k} a_i u^i = a_0 \mathrm{Id}_E + a_1 u + ... + a_k u^k \in \mathcal{L}(E),$$

et

$$P(A) = \sum_{i=0}^{k} a_i A^i = a_0 I_n + a_1 A + ... + a_k A^k \in M_n(\mathbb{K}).$$

On note, pour $u \in \mathcal{L}(E)$ et $A \in M_n(\mathbb{K})$:

$$\mathbb{K}[u] = \{P(u) \mid P \in \mathbb{K}[X]\} \quad \text{et} \quad \mathbb{K}[A] = \{P(A) \mid P \in \mathbb{K}[X]\}$$

On dit que $P \in \mathbb{K}[X]$ est un **polynôme annulateur** pour u (resp. pour A) si $P(u) = 0$ (resp. si $P(A) = 0_n$).

b. Polynômes annulateurs et éléments propres

Proposition 32.

Soit $u \in \mathcal{L}(E)$.

i) Pour tous $P, Q \in \mathbb{K}[X]$, $P(u)$ et $Q(u)$ commutent.

ii) Pour tout $P \in \mathbb{K}[X]$, $\mathrm{Ker}(P(u))$ et $\mathrm{Im}(P(u))$ sont stables par u.

Démonstration.

i) On a, pour tous $i, j \in \mathbb{N}$, $u^i \circ u^j = u^j \circ u^i$ et u^i est linéaire donc en déduit que, pour tous $P, Q \in \mathbb{K}[X]$:

$$\begin{aligned} P(u) \circ Q(u) &= \sum_{i=0}^{k} a_i u^i \circ (\sum_{j=0}^{l} b_j u^j) \\ &= \sum_{i=0}^{k} a_i \sum_{j=0}^{l} b_j u^i \circ u^j \\ &= \sum_{j=0}^{l} b_j \sum_{i=0}^{k} a_i u^j \circ u^i \\ &= \sum_{j=0}^{l} b_j u^j \circ (\sum_{i=0}^{k} a_i u^i) \\ &= Q(u) \circ P(u) \end{aligned}$$

ii) Soit $P \in \mathbb{K}[X]$. On a $u = Q(u)$ avec $Q = X$, donc u et $P(u)$ commutant d'après i). Donc, d'après la proposition 7, $\mathrm{Ker}(P(u))$ et $\mathrm{Im}(P(u))$ sont stables par u.

□

Proposition 33.

Soit $u \in \mathcal{L}(E)$, $A \in M_n(K)$, $\lambda \in \mathbb{K}$ et $P \in \mathbb{K}[X]$.
- Soit $x \in E$. Si $u(x) = \lambda x$ alors $P(u)(x) = P(\lambda)x$.
- Soit $X \in M_{n,1}(\mathbb{K})$. Si $AX = \lambda X$ alors $P(A)X = P(\lambda)X$.

Démonstration.

- On suppose $u(x) = \lambda x$. Alors on a :

$$P(u)(x) = \sum_{i=0}^{k} a_i u^i(x) = \sum_{i=0}^{k} a_i(\lambda^i x) = (\sum_{i=0}^{k} a_i \lambda^i)x = P(\lambda)x.$$

- On fixe une base $\mathcal{B}$ de E et on raisonne comme pour pour le point précédent en considérant A et X comme les matrices dans la base $\mathcal{B}$ de u et x respectivement.

□

Corollaire 11.

Soit $u \in \mathcal{L}(E)$, $A \in M_n(K)$, $\lambda \in \mathbb{K}$ et $P \in \mathbb{K}[X]$.
- Si λ est une valeur propre de u, alors $P(\lambda)$ est une valeur propre de $P(u)$ et

$$E_\lambda(u) = E_{P(\lambda)}(P(u)).$$

- Si λ est une valeur propre de A, alors $P(\lambda)$ est une valeur propre de $P(A)$ et

$$E_\lambda(A) = E_{P(\lambda)}(P(A)).$$

Proposition 34.

Soit $u \in \mathcal{L}(E)$, $A \in M_n(\mathbb{K})$ et $P \in \mathbb{K}[X]$.
- Si P est un polynôme annulateur de u, alors toute valeur propre de u est une racine de P ; autrement dit, pour $\lambda \in \mathrm{Sp}(u)$, $P(\lambda) = 0$.
- Si P est un polynôme annulateur de A, alors toute valeur propre de A est une racine de P ; autrement dit, pour $\lambda \in \mathrm{Sp}(A)$, $P(\lambda) = 0$.

Démonstration.

On suppose que P est un polynôme annulateur de u, i.e. $P(u) = 0$. Alors, pour $\lambda \in \mathbb{K}$ une valeur propre de u, d'après la proposition précédente, $\mathbb{P}(\lambda)$ est un valeur propre de l'endomorphisme $P(u)$ qui est l'endomorphisme nul. Or 0 est l'unique valeur propre de $0 \in \mathcal{L}(E)$. D'où $P(\lambda) = 0$ i.e. λ est une racine de P. $\square$

Remarque 14.

ATTENTION, la réciproque de la proposition précédente est fausse ! Par exemple, $X^2(X-1)$ est un polynôme annulateur pour la matrice I_n mais 0 n'est pas valeur propre de I_n.

Exemple 6.

- Soit p un projecteur. Alors $p^2 = p$ donc $X^2 - X = (X-1)X$ est une polynôme annulateur de p et on a bien $\mathrm{Sp}(p) \subset \{0, 1\}$.
- Soit s une symétrie. Alors $s^2 = \mathrm{Id}_E$ donc $X^2 - 1 = (X-1)(X+1)$ est une polynôme annulateur de s et on a bien $\mathrm{Sp}(s) \subset \{-1, 1\}$.

Proposition 35.

Soit $u \in \mathcal{L}(E)$, $A \in M_n(\mathbb{K})$ et $P \in \mathbb{K}[X]$.
- Si P est un polynôme annulateur de u et $P(0) \neq 0$, alors u est injectif (et donc bijectif car $\dim(E)$ est finie).
- Si P est un polynôme annulateur de A et $P(0) \neq 0$, alors A est inversible.

Démonstration.

On suppose que $P = \sum_{i=0}^k a_i X^i$ est un polynôme annulateur de u et que $P(0) \neq 0$. Alors $a_0 \neq 0$ et on a :

$$a_0 \mathrm{Id}_E + \left(\sum_{i=1}^k a_i u^{i-1}\right) \circ u = P(u) = 0,$$

donc $\left(\frac{-1}{a_0} \sum_{i=1}^k a_i u^{i-1}\right) \circ u = \mathrm{Id}_E$. Par suite, u est inversible et son inverse est $\frac{-1}{a_0} \sum_{i=1}^k a_i u^{i-1}$.

On raisonne de même pour $A \in M_n(\mathbb{K})$ pour démontrer que A est inversible et que son inverse est $\frac{-1}{a_0} \sum_{i=1}^k a_i A^{i-1}$. $\square$

Méthode : Calcul d'une inverse grâce à un polynôme annulateur. La démonstration précédente nous donne un moyen pratique de détermination de l'inverse d'un endomorphisme (en dimension finie) ou d'une matrice quand on a un polynôme annulateur P tel que $P(0) \neq 0$. En effet, pour $P = \sum_{i=0}^{k} a_i X^i$ avec $a_0 \neq 0$ un polynôme annulateur de u (resp. de A), on a :

$$u^{-1} = \frac{-1}{a_0} \sum_{i=1}^{k} a_i u^{i-1};$$

respectivement,

$$A^{-1} = \frac{-1}{a_0} \sum_{i=1}^{k} a_i A^{i-1}.$$

Exercice 19.

1. Déterminer l'inverse de $f : (x, y) \mapsto (2x + y, x + 2y)$
2. Déterminer l'inverse de $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

Correction.

1. On a $f^2 - 3f = -3\mathrm{Id}_E$, donc $f^{-1} = \frac{-1}{3}(f - 3\mathrm{Id}_E)$.
2. On a $A^2 - 2A = 3I_n$, donc $A^{-1} = \frac{1}{3}(A - 3I_n)$.

Proposition 36.

Soit $u \in \mathcal{L}(E)$ et $P \in \mathbb{K}[X]$. Si u_F est l'endomorphisme induit par u sur $F = \mathrm{Ker}(P(u))$, alors P est un polynôme annulateur de u_F.

Correction.

Soit $P = \sum_{k=0}^{m} a_k X^k$ la décomposition de P dans la base canonique de $\mathbb{K}[X]$. On suppose que u_F est l'endomorphisme induit par u sur $F = \mathrm{Ker}(P(u))$. Alors, pour tout $x \in F$, on a $P(u)(x) = 0_E$ et :

$$P(u_F)(x) = \sum_{k=0}^{m} a_k u_F^k(x) = \sum_{k=0}^{m} a_k u^k(x) = P(u)(x) = 0_E$$

Par suite, $P(u_F) = 0_{\mathcal{L}(F)}$ i.e. P est annulateur de u_F.

2. Polynôme minimal

Dans le chapitre précédent, on a introduit la notion de polynôme minimal d'un élément d'une algèbre, on rappelle ici les principaux points de ce concept dans le contexte des algèbres de dimension finie $\mathcal{L}(E)$ et $M_n(\mathbb{K})$.

Proposition 37. *Idéal annulateur*

— Soit $u \in \mathcal{L}(E)$. L'ensemble $I_u = \{P \in \mathbb{K}[X] \mid P(u) = 0\}$ appelé **idéal annulateur de** u est un idéal de $\mathbb{K}[X]$ non réduit à $\{0\}$.
— Soit $A \in M_n(\mathbb{K})$. L'ensemble $I_A = \{P \in \mathbb{K}[X] \mid P(A) = 0_n\}$ appelé **idéal annulateur de** A est un idéal de $\mathbb{K}[X]$ non réduit à $\{0\}$.

Démonstration.

On a déjà démontré ce résultat dans la partie Algèbres du chapitre Structures algébriques. On rappelle tout de même la démonstration dans notre contexte :

I_u est un idéal de $K[X]$ comme noyau du morphisme d'anneaux $f : P \mapsto P(u)$ de $\mathbb{K}[X]$ dans $\mathcal{L}(E)$.
Montrons que u possède un polynôme annulateur non nul. Comme $\mathcal{L}(E)$ est de dimension finie égale à n^2, la famille $(\mathrm{Id}_E, u, ..., u^{n^2})$ est liée car composée de n^2+1 vecteurs dans un espace de dimension n^2 donc il existe $\lambda_0, ..., \lambda_{n^2} \in \mathbb{K}$ non tous nuls tels que

$$\sum_{i=0}^{n^2} \lambda_i u^i = 0.$$

Par suite $P = \sum_{i=0}^{n^2} \lambda_i X^i$ est un polynôme annulateur non nul de u d'où $I_u \neq \{0\}$. □

Définition 20. *Polynôme minimal*

— Soit $u \in \mathcal{L}(E)$. On appelle **polynôme minimal de** u et on note π_u le générateur unitaire de l'idéal annulateur I_u de u. En particulier, $I_u = \pi_u \mathbb{K}[X]$.
— Soit $A \in M_n(\mathbb{K})$. On appelle **polynôme minimal de** A et on note π_A le générateur unitaire de l'idéal annulateur I_A de A. En particulier, $I_A = \pi_A \mathbb{K}[X]$.

Remarque 15.

Soit $u \in \mathcal{L}(E)$.
— Si E est de dimension $n \geq 1$, $\deg(\pi_u) \geq 1$.
— On a $\deg(\pi_u) = 1$ si, et seulement si, u est une homothétie.

Proposition 38.

— Soit $u \in \mathcal{L}(E)$ et $d = \deg(\pi_u)$. La famille $(\mathrm{Id}_E, u, ..., u^{d-1})$ est une base de l'algèbre $\mathbb{K}[u] = \{P(u) \mid P \in \mathbb{K}[X]\}$ engendré par u.
— Soit $A \in M_n(\mathbb{K})$ et $d = \deg(\pi_A)$. La famille $(I_n, A, ..., A^{d-1})$ est une base de l'algèbre $\mathbb{K}[A] = \{P(A) \mid P \in \mathbb{K}[X]\}$ engendré par A.

Démonstration.

On a déjà démontré ce résultat dans la partie Algèbres du chapitre Structures algébriques. On rappelle tout de même la démonstration dans notre contexte :

On suppose que u admet un polynôme minimal π_u avec $d = \deg(\pi_u)$. Montrons que $\left(u^k\right)_{0\leq k\leq d-1}$ est une base de $\mathbb{K}[u]$.

- *Famille libre :* soit $\lambda_0, \ldots, \lambda_{d-1} \in \mathbb{K}$ des scalaires tels que $\sum_{k=1}^{d-1} \lambda_k u^k = 0$. S'il existe $i \in [\![0, d-1]\!]$ tel que $\lambda_i \neq 0$, $\sum_{k=1}^{d-1} \lambda_k X^k$ est un polynôme annulateur non nul de u de degré $< d = \pi_u$. Contradiction car π_u est de degré minimal parmi les polynômes annulateurs. Donc pour tout $i \in [\![0, d-1]\!]$, $\lambda_i = 0$. Donc la famille $\left(u^k\right)_{0\leq k\leq d-1}$ est libre.
- *Famille génératrice :* Soit $P(u) \in \mathbb{K}[u]$. Alors $P \in \mathbb{K}[X]$ et en faisant la division euclidienne de ce polynôme par π_u, on obtient qu'il existe $Q, R \in \mathbb{K}[X]$ tels que $P = \pi_u Q + R$ et $\deg(R) < d-1$. Par suite,
$$P(u) = \underbrace{\pi_u(u)}_{=0_A} Q(u) + R(u) = R(u).$$
et R est de degré $\leq d-1$ donc il existe $\lambda_0, \ldots, \lambda_{d-1} \in \mathbb{K}$ tels que $R = \sum_{k=1}^{d-1} \lambda_k X^k$. Il en résulte que :
$$P(u) = R(u) = \sum_{k=1}^{d-1} \lambda_k X^k \in \mathrm{Vect}\left(u^k\right)_{0\leq k\leq d-1}.$$

Donc $\left(u^k\right)_{0\leq k\leq d-1}$ est une base de $\mathbb{K}[u]$. □

Proposition 39.

— Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Alors $\lambda \in \mathrm{Sp}(u)$ si, et seulement si, $\pi_u(\lambda) = 0$.

— Soit $A \in M_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. Alors $\lambda \in \mathrm{Sp}(A)$ si, et seulement si, $\pi_A(\lambda) = 0$.

Démonstration.

- ($\Rightarrow$). Si λ est une valeur propre de u, alors λ est racine de tout polynôme annulateur de u. Or π_u est un polynôme annulateur de u. Donc $\pi_u(\lambda) = 0$.
- ($\Leftarrow$). On suppose $\pi_u(\lambda) = 0$. Alors on a la factorisation
$$\pi_u = (X - \lambda)P,$$
où $P \in \mathbb{K}[X]$ et $\deg(P) = \deg(\pi_u) - 1 < \deg(\pi_u)$. Par suite,
$$0 = \pi_u(u) = (u - \lambda \mathrm{Id}_E) \circ P(u).$$
Supposons par l'absurde que λ n'est pas valeur propre de u. Alors $u - \lambda\mathrm{Id}_E$ est injective et donc bijective car E est de dimension finie ; d'où $P(u) = 0$. Ce qui est impossible par minimalité du degré du polynôme minimal parmi les polynômes annulateurs de u.
Il en résulte que λ est une valeur propre de u.

□

Corollaire 12.

- Soit $u \in \mathcal{L}(E)$. Alors π_u et χ_u ont les mêmes facteurs irréductibles dans $\mathbb{K}[X]$. En particulier, ils ont les mêmes racines.
- Soit $A \in M_n(\mathbb{K})$. Alors π_A et χ_A ont les mêmes facteurs irréductibles dans $\mathbb{K}[X]$. En particulier, ils ont les mêmes racines.

Démonstration.

Traitons le cas matriciel. Soit $A \in M_n(\mathbb{K}) \subset M_n(\mathbb{C})$. D'après la proposition précédente, les racines de π_A dans $\mathbb{C}$ sont exactement les valeurs propres de u qui sont également les racines de χ_A dans $\mathbb{C}$. Les facteurs irréductibles de $\mathbb{C}[X]$ étant seulement les polynômes de degré 1, il en résulte que π_A et χ_A ont les mêmes facteurs irréductibles dans $\mathbb{C}[X]$.
Ainsi, si $\mathbb{K} = \mathbb{C}$, on a le résultat. Dans le cas $\mathbb{K} = \mathbb{R}$, il reste donc à traiter le cas où P est un facteur irréductible de degré 2 de π_A ou de χ_A. Alors $P = (X-\lambda)(X-\overline{\lambda})$ avec $\lambda \in \mathbb{C} \setminus \mathbb{R}$ i.e. P possède deux racines complexes non réelles conjuguées. Ces deux racines $\lambda, \overline{\lambda}$ sont donc des valeurs propres de A vu comme une matrice à coefficients complexes. Par suite, ce sont des racines communes de π_A et χ_A d'où $P = (X-\lambda)(X-\overline{\lambda})$ est un facteur irréductible dans $\mathbb{R}[X]$ commun à π_A et χ_A. $\square$

Corollaire 13.

Soit $u \in \mathcal{L}(E)$. Alors π_u est scindé si, et seulement si, χ_u est scindé.
De même dans le cas matriciel.

3. Lemme de décomposition des noyaux

Théorème 7. *Lemme de décomposition des noyaux*

Soit $u \in \mathcal{L}(E)$ et $P_1, ..., P_k \in \mathbb{K}[X]$ deux à deux premiers entre eux. Alors, pour $P = P_1 ... P_k$, on a :

$$\mathrm{Ker} P(u) = \bigoplus_{i=1}^{k} \mathrm{Ker} P_i(u).$$

Démonstration.

On montre le résultat par récurrence sur $k \in \mathbb{N}$, $k \geq 2$.

- *Initialisation :* $k = 2$ Soit $A, B \in \mathbb{K}[X]$ premiers entre eux et $P = AB$. On procède par double inclusion.
  Comme $\mathbb{K}[u]$ est commutatif, on a :
  $$A(u) \circ B(u) = P(u) = B(u) \circ A(u),$$
  d'où $\mathrm{Ker}(A(u)) \subset \mathrm{Ker}(P(u))$ et $\mathrm{Ker}(B(u)) \subset \mathrm{Ker}(P(u))$. Par suite,
  $$\mathrm{Ker}(A(u)) + \mathrm{Ker}(B(u)) \subset \mathrm{Ker}(P(u))$$
  D'après le théorème de Bézout, il existe $U, V \in \mathbb{K}[X]$ tels que $AU + BV = 1$, donc

$$A(u) \circ U(u) + B(u) \circ V(u) = \mathrm{Id}_E. \quad (*)$$

- Montrons tout d'abord que la somme $\mathrm{Ker}(A(u))+\mathrm{Ker}(B(u))$ est directe i.e. $\mathrm{Ker}(A(u)) \cap \mathrm{Ker}(B(u)) = \{0_E\}$.
  Soit $x \in \mathrm{Ker}(A(u)) \cap \mathrm{Ker}(B(u))$. Alors on a, d'après $(*)$ et en utilisant le fait que $\mathbb{K}[u]$ est une algèbre commutative :
  $$x = U(u)\left(\underbrace{A(u)(x)}_{=0_E}\right) + V(u)\left(\underbrace{B(u)(x)}_{=0_E}\right) = 0_E.$$
  Par suite, $\mathrm{Ker}(A(u)) \cap \mathrm{Ker}(B(u)) = \{0_E\}$.
- Soit $x \in \mathrm{Ker}(P(u))$. On note $y = A(u) \circ U(u)(x)$ et $z = B(u) \circ V(u)(x)$. D'après $(*)$, $x = y + z$, et de plus, on a :
  $$A(u)(z) = A(u) \circ B(u) \circ U(u)(x) = P(u) \circ U(u)(x) = U(u)\left(\underbrace{P(u)(x)}_{=0_E}\right) = 0_E$$
  donc $z \in \mathrm{Ker}(A(u))$ et, par le même raisonnement, on obtient $y \in \mathrm{Ker}(B(u))$.
  Par suite, $x = z + y \in \mathrm{Ker}(A(u)) \oplus \mathrm{Ker}(B(u))$.
  Il en résulte que
  $$\mathrm{Ker}(A(u)) \oplus \mathrm{Ker}(B(u)) = \mathrm{Ker}(P(u)).$$

— *Hérédité :* Soit $k \geq 2$. On suppose le propriété vraie pour k. Soit $P = P_1...P_{k+1}$ avec $P_1, ..., P_{k+1}$ premiers entre eux deux à deux.
On pose $A = P_1...P_k$ et $B = P_{k+1}$. Alors A et B sont premiers entre eux, $P = AB$. D'après le raisonnement effectué pour l'initialisation, on a $\mathrm{Ker}(P(u)) = \mathrm{Ker}(A(u)) + \mathrm{Ker}(B(u))$. Par suite, d'après l'hypothèse de récurrence, on a :

$$\mathrm{Ker}(P(u)) = \mathrm{Ker}(A(u)) + \mathrm{Ker}(B(u)) = \bigoplus_{i=1}^{k} \mathrm{Ker} P_i(u) \oplus \mathrm{Ker} P_{k+1}(u).$$

Ce qui achève le raisonnement par récurrence.

□

Corollaire 14.

Soit $u \in \mathcal{L}(E)$ et $P_1, ..., P_k \in \mathbb{K}[X]$ deux à deux premiers entre eux. Alors $P = P_1...P_k$ est un polynôme annulateur de u si, et seulement si,

$$E = \bigoplus_{i=1}^{k} \mathrm{Ker} P_i(u).$$

Démonstration.

On applique le lemme de décomposition des noyaux pour obtenir :

$$\mathrm{Ker}P(u) = \bigoplus_{i=1}^{k} \mathrm{Ker}P_i(u).$$

Or P est un polynôme annulateur de u si, et seulement si, on a $\mathrm{Ker}P(u) = E$, d'où le résultat. □

Remarque 16.

On retrouve que pour p un projecteur et s une symétrie, $\mathrm{Ker}(p)$ et $\mathrm{Ker}(p - \mathrm{Id}_E)$ sont supplémentaires ainsi que $\mathrm{Ker}(s - \mathrm{Id}_E)$ et $\mathrm{Ker}(s + \mathrm{Id}_E)$.

Exercice 20.

1. Soit $a, b \in \mathbb{R}$ tels que $a^2 + 4b > 0$ et $\mathcal{U}$ l'ensemble des suites récurrentes doubles dont le terme général vérifie $u_{n+2} = au_{n+1} + bu_n$. Montrer que $\mathcal{U}$ est le noyau de l'endomorphisme $P(s)$ où
$$P = X^2 - aX - b \quad \text{et} \quad s : u = (u_n)_{n\in\mathbb{N}} \mapsto (u_{n+1})_{n\in\mathbb{N}}.$$
En déduire une expression explicite de $\mathcal{U}$.
2. Soit $a, b \in \mathbb{R}$ tels que $a^2 + 4b > 0$ et $\mathcal{S}$ l'ensemble des fonctions f de $C^\infty(\mathbb{R})$ telles que $f'' = af' + bf$. Montrer que $\mathcal{S}$ est le noyau de l'endomorphisme $P(D)$ où
$$P = X^2 - aX - b \quad \text{et} \quad D : f \mapsto f'.$$
En déduire une expression explicite de $\mathcal{S}$.

Correction.

1. On a, pour $u = (u_n) \in \mathbb{R}^{\mathbb{N}}$,
$$P(s)(u) = s^2(u) - as(u) - b\mathrm{id}(u) = u_{n+2} - au_{n+1} - bu_n$$
Donc on a bien $\mathcal{U} = \mathrm{Ker}(P(s)) = \{u \in \mathbb{R}^{\mathbb{N}} \mid P(s)(u) = 0\}$. On a $\Delta(P) = a^2 + 4b > 0$ par hypothèse, donc P possède deux racines $r_1, r_2 \in \mathbb{R}$ distinctes. Ainsi $P = (X - r_1)(X - r_2)$ et $X - r_1, X - r_2$ sont premiers entre eux, donc, d'après le lemme de décomposition des noyaux :
$$\mathcal{U} = \mathrm{Ker}(P(s)) = \mathrm{Ker}(s - r_1\mathrm{id}) \oplus \mathrm{Ker}(s - r_2\mathrm{id}).$$
Or, pour $r \in \mathbb{R}$,
$$\mathrm{Ker}(s - r\mathrm{id}) = \{u = (u_n) \mid \forall n \in \mathbb{N},\ u_{n+1} = ru_n\} = \{(Ar^n)_{n\in\mathbb{N}} \mid A \in \mathbb{R}\}.$$
Il en résulte que :
$$\mathcal{U} = \{(Ar_1^n + Br_2^n)_{n\in\mathbb{N}} \mid A, B \in \mathbb{R}\}.$$
2. Par un raisonnement similaire, on obtient, pour r_1, r_2 les deux racines réelles distinctes de P :
$$\mathcal{S} = \{t \mapsto \alpha e^{r_1 t} + \beta e^{r_2 t} \mid \alpha, \beta \in \mathbb{R}\},$$

en remarquant que :

$$\mathrm{Ker}(D - r\mathrm{id}) = \{f \in C^\infty(\mathbb{R}) \mid f' = rf\} = \{t \mapsto \alpha e^{rt} \mid \alpha \in \mathbb{R}\}.$$

4. Polynômes annulateurs et réduction

Théorème 8.

Soit $u \in \mathcal{L}(E)$. On a équivalence entre les assertions :

i) u est diagonalisable ;

ii) u possède un polynôme annulateur scindé à racines simples ;

iii) le polynôme minimal π_u de u est scindé à racines simples.

Le même résultat est valable pour $A \in M_n(\mathbb{K})$.

Démonstration.

On démontre i)⇒ ii) ⇒ iii) ⇒ i)

- i)⇒ ii). On suppose u diagonalisable et on note $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ avec $\lambda_1, ..., \lambda_k$ deux à deux distinctes. Alors on a

$$E = \bigoplus_{i=1}^{k} E_{\lambda_i}(u) = \bigoplus_{i=1}^{k} \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E).$$

  Par suite, d'après le corollaire 14, le polynôme scindé à racines simples $P = \prod_{i=1}^{k}(X - \lambda_i)$ est annulateur de u.
- ii)⇒ iii). Si u possède un polynôme annulateur scindé à racines simples P alors $\pi_u | P$ et donc π_u est scindé à racines simples.
- iii)⇒ i). On suppose que π_u est scindé à racines simples i.e. $\pi_u = \prod_{i=1}^{k}(X - \lambda_i)$ avec $\lambda_1, ..., \lambda_k$ deux à deux distincts. Alors $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_k\}$ et d'après le corollaire 14, on a : $E = \bigoplus_{i=1}^{k} \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E) = \bigoplus_{i=1}^{k} E_{\lambda_i}(u)$. Par suite, u est diagonalisable.

□

Corollaire 15.

Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E stable par u. Si u est diagonalisable, alors l'endomorphisme $u_F \in \mathcal{L}(F)$ induit par u sur F est diagonalisable.

Démonstration.

Si u est diagonalisable, alors π_u est scindé à racines simples d'après le théorème précédent. Or on a $\pi_u(u_F) = 0$: en effet, pour tout $x \in F$, $\pi_u(u_F)(x) = \pi_u(u)(x) = 0_E$; par suite, u_F possède un polynôme annulateur scindé à racines simples. D'après le théorème précédent, u_F est diagonalisable. □

Exemple 7.

Soit $u \in \mathcal{L}(E)$ tel que $u^3 + 11u = 6u^2 + 6\mathrm{Id}_E$. Alors u est diagonalisable.

> En effet, le polynôme $P = X^3 - 3X^2 + 11X - 6$ est annulateur de u et on remarque que $P = (X-1)(X-2)(X-3)$ est scindé à racines simples. Ainsi, u est diagonalisable d'après le théorème précédent.

Exercice 21.

Soit $n \in \mathbb{N}^*$ et $M \in M_n(\mathbb{R})$ telle que $M^3 - M^2 + M = I_n$. Déterminer le déterminant de M et montrer que sa trace est un entier naturel inférieur ou égal à n.

Correction.

On remarque que $P = X^3 - X^2 + X - 1 = (X-1)(X-i)(X+i)$ est un polynôme annulateur scindé à racines simples de M. Par suite, M est diagonalisable dans $\mathbb{C}$ avec $\mathrm{Sp}_{\mathbb{C}}(M) \subset \{1, i, -i\}$ et ainsi, on a :

$$\det(M) = 1^{m(1)} \times i^{m(i)} \times (-i)^{m(-i)} \text{ et } \mathrm{tr}(M) = m(1) + m(i)i - m(-i)i$$

où, si λ n'est pas valeur propre de M, $m(\lambda) = 0$ et si λ est valeur propre de M, $m(\lambda)$ désigne sa multiplicité.

De plus, M étant à coefficients dans $\mathbb{R}$, ses valeurs propres non réelles conjuguées (potentielles) i et $-i$ ont la même multiplicité i.e. $m(i) = m(-i)$. Ainsi :

$$\det(M) = (i \times (-i))^{m(i)} = 1^{m(i)} = 1 \text{ et } \mathrm{tr}(M) = m(1) + m(i)(i - i) = m(1) \in [\![0, n]\!].$$

Exercice 22.

Soit $n \in \mathbb{N}^*$. On pose $J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in M_n(\mathbb{R})$ et $A = \begin{pmatrix} n+1 & 1 & \dots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & n+1 \end{pmatrix} \in M_n(\mathbb{R})$.

1. Exprimer J^2 en fonction de J.
2. Exprimer A en fonction de J et I_n puis en déduire un polynôme annulateur de degré 2 de A.
3. En déduire que A est diagonalisable sur $\mathbb{R}$.

Correction.

1. On a $J^2 = nJ$.
2. On a $A = J + nI_n$ et comme J et I_n commutent, on a :

$$A^2 = (J + nIn)^2 = J^2 + 2nJ + n^2 I_n = 3nJ + n^2 I_n = 3nA - 2n^2 I_n$$

Par suite, $P = X^2 - 3nX + 2n^2$ est annulateur de A

3. On remarque que $P = (X-n)(X-2n)$ est un polynôme annulateur de A scindé à racines simples, donc A est diagonalisable.

Exercice 23.

Soit $n \in \mathbb{N}^*$, $A \in M_n(\mathbb{R})$ et $B = \left(\begin{array}{c|c} A & A \\ \hline 0_n & A \end{array}\right) \in M_{2n}(\mathbb{R})$.
Montrer que si B est diagonalisable, alors A l'est aussi puis, sous la même hypothèse, que $A = 0_n$.

Correction.

On remarque tout d'abord que pour $P \in \mathbb{R}[X]$, $P(B) = \left(\begin{array}{c|c} P(A) & AP'(A) \\ \hline 0_n & P(A) \end{array}\right)$ car, pour tout $k \in \mathbb{N}^*$, $B^k == \left(\begin{array}{c|c} A^k & kA^k \\ \hline 0_n & A^k \end{array}\right)$.
On suppose B diagonalisable. D'après la caractérisation de la diagonalisabilité par les polynômes annulateurs, B possède un polynôme annulateur P scindé à racines simples. Ainsi, d'après la remarque précédente, P est donc également annulateur de A, d'où A est diagonalisable, toujours d'après la caractérisation utilisée précédemment.
De plus, on a également XP' qui est annulateur de A. Ainsi, les valeurs propres de A font partie des racines communes de XP' et de P. Or, comme P est à racines simples, P et P' n'ont pas de racine commune ; donc les valeurs propres de A font partie des racines communes de X et P... et ils en ont nécessairement au moins une en commun et il s'agit bien-sûr seulement de 0, car A possède au moins une valeur propre ! Ainsi, 0 est la seule valeur propre de A qui est diagonalisable : par suite, $A = 0_n$.

Théorème 9.

Soit $u \in \mathcal{L}(E)$. On a équivalence entre les assertions :

i) u est trigonalisable ;

ii) u possède un polynôme annulateur scindé ;

iii) le polynôme minimal π_u de u est scindé.

Le même résultat est valable pour $A \in M_n(\mathbb{K})$.

Démonstration.

On démontre i)$\Rightarrow$ ii) $\Rightarrow$ iii) $\Rightarrow$ i)

- i)$\Rightarrow$ ii). On suppose u trigonalisable. Alors d'après le théorème 6, le polynôme caractéristique χ_u de u est scindé. Donc, d'après le corollaire 13, le polynôme minimal π_u est scindé ; or, celui-ci est annulateur ; d'où ii).
- ii)$\Rightarrow$ iii). Si u possède un polynôme annulateur scindé P alors $\pi_u | P$ et donc π_u est scindé.

- iii)$\Rightarrow$ i). Pour $n \in \mathbb{N}^*$, on considère la propriété :

$$\mathcal{P}_n : \text{"}\forall A \in M_n(\mathbb{K}),\ \pi_A \text{ est scindé } \Rightarrow A \text{ est trigonalisable."}$$

Montrons que, pour tout $n \in \mathbb{N}^*$, $\mathcal{P}_n$ est vraie par récurrence $n \in \mathbb{N}^*$.

— *Initialisation.* Pour $n = 1$, la propriété $\mathcal{P}_0$ est triviale : toute matrice de dimension 1 est triangulaire.

— *Hérédité.* Soit $n \in \mathbb{N}^*$. On suppose la propriété $\mathcal{P}_n$ vraie.

Soit $A \in M_{n+1}(\mathbb{K})$. On suppose que on polynôme minimal π_A est scindé. Par suite, π_A admet au moins une racine λ qui est valeur propre de A. Soit $C_1 \in M_{n+1,1}(\mathbb{K})$ un vecteur propre de A associé à λ. On complète C_1 en une base $\mathcal{B} = \{C_1, C_2, ..., C_{n+1}\}$ de $M_{n+1,1}(\mathbb{K})$. Alors, en posant $Q = \left(\begin{array}{c|c|c} C_1 & \dots & C_{n+1}\end{array}\right)$ i.e. Q est la matrice de passage de la base canonique de $M_{n+1,1}(\mathbb{K})$ vers $\mathcal{B}$, on a

$$Q^{-1}AQ = \left(\begin{array}{c|c} \lambda & B \\ \hline 0 & C \end{array}\right).$$

où $B \in M_{1,n}(\mathbb{K})$ et $C \in M_n(\mathbb{K})$.

De plus, on a :

$$\pi_A(Q^{-1}AQ) = Q^{-1}\underbrace{\pi_A(A)}_{=0_{n+1}}Q = 0_{n+1}.$$

Or,

$$\pi_A(Q^{-1}AQ) = \pi_A\left(\left(\begin{array}{c|c} \lambda & B \\ \hline 0 & C \end{array}\right)\right) = \left(\begin{array}{c|c} \pi_A(\lambda) & B' \\ \hline 0 & \pi_A(C) \end{array}\right) \text{ où } B' \in M_{1,n}(\mathbb{K}).$$

Par suite, $\pi_A(C) = 0_n$ i.e. π_A est un polynôme annulateur de C qui est scindé par hypothèse ; donc le polynôme minimal π_C de C est scindé car il divise π_A. Ainsi, par hypothèse de récurrence, C est trigonalisable. Par suite, il existe $T' \in M_n(\mathbb{K})$ triangulaire et $R \in GL_n(\mathbb{K})$ tels que $C = RT'R^{-1}$. Alors, si on pose :

$$P' = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & R \end{array}\right) \quad \text{et} \quad P = QP'$$

on obtient :

$$P^{-1}AP = P'^{-1}Q^{-1}AQP' = P'^{-1}\left(\begin{array}{c|c} \lambda & B \\ \hline 0 & C \end{array}\right)P' = \left(\begin{array}{c|c} \lambda & BR \\ \hline 0 & T' \end{array}\right).$$

Donc $T = P^{-1}AP$ est triangulaire ; d'où A est trigonalisable. Par suite, $\mathcal{P}_{n+1}$ est vraie. Ce qui achève la récurrence.

Montrons maintenant l'implication iii) $\Rightarrow$ i). Supposons π_u scindé. Soit $\mathcal{B}$ une base de E et $A = \mathrm{Mat}_{\mathcal{B}}(u)$. Comme $\pi_A = \pi_u$, π_A est scindé, et donc, d'après le résultat précédent, A est trigonalisable, ce qui implique que u l'est aussi.

$\square$

Corollaire 16.

Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E stable par u. Si u est trigonalisable, alors l'endomorphisme $u_F \in \mathcal{L}(F)$ induit par u sur F est trigonalisable.

Démonstration.

Si u est trigonalisable, alors π_u est scindé d'après le théorème précédent. Or on a $\pi_u(u_F) = 0$: en effet, pour tout $x \in F$, $\pi_u(u_F)(x) = \pi_u(u)(x) = 0_E$; par suite, u_F possède un polynôme annulateur scindé. D'après le théorème précédent, u_F est trigonalisable. □

5. Théorème de Cayley-Hamilton

Théorème 10. *Théorème de Cayley-Hamilton*

— Soit $u \in \mathcal{L}(E)$. Le polynôme caractéristique χ_u de u est un polynôme annulateur de u i.e.

$$\chi_u(u) = 0.$$

Autrement dit, π_u divise χ_u.

— Soit $A \in M_n(\mathbb{K})$. Le polynôme caractéristique χ_A de A est un polynôme annulateur de A i.e.

$$\chi_A(A) = 0_n.$$

Autrement dit, π_A divise χ_A.

Démonstration Non exigible.

On suppose u trigonalisable. Soit $\mathcal{B} = (e_1, ..., e_n)$ une base de trigonalisation de u. On note $T = \mathrm{Mat}_{\mathcal{B}}(u)$ et on a, pour $\lambda_1, ..., \lambda_n$ les n valeurs propres de u (pas forcément distinctes donc) :

$$T = (t_{ij}) = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

Par suite, on a $\chi_u = \chi_T = \prod_{i=1}^{n}(X - \lambda_i)$. Pour $i = 1, ..., n$, on note $F_i = \mathrm{Vect}(e_1, ..., e_i)$, $F_0 = \{0_E\}$ et $P_i = X - \lambda_i$. Alors on a :

$$\chi_u(u) = P_1(u) \circ ... \circ P_n(u).$$

Montrons que pour $i \in [\![1, n]\!]$, que $P_i(u)(F_i) \subset F_{i-1}$.

- Cas $i = 1$. On a
$$P_1(u)(e_1) = u(e_1) - \lambda_1 e_1 = \lambda_1 e_1 - \lambda_1 e_1 = 0_E,$$
d'où $P_1(u)(F_1) \subset \{0_E\} = F_0$.
- Cas $i \in [\![1, n-1]\!]$. Pour $x \in F_i$, on a $x = \alpha e_i + \underbrace{x_{i-1}}_{\in F_{i-1}}$.

  Comme F_{i-1} est stable par u, alors F_{i-1} est stable par $P_i(u)$ d'où $P_i(u)(x_{i-1}) \in F_{i-1}$. De plus, on a :

$$\begin{aligned} P_i(u)(e_i) &= u(e_i) - \lambda_i e_i \\ &= (t_{1i}e_1 + ... + t_{i-1,i}e_{i-1} + \underbrace{t_{ii}}_{=\lambda_i} e_i) - \lambda_i e_i \\ &= t_{1i}e_1 + ... + t_{i-1,i}e_{i-1} \in F_{j-1}. \end{aligned}$$

Donc $P_i(u)(x) = \alpha P_i(u)(e_i) + P_i(u)(x_{i-1}) \in F_{j-1}$.

Alors, on a bien, pour tout $i \in [\![1, n]\!]$, que $P_i(u)(F_i) \subset F_{i-1}$.
Ainsi, on a :

$$\begin{aligned} \chi_u(u)(E) &= P_1(u) \circ ... \circ P_{n-1}(u) \circ P_n(u)(F_n) \\ &\subset P_1(u) \circ ... \circ P_{n-1}(u)(F_{n-1}) \\ &\subset \\ &\vdots \\ &\subset \\ &\subset P_1(u)(F_1) \\ &\subset F_0 = \{0_E\}. \end{aligned}$$

Et donc :

$$\chi_u(u)(E) = \{0_E\} \text{ i.e. } \chi_u(u) = 0.$$

Si u n'est pas trigonalisable, alors on considère une matrice $A \in M_n(\mathbb{K})$ représentant u comme une matrice de $M_n(\mathbb{C})$. Alors A est trigonalisable dans $M_n(\mathbb{C})$ et donc l'endomorphisme u' de $\mathbb{C}^n$ canoniquement associé à A l'est aussi. D'après ce qui précède, on a $\chi_{u'}(u') = 0$ et donc $\chi_A(A) = 0_n$. Le polynôme caractéristique de A est à coefficients dans $\mathbb{K}$, donc on a également $\chi_u(u) = 0$.

Dans tous les cas, on $\chi_u(u) = 0$. □

Pour ce corollaire, on rappelle qu'on a prouvé précedemment que les polynômes caractéristique et minimal possèdent les mêmes facteurs irréductibles dans $\mathbb{K}[X]$.

Corollaire 17.

Soit $u \in \mathcal{L}(E)$. Soit $P_1, ..., P_k$ la liste facteurs irréductibles distincts des polynômes minimal π_u et caractéristique χ_u de u.
Si $\pi_u = \prod_{i=1}^{k} P_i^{p_i}$ et $\chi_u = \prod_{i=1}^{k} P_i^{m_i}$ sont les décompositions en facteurs irréductibles de π_u et χ_u dans $\mathbb{K}[X]$ alors :

$$\text{pour tout } i \in [\![1, k]\!],\ p_i \leq m_i.$$

De même dans le cas matriciel.

Démonstration.

D'après le théorème de Cayley-Hamilton, on a $\pi_u|\chi_u$. Soit $i \in [\![1, k]\!]$. Alors $P_i^{n_i}|\prod_{j=1}^{k} P_j^{n_j} = \pi_u = \chi_u|\chi_u = \prod_{j=1}^{k} P_j^{m_j}$; donc par transitivité, $P_i^{n_i}|\prod_{i=1}^{k} P_i^{m_i}$. Or, pour tout $j \neq i$, $P_i^{n_i}, P_j^{m_j}$ sont premiers entre eux car P_i, P_j sont polynômes irréductibles distincts ; donc, d'après le lemme de Gauss, $P_i^{n_i} | P_i^{m_i}$ d'où $n_i \leq m_i$. $\square$

Exercice 24.

Soit $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}$. Montrer que A est inversible et exprimer A^{-1} en fonction de puissances de A.

Correction.

On a $\chi_A = X^3 - 3X^2 + 4X - 2$, donc d'après le théorème de Cayley-Hamilton,

$$0_n = \chi_A(A) = A^3 - 3A^2 + 4A - 2.$$

Par suite,

$$A^{-1} = \frac{1}{2}(A^2 - 3A + 4I_3).$$

Exercice 25.

Soit $A \in M_n(\mathbb{R})$ une matrice nilpotente d'indice $p \in \mathbb{N}^*$. Retrouver grâce au théorème de Cayley-Hamilton, que $p \leq n$.

Correction.

Comme A est nilpotente d'indice p, son polynôme minimal est $\pi_A = X^p$. Celui-ci est scindé, donc χ_A l'est aussi, et comme ils ont les même facteurs irréductibles et $\deg(\chi_A) = n$, on a $\chi_A = X^n$. Or, d'après le theorème de Cayley-Hamilton, $X^p \pi_A | \chi_A = X^n$ d'où $p \leq n$.

6. Sous-espaces caractéristiques

Tout les énoncés suivants sont directement transposables au cas matriciel.

Définition 21. *Sous-espace caractéristique*

Soit $u \in \mathcal{L}(E)$. Pour $\lambda \in \mathbb{K}$ une valeur propre de u, on appelle **sous-espace caractéristique de u associé à** λ et on note $C_\lambda(u)$ le sous-espace vectoriel de E :

$$C_\lambda(u) = \operatorname{Ker}\left((u - \lambda \mathrm{Id}_E)^{m(\lambda)}\right)$$

où $m(\lambda)$ désigne la multiplicité de la valeur propre λ.

Exercice 26.

Déterminer les sous-espaces caractéristiques de

$$A = \begin{pmatrix} 6 & -6 & 5 \\ -4 & -1 & 10 \\ 7 & -6 & 4 \end{pmatrix}$$

Proposition 40.

Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. Si λ est une valeur propre de u de multiplicité $m(\lambda)$, alors :

$$\dim(C_\lambda(u)) = m(\lambda).$$

Démonstration.

On suppose que $\lambda \in \mathrm{Sp}(u)$. On note $m = m(\lambda)$ et $p = \dim(C_\lambda(u))$. Alors on a $\chi_u = P(X-\lambda)^m$ où $(X-\lambda)$ et P sont premiers entre eux. Comme χ_u est annulateur de u d'après le Théorème de Cayley-Hamilton, on a, d'après le lemme des noyaux :

$$E = \mathrm{Ker}(P(u)) \oplus C_\lambda(u).$$

De plus, comme $P(u)$ et $(u-\lambda \mathrm{Id}_E)^m$ commutent avec u, $\mathrm{Ker}(P(u))$ et $C_\lambda(u)$ sont stables par u. On considère alors les endomorphismes induits v et w par u sur $\mathrm{Ker}(P(u))$ et $C_\lambda(u)$ respectivement. Comme Q est un polynôme annulateur de v et que λ n'est pas racine de Q, alors λ n'est pas valeur propre de v. Ainsi, $(X-\lambda)$ et χ_v sont premier entre eux.
Comme $(X-\lambda)^m$ est annulateur de w, alors λ est la seule racine de w et donc $\chi_w = (X-\lambda)^p$. Par suite, on a $P(X-\lambda)^m = \chi_u = \chi_v{\cdot}\chi_w = \chi_v{\cdot}(X-\lambda)^p$. Comme $(X-\lambda)^p$ et P sont premiers entre eux, $(X-\lambda)^p | (X-\lambda)^m$ d'où $p \le m$ puis, comme $(X-\lambda)^p$ et χ_v sont premiers entre eux, $(X-\lambda)^m | (X-\lambda)^p$ d'où $m \le p$.
Il en résulte que $m = p$. □

Proposition 41.

Soit $u \in \mathcal{L}(E)$ tel que χ_u est scindé. Si $\pi_u = \prod_{i=1}^{k}(X-\lambda_i)^{p_i}$ où $\lambda_1, .., \lambda_k$ sont exactement toutes les valeurs propres distinctes de u, alors, pour tout $i \in [\![1,k]\!]$:

$$\mathrm{Ker}\left((u-\lambda_i \mathrm{Id}_E)^{p_i}\right) = C_{\lambda_i}(u) \; \left(= \mathrm{Ker}\left((u-\lambda_i\mathrm{Id}_E)^{m(\lambda_i)}\right)\right)$$

Correction.

D'après le théorème de Cayley-Hamilton, $\pi_u | \chi_u$ donc pour tout $i \in [\![1,k]\!]$, $p_i \le m(\lambda_i)$. Par suite, $\mathrm{Ker}\left((u-\lambda_i\mathrm{Id}_E)^{p_i}\right) \subset C_{\lambda_i}(u)$; en effet, pour $f \in \mathcal{L}(E)$, si $p \le q$, $\mathrm{Ker}(f^p) \subset \mathrm{Ker}(f^q)$.
En notant $d_i = \dim(\mathrm{Ker}\left((u-\lambda_i\mathrm{Id}_E)^{p_i}\right))$, on a donc $d_i \le m(\lambda_i) = \dim(C_{\lambda_i}(u))$

De plus, comme π_u et χ_u sont annulateurs de u, on a, d'après le lemme des noyaux :

$$\bigoplus_{i=1}^{k} \operatorname{Ker}\left((u-\lambda_i \mathrm{Id}_E)^{p_i}\right) = E = \bigoplus_{i=1}^{k} C_{\lambda_i}(u).$$

Ainsi,

$$\sum_{i=1}^{k} d_i = n = \sum_{i=1}^{k} m(\lambda_i).$$

Par suite, on a pour tout $i \in [\![1,k]\!]$, $d_i = m(\lambda_i)$: en effet, comme pour tout $i \in [\![1,k]\!]$, $d_i \leq m(\lambda_i)$, si, par l'absurde, il existe $i_0 \in [\![1,k]\!]$ tel que $d_{i_0} < m(\lambda_{i_0})$, alors $\sum_{i=1}^{k} d_i < \sum_{i=1}^{k} m(\lambda_i)$, contradiction.
Ainsi, pour tout $i \in [\![1,k]\!]$, comme $\operatorname{Ker}\left((u-\lambda_i \mathrm{Id}_E)^{p_i}\right)$ est un sous-espace vectoriel de $C_{\lambda_i}(u)$ et qu'ils otn la même dimension (finie), on en déduit $\operatorname{Ker}\left((u-\lambda_i \mathrm{Id}_E)^{p_i}\right) = C_{\lambda_i}(u)$.

Théorème 11.

Soit $u \in \mathcal{L}(E)$ et π_u son polynôme minimal. Si π_u est scindé de racines (pas forcément simples) deux à deux distinctes $\lambda_1, ..., \lambda_k \in \mathbb{K}$, alors, pour $i \in [\![1,k]\!]$,

$$u_{C_{\lambda_i}(u)} = \lambda_i \mathrm{Id}_{C_{\lambda_i}(u)} + n_i,$$

où n_i est un endomorphisme nilpotent de $C_{\lambda_i}(u)$. De plus, dans une base $\mathcal{B}$ adaptée de E à la somme directe $\bigoplus_{i=1}^{k} C_{\lambda_i}(u)$, on a :

$$\operatorname{Mat}_{BB}(u) = \begin{pmatrix} \lambda_1 I_{m(\lambda_1)} + N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k I_{m(\lambda_k)} + N_k \end{pmatrix}$$

où pour tout $i \in [\![1,k]\!]$, N_i est la matrice de n_i dans la base de $C_{\lambda_i}(u)$ extraite de la base $\mathcal{B}$.
Dans le cas matriciel, cet énoncé se résume à : toute matrice de polynôme minimal scindé est semblable à une matrice de la forme précédente.

Démonstration.

On suppose que $\pi_u = \prod_{i=1}^{k}(X-\lambda_i)^{p_i}$. Alors, comme $\pi_u(u) = 0$, d'après le lemme de décomposition des noyaux, on a :

$$E = \operatorname{Ker}(\pi_u) = \bigoplus_{i=1}^{k} \operatorname{Ker}\left((u-\lambda_i \mathrm{Id}_E)^{p_i}\right).$$

Pour $i = 1, ..., k$, on pose $F_i = \operatorname{Ker}\left((u-\lambda_i \mathrm{Id}_E)^{p_i}\right)$. Soit $i \in [\![1,k]\!]$. Alors F_i est stable par u car u et le polynôme en u, $(u-\lambda_i \mathrm{Id}_E)^{p_i}$ commutent.
On note alors u_{F_i} l'endomorphisme induit par u sur F_i. On a, pour tout $x \in F_i$:

$$(u_{F_i} - \lambda_i \mathrm{Id}_{F_i})^{p_i}(x) = (u-\lambda_i \mathrm{Id}_E)^{p_i}(x) = 0_E$$

donc $(u_{F_i} - \lambda_i \mathrm{Id}_{F_i})^{p_i} = 0 \in \mathcal{L}(F_i)$. Par suite $n_i = u_{F_i} - \lambda_i \mathrm{Id}_{F_i} \in \mathcal{L}(F_i)$ est nilpotent et on a bien :

$$u_{F_i} = \lambda_i \mathrm{Id}_{F_i} + n_i.$$

$\square$

Exercice 27. *Projecteurs Spectraux*

Soit $u \in \mathcal{L}(E)$. On suppose que son polynôme minimal π_u est scindé de la forme $\pi_u = \prod_{i=1}^{k}(X - \lambda_i)^{q_i}$.

— On note, pour $i \in [\![1, k]\!]$, $P_i = \prod_{\substack{j=1 \\ j \neq i}}^{k}(X - \lambda_j)^{q_j} = \dfrac{\pi_u}{(X - \lambda_i)^{q_i}}$.

— On note $p_1, ..., p_k$ les projecteurs associée à la somme directe des sous-espaces caractéristiques ; ceux-ci sont appelés *les projecteurs spectraux* de u.

1. Montrer que, pour tout $i \in [\![1, k]\!]$, il existe $U_i \in \mathbb{R}[X]$ tel que

$$U_i P_i \equiv 1 \bmod (X - \lambda_i)^{q_i}.$$

2. En déduire que $(U_i P_i)(u) = p_i$.

3. Avec un "bon" choix de chaque U_i, montrer que $\dfrac{1}{\pi_u} = \displaystyle\sum_{i=1}^{k} \frac{U_i}{(X - \lambda_i)^{q_i}}$.

Application : On pose $u \in \mathcal{L}(\mathbb{R}^4)$ l'endomorphisme canoniquement associé à $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$

1. Déterminer π_u puis effectuer la décomposition en éléments simple de $\frac{1}{\pi_u}$.
2. En déduire une expression des projecteurs spectraux de u comme sous forme de polynômes en u.

Correction.

1. Comme les λ_j sont distincts, P_i et $(X - \lambda_i)^{q_i}$ sont premiers entre eux. Par suite, d'après le théorème de Bézout, il existe $U_i, V_i \in \mathbb{K}[X]$ tels que :

$$U_i P_i + V_i (X - \lambda_i)^{q_i} = 1$$

Par suite,

$$U_i P_i \equiv 1 \bmod (X - \lambda_i)^{q_i}$$

2. On va utiliser le fait que, pour $p_1, ..., p_k \in \mathcal{L}(E)$ et $F_1, ..., F_k$ des sous-espaces vectoriels de E en somme directe :

$p_1, ...p_k$ sont les projecteurs associés à $\bigoplus_{i=1}^{k} F_i$ si, et seulement si,

— $\sum_{i=1}^{k} p_i = \mathrm{Id}_E$; $p_i \circ p_j = \mathbf{0}$ pour tous $i \neq j$ et,

— pour tout i, $\mathrm{Im}(p_i) = F_i$; $\mathrm{Ker}(p_i) = \bigoplus_{\substack{j=1 \\ j\neq i}}^{k} F_j$.

Allons-y :

— Pour tous $j \neq i$, on a $U_j P_j \equiv 0 \bmod (X-\lambda_i)^{q_i}$, donc

$$S = \sum_{j=1}^{k} U_j P_j \equiv 1 \bmod (X-\lambda_i)^{q_i}$$

D'où, pour tout $i \in [\![1,k]\!]$, $(X-\lambda_i)^{q_i} | S-1$. Or les $(X-\lambda_i)^{q_i}$ sont premiers entre eux, donc leur produit, qui vaut π_u divise $S-1$ et ainsi :

$$\sum_{j=1}^{k} U_j P_j = S \equiv 1 \bmod \pi_u$$

Il en résulte qu'il existe $Q \in \mathbb{K}[X]$ tel que :

$$\sum_{j=1}^{k} U_j P_j = 1 + \pi_u . Q$$

donc :

$$\sum_{j=1}^{k} (U_j P_j)(u) = \mathrm{Id}_E + \underbrace{\pi_u(u)}_{=\mathbf{0}} \circ Q(u) = \mathrm{Id}_E$$

De plus, pour tous $i, j \in [\![1,k]\!]$ avec $i \neq j$, on a :

$$(U_i P_i)(U_j P_j) = U_i U_j R \pi_u \text{ où } R = \prod_{\substack{m=1 \\ m \notin \{i,j\}}}^{k} (X-\lambda_m)^{q_m}$$

Par suite,

$$(U_i P_i)(u) \circ (U_j P_j)(u) = (U_i U_j R)(u) \circ \underbrace{\pi_u(u)}_{=\mathbf{0}} = \mathbf{0}.$$

— Soit $i \in [\![1,k]\!]$. Notons $f_i = (U_i P_i)(u)$. Rappelons que $C_{\lambda_i}(u) = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{q_i}$ (proposition 41).

- Montrons $\mathrm{Im}(f_i) = C_{\lambda_i}(u)$ par double inclusion.
  Soit $y \in \mathrm{Im}(f_i)$. Alors il existe $x \in E$ tel que $y = f_i(x)$. On a :

  $(u-\lambda_i \mathrm{Id}_E)^{q_i}(y) = (u-\lambda_i \mathrm{Id}_E)^{q_i} \circ f_i(x) = ((X-\lambda_i)^{q_i} U_i P_i)(u)(x) = U_i(u) \circ \pi_u(u)(x) = 0_E$.

  D'où $y \in C_{\lambda_i}(u)$. Ainsi $\mathrm{Im}(f_i) \subset C_{\lambda_i}(u)$.

  Réciproquement, soit $x \in C_{\lambda_i}(u)$. Alors, comme pour tous $j \neq i$, $(X-\lambda_i)^{q_i}$ divise P_j et donc $U_j P_j$, on a :

  $$(U_j P_j)(u)(x) = 0_E \text{ car } x \in C_{\lambda_i}(u) = \mathrm{Ker}(u - \lambda_i \mathrm{Id}_E)^{q_i}.$$

  Or, d'après ce qui précède, on a $\mathrm{Id}_E = \sum_{j=1}^{k} (U_j P_j)(u)$, donc :

  $$x = \sum_{j=1}^{k} (U_j P_j)(u)(x) = (U_i P_i)(u)(x) \in \mathrm{Im}(f_i)$$

  Ainsi $C_{\lambda_i}(u) \subset \mathrm{Im}(f_i)$.

- Montrons $\mathrm{Ker}(f_i) = \bigoplus_{\substack{j=1 \\ j \neq i}}^{k} C_{\lambda_j}(u)$ par double inclusion.

  Soit $x \in \mathrm{Ker}(f_i)$. Comme dans le point précédent, on a $\mathrm{Id}_E = \sum_{j=1}^{k}(U_jP_j)(u)$ donc :

$$x = \sum_{j=1}^{k}(U_jP_j)(u)(x) = \sum_{\substack{j=1 \\ j \neq i}}^{k}(U_jP_j)(u)(x) \in \bigoplus_{\substack{j=1 \\ j \neq i}}^{k} C_{\lambda_j}(u)$$

  car pour tous j, $\mathrm{Im}(f_j) \subset C_{\lambda_j}(u)$. Ainsi $\mathrm{Ker}(f_i) \subset \bigoplus_{\substack{j=1 \\ j \neq i}}^{k} C_{\lambda_j}(u)$.

  Soit $j \in [\![1,k]\!]$ tel que $j \neq i$. Soit $x \in C_{\lambda_j}(u)$. On a $(X-\lambda_j)^{q_j}$ divise P_i et donc U_iP_i, d'où :

$$(U_iP_i)(u)(x) = 0_E \text{ car } x \in C_{\lambda_j}(u) = \mathrm{Ker}(u-\lambda_j\mathrm{Id}_E)^{q_j}.$$

  Par suite, pour tous $j \in [\![1,k]\!]$, $C_{\lambda_j}(u) \subset \mathrm{Ker}(f_i)$ et ainsi, $\bigoplus_{\substack{j=1 \\ j \neq i}}^{k} C_{\lambda_j}(u) \subset \mathrm{Ker}(f_i)$.

Il en résulte, d'après la caractérisation des projecteurs associés à une somme directe, que, pour tous $i \in [\![1,k]\!]$,

$$(U_iP_i)(u) = p_i.$$

— On a montrer précédemment qu'il existe $Q \in \mathbb{K}[X]$ tel que :

$$\sum_{j=1}^{k} U_jP_j = 1 + \pi_u.Q$$

Et comme $U_iP_i \equiv 1 \bmod (X-\lambda_i)^{q_i}$, quitte à prendre le reste de la division euclidienne de U_i par $X-\lambda_i)^{q_i}$, on peut supposer que $\deg(U_i) < q_i$.
Ainsi, on a :

$$\deg(\sum_{j=1}^{k} U_jP_j) \leq \max_{j=1,\ldots,k}(\underbrace{\deg(U_j)}_{<q_j} + \underbrace{\deg(P_j)}_{=\deg(\pi_u)-q_j}) < \deg(\pi_u)$$

Ainsi, par comparaison des degré dans l'égalité précédente, on obtient $Q = 0$ et donc :

$$\sum_{j=1}^{k} U_jP_j = 1$$

Remarque : on aurait également pu obtenir cette relation et donc les bons U_i dès le début en utilisant le théorème de Bézout appliqué aux polynômes P_i qui sont premiers entre eux (dans leur ensemble mais pas deux à deux !

Ainsi, on obtient le résultat :

$$\frac{1}{\pi_u} = \frac{\sum_{i=1}^{k} U_iP_i}{\pi_u} = \sum_{i=1}^{k} \frac{U_i}{(X-\lambda_i)^{q_i}}$$

Application :

1. On a $\pi_u = (X-2)^2(X+1)$ et ainsi :

$$\frac{1}{\pi_u} = \frac{-\frac{1}{9}X + \frac{5}{9}}{(X-2)^2} + \frac{\frac{1}{9}}{X+1}$$

2. En utilisant les notations de la partie précédente, on a $P_1 = (X-2)^2$, $P_2 = (X+1)$ puis, avec les calculs précédents, $U_1 = -\frac{1}{9}X + \frac{5}{9}$ et $U_2 = \frac{1}{9}$. Par suite, on obtient :

$$p_1 = (U_1P_1)(u) = \frac{1}{9}(-u + 5\mathrm{Id}_E) \circ (u - 2\mathrm{Id}_E)^2 \text{ et } p_2 = (U_2P_2)(u) = \frac{1}{9}(u + \mathrm{Id}_E)$$

Annexe

Matrices symétriques

Nous démontrerons le résultat suivant dans le chapitre relatif aux espaces préhilbertiens :

Théorème.

Soit $A \in M_n(\mathbb{R})$ une matrice à **coefficients réels** et **symétrique** i.e. ${}^tA = A$.
Alors il existe $P \in GL_n(\mathbb{R})$ orthogonale i.e. ${}^tP = P^{-1}$ et une matrice diagonale $D \in M_n(\mathbb{R})$ telle que :

$$A = PD{}^tP.$$

Ainsi, en particulier, A est diagonalisable.

Remarque.

Attention ce résultat est faux pour les matrices à coefficients complexes.

Exercice 28.

Diagonaliser à l'aide d'une matrice orthogonale la matrice $\begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$.

Correction.

On sait que A est diagonalisable car A est symétrique à coefficients réels. Mais redémontrons le tout de même en déterminant les sous-espaces propres.

On remarque tout d'abord que $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est un vecteur propre associé à la valeur propre 6 dont le sous-espace propre $E_6(A)$ est donc de dimension au moins 1. De plus, les 2ème et 3ème colonnes A_2 et A_3 de A étant égale à la première A_1, le rang de A est égal à 1 et donc son noyau est de dimension 2 d'après le théorème du rang. Ainsi, 0 est valeur propre et le sous-espace propre associé $E_0(A)$ est de dimension 2 car il est égal au noyau de A. La somme des dimensions des sous-espaces propres étant inférieure à 3, on en déduit que $\dim(E_6(A)) = 1$ et donc $E_6(A) = \text{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right)$.

Comme $\dim(E_6(A)) + \dim(E_0(A)) = 3$, alors A est diagonalisable. Il ne reste plus qu'à déterminer $E_0(A)$ en déterminant deux vecteurs qui l'engendrent (et si, qui plus est, ils sont orthogonaux, ça arrange nos affaires pour la suite !).

On remarque que $A\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = A_2 - A_3 = 0_{3,1}$ et $A\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} = 2A_1 - A_2 - A_3 = 0_{3,1}$ et ainsi,

$E_0(A) = \text{Vect}\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}\right)$.

Ces trois vecteurs propres formant une famille orthogonale de $M_{3,1}(\mathbb{R})$ (muni de son produit scalaire canonique), il suffit alors de les normer pour former une bon de $M_{3,1}(\mathbb{R})$ et ainsi produire une matrice de passage P qui est orthogonale. Ce qui nous donne au final :

$$A = PD^tP \text{ où } D = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et } P = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \end{pmatrix}$$

Chapitre V

Séries numériques et vectorielles

Table des matières

Dans ce chapitre, E désigne un espace vectoriel normé sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. On notera $\|\cdot\|$ la norme dont est munie E.

Partie A

Séries de vecteurs

1. Définitions

Définition 1. *Série*

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. On appelle **série de terme général** u_n et on note $\sum u_n$ la **suite** $(S_n)_{n\in\mathbb{N}}$ où, pour $n \in \mathbb{N}$:

$$S_n = \sum_{i=0}^{n} u_i$$

Pour chaque $n \in \mathbb{N}$, le terme $S_n = \sum_{i=0}^{n} u_i$ est appelé **somme partielle** d'ordre n de la série $\sum u_n$.

Remarque 1.

— Une série est une **suite** particulière ! Toutes les méthodes connues pour les suites peuvent donc s'appliquer pour les séries ! De plus, leurs particularités parmi les suites nous permettent de développer des outils adaptés à leur étude. Il est donc a priori plus aisé d'étudier une série quelconque qu'une suite quelconque !

— Une série est donc la suite de ses sommes partielles.

— Si le terme général u_n d'une série n'est défini qu'à partir du rang $n_0 \in \mathbb{N}$, on notera $\displaystyle\sum_{n\geq n_0} u_n$ la série de somme partielle d'ordre $n \geq n_0$: $S_n = \displaystyle\sum_{i=n_0}^{n} u_i$.

Définition 2. *Somme d'une série*

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. On dit que la série $\sum u_n$ **converge** dans $(E, \|\cdot\|)$ si la suite $(S_n)_{n\in\mathbb{N}}$ de ses sommes partielles converge dans $(E, \|\cdot\|)$.

Dans ce cas, on appelle **somme de la série** $\sum u_n$ et on note $\sum_{n=0}^{+\infty} u_n$ la quantité :

$$\sum_{n=0}^{+\infty} u_n = \lim_{N\to+\infty} S_N = \lim_{N\to+\infty} \sum_{n=0}^{N} u_n.$$

Si la série $\sum u_n$ ne converge pas, on dira qu'elle **diverge**.

Dans la suite, $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ désignent des suites à valeurs dans E.

Proposition 1.

Si la série $\sum u_n$ converge, alors la suite $(u_n)_{n\in\mathbb{N}}$ converge vers 0_E.

Démonstration.

On suppose que $\sum u_n$ converge. Alors la suite $(S_n)_{n\in\mathbb{N}}$ converge vers $S \in E$ et donc, pour tout $n \in \mathbb{N}^*$, on a

$$u_n = S_n - S_{n-1} \xrightarrow[n\to+\infty]{} S - S = 0_E.$$

□

Remarque 2.

La réciproque de cette proposition est fausse : la suite $(\frac{1}{n})_{n\in\mathbb{N}^*}$ converge vers 0 et la série harmonique $\displaystyle\sum_{n\geq 1} \frac{1}{n}$ diverge !

Cette proposition motive donc la définition suivante :

Définition 3. *Divergence grossière*

Si la suite $(u_n)_{n\in\mathbb{N}}$ ne converge pas vers 0_E, on dit que la série $\sum u_n$ **diverge grossièrement**.

Définition-Proposition 4. *Reste d'une série*

Si la série $\sum u_n$ converge, alors pour tout $n \in \mathbb{N}$, la série $\sum_{k\geq n+1} u_k$ converge. Ainsi, pour $n \in \mathbb{N}$, on appelle **reste d'ordre** n de la série $\sum u_n$ et on note R_n la somme de la série $\sum_{k\geq n+1} u_k$ i.e.

$$R_n = \sum_{k=n+1}^{+\infty} u_k.$$

Remarque 3.

On ne parle de reste d'une série que si elle converge. De plus, dans ce cas, on remarque que pour tout $n \in \mathbb{N}$,

$$R_n = S - S_n \text{ où } S = \sum_{n=0}^{+\infty} u_n.$$

Exercice 1.

Soit $\lambda \in \mathbb{K} \setminus \{1\}$ et $x \in E$. On considère la série $\sum u_n$ de terme général $u_n = \lambda^n x$. Discuter des cas évidents de convergence et de divergence grossière de $\sum u_n$.
En cas de convergence, déterminer la somme et les restes de la série.

Correction.

Soit $n \in \mathbb{N}$. On a :

$$S_n = \sum_{i=1}^{n} u_i = \sum_{i=1}^{n} \lambda^i x = \left(\sum_{i=1}^{n} \lambda_i\right) x = \frac{1-\lambda^{n+1}}{1-\lambda} x.$$

Ainsi, si $|\lambda| < 1$, $\sum u_n$ converge et on a :

$$S = \sum_{n=0}^{+\infty} u_n = \lim_{N \to +\infty} S_N = \frac{1}{1-\lambda} x,$$

et, pour tout $n \in \mathbb{N}$

$$R_n = S - S_n = \frac{1}{1-\lambda} x - \frac{1-\lambda^{n+1}}{1-\lambda} x = \frac{\lambda^{n+1}}{1-\lambda} x.$$

Exercice 2.

Soit $A \in M_n(\mathbb{K})$ une matrice nilpotente d'indice $p \in \mathbb{N}^*$. Déterminer la nature de la série $\sum A^n$ et en cas de convergence, déterminer sa somme.

Correction.

Pour tout $n \geq p$, $A^n = 0_n$, donc, pour $N \geq p$, on a :

$$\sum_{n=0}^{N} A^n = \sum_{n=0}^{p-1} A^n + \sum_{n=p}^{N} \underbrace{A^n}_{=0_n} = \sum_{n=0}^{p-1} A^n.$$

Ainsi, $\sum A^n$ est stationnaire en $\sum_{n=0}^{p-1} A^n$, d'où $\sum A^n$ converge et

$$\sum_{n=0}^{+\infty} A^n = \sum_{n=0}^{p-1} A^n = I_n + A + ... + A^{p-1}.$$

Proposition 2.

La suite $(R_n)_{n \in \mathbb{N}}$ des restes de la série $\sum u_n$ converge vers 0_E.

Démonstration.

On a, avec S la somme de la série et S_n la somme partielle d'ordre $n \in \mathbb{N}$ de la série :

$$\lim_{n \to +\infty} R_n = \lim_{n \to +\infty} (S - S_n) = S - S = 0_E.$$

□

Définition 5. *Série télescopique*

On dit que la série $\sum u_n$ est une **série télescopique** si le terme général u_n de la série s'écrit, pour tout $n \in \mathbb{N}$,

$$u_n = v_{n+1} - v_n.$$

Proposition 3.

On suppose que la série $\sum u_n$ est télescopique avec, pour tout $n \in \mathbb{N}$, $u_n = v_{n+1} - v_n$. Alors la suite $(v_n)_{n \in \mathbb{N}}$ et la série $\sum u_n$ sont de même nature i.e. $(v_n)_{n \in \mathbb{N}}$ converge si, et seulement si, $\sum u_n$ converge. De plus, en cas de convergence, on a :

$$\sum_{n=0}^{+\infty} u_n = \lim_{n \to +\infty} v_n - v_0.$$

Démonstration.

On a, pour tout $n \in \mathbb{N}$,

$$\sum_{i=0}^{n} u_i = \sum_{i=0}^{n} (v_{i+1} - v_i) = (v_{n+1} + \sum_{i=1}^{n} v_i) - ((\sum_{i=1}^{n} v_i) - v_0) = v_n - v_0.$$

d'où le résultat en passant à la limite quand $n \to +\infty$ dans cette égalité. □

2. Espace vectoriel des séries convergentes

Proposition 4.

L'ensemble des séries convergentes de terme général dans E est un sous-espace vectoriel de l'espace vectoriel des suites à valeurs dans E.

De plus, l'application qui à une série convergente associe sa somme est une application linéaire i.e. pour tous $\lambda, \mu \in \mathbb{K}$ et toutes séries $\sum u_n$, $\sum v_n$:

$$\sum_{n=0}^{+\infty} (\lambda u_n + \mu v_n) = \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n.$$

Démonstration.

On note $\mathcal{S}(E)$ l'ensemble des séries de terme général dans E et $\mathcal{S}_c(E)$ l'ensemble des séries convergentes.

> *Exercice :* Montrer que $\mathcal{S}(E)$ est un sous-espace vectoriel de $E^{\mathbb{N}}$ (ensemble des suites à valeurs dans E) - on démontrera (entre autres) que pour $\lambda, \mu \in \mathbb{K}$ et $\sum u_n, \sum v_n \in \mathcal{S}(E)$,
> $$\lambda \sum u_n + \mu \sum v_n = \sum (\lambda u_n + \mu v_n).$$

On rappelle (voire le chapitre Topologie des e.v.n.) que l'ensemble $c(E)$ des suites à valeurs dans E convergentes est un sous-espace vectoriel de $E^{\mathbb{N}}$.
On remarque que $\mathcal{S}_c(E) = \mathcal{S}(E) \cap c(E)$ donc $\mathcal{S}_c(E)$ est un sous-espace vectoriel de $E^{\mathbb{N}}$ comme intersection de sous-espaces vectoriels de $E^{\mathbb{N}}$.

De plus, considérons l'application $f : \mathcal{S}_c(E) \to \mathbb{K}$ telle que :

$$f : \sum u_n \mapsto \sum_{n=0}^{+\infty} u_n.$$

Soit $\lambda, \mu \in \mathbb{K}$ et $\sum u_n, \sum v_n \in \mathcal{S}_c(E)$. Alors,

$$\begin{aligned}
f(\lambda \sum u_n + \mu \sum v_n) &= f(\sum (\lambda u_n + \mu v_n)) \\
&= \sum_{n=0}^{+\infty} (\lambda u_n + \mu v_n) \\
&= \lim_{N \to +\infty} \sum_{n=0}^{N} (\lambda u_n + \mu v_n) \\
&= \lim_{N \to +\infty} \left(\lambda \sum_{n=0}^{N} u_n + \mu \sum_{n=0}^{N} v_n \right) \\
&= \lambda \lim_{N \to +\infty} \sum_{n=0}^{N} u_n + \mu \lim_{N \to +\infty} \sum_{n=0}^{N} v_n \\
&= \lambda \sum_{n=0}^{+\infty} u_n + \mu \sum_{n=0}^{+\infty} v_n \\
&= \lambda f(\sum u_n) + \mu f(\sum v_n)
\end{aligned}$$

Donc f est une application linéaire.

□

Exercice 3.

Soit $A = \frac{1}{2} \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \in M_2(\mathbb{R})$. Discuter de la nature de la série $\sum A^n$.

Correction.

On a $A = \frac{1}{2}(I_2 + J_a)$ où $J_a = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$ est nilpotente d'indice 2 et commute avec I_2. Ainsi en appliquant la formule du binôme, on obtient pour $n \geq 2$:

$$A^n = \frac{1}{2^n}(I_2 + J_a)^n = \sum_{i=0}^{n} \binom{n}{i} J_a^i I_2^{n-i} = I_2 + nJ_a;$$

par suite,

$$S_N = \sum_{n=0}^{N} A^n = \left(\sum_{n=0}^{N} \frac{1}{2^n}\right) I_2 + \left(\sum_{n=0}^{N} \frac{n}{2^n}\right) J_a$$

Donc $\sum A^n$ converge car les séries numériques $\sum \frac{1}{2^n}$ et $\sum \frac{n}{2^n}$ converge. En effet, $\frac{1}{2} < 1$ et $\frac{3}{4} < 1$ donc les séries géométriques $\sum \frac{1}{2^n}$ et $\sum (\frac{3}{4})^n$ convergent et on a $\frac{n}{2^n} = o((\frac{3}{4})^n)$ donc, par comparaison avec le terme général d'une série convergente, on en déduit que $\sum \frac{n}{2^n}$ converge (on aurait pu utiliser la règle de D'Alembert, que l'on verra dans la prochaine partie, pour conclure la convergence de $\sum \frac{n}{2^n}$).

Proposition 5. *Utilisation d'une base*

On suppose que E est de dimension finie p. Soit $\mathcal{B} = (e_1, ..., e_p)$ une base de E et $\sum u_n$ la série de terme général $u_n = \sum_{i=1}^n u_n^{(i)} e_i \in E$. Alors la série $\sum u_n$ converge si, et seulement si, pour tout $i \in [\![1, p]\!]$, la série numérique $\sum u_n^{(i)}$ converge. Dans ce cas, on a :

$$\sum_{n=0}^{+\infty} u_n = \sum_{i=1}^{p} \left(\sum_{n=0}^{+\infty} u_n^{(i)} \right) e_i.$$

Exercice 4.

1. Soit $p \in \mathbb{N}$ et $A \in M_p(\mathbb{K})$ une matrice diagonalisable dans $M_p(\mathbb{K})$ telle que pour tout $\lambda \in \mathrm{Sp}(A)$, $|\lambda| < 1$. Démontrer que $\sum A^n$ est convergente et déterminer sa somme.

2. Application : déterminer la somme de la série $\sum \left(\frac{1}{8} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 4 & 0 \\ 3 & 0 & 1 \end{pmatrix} \right)^n$.

Correction.

1. Comme A est diagonalisable, il existe $P \in GL_n(\mathbb{K})$ et $D = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ avec $\alpha_1, ..., \alpha_p \in \mathrm{Sp}(A)$ tels que $A = PDP^{-1}$. On a $|\alpha_i| < 1$, pour tout $i \in [\![1, p]\!]$ et donc, pour tout $N \in \mathbb{N}$,

$$\begin{aligned} S_N &= \sum_{n=0}^{N} A^n = P\left(\sum_{n=0}^{N} D^n\right)P^{-1} \\ &= P\ \mathrm{diag}\left(\frac{1-\alpha_1^{N+1}}{1-\alpha_1}, \dots, \frac{1-\alpha_p^{N+1}}{1-\alpha_p}\right)P^{-1} \\ &\xrightarrow[N\to+\infty]{} P\ \mathrm{diag}\left(\frac{1}{1-\alpha_1}, \dots, \frac{1}{1-\alpha_p}\right)P^{-1}. \end{aligned}$$

Il en résulte que $\sum A^n$ converge et $\displaystyle\sum_{n=0}^{+\infty} A^n = P\ \mathrm{diag}\left(\frac{1}{1-\alpha_1}, \dots, \frac{1}{1-\alpha_p}\right)P^{-1}$.

2. On a, pour $n \in \mathbb{N}$,

$$A = \frac{1}{8}\begin{pmatrix} 1 & 0 & 3 \\ 0 & 4 & 0 \\ 3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{-1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Donc

$$\sum_{n=0}^{+\infty} A^n = P\begin{pmatrix} \frac{4}{5} & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} {}^tP = \begin{pmatrix} \frac{7}{5} & 0 & \frac{3}{5} \\ 0 & 2 & 0 \\ \frac{3}{5} & 0 & \frac{7}{5} \end{pmatrix}$$

3. Série absolument convergente

Dans ce paragraphe, on suppose que E est de **dimension finie** p.

Définition 6. *Série absolument convergente*

On dit que la série $\sum u_n$ de terme général $u_n \in E$ est **absolument convergente** si la série à termes positifs $\sum \|u_n\|$ est convergente.

Théorème 1.

Si la série $\sum u_n$ converge absolument, alors $\sum u_n$ converge. De plus, on a :

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

Démonstration.

On suppose que $\sum u_n$ converge absolument. Alors $\sum \|u_n\|$ converge. Soit $\mathcal{B} = (e_1, ..., e_p)$ un base de E. Comme E est de dimension finie, toutes les normes sont équivalentes sur E ; ainsi, il existe $c \in \mathbb{R}_+$ tel que, pour tout $x \in E$, $\|x\|_\infty \leq c\|x\|$. Par suite, on a, pour tout $i \in [\![1, p]\!]$:

$$|u_n^{(i)}| \leq \|u_n\|_\infty \leq c\|u_n\|,$$

donc $\sum u_n^{(i)}$ est une suite numérique absolument convergente, donc convergente. Ainsi, les séries des composantes de $\sum u_n$ convergent, donc $\sum u_n$ converge. $\square$

Exercice 5.

Soit $p \in \mathbb{N}^*$. On munit l'espace vectoriel $M_p(\mathbb{K})$ d'une norme $\|\cdot\|$ qui est sous-multiplicative i.e. $\forall\, A, B \in M_p(\mathbb{K})$, $\|AB\| \leq \|A\|.\|B\|$.

Soit $A \in M_p(\mathbb{K})$ telle que $\|A\| < 1$.

1. Montrer que $\sum A^n$ est absolument convergente. Vers quelle limite converge la suite $(A^n)_{n\in\mathbb{N}}$?
2. En déduire que $I_p - A$ est inversible en déterminant son inverse.

Correction.

1. On a, par sous-multiplicativité de la norme $\|\cdot\|$, pour tout $n \in \mathbb{N}$, $\|A^n\| \leq \|A\|^n$. Or $\sum \|A\|^n$ est une série géométrique de raison $\|A\| < 1$, donc le terme général $\|A^n\|$ de la série à termes positifs est majoré par le terme général d'une série convergente. Par suite, $\sum \|A^n\|$ converge ; d'où la série $\sum A^n$ est absolument convergente.
   Comme $\sum A^n$ est absolument convergente, $\sum A^n$ converge. Par suite, $\lim_{n\to+\infty} A^n = 0_n$.

2. On a, pour $N \in \mathbb{N}$,
   $$(I_p - A)S_N = (I_p - A)\left(I_p + A + A^2 + \ldots + A^N\right) = I_p - A^{N+1},$$
   Par suite,
   $$(I_p - A)\sum_{n=0}^{+\infty} A^n = \lim_{n\to+\infty}(I_p - A)S_N = I_p - \lim_{n\to+\infty} A^{N+1} = I_p.$$
   Ainsi, $I_p - A$ est inversible et $(I_p - A)^{-1} = \sum_{n=0}^{+\infty} A^n$.

Partie B
Compléments sur les séries numériques

Dans cette partie, on considérera seulement des séries à termes **réels**.

1. Règle de D'Alembert

Proposition 6. *Comparaison logarithmique*

Soit (u_n) et v_n des suites de réels strictement positifs à partir d'un certain rang n_0 telles que pour tout $n \geq n_0$,

$$\frac{u_{n+1}}{u_n} \leq \frac{v_{n+1}}{v_n}.$$

Alors,
- si $\sum v_n$ converge, $\sum u_n$ converge ;
- si $\sum u_n$ diverge, $\sum v_n$ diverge.

Démonstration.

On a remarque que pour tout $n \geq n_0$, on a :

$$\frac{u_n}{u_{n_0}} = \prod_{i=n_0}^{n-1} \frac{u_{i+1}}{u_i} \leq \prod_{i=n_0}^{n-1} \frac{v_{i+1}}{u_i} = \frac{v_n}{v_{n_0}}$$

d'où :

$$u_n \leq \frac{u_{n_0}}{v_{n_0}} v_n.$$

Par suite, par comparaison de séries à termes positifs, on obtient que si $\sum v_n$ converge, alors $\sum u_n$ converge (et donc également la contraposée).

□

Corollaire 1.

Soit $\sum u_n$ une série à termes strictement positifs. Si, à partir d'un certain rang n_0 :
- il existe $M \in]0, 1[$ tel que pour tout $n \geq n_0$, $\dfrac{u_{n+1}}{u_n} \leq M$, alors $\sum u_n$ converge.
- pour tout $n \geq n_0$, $\dfrac{u_{n+1}}{u_n} \geq 1$, alors $\sum u_n$ diverge.

Démonstration.

- On considère la suite géométrique $(v_n)_{n\in\mathbb{N}}$ de raison $0 < M < 1$. Alors $\sum v_n$ converge, et, à partir du certain rang n_0, on a :

$$\frac{u_{n+1}}{u_n} \leq M = \frac{v_{n+1}}{v_n}.$$

Par suite, d'après la proposition précédente, $\sum u_n$ converge.

— La suite (u_n) à termes strictement positifs est croissante à partir du rang n_0, donc (u_n) ne converge pas vers 0. Ainsi $\sum u_n$ diverge grossièrement. (on aurait pu, comme dans le cas précédent, comparer $\sum u_n$ avec la série géométrique de raison 1).

□

Théorème 2. *Règle de D'Alembert*

Soit $\sum u_n$ une série à termes strictement positifs. On suppose que la suite $\left(\frac{u_{n+1}}{u_n}\right)_{n\in\mathbb{N}}$ admet une limite $\ell \in \mathbb{R}_+ \cup \{+\infty\}$. Alors :

— Si $\ell < 1$, alors $\sum u_n$ converge.

— Si $\ell > 1$, alors $\sum u_n$ diverge.

Démonstration.

- 1er cas : $\ell \in \mathbb{R}_+$:
  Comme $\frac{u_{n+1}}{u_n} \xrightarrow[n\to+\infty]{} \ell$, alors, pour tout $\varepsilon > 0$, il existe un certain rang n_0, tel que, pour tout $n \geq n_0$, $\frac{u_{n+1}}{u_n} \in [\ell - \varepsilon, \ell + \varepsilon]$.
  — On suppose $\ell < 1$. Alors pour $\varepsilon < 1 - \ell$, on a, pour tout $n \geq n_0$, $\frac{u_{n+1}}{u_n} \leq \ell + \varepsilon$ et on applique le corollaire précédent en posant $M = \ell + \varepsilon < 1$. Par suite, $\sum u_n$ converge.
  — On suppose $\ell > 1$. Alors pour $\varepsilon = \ell - 1$, on a, pour tout $n \geq n_0$, $\frac{u_{n+1}}{u_n} \geq \ell - \varepsilon = 1$ et on applique le corollaire précédent. Par suite, $\sum u_n$ diverge.
  - 2nd cas : $\ell = +\infty$:
    Comme $\frac{u_{n+1}}{u_n} \xrightarrow[n\to+\infty]{} \ell$, alors, pour tout $A \geq 0$, il existe un certain rang n_0, tel que, pour tout $n \geq n_0$, $\frac{u_{n+1}}{u_n} \geq A$.
    Ainsi, pour $A = 1$, il existe un rang $n_0 \in \mathbb{N}$ tel que pour tout entier $n \geq n_0$, $\frac{u_{n+1}}{u_n} \geq 1$.
    D'après le corollaire précédent, $\sum u_n$ diverge.

□

Remarque 4.

— On peut voir, en regardant la construction de la démonstration de cette règle, que l'on compare la série $\sum u_n$ à une série géométrique. Ainsi, cette règle ne peut servir à étudier des suites dont la croissance est moins forte que celle d'une série géométrique.

— Ainsi dans le cas $\ell = 1$, on ne peut pas conclure comme en témoigne par exemple $\sum \frac{1}{n}$ et $\sum \frac{1}{n^2}$.

Exercice 6. *Applications de la règle de D'Alembert*

1. Montrer que $\sum \dfrac{x^n}{n!}$ converge pour tout $x \in \mathbb{R}$.
2. Étudier la nature de la série $\sum \dfrac{n!}{n^n}$.
3. Soit $x \in \mathbb{R}_+$. Discuter de la nature de la série $\sum \dfrac{x^n}{\binom{2n}{n}}$ en fonction de x.

Correction.

1. Pour $n \in \mathbb{N}$, on pose $u_n = \left|\frac{x^n}{n!}\right|$. Alors on a, pour tout $n \in \mathbb{N}$:
$$\frac{u_{n+1}}{u_n} = \frac{|x|^{n+1}}{|x|^n}.\frac{n!}{(n+1)!} = \frac{|x|}{n+1} \xrightarrow[n\to+\infty]{} 0 < 1.$$
Ainsi, d'après la règle de D'Alembert, $\sum \dfrac{x^n}{n!}$ est absolument convergente et donc convergente.
2. On a, pour tout $n \in \mathbb{N}$:
$$\frac{u_{n+1}}{u_n} = \frac{n^n}{(n+1)^{n+1}}.\frac{(n+1)!}{n!} = \left(\frac{n}{n+1}\right)^n = \frac{1}{\left(1+\frac{1}{n}\right)^n} \xrightarrow[n\to+\infty]{} \frac{1}{e} < 1.$$
Ainsi, d'après la règle de D'Alembert, $\sum \dfrac{n!}{n^n}$ converge.
3. Soit $x \in \mathbb{R}_+$. Pour $n \in \mathbb{N}$, on pose $u_n = \frac{x^n}{\binom{2n}{n}}$. Alors on a, pour tout $n \in \mathbb{N}$:
$$\frac{u_{n+1}}{u_n} = \frac{x(n+1)^2}{(2n+2)(2n+1)} \xrightarrow[n\to+\infty]{} \frac{1}{4}x \begin{cases} < 1 & \text{si } x < 4 \\ = 1 & \text{si } x = 4 \\ > 1 & \text{si } x > 4 \end{cases}$$
D'après la règle de D'Alembert, $\sum u_n$ converge si $x < 4$ et diverge si $x > 4$. Si $x = 4$, on a
$$\frac{u_{n+1}}{u_n} = \frac{4n+4}{4n+2} \geq 1.$$
Par suite, $(u_n)_{n\in\mathbb{N}}$ est croissante (et à valeurs positives) donc elle ne converge pas vers 0. Ainsi, $\frac{4^n}{\binom{2n}{n}}$ diverge grossièrement.

Exercice 7. *Exponentielle de matrice*

Soit $p \in \mathbb{N}^*$. Soit $A \in M_p(\mathbb{K})$ muni d'une norme sous-multiplicative.

1. Montrer que la série $\sum \dfrac{A^n}{n!}$ est convergente. On note alors
$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$
2. Calculer $\exp(A)$ pour $A = \begin{pmatrix} -7 & -4 & -4 \\ 4 & 1 & 4 \\ 4 & 4 & 1 \end{pmatrix}$.

Correction.

1. Comme $\|\cdot\|$ est sous-multiplicative, on a, pour $n \in \mathbb{N}$, $\left\| \dfrac{A^n}{n!} \right\| \leq \dfrac{\|A\|^n}{n!}$ qui est le terme général d'une série géométrique de raison $\|A\| < 1$ et donc convergente. Ainsi, par majoration du terme général d'une série à termes positifs par le terme général d'une série convergente, $\sum \dfrac{A^n}{n!}$ est absolument convergente. Par suite, $\sum \dfrac{A^n}{n!}$ est convergente.
2. On remarque que si A est diagonalisable avec $A = PDP^{-1}$ où $P \in \mathrm{GL}_p(\mathbb{K})$ et $D = \mathrm{diag}(\alpha_1, ..., \alpha_p)$, on a :
$$\exp(A) = P \exp(D) P^{-1}$$
et
$$\exp(D) = \mathrm{diag}(e^{\alpha_1}, ..., e^{\alpha_p}).$$
Ainsi, dans notre cas, on a :
$$A = PDP^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$
d'où
$$\begin{aligned} \exp(A) = P \exp(D) P^{-1} &= \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} e & 0 & 0 \\ 0 & e^{-3} & 0 \\ 0 & 0 & e^{-3} \end{pmatrix} \begin{pmatrix} -1 & -1 & -1 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -e + 2e^{-3} & -e + e^{-3} & -e + e^{-3} \\ e - e^{-3} & e & e - e^{-3} \\ e - e^{-3} & e - e^{-3} & e \end{pmatrix} \end{aligned}$$

2. Comparaison avec une intégrale

Théorème 3. *Comparaison série intégrale*

Soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction **continue par morceaux et décroissante**. Alors la série :

$$\sum_{n\geq 1}\left(\int_{n-1}^{n} f(t)dt - f(n)\right)$$

converge.

Démonstration.

Soit $n \in \mathbb{N}^*$. Alors on a :

$$0 = f(n) - f(n) \leq \int_{n-1}^{n} f(t)dt - f(n) \leq f(n-1) - f(n)$$

Donc la série est à termes positifs et on a, pour tout $N \geq 1$:

$$S_N = \sum_{n=1}^{N}\left(\int_{n-1}^{n} f(t)dt - f(n)\right) \leq \sum_{n=1}^{N}(f(n-1) - f(n)) = f(0) - f(N).$$

Or comme f est décroissante et minorée (par 0), f admet une limite $\ell \in \mathbb{R}_+$ en $+\infty$; ainsi, pour tout $N \geq 1$, $S_N \leq f(0) - \ell$. Par suite, par majoration uniforme des sommes partielles d'une série à termes positifs, la série converge. □

Corollaire 2.

Soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction **continue par morceaux et décroissante**. Alors $\sum f(n)$ converge si, et seulement si, $\int_0^{+\infty} f(t)dt$ converge.

Démonstration.

On sait que $\sum_{n\geq 1}\left(\int_{n-1}^{n} f(t)dt - f(n)\right)$ converge, donc $\sum f(n)$ converge si, et seulement si, $\sum_{n\geq 1}\int_{n-1}^{n} f(t)dt$ converge. Or on a $\sum_{n\geq 1}\int_{n-1}^{n} f(t)dt$ converge si, et seulement si

$$\int_0^{\infty} f(t)dt = \lim_{N\to+\infty}\int_0^{N} f(t)dt = \lim_{N\to+\infty}\sum_{n=1}^{N}\int_{n-1}^{n} f(t)dt$$

existe dans $\mathbb{R}_+$. D'où le résultat. □

Proposition 7. *Série-Intégrale : estimation des sommes partielles et des restes*

Soit $f : \mathbb{R}_+ \to \mathbb{R}_+$ une fonction **continue par morceaux et décroissante** et $n \in \mathbb{N}$. Alors les sommes partielles de la série $\sum f(n)$ vérifient pour tout $m \in \mathbb{N}$:

$$\int_{n+1}^{n+m+1} f(t)dt \leq S_{n+m} - S_n = f(n+1) + \ldots + f(n+m) \leq \int_{n}^{n+m} f(t)dt.$$

Et en cas de convergence de la série $\sum f(n)$, les restes de la série vérifient :

$$\int_{n+1}^{+\infty} f(t)dt \leq R_n \leq \int_{n}^{+\infty} f(t)dt.$$

Démonstration.

On a, pour tout $i \in \mathbb{N}^*$,

$$\int_{i}^{i+1} f(t)dt \leq f(i) \leq \int_{i-1}^{i} f(t)dt.$$

Donc, pour $m \in \mathbb{N}$, on a :

$$\sum_{i=n+1}^{n+m} \int_{i}^{i+1} f(t)dt \leq \sum_{i=n+1}^{n+m} f(i) \leq \sum_{i=n+1}^{n+m} \int_{i-1}^{i} f(t)dt,$$

et ainsi, on obtient le résultat concernant les sommes partielles. Si $\sum f(n)$ converge, alors $\int_0^{+\infty} f(t)dt$ converge et donc, les mêmes intégrales mais à partir de n et de $n+1$ convergent. De plus, d'après les inégalités précédentes, on a :

$$\underbrace{\lim_{m\to+\infty} \int_{n+1}^{n+m+1} f(t)dt}_{=\int_{n+1}^{+\infty} f(t)dt} \leq \underbrace{\lim_{m\to+\infty} S_{n+m} - S_n}_{=R_n} \leq \underbrace{\lim_{m\to+\infty} \int_{n}^{n+m} f(t)dt}_{=\int_{n}^{+\infty} f(t)dt}.$$

$\square$

Remarque 5.

En particulier, on a l'inégalité $\int_1^{N+1} f(t)dt \leq \sum_{n=1}^{N} f(n) \leq \int_0^N f(t)dt$.

Exercice 8.

Donner, selon la valeur de $\alpha \in \mathbb{R}_+^*$, la nature des séries suivantes. En cas de divergence, donner un équivalent simple des sommes partielles et en cas de convergence, donner un équivalent simple des restes.

1. $\displaystyle\sum_{n\geq 1} \frac{1}{n^\alpha}$.

2. $\displaystyle\sum_{n\geq 2} \frac{1}{n(\ln(n))^\alpha}$.

Correction.

1. Pour $\alpha > 0$, la fonction $f : t \mapsto \frac{1}{t^\alpha}$ est continue et décroissante sur de $\mathbb{R}_+^*$ dans lui-même. Ainsi, d'après le théorème de comparaison série-intégrale (voire Corollaire 2), $\sum_{n\geq 1} \frac{1}{n^\alpha}$ converge si, et seulement si, $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge. Or on a :

$$\int_1^x \frac{1}{t^\alpha} dt = \begin{cases} \frac{1}{1-\alpha}\left(x^{1-\alpha} - 1\right) & \text{si } \alpha \neq 1 \\ \ln(x) & \text{si } \alpha = 1 \end{cases}.$$

d'où les trois cas suivants :

1er cas : $0 < \alpha < 1$. Dans ce cas, $1 - \alpha > 0$, donc

$$\int_1^x \frac{1}{t^\alpha} dt = \frac{1}{1-\alpha}\left(x^{1-\alpha} - 1\right) \xrightarrow[x\to+\infty]{} +\infty.$$

Par suite, $\sum_{n\geq 1} \frac{1}{n^\alpha}$ diverge.

2eme cas : $\alpha = 1$. On a

$$\int_1^x \frac{1}{t} dt = \ln(x) \xrightarrow[x\to+\infty]{} +\infty,$$

Par suite, $\sum_{n\geq 1} \frac{1}{n}$ diverge.

3eme cas : $\alpha > 1$. Dans ce cas, $\alpha - 1 > 0$, donc

$$\int_1^x \frac{1}{t^\alpha} dt = \frac{1}{\alpha - 1}\left(1 - \frac{1}{x^{\alpha-1}}\right) \xrightarrow[x\to+\infty]{} \frac{1}{\alpha-1}.$$

Par suite, $\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge et ainsi $\sum_{n\geq 1} \frac{1}{n^\alpha}$ converge.

On a donc bien le critère de Riemann : $\sum_{n\geq 1} \frac{1}{n^\alpha}$ converge si, et seulement si, $\alpha > 1$. (Pour $\alpha \leq 0$, on remarque que la série diverge grossièrement).

Intéressons nous désormais aux équivalents des restes et des sommes partielles :
On a, pour tout $i \geq 2$,

$$\int_{i-1}^{i} \frac{1}{t^\alpha} dt \leq \frac{1}{i^\alpha} \leq \int_i^{i+1} \frac{1}{t^\alpha} dt,$$

1er cas : $0 < \alpha < 1$. La série diverge, donc on s'intéresse à un équivalent des sommes partielles. On a

$$\int_1^{n+1} \frac{1}{t^\alpha} dt \leq 1 + \int_2^{n+1} \frac{1}{t^\alpha} dt \leq S_n(\alpha) = \sum_{i=1}^n \frac{1}{i^\alpha} \leq 1 + \int_1^n \frac{1}{t^\alpha} dt \qquad (*)$$

Ainsi, on obtient :

$$\frac{1}{1-\alpha}\left((n+1)^{1-\alpha} - 1\right) \leq S_n(\alpha) \leq 1 + \frac{1}{1-\alpha}\left(n^{1-\alpha} - 1\right).$$

On remarque alors que :

$$\frac{1}{1-\alpha}\left((n+1)^{1-\alpha}-1\right) \underset{n\to+\infty}{\sim} \frac{n^{1-\alpha}}{1-\alpha} \underset{n\to+\infty}{\sim} 1+\frac{1}{1-\alpha}\left(n^{1-\alpha}-1\right);$$

par suite, d'après le théorème des gendarmes, $S_n(\alpha)\dfrac{1-\alpha}{n^{1-\alpha}} \xrightarrow[n\to+\infty]{} 1$. Il en résulte que :

$$S_n(\alpha) \underset{n\to+\infty}{\sim} \frac{n^{1-\alpha}}{1-\alpha}.$$

2eme cas : $\alpha = 1$. La série diverge, donc on s'intéresse de nouveau à un équivalent des sommes partielles. D'après $(*)$, on obtient :

$$\ln(n+1) \leq S_n(1) \leq 1+\ln(n).$$

Par suite,

$$S_n(1) \underset{n\to+\infty}{\sim} \ln(n).$$

3eme cas : $\alpha > 1$. La série converge, donc on s'intéresse à un équivalent des restes. On a, pour $n \geq 1$:

$$\int_{n+1}^{+\infty} \frac{1}{t^\alpha}dt \leq R_n(\alpha) \leq \int_{n}^{+\infty} \frac{1}{t^\alpha}dt.$$

Ainsi,

$$\frac{1}{\alpha-1}\frac{1}{(n+1)^{\alpha-1}} \leq R_n(\alpha) \leq \frac{1}{\alpha-1}\frac{1}{n^{\alpha-1}}.$$

D'où, finalement,

$$R_n(\alpha) \underset{n\to+\infty}{\sim} \frac{1}{\alpha-1}\frac{1}{n^{\alpha-1}}.$$

2. On réitère les mêmes techniques avec la fonction $f : x \mapsto \dfrac{1}{x(\ln(x))^\alpha}$.

3. Rappels : Critère des séries alternées

Définition 7. *Série alternée*

Soit $(v_n)_{n\in\mathbb{N}}$ une suite de réels. On dit que la série $\sum v_n$ est **alternée** s'il existe une suite $(u_n)_{n\in\mathbb{N}}$ de réels tous de même signe telle que

$$\sum v_n = \sum (-1)^n u_n.$$

Exemple 1.

$\sum \dfrac{(-1)^n}{n}$ ou $\sum \sin\left(\pi(\sqrt{n^2+1})\right)$ sont des séries alternées.

Théorème 4. *Critère des séries alternées*

Soit $\sum v_n$ une série alternée. Si :

i) la suite $(|v_n|)_{n\in\mathbb{N}}$ est décroissante et

ii) $(v_n)_{n\in\mathbb{N}}$ converge vers 0,

alors $\sum v_n$ converge et on a les propriétés suivantes :

— si $v_0 \geq 0$, pour tout $k \in \mathbb{N}$,

$$S_{2k+1} \leq \sum_{n=0}^{+\infty} v_n \leq S_{2k}.$$

— pour tout $n \in \mathbb{N}$, R_n est du signe de v_{n+1} et

$$|R_n| \leq |v_{n+1}|.$$

Démonstration.

Quitte à échanger la suite (v_n) par la suite $(-v_n)$, on suppose $v_0 \geq 0$. Cela implique que pour tout $k \in \mathbb{N}$, $v_{2k} \geq 0$ et $v_{2k+1} \leq 0$ car la série est alternée.

On considère les suites $(a_k)_{k\in\mathbb{N}} = (S_{2k})_{k\in\mathbb{N}}$ et $(b_k)_{k\in\mathbb{N}} = (S_{2k+1})_{k\in\mathbb{N}}$. Montrons qu'elles sont adjacentes i.e. (a_k) est décroissante, (b_k) est croissante et $(a_k - b_k)$ converge vers 0.

— Soit $k \in \mathbb{N}$. On a

$$a_{k+1} - a_k = S_{2k+2} - S_{2k} = v_{2k+2} + v_{2k+1} = |v_{2k+2}| - |v_{2k+1}| \leq 0$$

car $(|v_n|)$ est décroissante d'après ii) ; et, grâce à un argument similaire, on a :

$$b_{k+1} - b_k = S_{2k+3} - S_{2k+1} = v_{2k+3} + v_{2k+2} = |v_{2k+2}| - |v_{2k+3}| \geq 0.$$

Par suite, (a_k) est décroissante et (b_k) est croissante.

— Soit $k \in \mathbb{N}$. On a

$$a_k - b_k = S_{2k} - S_{2k+1} = -v_{2k+1} = |v_{2k+1}| \xrightarrow[k\to+\infty]{} 0$$

Il en résulte que $(a_k)_{k\in\mathbb{N}}$ et $(b_k)_{k\in\mathbb{N}}$ sont adjacentes et converge ainsi vers une même limite notée S. Or, pour tout $n \in \mathbb{N}$, on a, pour $k = E(\frac{n}{2})$:

$$b_k = S_{2k+1} \leq S_n \leq S_{2_k} = a_k,$$

avec égalité à gauche lorsque n est impair, et égalité à droite lorsque n est pair. Ainsi, en passant à la limite quand n tend vers $+\infty$ dans l'inégalité précédente, et en remarquant que $k = E(\frac{n}{2}) \xrightarrow[n\to+\infty]{} +\infty$, on obtient que (S_n) converge vers S, d'où l'existence de la somme de la série $\sum v_n$ et on a donc, pour tout $k \in \mathbb{N}$:

$$S_{2k+1} \leq S = \sum_{n=0}^{+\infty} v_n \leq S_{2_k}.$$

□

Remarque 6.

Sous les hypothèses du critère des séries alternées, en particulier, $S = \sum_{n=0}^{+\infty} v_n$ est du même signe que v_0 et lorsque $v_0 \geq 0$, $S \leq v_0$.

Exercice 9.

1. Étudier la nature $\sum \dfrac{(-1)^n}{n^\alpha}$ selon les valeurs de $\alpha \in \mathbb{R}$.
2. Montrer que le reste R_n de la série $\sum \dfrac{(-1)^n}{n^2}$ est le terme général d'une série convergente.
3. On considère la suite $\sum u_n$ de terme général :
$$u_n = \begin{cases} \frac{-1}{k^2} \text{ si } n = 2k \text{ est pair} \\ \frac{1}{k} \text{ si } n = 2k+1 \text{ est impair} \end{cases}$$
Quelle est la nature de cette série ? Qu'en conclure sur les hypothèses du critère des séries alternées ?

Correction.

1. *1er cas* $\alpha \leq 0$: on a $\frac{(-1)^n}{n^\alpha} = (-1)^n n^{-\alpha} \nrightarrow 0$ car $-\alpha \geq 0$ donc $\sum \dfrac{(-1)^n}{n^\alpha}$ diverge grossièrement.
2eme cas $\alpha > 0$. Alors $(\frac{1}{n^\alpha})_{n\in\mathbb{N}^*}$ est décroissante et tend vers 0, donc d'après le critère des séries alternées, $\sum \dfrac{(-1)^n}{n^\alpha}$ converge.

2. D'après la question précédente, $\sum \dfrac{(-1)^n}{n^2}$ converge ; on peut donc s'intéresser aux restes de cette série.
Soit $n \in \mathbb{N}^*$. On a :
$$R_n = \sum_{k=n+1}^{+\infty} \frac{(-1)^k}{k^2}.$$
Toujours d'après le critère des séries alternées, on a :
$$|R_n| \leq \left|\frac{(-1)^{n+1}}{(n+1)^2}\right| = \frac{1}{(n+1)^2}$$
Ainsi, $|R_n|$ est bien le terme général d'une série convergente d'après le critère de Riemann car $\dfrac{1}{(n+1)^2} = o(\frac{1}{n^2})$.
Par suite, $\sum R_n$ est absolument convergente, et donc convergente.

3. Tout d'abord, on remarque que (u_n) est bien le terme général d'une série alternée qui vérifie $u_n \to 0$. Par contre, $(|u_n|)$ n'est pas décroissante...
On a, pour $2 \leq n = 2k$ un entier pair :
$$u_n + u_{n+1} = u_{2k} + u_{2k+1} = \frac{-1}{k^2} + \frac{1}{k} = \frac{k-1}{k^2}.$$

Ainsi, pour $N \in \mathbb{N}$, on a :

$$S_{2N+1} = \sum_{n=1}^{2N+1} u_n = 1 + \sum_{k=1}^{N}(u_{2k} + u_{2k+1}) = 1 + \sum_{k=1}^{N} \frac{k-1}{k^2}$$

Or, on a $\frac{k-1}{k^2} = o(\frac{1}{k})$ donc d'après le critère de Riemann, la série $\sum \frac{k-1}{k^2}$ diverge. Par suite, la suite $(S_{2N+1})_{N\in\mathbb{N}}$ diverge également !
Il en résulte que $\sum u_n$ diverge.

On en conclut que l'hypothèse de décroissance de la suite $(|u_n|)$ est nécessaire ! Si on l'omet, le théorème devient faux comme en atteste cet exemple.

4. Sommation des relations de comparaison

a. En cas de convergence

Proposition 8.

Soit $\sum u_n$ et $\sum v_n$ des séries numériques avec $\sum v_n$ à termes positifs. On suppose que $\sum u_n$ et $\sum v_n$ convergent. Alors les restes $R_n(u)$ et $R_n(v)$ respectifs des séries associées aux suites $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ vérifient quand $n \to +\infty$:

— Si $u_n = O(v_n)$, alors $R_n(u) = O(R_n(v))$.

— Si $u_n = o(v_n)$, alors $R_n(u) = o(R_n(v))$.

— Si $u_n \sim v_n$, alors $R_n(u) \sim R_n(v)$.

Démonstration.

On remarque que dans chacun des cas précédent, la suite $(|u_n|)_{n\in\mathbb{N}}$ verifie la même relation de comparaison que la suite $(u_n)_{n\in\mathbb{N}}$ vis-à-vis de la suite $(v_n)_{n\in\mathbb{N}}$. Par suite, comme $\sum v_n$ converge, par comparaison, $\sum |u_n|$ converge et donc ses restes $R_n(|u|)$ sont bien définis et on a, pour tout $n \in \mathbb{N}$:

$$|R_n(u)| = |\sum_{k=n+1}^{+\infty} u_n| \leq \sum_{k=n+1}^{+\infty} |u_n| = R_n(|u|).$$

— On suppose $u_n \underset{n\to+\infty}{=} O(v_n)$. Alors il existe $M \in \mathbb{R}_+$ tel que, pour tout $n \in \mathbb{N}$, $|u_n| \leq M v_n$. Par suite, on a, pour tout $n \in \mathbb{N}$,

$$|R_n(u)| \leq R_n(|u|) = \sum_{k=n+1}^{+\infty} \underbrace{|u_n|}_{\leq M v_n} \leq M \sum_{k=n+1}^{+\infty} v_n = M R_n(v).$$

D'où $R_n(u) \underset{n\to+\infty}{=} O(R_n(v))$.

— On procède de manière similaire : on suppose $u_n \underset{n\to+\infty}{=} o(v_n)$. Soit $\varepsilon > 0$ <. Alors, il existe $N \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, $n \geq N$ implique $|u_n| \leq \varepsilon v_n$.

Par suite, pour tous $n \geq N$ et $k \geq n+1$, on a $k \geq N$ donc :

$$|R_n(u)| \leq R_n(|u|) = \sum_{k=n+1}^{+\infty} \underbrace{|u_n|}_{\leq \varepsilon v_n} \leq \varepsilon \sum_{k=n+1}^{+\infty} v_n = \varepsilon R_n(v).$$

D'où $R_n(u) = \underset{n\to+\infty}{o}(R_n(v))$.

— On suppose $u_n \underset{n\to+\infty}{\sim} v_n$. Alors par définition, $u_n - v_n = \underset{n\to+\infty}{o}(v_n)$. On pose, pour tout $n \in \mathbb{N}$, $w_n = u_n - v_n$. Alors, par comparaison (ou par combinaison linéaire de série convergente), $\sum w_n$ est convergente et ses restes $R_n(w)$ vérifient, d'après le point précédent, $R_n(w) = \underset{n\to+\infty}{o}(R_n(v))$.
Or, pour tout $n \in \mathbb{N}$, $R_n(w) = R_n(u) - R_n(v)$, donc :

$$R_n(u) - R_n(v) = \underset{n\to+\infty}{o}(R_n(v)).$$

D'où $R_n(u) = \underset{n\to+\infty}{\sim} R_n(v)$.

□

b. En cas de divergence

Proposition 9.

Soit $\sum u_n$ et $\sum v_n$ des séries numériques avec $\sum v_n$ à termes positifs. On suppose que $\sum v_n$ diverge. Alors les sommes partielles $S_n(u)$ et $S_n(v)$ respectifs des séries associées aux suites $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ vérifient quand $n \to +\infty$:

— Si $u_n = O(v_n)$, alors $S_n(u) = O(S_n(v))$.

— Si $u_n = o(v_n)$, alors $S_n(u) = o(S_n(v))$.

— Si $u_n \sim v_n$, alors $S_n(u) \sim S_n(v)$.

Démonstration.

Les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ étant à termes positifs, donc les séries associées sont des suites croissantes. De plus, comme $\sum v_n$ diverge, alors $(S_n(v))_{n\in\mathbb{N}}$ tend vers $+\infty$, i.e. pour tout $A \geq 0$, il existe $N_1 \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, $n \geq N_1$ implique $S_n(v) \geq A$.

— On suppose $u_n = \underset{n\to+\infty}{O}(v_n)$. Alors il existe $M \in \mathbb{R}_+$ tel que, pour tout $n \in \mathbb{N}$, $u_n \leq M v_n$.
Par suite, on a, pour tout $n \in \mathbb{N}$,

$$S_n(u) = \sum_{k=0}^{n} \underbrace{u_n}_{\leq M v_n} \leq M \sum_{k=0}^{n} v_n = M S_n(v).$$

D'où $S(u) = \underset{n\to+\infty}{O}(S_n(v))$.

— on suppose $u_n = \underset{n\to+\infty}{o}(v_n)$. Soit $\varepsilon > 0$. Alors, il existe $N \in \mathbb{N}$ tel que, pour tout $n \in \mathbb{N}$, $n \geq N$ implique $u_n \leq \frac{\varepsilon}{2} v_n$.
De plus, pour $A = \frac{2S_N(u)}{\varepsilon} \geq 0$, d'après la remarque initiale, il existe $N_1 \in \mathbb{N}$ tel que pour

tout entier $n \geq N_1$, $S_n(v) \geq A$ i.e. $S_N(u) \leq \frac{\varepsilon}{2} S_n(v)$.
Soit un entier $n \geq \max(N, N_1)$. Alors :

$$S_n(u) = \sum_{k=0}^{N} u_n + \sum_{k=N+1}^{n} \underbrace{u_n}_{\leq \frac{\varepsilon}{2} v_n} \leq \underbrace{S_N(u)}_{\leq \frac{\varepsilon}{2} S_n(v)} + \frac{\varepsilon}{2} \underbrace{\sum_{k=N+1}^{n} v_n}_{\leq S_n(v)} = \varepsilon S_n(v).$$

D'où $S_n(u) \underset{n\to+\infty}{=} o(S_n(v))$.

— On suppose $u_n \underset{n\to+\infty}{\sim} v_n$. Alors par définition, $u_n - v_n \underset{n\to+\infty}{=} o(v_n)$. On pose, pour tout $n \in \mathbb{N}$, $w_n = u_n - v_n$, et $S_n(w)$ la somme partielle d'ordre n de $\sum w_n$. Alors on a, pour tout $n \in \mathbb{N}$, $S_n(w) = S_n(u) - S_n(v)$ et d'après le cas précédent, comme $w_n \underset{n\to+\infty}{=} o(v_n)$,

$$S_n(u) - S_n(v) = S_n(w) \underset{n\to+\infty}{=} o(S_n(v)).$$

D'où $S_n(u) \underset{n\to+\infty}{=\sim} S_n(v)$.

□

Corollaire 3. *Lemme de Cesàro*

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs complexes. Si $(u_n)_{n\in\mathbb{N}}$ converge vers $\ell \in \mathbb{C}$, alors :

$$\frac{1}{n+1} \sum_{k=0}^{n} u_k \xrightarrow[n\to+\infty]{} \ell.$$

Correction.

On suppose que $(u_n)_{n\in\mathbb{N}}$ converge vers $\ell \in \mathbb{C}$ et pour $n \in \mathbb{N}$, on pose $v_n = u_n - \ell$. Alors on a $v_n = u_n - \ell \underset{n\to+\infty}{=} o(1)$. Comme la série $\sum 1$ diverge, d'après la proposition précédente, la somme partielle S_n d'indice n de la série $\sum v_n$ vérifie :

$$S_n \underset{n\to+\infty}{=} o\left(\sum_{k=0}^{n} 1\right) \underset{n\to+\infty}{=} o(n)$$

Or on a $S_n = \sum_{k=0}^{n} v_n = \sum_{k=0}^{n} u_n - \sum_{k=0}^{n} \ell = \sum_{k=0}^{n} u_n - (n+1)\ell$. Par suite,

$$\sum_{k=0}^{n} u_n = (n+1)\ell + \underset{n\to+\infty}{o}(n),$$

d'où le résultat en quotientant cette dernière égalité par $n+1$.

c. Applications

Exercice 10.

Retrouver un équivalent simple des restes de $\sum \frac{1}{n^2}$ grâce à une série télescopique.

Correction.

On a $\dfrac{1}{n^2} \underset{n\to+\infty}{\sim} \dfrac{1}{n(n+1)}$. De plus $\sum \dfrac{1}{n^2}$ converge donc :

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^2} \underset{n\to+\infty}{\sim} \sum_{k=n+1}^{+\infty} \frac{1}{k(k+1)}$$

Or, pour tout $k \in \mathbb{N}^*$, $\dfrac{1}{k(k+1)} = \dfrac{1}{k} - \dfrac{1}{k+1}$ donc, par "télescopage",

$$\sum_{k=n+1}^{+\infty} \frac{1}{k(k+1)} = \frac{1}{n+1}.$$

Il en résulte que :

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^2} \underset{n\to+\infty}{\sim} \frac{1}{n+1} \underset{n\to+\infty}{\sim} \frac{1}{n}.$$

Exercice 11. *Développement asymptotique d'une suite récurrente*

On considère la suite récurrente $(u_n)_{n\in\mathbb{N}}$ définie par :

$$\begin{cases} u_0 \in]0,1] \\ u_{n+1} = \sin(u_n) \quad \forall\, n \in \mathbb{N} \end{cases}$$

1. Montrer que, pour tout $x \in]0,1]$, $0 < \sin(x) < x$. En déduire que (u_n) admet une limite $\lambda \in \mathbb{R}$ et la déterminer.
2. Chercher $\alpha > 0$ tel que la suite $\left(u_{n+1}^{-\alpha} - u_n^{-\alpha}\right)$ converge vers une limite non nulle et déterminer cette limite.
3. En déduire un équivalent en $+\infty$ de $\frac{1}{u_n^2}$. Déterminer ainsi un équivalent simple en $+\infty$ de u_n.

Correction.

1. Pour $x \in]0,1]$, $\sin(x) > 0$ car sin est strictement positive sur $]0,\pi[$. De plus, en étudiant la fonction $f : t \mapsto t - sin(t)$ sur $\mathbb{R}$, on obtient la stricte positivité de f sur $]0,1]$.

   Il en résulte que (u_n) est à valeurs dans $]0,1]$ et est décroissante. Ainsi, (u_n) est décroissante minorée et donc converge vers une limite $\ell \in [0,1]$. Par passage à la limite dans $u_{n+1} = \sin(u_n)$ (sin étant continue), on obtient $\ell = \sin(\ell)$ et donc $\ell = 0$ car c'est l'unique solution de $x = \sin(x)$ sur $[0,1]$.
   En conclusion, (u_n) converge vers 0.

2. Soit $\alpha > 0$. Comme $u_n \to 0$, on a $\sin(u_n) = u_n - \frac{u_n^3}{6} + o(u_n^3)$ et donc :

$$\begin{array}{rcl} u_{n+1}^{-\alpha} - u_n^{-\alpha} & = & (\sin(u_n))^{-\alpha} - u_n^{-\alpha} \\ & = & (u_n - \frac{u_n^3}{6} + o(u_n^3))^{-\alpha} - u_n^{-\alpha} \\ & = & u_n^{-\alpha}(1 - \frac{u_n^2}{6} + o(u_n^2))^{-\alpha} - u_n^{-\alpha} \\ & = & u_n^{-\alpha}\left((1 - \dfrac{u_n^2}{6} + o(u_n^2))^{-\alpha} - 1\right) \end{array}$$

Or, on a $(1-x)^{-\alpha} = 1 + \alpha x + o(x)$ quand $x \to 0$, donc, en prenant $x = \frac{u_n^2}{6} + o(u_n^2)$, on obtient :

$$u_{n+1}^{-\alpha} - u_n^{-\alpha} = u_n^{-\alpha}\left(1 + \alpha\frac{u_n^2}{6} + o(u_n^2) - 1\right) = \alpha\frac{u_n^{2-\alpha}}{6} + o(u_n^{2-\alpha}).$$

Ainsi, pour $\alpha = 2$, $(u_{n+1}^{-\alpha} - u_n^{-\alpha})$ converge et :

$$\lim_{n\to+\infty} u_{n+1}^{-2} - u_n^{-2} = \frac{2}{6} = \frac{1}{3}.$$

3. D'après la question précédente, $\sum(u_{n+1}^{-2} - u_n^{-2})$ diverge grossièrement car $u_{n+1}^{-2} - u_n^{-2} \sim \frac{1}{3}$. Par suite, la somme partielle S_n de cette série vérifie :

$$S_n \underset{n\to+\infty}{\sim} \sum_{k=0}^{n} \frac{1}{3} = \frac{n+1}{3} \underset{n\to+\infty}{\sim} \frac{n}{3}.$$

Or, par télescopage, $S_n = u_n^2 - u_0^{-2}$, donc

$$\frac{3}{n}\frac{1}{u_n^2} = \frac{3}{n}S_n + \frac{3}{n}u_0^{-2} \xrightarrow[n\to+\infty]{} 1 + 0 = 1.$$

Donc $\dfrac{1}{u_n^2} \underset{n\to+\infty}{\sim} \dfrac{n}{3}$.

Il en résulte que :

$$u_n \underset{n\to+\infty}{\sim} \sqrt{\frac{3}{n}}.$$

> *Pour aller plus loin :* on peut également montrer que :
>
> $$u_n = \sqrt{\frac{3}{n}} + \frac{3\sqrt{3}\ln(n)}{10n^{3/2}} + o\left(\frac{\ln(n)}{n^{3/2}}\right)$$
>
> en utilisant les mêmes techniques que précédemment et un développement limité à l'ordre 2 en 0 de $x \mapsto \frac{1}{sin^2(x)} - \frac{1}{x^2}$.

Exercice 12. *Développement asymptotique à 3 termes de la série harmonique*

On considère la série harmonique $\sum_{n\geq 1} \frac{1}{n}$ et pour $n \in \mathbb{N}^*$, on note H_n sa somme partielle d'indice n et $d_n = H_n - \ln(n)$.

1. Montrer que pour tout entier $n \geq 2$,
$$d_n - d_{n-1} \underset{n\to+\infty}{\sim} -\frac{1}{2n^2}.$$
puis en déduire que $(d_n)_{n\in\mathbb{N}^*}$ converge.
Dans la suite, on notera γ la limite de $(d_n)_{n\in\mathbb{N}^*}$ qui par ailleurs répond au doux nom de *constant d'Euler*.

2. En comparant les restes des séries (convergentes !) de termes généraux $d_{n-1} - d_n$ et $\frac{1}{2n^2}$, montrer que :
$$H_n = \ln(n) + \gamma + \frac{1}{2n} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right).$$

Correction.

1. Soit $n \geq 2$ un entier. Comme $\ln(1 - \frac{1}{n}) = -\frac{1}{n} - \frac{1}{2n^2} + o(\frac{1}{n^2})$, on a :
$$d_n - d_{n-1} = (H_n - H_{n-1}) - (\ln(n) - \ln(n-1)) = \frac{1}{n} + \ln(1 - \frac{1}{n}) = -\frac{1}{2n^2} + o(\frac{1}{n^2}) \underset{n\to+\infty}{\sim} -\frac{1}{2n^2}.$$
Ainsi, la série télescopique $\sum_{n\geq 2}(d_n - d_{n-1})$ converge absolument par comparaison à une série de Riemann convergente et donc, elle converge. D'après la proposition 3, il en résulte que la suite $(d_n)_{n\in\mathbb{N}^*}$ converge.

2. Comme $d_{n-1} - d_n \underset{n\to+\infty}{\sim} \frac{1}{2n^2}$, par comparaison des restes de séries de termes généraux équivalents, on a, en notant $R_n(d)$ et $R_n(r)$ les restes respectifs de $\sum_{n\geq 2}(d_{n-1} - d_n)$ et $\sum_{n\geq 2} \frac{1}{2n^2}$:
$$R_n(d) \underset{n\to+\infty}{\sim} R_n(r) = \frac{1}{2} \sum_{k=n+1}^{+\infty} \frac{1}{n^2} \underset{n\to+\infty}{\sim} \frac{1}{2n},$$
ce dernier équivalent étant obtenu par comparaison série/intégrale ou par l'exercice 10.
Or, par télescopage, $R_n(d) = d_n - \gamma$, donc $d_n = \gamma + \frac{1}{2n} + o(\frac{1}{n})$ et ainsi,
$$H_n = d_n + \ln(n) = \ln(n) + \gamma + \frac{1}{2n} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right).$$

Chapitre VI

Suites et Séries de fonctions

Table des matières

Dans ce chapitre, E et F désigne un espace vectoriel normé de dimension finie sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Toutes les fonctions que l'on considère dans ce chapitre sont, sauf indication contraire, des fonctions définies sur une partie A de E et à valeurs dans F.

Partie A

Convergences des suites et des séries de fonctions

1. Convergence simple

a. Suites de fonctions

Définition 1. *Convergence simple d'une suite de fonctions*

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de $A \subset E$ dans F.

On dit que la suite $(f_n)_{n\in\mathbb{N}}$ **converge simplement sur** A si, pour tout $x \in A$, la suite $(f_n(x))_{n\in\mathbb{N}}$ à valeurs dans F est convergente.

Dans ce cas, on peut considérer la fonction $f : x \mapsto \lim_{n\to+\infty} f_n(x)$ de A dans F et on dit que la suite $(f_n)_{n\in\mathbb{N}}$ **converge simplement vers** f.

Exemple 1.

La suite $(f_n)_{n\in\mathbb{N}}$ de $\mathbb{R}_+$ dans $\mathbb{R}_+$ définie, pour $n \in \mathbb{N}$, par $f_n : t \mapsto t^n$ converge simplement sur $[0,1]$ vers la fonction

$$f : x \mapsto \begin{cases} 0 & \text{si } x \in [0,1[\\ 1 & \text{si } x = 1. \end{cases}$$

Remarque 1.

L'exemple précédent nous montre que si une suite de fonctions continues converge simplement vers f, f n'est pas continue en général !

Exercice 1.

1. Écrire la définition de la convergence simple vers une fonction en termes "epsilonesques".
2. Pour $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ des suites de fonctions qui converge simplement sur A vers des fonctions f et g respectivement, étudier la convergence simple des suites de fonctions $(\lambda f_n + \mu g_n)_{n\in\mathbb{N}}$ (où $\lambda, \mu \in \mathbb{K}$) et $(f_n g_n)_{n\in\mathbb{N}}$.
3. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions qui converge simplement sur A vers une fonction f. Que dire de f lorsqu'à partir d'un certain rang, les f_n sont : positives ? croissantes ? dérivables ?

périodiques (de même période T) ? strictement positives ?

Correction.

1. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de A dans F et $f : A \to F$. Alors $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur A si, et seulement si,
$$\forall \varepsilon > 0,\ \forall x \in A,\ \exists N \in \mathbb{N},\ \forall n \geq N, \quad \|f_n(x) - f(x)\|_F \leq \varepsilon.$$
2. On suppose que $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ convergent simplement vers f et g respectivement sur A. Soit $x \in A$. On a :
$$(\lambda f_n + \mu g_n)(x) = \lambda f_n(x) + \mu g_n(x) \xrightarrow[n\to+\infty]{} \lambda f(x) + \mu g(x) = (\lambda f + \mu g)(x),$$
par linéarité du passage à la limite. Ainsi, $(\lambda f_n + \mu g_n)$ converge simplement vers $\lambda f + \mu g$ sur A.
De plus, on a, pour $x \in A$:
$$(f_n g_n)(x) = f_n(x) g_n(x) \xrightarrow[n\to+\infty]{} f(x)g(x) = (fg)(x),$$
donc $(f_n g_n)_{n\in\mathbb{N}}$ converge simplement vers fg sur A.
3. Par les propriétés de la limite, on remarque que la positivité, la croissance et la périodicité (avec une période commune) "passent" à la convergence simple.
Par contre, la dérivabilité ne passe pas (voire $f_n : t \to t^n$ sur $[0,1]$) et la stricte positivité ne passe pas non plus (voire $f_n : t \to t^n$ sur $]0,1]$).

Exercice 2.

Étudier la convergence simple de la suite de fonctions (f_n) définie, pour $n \in \mathbb{N}$, par :
$$a)\ f_n : t \mapsto \frac{1+nt^2}{1+nt}. \qquad b)\ f_n : t \mapsto \sin^n(t)\cos(t).$$
$$c)\ f_n : x \mapsto f_n(x) = \begin{cases} nx^2 & \text{si } x \in [-\frac{1}{n}, \frac{1}{n}] \\ 6n\left(\frac{1}{\sin(\frac{1}{nx})} - \frac{1}{\tan(\frac{1}{nx})}\right) & \text{si } x \notin [-\frac{1}{n}, \frac{1}{n}] \end{cases}$$

Correction.

a) Étudions la convergence simple de la suite $(f_n)_{n\in\mathbb{N}^*}$ sur $\mathbb{R}_+$:
Soit $x \in \mathbb{R}_+$. On a :
$$f_n(x) = \frac{1+nx^2}{1+nx} \xrightarrow[n\to+\infty]{} \begin{cases} 1 & \text{si } x = 0 \\ t & \text{si } x > 0 \end{cases}$$
Donc (f_n) converge simplement vers $f : t \mapsto \begin{cases} 1 & \text{si } t = 0 \\ t & \text{si } t > 0 \end{cases}$ sur $\mathbb{R}_+$.

b) Étudions la convergence simple de la suite $(f_n)_{n\in\mathbb{N}}$ sur R.
Soit $t \in \mathbb{R}$. On a :

$$f_n(t) = \sin^n(t)cos(t) \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{si } t \in \{\frac{\pi}{2}+k\pi \mid k \in \mathbb{Z}\} \text{ car } cos(t)=0 \\ 0 & \text{si } t \notin \{\frac{\pi}{2}+k\pi \mid k \in \mathbb{Z}\} \text{ car } |sin(t)|<1 \end{cases}$$

Donc $(f_n)_{n\in\mathbb{N}}$ converge simplement sur $\mathbb{R}$ vers la fonction nulle.

c) Étudions la convergence simple de la suite $(f_n)_{n\in\mathbb{N}^*}$ sur $\mathbb{R}$. Soit $x \in \mathbb{R}$.
Traitons tout d'abord le cas $x \neq 0$. Alors, à partir d'un certain rang N ($N = E(1/|x|)+1$ par exemple), pour tout $n \geq N$, $x \notin [-\frac{1}{n}, \frac{1}{n}]$. Par suite, pour tout $n \geq N$,

$$f_n(x) = 6n\left(\frac{1}{\sin(\frac{1}{nx})} - \frac{1}{\tan(\frac{1}{nx})}\right).$$

Or quand $n \to +\infty$, on a :

$$\begin{aligned} \frac{1}{\sin(\frac{1}{nx})} - \frac{1}{\tan(\frac{1}{nx})} &= \frac{\tan(\frac{1}{nx}) - \sin(\frac{1}{nx})}{\sin(\frac{1}{nx})\tan(\frac{1}{nx})} \\ &= \frac{\frac{1}{nx} + \frac{1}{3n^3x^3} + o(\frac{1}{n^3}) - (\frac{1}{nx} - \frac{1}{6n^3x^3} + o(\frac{1}{n^3}))}{\frac{1}{n^2x^2} + o(\frac{1}{n^2})} \\ &= \frac{\frac{1}{2n^3x^3} + o(\frac{1}{n^3})}{\frac{1}{n^2x^2} + o(\frac{1}{n^2})} \\ &\underset{n\to+\infty}{\sim} \frac{1}{2nx} \end{aligned}$$

Ainsi, toujours pour $x \neq 0$ et $n \geq N$,

$$f_n(x) = 6n\left(\frac{1}{\sin(\frac{1}{nx})} - \frac{1}{\tan(\frac{1}{nx})}\right) \underset{n\to+\infty}{\sim} \frac{6n}{3nx} = \frac{3}{x}.$$

Il en résulte que :

$$f_n(t) \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{si } x = 0 \\ \frac{3}{x} & \text{si } x \in \mathbb{R}^* \end{cases}$$

Donc (f_n) converge simplement sur $\mathbb{R}$ vers la fonction $f : x \mapsto \begin{cases} 0 & \text{si } x = 0 \\ \frac{3}{x} & \text{si } x \in \mathbb{R}^* \end{cases}$.

La définition suivante est une simple reformulation de la précédente dans le cas particulier des séries :

Définition 2. *Convergence simple d'une série de fonctions*

Soit $\sum f_n$ une série de fonctions de $A \subset E$ dans F. On dit que $\sum f_n$ **converge simplement sur** A si la suite de ses sommes partielles converge simplement sur A i.e. si pour tout $x \in A$, $\sum f_n(x)$ est convergente. Dans ce cas, on note $S : A \to F$ la fonction :

$$S : x \mapsto \sum_{n=0}^{+\infty} f_n(x).$$

Exemple 2.

— La série de fonctions $\sum f_n$ où $f_n : t \mapsto t^n$ converge simplement vers $t \mapsto \frac{1}{1-t}$ sur $]-1,1[$.

> En effet : soit $t \in]-1,1[$. On a :
>
> $$S_N = \sum_{n=0}^{N} t^n = \frac{1-t^{N+1}}{1-t} \xrightarrow[N\to+\infty]{} \frac{1}{1-t} \text{ car } |t|<1.$$

— La série de fonctions $\sum f_n$ où $f_n : t \mapsto \dfrac{t^n}{n!}$ converge simplement vers $t \mapsto e^t$ sur $\mathbb{R}$.

> En effet : soit $t \in \mathbb{R}$. En appliquant la règle de D'Alembert, on montre que $\sum \frac{t^n}{n!}$ est une série absolument convergente et donc convergente ; et de plus, sa somme vérifie
>
> $$\sum_{n=0}^{+\infty} \frac{t^n}{n!} = e^t.$$

2. Convergence uniforme

a. Définition et premières propriétés

Définition 3. *Convergence uniforme d'une suite de fonctions*

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de $A \subset E$ dans F.
On dit que la suite $(f_n)_{n\in\mathbb{N}}$ **converge uniformément vers** f **sur** A si :

$$\forall\, \varepsilon > 0,\ \exists\, N \in \mathbb{N},\ \forall\, x \in A,\ \forall\, n \geq N, \quad \|f_n(x) - f(x)\|_F \leq \varepsilon.$$

On dit que $(f_n)_{n\in\mathbb{N}}$ **converge uniformément sur** A s'il existe f telle que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A.

Question 1.

Comparer les définitions des convergences simple et uniforme en termes epsilonesques ! Quelle est la différence ?

On rappelle la notation suivante :

Notation 1. *Norme de la convergence uniforme*

Soit $\mathcal{F}_b(A,F)$ (ou encore $\mathcal{B}(A,F)$) l'espace vectoriel sur $\mathbb{K}$ des fonctions bornées de $A \subset E$ dans

F. On note $\|\cdot\|_\infty$ la **norme de la convergence uniforme** sur $\mathcal{F}_b(A, F)$ i.e. pour $f \in \mathcal{F}_b(A, F)$:

$$\|f\|_\infty = \sup_{x \in A} \|f(x)\|_F.$$

Proposition 1.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de A dans F. La suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A si, et seulement si :

i) Les fonctions $f_n - f$ sont bornées sur A à partir d'un certain rang N et,

ii) $\|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0$.

Démonstration.

- $(\Rightarrow)$. On suppose que (f_n) converge uniformément vers f sur A. Alors :

$$\forall\, \varepsilon > 0,\ \exists\, N \in \mathbb{N},\ \forall\, x \in A,\ \forall\, n \geq N, \quad \|f_n(x) - f(x)\|_F \leq \varepsilon.$$

 Prenons $\varepsilon = 1$. Alors il existe un rang $N_1 \in \mathbb{N}$ tel que pour tout $n \geq N$, on a, pour tout $x \in A$, $\|f_n(x) - f(x)\|_F \leq 1$.
 Ainsi, à partir d'un certain rang, $f_n - f$ est bornée sur A par 1.
 Montrons désormais que $\|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0$.
 Soit $\varepsilon > 0$. Alors

$$\exists\, N \in \mathbb{N},\ \forall\, x \in A,\ \forall\, n \geq N, \quad \|f_n(x) - f(x)\|_F \leq \varepsilon.$$

 Par suite, comme l'inégalité précédente est vraie pour tout $x \in A$, on a, pour tout $n \geq N$,

$$\|f_n - f\|_\infty = \sup_{x\in A} \|f_n(x) - f(x)\|_F \leq \varepsilon.$$

 Il en résulte que $\|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0$.

- $(\Leftarrow)$. On suppose $i)$ et $ii)$. Du fait de $i)$, quitte à supprimer une nombre fini de termes de la suite (f_n) on peut supposer que, pour tout $n \in \mathbb{N}$, $f_n - f$ est bornée sur A. Ainsi, on peut considérer, pour chaque $n \in \mathbb{N}$, la quantité $\|f_n - f\|_\infty = \sup_{t\in A} \|f_n(t) - f(t)\|_F$.
 Soit $\varepsilon > 0$. D'après $ii)$, $\|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0$, il existe un rang $N \in \mathbb{N}$ tel que pour tout entier $n \geq N$,

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

 Soit $n \in \mathbb{N}$ avec $n \geq N$ et $x \in A$. Alors on a :

$$\|f_n(x) - f(x)\|_F \leq \sup_{t\in A} \|f_n(t) - f(t)\|_F = \|f_n - f\|_\infty \leq \varepsilon.$$

 Il en résulte que la suite $(f_n - f)_{n\in\mathbb{N}}$ converge uniformément vers f.

□

Proposition 2.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de A dans F. Si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A, alors $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur A.

Démonstration.

On suppose que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A. alors d'après la proposition précédente, il existe un rang $N \in \mathbb{N}$ tel que pour tout $n \geq N$, $f_n - f$ est bornée. Soit $x \in A$. Pour $n \geq N$, on a :

$$|f_n(x) - f(x)| \leq \|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0,$$

d'où $f_n(x) \xrightarrow[n\to+\infty]{} f(x)$. Il en résulte que $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur A. $\square$

Exemple 3.

— La suite $(f_n)_{n\in\mathbb{N}}$ telle que, pour $n \in \mathbb{N}$, $f_n : x \mapsto \sqrt{x + \frac{1}{n}}$ converge uniformément vers $x \mapsto \sqrt{x}$ sur $\mathbb{R}_+$.

> En effet : soit $n \in \mathbb{N}^*$ et on note $f : x \mapsto \sqrt{x}$. On a, pour tout $x \in \mathbb{R}_+$,
>
> $$|f_n(x) - f(x)| = |\sqrt{x + \frac{1}{n}} - \sqrt{x}| = \frac{\frac{1}{n}}{\sqrt{x + \frac{1}{n}} + \sqrt{x}} \leq \frac{\frac{1}{n}}{\sqrt{0 + \frac{1}{n}} + \sqrt{0}} = \frac{\frac{1}{n}}{\frac{1}{\sqrt{n}}} = \frac{1}{\sqrt{n}}.$$
>
> Par suite, les fonctions $f_n - f$ sont bornées sur $\mathbb{R}_+$ et :
>
> $$\|f_n - f\|_\infty = \sup_{x\in\mathbb{R}_+} |f_n(x) - f(x)| \leq \frac{1}{\sqrt{n}} \xrightarrow[n\to+\infty]{} 0$$

— La suite de fonctions de terme général $t \mapsto t^n$ ne converge pas uniformément sur $[0, 1[$.

> En effet, on remarque tout d'abord que (f_n) converge simplement vers le fonction nulle sur $[0, 1[$.
> Soit $n \in \mathbb{N}$ et on note $f : x \mapsto 0$. Comme $f_n : t \mapsto t^n$ est strictement croissante sur $[0, 1[$ (pour $n > 0$), on a :
>
> $$\|f_n - f\|_\infty = \sup_{t\in[0,1[} t^n = \lim_{t\to 1^-} t^n = 1$$
>
> Par suite,
>
> $$\|f_n - f\|_\infty = 1 \underset{n\to+\infty}{\not\to} 0.$$

Proposition 3.

Soit $(f_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ des suites de fonctions de A dans F et $\lambda,\mu\in\mathbb{K}$. Si $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ convergent uniformément vers respectivement f et g sur A, alors $(\lambda f_n+\mu g_n)_{n\in\mathbb{N}}$ converge uniformément vers $\lambda f+\mu g$ sur A.

Démonstration.

On suppose $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ convergent uniformément vers respectivement f et g sur A. Alors, à partir d'un certain rang N_1 (resp. N_2), pour tout $n\geq N_1$ (resp. $n\geq N_2$), f_n-f (resp. g_n-g) est bornée sur A. Par suite, comme l'ensemble des fonctions bornées sur A est un espace vectoriel, à partir du rang $N=\max(N_1,N_2)$, pour tout $n\geq N$, $\lambda f_n+\mu g_n-(\lambda f+\mu g)$ est bornée sur A. Et de plus, on a :

$$\|\lambda f_n+\mu g_n-(\lambda f+\mu g)\|_\infty\leq|\lambda|\|f_n-f\|_\infty+|\mu|\|g_n-g\|_\infty\xrightarrow[n\to+\infty]{}0,$$

car $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ convergent uniformément vers respectivement f et g sur A. Il en résulte que $(\lambda f_n+\mu g_n)_{n\in\mathbb{N}}$ converge uniformément vers $\lambda f+\mu g$ sur A. □

Méthode : Montrer qu'une suite de fonctions converge uniformément.

- *Limite potentielle :* On étudie la convergence simple de la suite $(f_n)_{n\in\mathbb{N}}$. S'il y a convergence simple vers une fonction f sur A, on étudie alors la convergence uniforme vers f sur A.
- *Convergence uniforme vers la limite :* Pour montrer la convergence uniforme de $(f_n)_{n\in\mathbb{N}}$, on cherche à obtenir une majoration de $\|f_n-f\|_\infty$ qui tende vers 0 i.e. une majoration **indépendante de** $x\in A$ du type (à partir d'un certain rang)

$$\|f_n(x)-f(x)\|_F\leq u_n\xrightarrow[n\to+\infty]{}0$$

où (u_n) est une suite de réels positifs qui tend vers 0 (**et qui ne dépend pas de** $x\in A$ **!!!**).

La suite (u_n) s'obtient la plus souvent par une majoration simple, quand c'est possible, de $\|f_n(x)-f(x)\|_F$ ou par une étude des extrema de la fonction $x\mapsto\|f_n(x)-f(x)\|_F$ (à n fixé).

Exercice 3.

Étudier la convergence uniforme des suites de fonctions de termes généraux suivants :

1. pour $n\in\mathbb{N}$, $f_n:[-\frac{1}{2},\frac{1}{2}]\to\mathbb{R}$ tel que $f_n:x\mapsto x^n$.
2. pour $n\in\mathbb{N}$, $f_n:[0,1]\to\mathbb{R}$ tel que $f_n:x\mapsto x^n(1-x)$.
3. pour $n\in\mathbb{N}$, $f_n:\mathbb{R}\to\mathbb{R}$ tel que $f_n:x\mapsto\sin(x+\frac{1}{n})$.
4. pour $n\in\mathbb{N}$, $f_n:\mathbb{R}_+\to\mathbb{R}$ tel que $f_n:x\mapsto\frac{x}{n(1+x^n)}$.

Correction.

1. — CVS sur $[-\frac{1}{2}, \frac{1}{2}]$. Soit $x \in [-\frac{1}{2}, \frac{1}{2}]$. On a :

$$f_n(x) = x^n \xrightarrow[n \to +\infty]{} 0$$

car $|x| \leq \frac{1}{2} < 1$.
Par suite, $(f_n)_{n \in \mathbb{N}}$ converge simplement vers $\mathbf{0}$ (la fonction nulle) sur $[-\frac{1}{2}, \frac{1}{2}]$.

— CVU sur $[-\frac{1}{2}, \frac{1}{2}]$. Soit $n \in \mathbb{N}$. Pour tout $x \in [-\frac{1}{2}, \frac{1}{2}]$, on a :

$$|f_n(x) - f(x)| = |x|^n \leq \frac{1}{2^n},$$

D'où $f_n - f$ est borné sur $[-\frac{1}{2}, \frac{1}{2}]$ et :

$$\|f_n - f\|_\infty \leq \frac{1}{2^n} \xrightarrow[n \to +\infty]{} 0.$$

Remarque : l'inégalité précédente est en fait une égalité.

Il en résulte que $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers $\mathbf{0}$ sur $[-\frac{1}{2}, \frac{1}{2}]$.

2. — CVS sur $[0, 1]$. Soit $x \in [0, 1]$. On a :

$$f_n(x) = x^n(1-x) \xrightarrow[n \to +\infty]{} \begin{cases} 0 \times (1-x) = 0 & \text{si } 0 \leq x < 1 \\ 1 \times 0 = 0 & \text{si } x = 1 \end{cases}$$

Par suite, $(f_n)_{n \in \mathbb{N}}$ converge simplement vers $\mathbf{0}$ (la fonction nulle) sur $[0, 1]$.

— CVU sur $[0, 1]$. Soit $n \in \mathbb{N}^*$. On étudie sur $[0, 1]$ la fonction :

$$x \mapsto g_n(x) = |f_n(x) - f(x)| = x^n(1-x).$$

La fonction g_n est dérivable, et on a $g_n'(x) = x^{n-1}(n - (n+1)x)$.

x	0		$\frac{n}{n+1}$		1
$g_n'(x)$	0	$+$	0	$-$	
$g_n(x)$	0	$\nearrow$	$g_n(\frac{n}{n+1})$	$\searrow$	0

Par suite, on a :

$$\|f_n - f\|_\infty = g_n(\frac{n}{n+1}) = \left(\frac{n}{n+1}\right)^n (1 - \frac{n}{n+1}) \xrightarrow[n \to +\infty]{} \frac{1}{e} \times 0 = 0.$$

Il en résulte que $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers $\mathbf{0}$ sur $[0, 1]$.

3. — CVS sur $\mathbb{R}$. Soit $x \in \mathbb{R}$. On a, par continuité de la fonction sin sur $\mathbb{R}$ et donc en x :

$$f_n(x) = \sin(x + \frac{1}{n}) \xrightarrow[n \to +\infty]{} \sin(x)$$

Par suite, $(f_n)_{n \in \mathbb{N}}$ converge simplement vers $f = \sin$ sur $\mathbb{R}$.

— <u>CVU sur $\mathbb{R}$</u>. Soit $n \in \mathbb{N}^*$. Pour tout $x \in \mathbb{R}$, on a :

$$\begin{aligned}
|f_n(x) - f(x)| &= |\sin(x + \tfrac{1}{n}) - \sin(x)| \\
&= |\cos(x)\sin(\tfrac{1}{n}) + \sin(x)\left(\cos(\tfrac{1}{n}) - 1\right)| \\
&\leq |\cos(x)||\sin(\tfrac{1}{n})| + |\sin(x)||\cos(\tfrac{1}{n}) - 1| \\
&\leq |\sin(\tfrac{1}{n})| + |\cos(\tfrac{1}{n}) - 1|
\end{aligned}$$

D'où $f_n - f$ est borné par 2 sur $\mathbb{R}$ et :

$$\|f_n - f\|_\infty \leq |\sin(\frac{1}{n})| + |\cos(\frac{1}{n}) - 1| \xrightarrow[n\to+\infty]{} 0.$$

car $\lim_{t\to 0}\sin(t) = 0$ et $\lim_{t\to 0}\cos(t) = 1$.

Il en résulte que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers sin sur $\mathbb{R}$.

Correction suite.

4. — <u>CVS sur $\mathbb{R}_+$</u>. Soit $x \in \mathbb{R}_+$. On a :

$$f_n(x) = \frac{x}{n(1+x^n)} \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{si } x = 0 \\ 0 & \text{si } x > 0. \end{cases}$$

Par suite, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $\mathbf{0}$ (la fonction nulle) sur $\mathbb{R}_+$.

— <u>CVU sur $\mathbb{R}_+$</u>. Soit $n \geq 2$. On étudie sur $[0,1]$ la fonction :

$$x \mapsto g_n(x) = |f_n(x) - f(x)| = \frac{x}{n(1+x^n)}.$$

La fonction g_n est dérivable, et on a

$$g_n'(x) = \frac{n(1+x^n) - n^2x^n}{n^2(1+x^n)^2} = \frac{1-(n-1)x^n}{n(1+x^n)^2} = \frac{1-(n-1)x^n}{n(1+x^n)^2}$$

x	0		$\frac{1}{\sqrt[n]{n-1}}$		$+\infty$
$g_n'(x)$		$+$	0	$-$	
$g_n(x)$	0	↗	$g_n(\frac{1}{\sqrt[n]{n-1}})$	↘	0

Par suite, on a :

$$\|f_n - f\|_\infty = g_n(\frac{1}{\sqrt[n]{n-1}}) = \frac{1}{\sqrt[n]{n-1}}.\frac{1}{n(1+\dfrac{1}{n-1})} \xrightarrow[n\to+\infty]{} 1 \times 0 = 0.$$

Il en résulte que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers $\mathbf{0}$ sur $\mathbb{R}_+$.

Remarque : on pouvait conclure plus rapidement en remarquant l'inégalité suivante :

$$g_n(x) = \frac{x}{n(1+x^n)} \leq \frac{1}{n}.$$

En effet, pour $0 \leq x \leq 1$, $\frac{x}{1+x^n} \leq \dfrac{1}{1+x^n} \leq 1$;

et pour $x > 1$, $x^n > x$, *donc* $\frac{x}{1+x^n} \leq \dfrac{x}{1+x} \leq 1$.

Méthode : Montrer qu'une suite de fonctions ne converge pas uniformément.

S'il n'y a pas convergence simple sur A, il n'y pas convergence uniforme.
Mais si on a déterminé une limite f pour la convergence simple, pour montrer que la suite (f_n) ne converge pas uniformément vers f, on peut :

- montrer que la fonction $f_n - f$ n'est pas bornée sur A, ou
- exhiber une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans A telle que la suite de terme général

$$\|f_n(x_n) - f(x_n)\|_F \textbf{ ne tend pas vers } 0.$$

Exercice 4.

Étudier la convergence uniforme des suites de fonctions de termes généraux suivants :

1. pour $n \in \mathbb{N}$, $f_n : \mathbb{R} \to \mathbb{R}$ tel que $f_n : x \mapsto \frac{x^n}{n!}$.
2. pour $n \in \mathbb{N}$, $f_n : \mathbb{R} \to \mathbb{R}$ tel que $f_n : x \mapsto \frac{nx}{1+n^2x^2}$.
3. pour $n \in \mathbb{N}^*$, $f_n : \mathbb{R} \to \mathbb{R}$ tel que $f_n : x \mapsto \left(x + \frac{1}{n}\right)^2$.

Correction.

1. — CVS sur $\mathbb{R}$. Soit $x \in \mathbb{R}$. On a :

$$f_n(x) = \frac{x^n}{n!} \xrightarrow[n\to+\infty]{} 0.$$

Par suite, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $\mathbf{0}$ (la fonction nulle) sur $\mathbb{R}$.

— CVU sur $\mathbb{R}$. Soit $n \in \mathbb{N}^*$. Pour $x \in \mathbb{R}$, on a :

$$|f_n(x) - f(x)| = \frac{|x|^n}{n!} \xrightarrow[x\to\pm\infty]{} +\infty,$$

Par suite, pour tout $n \in \mathbb{N}^*$, $f_n - f$ n'est pas bornée sur $\mathbb{R}$.
Il en résulte que $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément vers $\mathbf{0}$ sur $\mathbb{R}$.

2. — CVS sur $\mathbb{R}$. Soit $x \in \mathbb{R}$. On a :

$$f_n(x) = \frac{nx}{1+n^2x^2} \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{si } x = 0 \\ 0 & \text{sinon.} \end{cases}$$

car, pour $x \neq 0$, $\dfrac{nx}{1+n^2x^2} \underset{n\to+\infty}{\sim} \dfrac{1}{nx}$.
Par suite, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $\mathbf{0}$ (la fonction nulle) sur $\mathbb{R}$.

— CVU sur $\mathbb{R}$. Soit $n \in \mathbb{N}$. On considère, pour $x \in \mathbb{R}$:

$$g_n(x) = |f_n(x) - f(x)| = \frac{n|x|}{1 + n^2x^2}.$$

On remarque que pour $x_n = \dfrac{1}{n}$, on a :

$$g_n(x_n) = \frac{n\frac{1}{n}}{1 + n^2(\frac{1}{n})^2} = \frac{1}{2}$$

Donc :

$$\|f_n - f\|_\infty = \sup_{x \in \mathbb{R}} g_n(x) \geq g_n(x_n) = \frac{1}{2} \underset{n \to +\infty}{\not\to} 0.$$

Il en résulte que $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément vers $\mathbf{0}$ sur $\mathbb{R}$.

3. — CVS sur $\mathbb{R}$. Soit $x \in \mathbb{R}$. On a, par continuité sur $\mathbb{R}$ et donc en x de la fonction carrée :

$$f_n(x) = \left(x + \frac{1}{n}\right)^2 \xrightarrow[n \to +\infty]{} x^2.$$

Par suite, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $f : x \mapsto x^2$ sur $\mathbb{R}$.

— CVU sur $\mathbb{R}$. Soit $n \in \mathbb{N}^*$. Pour $x \in \mathbb{R}$, on a :

$$|f_n(x) - f(x)| = |\left(x + \frac{1}{n}\right)^2 - x^2| = |\frac{2x}{n} + \frac{1}{n^2}| \xrightarrow[x \to \pm\infty]{} +\infty,$$

Par suite, pour tout $n \in \mathbb{N}^*$, $f_n - f$ n'est pas bornée sur $\mathbb{R}$.
Il en résulte que $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément vers f sur $\mathbb{R}$.

Exercice 5.

Soit $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ des suites de fonctions à valeurs réelles qui convergent uniformément vers f et g respectivement. Est-ce que $(f_ng_n)_{n\in\mathbb{N}}$ converge uniformément ?

Correction.

Non, car la suite $(f_n)_{n\in\mathbb{N}}$, définie par $f_n : x \mapsto x + \frac{1}{n}$ pour $n \in \mathbb{N}$, converge uniformément vers $f : x \mapsto x$ sur $\mathbb{R}$ et on a prouvé précédemment que la suite de terme général $f_n^2 : x \mapsto (x + \frac{1}{n})^2$ ne converge pas uniformément vers $x \mapsto x^2$ sur $\mathbb{R}$.

b. Convergence uniforme des suites de fonctions bornées

Proposition 4.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions bornées sur A. Si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers une fonction f sur A, alors f est bornée sur A.

Démonstration.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions bornées sur A. On suppose que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A.

Par convergence uniforme, à partir d'un certain rang N, $f_n - f$ est bornée sur A. Par suite, pour tout $x \in A$, on a, pour $n \geq N$:

$$\|f(x)\|_F \leq \|f_n(x)\|_F + \|f_n(x) - f(x)\|_F \leq \|f_n\|_\infty + \|f_n - f\|_\infty$$

Donc f est bornée sur A. □

Question 2.

Cela est-il vrai dans le cas de la converge simple ?

Réponse : Non ! Considérons la fonction $f : x \mapsto x$ et la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ définie, pour $n \in \mathbb{N}$, par :

$$f_n : x \mapsto \begin{cases} f(x) = x & \text{si } x \in [-n, n] \\ 0 & \text{sinon.} \end{cases}$$

Alors, $(f_n)_{n\in\mathbb{N}}$ est un suite de fonctions bornées sur $\mathbb{R}$ qui converge simplement vers f qui n'est pas bornée sur $\mathbb{R}$.
(Et bien-sur, il ne peut y avoir convergence uniforme vers f sur $\mathbb{R}$ en vertu de la proposition précédente !)

Proposition 5.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions bornées sur A. Alors $(f_n)_{n\in\mathbb{N}}$ converge uniformément si, et seulement si, (f_n) converge dans l'espace vectoriel normé $(\mathcal{F}_b(A, F), \|\cdot\|_\infty)$.

Démonstration.

Il s'agit simplement de deux formulations différentes de la même propriété. □

Exercice 6.

Soit $g : \mathbb{R} \to \mathbb{R}$ telle que $g : x \mapsto \dfrac{x}{1 + 27x^4}$

1. Calculer $\|g\|_\infty$.
2. On considère la suite de terme générale $f_n : x \mapsto g(nx)$. Étudier les convergences simple et uniforme de cette suite.

Correction.

1. La fonction g est une fonction impaire et dérivable sur $\mathbb{R}$. On effectue son étude sur $\mathbb{R}_+$. Pour $x \in \mathbb{R}_+$, on a :

$$g'(x) = \frac{(1+27x^4) - 108x^4}{(1+27x^4)^2} = \frac{1-81x^4}{(1+27x^4)^2},$$

x	0		$\frac{1}{3}$		$+\infty$
$g'(x)$		$+$	0	$-$	
$g(x)$	0	$\nearrow$	$\frac{1}{4}$	$\searrow$	0

Par suite, comme g est impaire, on a :

$$\|g\|_\infty = \sup_{x\in\mathbb{R}} |g(x)| = \sup_{x\in\mathbb{R}_+} |g(x)| = \frac{1}{4}.$$

2. — <u>CVS sur $\mathbb{R}$.</u> Soit $x \in \mathbb{R}$. On a :

$$f_n(x) = g(nx) = \frac{nx}{1+27n^4x^4} \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{si } x = 0 \\ 0 & \text{si } x \neq 0, \end{cases}$$

car pour $x \neq 0$, $\dfrac{nx}{1+27n^4x^4} \underset{n\to+\infty}{\sim} \dfrac{1}{27n^3x^3}$.
Par suite, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $f : x \mapsto 0$ sur $\mathbb{R}$.

— <u>CVS sur $\mathbb{R}$.</u> Soit $n \in \mathbb{N}^*$. La fonction $\varphi : x \mapsto nx$ est une bijection de $\mathbb{R}$ dans $\mathbb{R}$, donc :

$$\|f_n - f\|_\infty = \sup_{x\in\mathbb{R}} |g(nx)| = \sup_{y\in\mathbb{R}} |g(y)| = \frac{1}{4} \underset{n\to+\infty}{\not\to} 0.$$

Par suite, $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément vers f sur $\mathbb{R}$.

Exercice 7.

Soit $(f_n)_{n\in\mathbb{N}}$ et $(g_n)_{n\in\mathbb{N}}$ des suites de fonctions à valeurs réelles **bornées** qui convergent uniformément vers f et g respectivement. Est-ce que $(f_ng_n)_{n\in\mathbb{N}}$ converge uniformément ?

Correction.

Cette fois-ci, l'hypothèse "bornées" permet de conclure par l'affirmative.
En effet, d'après la proposition précédente, f et g sont bornées sur A. De plus, pour tout $n \in \mathbb{N}$ et pour tout $x \in A$, on a :

$$\begin{array}{rcl} |(f_n g_n - fg)(x)| & = & |f_n(x)g_n(x) - f(x)g_n(x) + f(x)g_n(x) - f(x)g(x)| \\ & \leq & |f_n(x)g_n(x) - f(x)g_n(x)| + |f(x)g_n(x) - f(x)g(x)| \\ & \leq & |g_n(x)|.|f_n(x) - f(x)| + |f(x)|.|g_n(x) - g(x)| \\ & \leq & |g_n(x)|.|f_n(x) - f(x)| + |f(x)|.|g_n(x) - g(x)| \\ & \leq & \|g_n\|_\infty \|f_n - f\|_\infty + \|f\|_\infty \|g_n - g\|_\infty \end{array}$$

Donc $f_n g_n - fg$ est bornée et on remarque que, par convergence uniforme de $(g_n)_{n\in\mathbb{N}}$ vers g, on a :

$$\|g_n\|_\infty = \|g_n - g + g\|_\infty \leq \|g_n - g\|_\infty + \|g\|_\infty \xrightarrow[n\to+\infty]{} \|g\|_\infty.$$

Ainsi, on a :

$$\|f_n g_n - fg\|_\infty \leq \|g_n\|_\infty \|f_n - f\|_\infty + \|f\|_\infty \|g_n - g\|_\infty \xrightarrow[n\to+\infty]{} \|g\|_\infty \times 0 + \|f\|_\infty \times 0 = 0.$$

Donc $(f_n g_n)_{n\in\mathbb{N}}$ converge uniformément vers fg.

c. Convergence uniforme sur une partie

Définition 4. *Convergence uniforme sur une partie*

Soit $B \subset A$ et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de A dans $\mathbb{R}$. On dit que $(f_n)_{n\in\mathbb{N}}$ **converge uniformément sur** $B \subset A$ si, pour $(f_{n|_B})_{n\in\mathbb{N}}$ converge uniformément sur B.

Exemple 4.

Pour tout $a > 0$, la suite de fonctions de terme général $f_n : t \mapsto \frac{1}{1+nt}$ converge uniformément sur $[a, +\infty[\subset \mathbb{R}_+^*$.

On remarque tout d'abord que $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément sur $\mathbb{R}_+^*$. En effet, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $f : x \mapsto 0$ sur $\mathbb{R}_+^*$ mais on a, pour tout $n \in \mathbb{N}$:

$$\|f_n - f\|_\infty = \sup_{t\in\mathbb{R}}(\frac{1}{1+nt}) = 1 \underset{n\to+\infty}{\not\to} 0,$$

car f_n est décroissante sur $\mathbb{R}_+^*$ et $\lim_{t\to 0} \dfrac{1}{1+nt} = 1$.
Donc $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément sur $\mathbb{R}_+^*$.
Remarque : on aurait aussi pu utiliser le fait que, pour $n \in \mathbb{N}^*$:

$$\|f_n - f\|_\infty = \|f_n\|_\infty \geq f_n(\frac{1}{n}) = \frac{1}{1+n\frac{1}{n}} = \frac{1}{2} \underset{n\to+\infty}{\not\to} 0.$$

Soit $a > 0$. Étudions la convergence uniforme sur $[a, +\infty[$.
Soit $n \in \mathbb{N}$. On a, pour tout $x \in [a, +\infty[$,

$$|f_n(t) - f(t)| = \frac{1}{1+nt} \leq \frac{1}{1+na}.$$

Ainsi, les $f_n - f$ sont bornés sur $[a, +\infty[$ et :

$$\|f_n - f\|_\infty \leq \frac{1}{1+na} \xrightarrow[n\to+\infty]{} 0.$$

Il en résulte que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[a, +\infty[$.

Exercice 8.

Étudier la converge uniforme de la suite de terme général $f_n : x \mapsto e^{-nx}\sin(nx)$ sur $\mathbb{R}_+$ puis sur les intervalles de la forme $[a, +\infty[$ avec $a > 0$.

Correction.

— CVS sur $\mathbb{R}_+$. Soit $x \in \mathbb{R}_+$. On a :

$$f_n(x) \xrightarrow[n\to+\infty]{} \begin{cases} 0 & \text{si } x = 0 \text{ car } \sin(0) = 0 \\ 0 & \text{sinon} \end{cases}$$

car, pour $x > 0$, $e^{-nx} \xrightarrow[n\to]{} 0$ et $|\sin(nx)| \leq 1$.
Par suite, $(f_n)_{n\in\mathbb{N}}$ converge simplement vers $f : x \mapsto 0$ sur $\mathbb{R}_+$.

— CVU sur $\mathbb{R}_+$. On considère la suite $(x_n)_{n\in\mathbb{N}^*}$ telle que, pour $n \in \mathbb{N}^*$, $x_n = \frac{1}{n}$. Alors, on a :

$$|f_n(x_n) - f(x_n)| = e^{-n\frac{1}{n}}\sin(n\frac{1}{n}) = e\sin(1).$$

Par suite,

$$\|f_n - f\|_\infty \geq e\sin(1) \underset{n\to+\infty}{\not\to} 0.$$

Donc $(f_n)_{n\in\mathbb{N}}$ ne converge pas uniformément sur $\mathbb{R}_+$.

— Soit $a > 0$. CVU sur $[a,+\infty[$. Soit $n \in \mathbb{N}$. Pour tout $x \in [a,+\infty[$,

$$|f_n(x) - f(x)| = e^{-nx}|\sin(nx)| \leq e^{-na}.$$

Par suite,

$$\|f_n - f\|_\infty \leq e^{-na} \xrightarrow[n\to+\infty]{} 0.$$

Il en résulte que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[a,+\infty[$.

Définition 5. *Convergence uniforme sur tout segment*

Soit I un intervalle d'intérieur non vide et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de I dans $\mathbb{R}$. On dit que $(f_n)_{n\in\mathbb{N}}$ **converge uniformément sur tout segment de** I si, pour tout segment $[a,b] \subset I$, $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur $[a,b]$.

Exemple 5.

La suite de fonctions de terme général $f_n : t \mapsto t^n$ converge uniformément sur tout segment de $]-1,1[$.

Exercice 9.

Étudier la converge uniforme de la suite de terme général $f_n : x \mapsto \frac{x^n}{n!}$ sur tout segment de $\mathbb{R}$.

Correction.

On a déjà prouvé que $(f_n)_{n\in\mathbb{N}}$ sur $\mathbb{R}$ où $f_n : x \mapsto \frac{x^n}{n!}$ pour $n \in \mathbb{N}$, converge simplement vers la fonction nulle - notée f ici - sur $\mathbb{R}$ et on a prouvé que cette suite ne converge pas uniformément sur $\mathbb{R}$. Montrons qu'il y a tout de même convergence uniforme sur tout segment de $\mathbb{R}$.

Soit $a > 0$.

CVU sur $[-a,a]$. Soit $n \in \mathbb{N}$. Pour tout $x \in [-a,a]$, on a :

$$|f_n(x) - f(x)| = |f_n(x)| = \frac{|x|^n}{n!} \leq \frac{a^n}{n!}$$

Donc la fonction $f_n - f$ est bornée sur $[-a,a]$ et on a :

$$\|f_n - f\|_\infty = \sup_{x\in[-a,a]} |f_n(x) - f(x)| \leq \frac{a^n}{n!} \xrightarrow[n\to+\infty]{} 0.$$

Il en résulte que $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur $[-a,a]$ vers la fonction nulle.

Comme tout segment de $\mathbb{R}$ est inclus dans un intervalle de la forme $[-a,a]$ avec $a > 0$, $(f_n)_{n\in\mathbb{N}}$ converge uniformément sur tout segment de $\mathbb{R}$ vers la fonction nulle.

3. Convergence uniforme des séries de fonctions

a. Généralités

Définition 6.

Soit $\sum f_n$ une série de fonctions de A dans F. On dit que $\sum f_n$ **converge uniformément sur** A si la suite $(S_n)_{n\in\mathbb{N}}$ de ses sommes partielles converge uniformément sur A.

Proposition 6.

Soit $\sum f_n$ une série de fonctions de A dans F. Alors $\sum f_n$ converge uniformément sur A si, et seulement si, $\sum f_n$ converge simplement sur A et la suite $(R_n)_{n\in\mathbb{N}}$ des restes de $\sum f_n$ converge uniformément vers la fonction nulle 0 sur A.

Démonstration.

$\sum f_n = (S_n)_{n\in\mathbb{N}}$ converge uniformément vers la fonction S sur A si, et seulement si, $(R_n)_{n\in\mathbb{N}} = (S - S_n)_{n\in\mathbb{N}}$ converge uniformément vers la fonction $S - S = 0$ sur A. □

Exemple 6.

— La suite de fonctions de terme général $f_n : x \mapsto x^n$ converge uniformément vers $x \mapsto \dfrac{1}{1-x}$ sur tout segment de $]-1,1[$ mais ne converge pas uniformément sur $]-1,1[$.

Exercice 10.

1. Étudier la convergence uniforme de la série $\sum \frac{(-1)^n}{n+1+x^2}$ sur $\mathbb{R}$.
2. Étudier la convergence uniforme de la série $\sum \frac{x^n}{n!}$ sur $\mathbb{R}$ puis sur tout segment de $\mathbb{R}$.

Correction.

1. Tout d'abord étudions la convergence simple de $\sum f_n$ sur $\mathbb{R}$ où $f_n : x \mapsto \frac{(-1)^n}{n+1+x^2}$ pour $n \in \mathbb{N}$.

   <u>CVS sur $\mathbb{R}$.</u> Soit $x \in \mathbb{R}$. On considère la suite $(u_n)_{n\in\mathbb{N}}$ telle que, pour $n \in \mathbb{N}$, $u_n = \frac{1}{n+1+x^2}$. Alors $(u_n)_{n\in\mathbb{N}}$ est décroissante et tend vers 0 quen $n \to +\infty$ donc, d'après le critère des séries alternées, $\sum \frac{(-1)^n}{n+1+x^2}$ converge.

   Il en résulte que $\sum f_n$ converge simplement vers $f : x \mapsto \displaystyle\sum_{n=0}^{+\infty} \frac{(-1)^n}{n+1+x^2}$ sur $\mathbb{R}$.

   <u>CVU sur $\mathbb{R}$.</u> Soit $n \in \mathbb{N}$. Pour tout $x \in \mathbb{R}$, comme la somme d'une série alternées convergente

est plus petite que son premier terme (en valeur absolue), on a :

$$|R_n(x)| = \left| \sum_{k=n+1}^{+\infty} \frac{(-1)^k}{k+1+x^2} \right| \leq \frac{1}{n+2+x^2} \leq \frac{1}{n+2}$$

Donc la fonction R_n est bornée et :

$$\|R_n\|_\infty = \sup_{x \in \mathbb{R}} |R_n(x)| \leq \frac{1}{n+2} \xrightarrow[n \to +\infty]{} 0.$$

Par suite, $\sum f_n$ converge uniformément sur $\mathbb{R}$ (vers f).

2. On a déjà prouvé que $\sum f_n$ sur $\mathbb{R}$ où $f_n : x \mapsto \frac{x^n}{n!}$ pour $n \in \mathbb{N}$, converge simplement vers exp sur $\mathbb{R}$ - et on a également montrer que $\sum f_n$ converge uniformément sur tout segment de $\mathbb{R}$.

   Montrons que $\sum f_n$ ne converge pas uniformément sur $\mathbb{R}$ tout entier.

   CVU sur $\mathbb{R}$. Soit $n \in \mathbb{N}$. Pour tout $x \in \mathbb{R}$, on a :

$$|R_n(x)| = \left| \sum_{k=n+1}^{+\infty} \frac{x^k}{k!} \right|$$

   Si $x > 0$, on a $\mapsto \frac{x^k}{k!} > 0$ pour tout $k \geq n+1$. Ainsi, pour tout $x > 0$:

$$|R_n(x)| = \sum_{k=n+1}^{+\infty} \frac{x^k}{k!} \geq \frac{x^{n+1}}{(n+1)!}$$

   Or la fonction $x \mapsto \frac{x^{n+1}}{(n+1)!}$ n'est pas bornée sur $\mathbb{R}$; en effet sa limite quand $x \to +\infty$ est $+\infty$, donc la fonction R_n n'est pas bornée sur $\mathbb{R}$.

   Il en résulte que $\sum f_n$ ne converge pas uniformément sur $\mathbb{R}$.

   Comme il n'y a pas convergence sur $\mathbb{R}$ et que le "problème" se situe en $+\infty$, on tente la convergence uniforme sur tout segment de $\mathbb{R}$. Soit $a \in \mathbb{R}$.

   CVU sur $[-a, a]$. Soit $n \in \mathbb{N}$. Pour tout $x \in [-a, a]$, en faisant le changement d'indice $k' = k - (n+1)$ et en remarquant que, pour tout $k \in \mathbb{N}$, $(k+n+1)! \geq k!(n+1)!$, on obtient :

$$|R_n(x)| \leq \sum_{k=n+1}^{+\infty} \frac{|x|^k}{k!} \leq \frac{|x|^{n+1}}{(n+1)!} \sum_{k=0}^{+\infty} \frac{|x|^k}{k!} = \frac{|x|^{n+1}}{(n+1)!} e^{|x|} \leq \frac{a^{n+1}}{(n+1)!} e^a$$

   Donc la fonction R_n est bornée et on a :

$$\|R_n\|_\infty = \sup_{x \in [-a,a]} |R_n(x)| \leq \frac{a^{n+1}}{(n+1)!} e^a \xrightarrow[n \to +\infty]{} 0$$

   On aurait également pu majorer (et c'est même une égalité) $\|R_n\|_\infty$ par $r_n = \sum_{k=n+1}^{+\infty} \frac{a^k}{k!}$ qui est le reste d'ordre n d'une série convergente et donc qui tend vers 0 quand n tend vers l'infini.

   Par suite, $\sum f_n$ converge uniformément sur $[-a, a]$ (vers exp).

   Tout segment de $\mathbb{R}$ étant inclus dans un segment de la forme $[-a, a]$ avec $a > 0$, on a donc la convergence uniforme sur tout segment de $\mathbb{R}$ de $\sum f_n$ vers exp.

b. Convergence normale des séries de fonctions

Définition 7. *Convergence normale*

Soit $\sum f_n$ une série de fonctions de A dans F. On dit que $\sum f_n$ **converge normalement sur** A si :

i) pour tout $n \in \mathbb{N}$, la fonction f_n est bornée ; et

ii) la série $\sum \|f_n\|_\infty$ est convergente.

Exemple 7.

La série $\sum \dfrac{\cos(n^3x)}{(n+1)^2}$ converge normalement sur $\mathbb{R}$.

En effet, on a, pour $f_n : x \mapsto \dfrac{\cos(n^3x)}{(n+1)^2}$:

$$\|f_n\|_\infty = \sup_{x\in\mathbb{R}}\left(\left|\frac{\cos(n^3x)}{(n+1)^2}\right|\right) \leq \frac{1}{(n+1)^2}$$

Or d'après le critère de Riemann, $\frac{1}{(n+1)^2}$ est le terme général d'une série convergente. Ainsi, $\sum f_n$ converge normalement.

Proposition 7.

Soit $\sum f_n$ une série de fonctions de A dans F. Si $\sum f_n$ converge normalement sur A, alors $\sum f_n$ converge uniformément sur A.

Démonstration.

Si $\sum f_n$ converge normalement, alors les f_n sont bornés et $\sum \|f_n\|_\infty$ converge. La convergence de $\sum \|f_n\|_\infty$ revient à la convergence absolue de $\sum f_n$. Par suite, $\sum f_n$ converge dans $\mathcal{F}_b(A, F)$ vers f et donc $\sum f_n$ converge uniformément vers f sur A. $\square$

Remarque 2.

Attention, la réciproque est fausse : chercher parmi les exemples précédents une série de fonctions qui converge uniformément mais pas normalement.

On a ainsi, pour une **série** de fonctions, le diagramme suivant :

$$\textbf{CVN} \underset{\not\Leftarrow}{\Rightarrow} \textbf{CVU} \underset{\not\Leftarrow}{\Rightarrow} \textbf{CVS}$$

Proposition 8.

Soit $\sum f_n$ une série de fonctions de A dans F. Si $\sum f_n$ converge normalement sur A, alors, pour tout $x \in A$, la série de vecteurs $\sum f_n(x)$ est absolument convergente, i.e. $\sum \|f_n(x)\|_F$ converge.

Démonstration.

Si $\sum f_n$ converge normalement, $\|f_n\|_\infty = \sup_{x\in A} \|f_n(x)\|_F$ est le terme général d'une série convergente. Soit $x \in A$. Comme pour tout $n \in \mathbb{N}$, on a :

$$\|f_n(x)\|_F \leq \|f_n\|_\infty,$$

par comparaison, la série à terme positifs $\sum \|f_n(x)\|_F$ est convergente i.e. $\sum f_n(x)$ est absolument convergente. □

Méthode : Montrer qu'une série de fonctions converge normalement

Pour montrer la convergence normale de $\sum f_n$, on cherche à obtenir une majoration de $\|f_n\|_\infty$ qui tende vers 0 i.e. une majoration **indépendante de** $x \in A$ du type (à partir d'un certain rang)

$$\|f_n(x)\|_F \leq u_n$$

telle que $\sum u_n$ est une série numérique convergente (**et** u_n **ne dépend pas de** $x \in A$ **!!!**).

Exercice 11.

1. Étudier la convergence normale/uniforme/simple de $\sum \dfrac{e^{-nx}}{n^2}$ sur $\mathbb{R}_+$.
2. Étudier la convergence normale/uniforme/simple de $\sum_{n\geq 1} \dfrac{1}{n^2+x^2}$ puis de $\sum_{n\geq 0} \dfrac{1}{n^2+x^2}$ sur $\mathbb{R}$.

Correction.

1. On pose, pour $n \in \mathbb{N}^*$, $f_n : x \mapsto \frac{e^{-nx}}{n^2}$. Commençons par la convergence normale :
   CVN sur $\mathbb{R}_+$: Soit $n \in \mathbb{N}^*$. On a, pour tout $x \in \mathbb{R}_+$:
   $$|f_n(x)| = \frac{e^{-nx}}{n^2} \leq \frac{1}{n^2}.$$
   Donc f_n est bornée sur $\mathbb{R}_+$ et l'inégalité précédente étant une égalité pour $x = 0$, on a
   $$\|f_n\|_\infty = \frac{1}{n^2}$$
   Or $\frac{1}{n^2}$ est le terme général d'une série convergente d'après le critère de Riemann $(2 > 1)$; par suite, $\sum \|f_n\|_\infty$ converge et donc $\sum f_n$ converge normalement sur $\mathbb{R}_+$.
   Comme pour une série de fonctions, CVN $\Rightarrow$ CVU $\Rightarrow$ CVS, la série $\sum f_n$ converge uniformément et simplement sur $\mathbb{R}_+$.

2. • On pose, pour $n \in \mathbb{N}^*$, $f_n : x \mapsto \frac{1}{n^2+x^2}$ qui est bien définie sur $\mathbb{R}$.
CVN sur $\mathbb{R}$: Soit $n \in \mathbb{N}^*$. On a, pour tout $x \in \mathbb{R}$:

$$|f_n(x)| = \frac{1}{n^2+x^2} \leq \frac{1}{n^2}.$$

Donc f_n est bornée sur $\mathbb{R}$ et l'inégalité précédente étant une égalité pour $x = 0$, on a :

$$\|f_n\|_\infty = \frac{1}{n^2}$$

Or $\frac{1}{n^2}$ est le terme général d'une série convergente d'après le critère de Riemann $(2 > 1)$; par suite, $\sum \|f_n\|_\infty$ converge et donc $\sum f_n$ converge normalement sur $\mathbb{R}$.
Comme pour une série de fonctions, CVN $\Rightarrow$ CVU $\Rightarrow$ CVS, la série $\sum f_n$ converge uniformément et simplement sur $\mathbb{R}$.

• On garde les mêmes notations que précédemmennt et on pose $f_0 : x \mapsto \frac{1}{x^2}$ qui est définie sur $\mathbb{R}^*$. On ne peut donc plus étudier les convergences sur $\mathbb{R}$ de $\sum_{n \geq 0} f_n$ du fait que f_0 n'est pas définie en 0 ! On doit donc faire notre étude sur les intervalles $\mathbb{R}_-^*$ et $\mathbb{R}_+^*$. Les f_n étant tous des fonctions paires, on peut restreindre l'étude à $\mathbb{R}_+^*$.
CVN sur $\mathbb{R}_+^*$: Soit $n \in \mathbb{N}$. On remarque que f_0 n'est pas bornée sur $\mathbb{R}_+^*$ car $f_0(x) \xrightarrow[x \to 0^+]{} +\infty$ donc sa norme infinie n'est pas définie et donc l'étude de la convergence normale n'est pas possible même si cela semblait de prime abord être exactement comme le cas précédent.
Il n'y a donc pas convergence normale sur $\mathbb{R}_+^*$ à cause d'un "détail" mais on se doute que la convergence uniforme va sûrement fonctionner puisque dans les restes, f_0 n'apparaîtra pas !

CVU sur $\mathbb{R}_+^*$: Soit $n \in \mathbb{N}$. Pour tout $x \in \mathbb{R}_+^*$, on a :

$$|R_n(x)| = \sum_{k=n+1}^{+\infty} \frac{1}{n^2+x^2} \leq \sum_{k=n+1}^{+\infty} \frac{1}{n^2} = r_n$$

Or la série $\sum_{n \geq 1} \frac{1}{n^2}$ converge donc la suite $(r_n)_{n \in \mathbb{N}^*}$ de ses restes converge vers 0. Ainsi, R_n est une fonction bornée et on a :

$$\|R_n\|_\infty \leq r_n \xrightarrow[n \to +\infty]{} 0.$$

Par suite, $\sum_{n \geq 0} f_n$ converge uniformément sur $\mathbb{R}_+^*$ (et donc également sur $\mathbb{R}_-^*$. Comme CVU $\Rightarrow$ CVS, on a la convergence simple de $\sum_{n \geq 0} f_n$ sur les mêmes intervalles.

Exercice 12. *Fonction Zêta de Riemann*

La fonction ζ de Riemann est définie, pour $s \in]1, +\infty[$ par :

$$\zeta(s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}.$$

Justifier la définition de la fonction ζ sur $]1, +\infty[$ en étudiant la convergence simple de la série de fonctions. Que dire de la convergence uniforme de la série vers la fonction ζ ?

Correction.

Pour $n \in \mathbb{N}^*$, on pose $f_n : s \mapsto \frac{1}{n^s}$ qui est bien définie sur $\mathbb{R}$.

- Déterminons le domaine de la fonction ζ. On remarque que $\zeta(s)$ existe si, et seulement si, la série numérique $\sum f_n(s)$ est convergente. On se ramène donc à trouver le plus grand intervalle sur lequel la série de fonction $\sum f_n$ converge simplement.

  CVS sur $\mathbb{R}$: Soit $s \in \mathbb{R}$. On a, d'après le critère de Riemann,
  $f_n(s) = \frac{1}{n^s}$ est le terme général d'une série convergente si, et seulement si, $s > 1$.
  Ainsi, la série $\sum f_n$ converge simplement sur $]1, +\infty[$ et ne converge pas simplement sur $]-\infty, 1]$.
  Il en résulte que le domaine de définition de ζ est $]1, +\infty[$.

  > **Remarque :** En prenant s complexe, on peut, de la même manière que précédemment, montrer que ζ est bien définie sur $\mathrm{DP} = \{s \in \mathbb{C} \mid \mathrm{Re}(s) > 1\}$. On montrera dans la suite que ζ est de classe C^∞ sur $]1, +\infty[$ et de manière similaire, on peut montrer que ζ est *holomorphe* (une généralisation de la dérivabilité réelle pour les complexes) sur DP.
  > Par des méthodes que nous ne détaillerons pas ici, on peut montrer que la fonction ζ peut être prolongée de manière *analytique* (nous parlerons de cela quand nous aborderons les développements en séries entières) et donc en particulier, ζ admet un prologement continu sur $\mathbb{C} \setminus \{1\}$.
  > Ce prolongement (dont on peut prouver qu'il est unique) donne la valeur :
  >
  > $$\zeta(-1) = -\frac{1}{12}$$
  >
  > d'où la célèbre confusion : $\displaystyle\sum_{n=1}^{+\infty} n = -\frac{1}{12}$!
  > Mais il ne faut pas s'y méprendre ! Comme nous l'avons définie, la série à termes positifs $\sum n$ diverge et donc tend vers $+\infty$. La formule $\sum_{n=1}^{+\infty} \frac{1}{n^s}$ de la fonction ζ n'est plus valable dans le sens "somme de série numérique" pour les complexes s en dehors de DP et il faut utiliser une autre formule - une de celles qui permettent de définir le prolongement - pour obtenir la valeur de $\zeta(s)$.

- On essaye la convergence normale dans un premier temps :

  CVN sur $]1, +\infty[$: Soit $n \in \mathbb{N}^*$. On a :

  $$\|f_n\|_\infty = \sup_{s \in]1,+\infty[} \left(\frac{1}{n^s}\right) = \frac{1}{n}$$

  qui le terme général d'une série divergente, donc il n'y a pas convergence normale sur $]1, +\infty[$.
  On voit que le problème se situe en 1 puisque le "$\frac{1}{n}$" est atteint en $s = 1$. On tente donc d'isoler ce problème :
  soit $a > 1$.

  CVN sur $[a, +\infty[$: Soit $n \in \mathbb{N}^*$. On a :

  $$\|f_n\|_\infty = \sup_{s \in [a,+\infty[} \left(\frac{1}{n^s}\right) = \frac{1}{n^a}$$

  qui est le terme général d'une série convergente d'après le critère de Riemann car $a > 1$.
  Ainsi, $\sum f_n$ converge normalement et donc uniformément vers ζ sur tout intervalle de la

forme $[a, +\infty[$ pour $a > 1$. Nous verrons dans la partie suivante que cela nous satisfait amplement pour en déduire la continuité de ζ sur $]1, +\infty[$.

Mais pour répondre à la question initiale, il nous reste à étudier la convergence uniforme sur $]1, +\infty[$:

CVU sur $]1, +\infty[$ **:** Soit $n \in \mathbb{N}^*$. On a, pour tout $s > 1$:

$$|R_n(s)| = \sum_{k=n+1}^{+\infty} \frac{1}{k^s}.$$

Or, par une comparaison série-intégrale que nous avons effectuée au chapitre précédent, on a :

$$\sum_{k=n+1}^{+\infty} \frac{1}{k^s} \geq \int_{n+1}^{+\infty} \frac{1}{t^s} dt = \frac{1}{s-1} \frac{1}{(n+1)^{s-1}} \xrightarrow[s \to 1^+]{} +\infty \times 1 = +\infty$$

Donc, d'après le théorème des gendarmes, $|R_n(s)| \xrightarrow[s \to 1^+]{} +\infty$ et ainsi, R_n n'est pas bornée sur $]1, +\infty[$.

Ainsi, $\sum f_n$ ne converge pas uniformément sur $]1, +\infty[$.

Partie B

Convergence uniforme : continuité, limites, intégration et dérivation

1. Continuité

Proposition 9.

Soit $a \in A$, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de A dans F. Si la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A et, pour tout $n \in \mathbb{N}$, f_n est continue en a, alors f est continue en a.

Démonstration.

Pour tout $x \in A$, on a pour tout $n \in \mathbb{N}$.

$$\|f(x) - f(a)\|_F \leq \|f(x) - f_n(x)\|_F + \|f_n(x) - f_n(a)\|_F + \|f_n(a) - f(a)\|_F.$$

Soit $\varepsilon > 0$.

- Comme $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A, alors il existe $N \in \mathbb{N}$ tel que pour tout $x \in A$ et pour tout $n \geq N$:

$$\|f(x) - f_n(x)\|_F \leq \|f_n - f\|_\infty \leq \frac{\varepsilon}{3}.$$

  et c'est donc en particulier vrai pour $x = a$ également :

$$\|f(a) - f_n(a)\|_F \leq \|f_n - f\|_\infty \leq \frac{\varepsilon}{3}.$$

- Soit $n \geq N$. Comme f_n est continue en a, alors il existe $\delta > 0$ tel que pour tout $x \in A$ avec $\|x - a\|_E \leq \delta$:

$$\|f_n(x) - f_n(a)\|_F \leq \|f_n - f\|_\infty \leq \frac{\varepsilon}{3}.$$

Ainsi, en choisissant $n \geq N$, on exhibe un δ tel que pour tout $x \in A$ avec $\|x - a\|_E \leq \delta$,

$$\|f(x) - f(a)\|_F \leq \underbrace{\|f(x) - f_n(x)\|_F}_{\leq \frac{\varepsilon}{3}} + \underbrace{\|f_n(x) - f_n(a)\|_F}_{\leq \frac{\varepsilon}{3}} + \underbrace{\|f_n(a) - f(a)\|_F}_{\leq \frac{\varepsilon}{3}} \leq \varepsilon.$$

Par suite, f est continue en a. $\square$

Corollaire 1.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de A dans F. Si la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur A et, pour tout $n \in \mathbb{N}$, f_n est continue sur A, alors f est continue sur

A.

Proposition 10. *Cas particulier des séries*

Soit $a \in A$, $\sum f_n$ une série de fonctions et S une fonction de A dans F. Si la série $\sum f_n$ converge uniformément vers S sur A et, pour tout $n \in \mathbb{N}$, f_n est continue en a (resp. sur A), alors S est continue en a (resp. sur A).

Remarque 3.

- Ainsi, on peut, après avoir établi la convergence simple d'une suite/série de fonctions continues, conclure directement à la non convergence uniforme si la fonction limite n'est pas continue !
- Comme la convergence normale d'une série de fonctions implique la convergence uniforme, une série de fonctions continues qui converge normalement converge vers une fonction S continue.

Théorème 1.

Soit $a \in A$, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de A dans F. Si la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur un voisinage de a et, pour tout $n \in \mathbb{N}$, f_n est continue en a, alors f est continue en a.

Soit $a \in A$, $\sum f_n$ une série de fonctions et S une fonction de A dans F. Si la série $\sum f_n$ converge uniformément vers S sur un voisinage de a et, pour tout $n \in \mathbb{N}$, f_n est continue en a, alors S est continue en a.

Corollaire 2.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions et f une fonction de A dans F. Si la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f au voisinage de tout point de A et, pour tout $n \in \mathbb{N}$, f_n est continue sur A, alors f est continue sur A.

Soit $\sum f_n$ une série de fonctions et S une fonction de A dans F. Si la série $\sum f_n$ converge uniformément vers S au voisinage de tout point de A et, pour tout $n \in \mathbb{N}$, f_n est continue sur A, alors S est continue sur A.

Exemple 8.

La série de fonction $\sum \frac{x^n}{n!}$ converge uniformément sur tout segment de $\mathbb{R}$ vers la fonction S : $x \mapsto \frac{x^n}{n!}$. Ainsi, la fonction S est continue sur $\mathbb{R}$ car $\sum \frac{x^n}{n!}$ converge uniformément au voisinage de tout point de $\mathbb{R}$.

Exercice 13.

1. Justifier que la fonction $S : x \mapsto \sum_{n=0}^{+\infty} \dfrac{1}{n^2+x^2}$ est bien définie sur $]0,+\infty[$ et qu'elle est continue sur cet intervalle.
2. Montrer que la fonction ζ de Riemann est continue sur $]1,+\infty[$.

Correction.

2. On pose, pour $n \in \mathbb{N}$ et $x \in \mathbb{R}$, $f_n(x) = \dfrac{1}{n^x}$. Montrons que $\zeta : x \mapsto \sum_{n=0}^{+\infty} f_n(x)$ est définie et continue sur $]1,+\infty[$.
   — Domaine de définition de ζ.
   Soit $x \in \mathbb{R}$. Alors, d'après le critère de Riemann, $f_n(x) = \dfrac{1}{n^x}$ est le terme général d'une série convergente si, et seulement si, $x > 1$.
   Ainsi, $\sum f_n$ converge simplement sur $]1,+\infty[$ d'où ζ est définie sur $]1,+\infty[$.
   — On applique le théorème de continuité de la somme d'une série de fonctions :
   - les f_n sont continues sur $]1,+\infty[$ car $f_n : x \mapsto \dfrac{1}{n^x} = e^{-x\ln(n)}$ est une fonction continue sur $\mathbb{R}$.
   - CVU sur tout segment de $\mathbb{R}$:
     Soit $a > 1$. On établit la Convergence Normale sur $[a,+\infty[$.
     Soit $n \in \mathbb{N}$. pour tout $x \in [a,+\infty[$,
     $$|f_n(x)| = \frac{1}{n^x} \leq \frac{1}{n^a}$$
     donc $\|f_n\|_\infty \leq \dfrac{1}{n^a}$ (c'est même une égalité) qui est le terme général d'une série convergente d'après le critère de Riemann.
     Ainsi, par comparaison, $\sum \|f_n\|_\infty$ converge. Il en résulte que, pour tout $a > 1$, $\sum f_n$ converge normalement sur $[a,+\infty[$, et donc que $\sum f_n$ converge uniformément sur tout segment de $]1,+\infty[$.

   Par suite, d'après le théorème de continuité de la somme d'une série de fonction, ζ est continue sur $]1,+\infty[$

2. Limites

a. Interversion de limites

Théorème 2. *Théorème de la double limite*

Soit $a \in \overline{A}$ et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions qui converge uniformément vers une fonction f sur A. Si, pour tout $n \in \mathbb{N}$, il existe $\ell_n \in F$ tel que $f_n(x) \xrightarrow[x\to a]{} \ell_n$ alors :

i) il existe $\ell \in F$ tel que la suite $(\ell_n)_{n\in\mathbb{N}}$ converge dans F vers ℓ et,

ii) $f(x) \xrightarrow[x\to a]{} \ell$.

Autrement dit,
$$\lim_{x\to a} f(x) = \lim_{x\to a}\left(\lim_{n\to+\infty} f_n(x)\right) = \lim_{n\to+\infty}\left(\lim_{x\to a} f_n(x)\right).$$

Remarque 4.

Ce théorème reste valable dans les cas suivants :
- — $A\subset\mathbb{R}$ et $a=\pm\infty$;
- — $F=\mathbb{R}$; à partie d'un certain rang, $\ell_n=\lim_{x\to a} f_n(x)=\pm\infty$; et $\ell=\pm\infty$.

b. Cas des séries

Théorème 3. *Inversion limite/somme*

Soit $a\in\overline{A}$ et $\sum f_n$ une série de fonctions qui converge uniformément sur A. Si, pour tout $n\in\mathbb{N}$, il existe $\ell_n\in F$ tel que $f_n(x)\xrightarrow[x\to a]{}\ell_n$ alors la série $\sum l_n$ converge et on a :
$$\lim_{x\to a}\sum_{n=0}^{+\infty} f_n(x) = \sum_{n=0}^{+\infty}\left(\lim_{x\to a} f_n(x)\right).$$

Remarque 5.

Comme pour les suites de fonctions, le théorème précédent reste valable dans les cas des limites "infinies".

Exercice 14.

1. Justifier l'existence et déterminer $\lim_{s\to+\infty}\zeta(s)$.
2. Justifier l'existence et déterminer $\lim_{x\to0^+} x\sum_{n=0}\dfrac{1}{n^2+x}$.

3. Intégration

Dans ce paragraphe, les fonctions considérées sont définies sur un intervalle d'intérieur non vide I de $\mathbb{R}$ et sont à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$; de plus, a,b sont des éléments de I tels que $a<b$.

a. Intégration sur un segment

On rappelle la notation suivante :

Notation 2. *Norme de la convergence en moyenne*

Soit $C([a,b],\mathbb{K})$ l'espace vectoriel sur $\mathbb{K}$ des fonctions continues de $[a,b]$ dans $\mathbb{K}$. On note $\|\cdot\|_1$ la **norme de la convergence en moyenne** sur $C([a,b],\mathbb{K})$ i.e. pour $f \in C([a,b],\mathbb{K})$:

$$\|f\|_1 = \int_a^b |f(t)|dt.$$

Proposition 11.

Soit (f_n) une suite de fonctions continues et f une fonction continue de $[a,b]$ dans $\mathbb{K}$. Si $(f_n)_{n\in\mathbb{N}}$ converge en moyenne vers f i.e. $\|f_n - f\|_1 \xrightarrow[n\to+\infty]{} 0$, alors

$$\lim_{n\to+\infty} \int_a^b f_n(t)dt = \int_a^b f(t)dt.$$

Démonstration.

On a :

$$\begin{aligned} \left|\int_a^b f_n(t)dt - \int_a^b f(t)dt\right| &= \left|\int_a^b (f_n(t) - f(t))\,dt\right| \\ &\leq \int_a^b |f_n(t) - f(t)|\,dt \\ &= \|f_n - f\|_1 \xrightarrow[n\to+\infty]{} 0. \end{aligned}$$

□

Lemme 1.

Sur $C([a,b],\mathbb{K})$, la norme de la convergence en moyenne est dominée par la norme de la convergence uniforme ; plus précisément, on a, pour tout $f \in C([a,b],\mathbb{K})$:

$$\|f\|_1 \leq (b-a)\|f\|_\infty.$$

Démonstration.

Soit $f \in C([a,b],\mathbb{K})$. On a :

$$\int_a^b \underbrace{|f(t)|}_{\leq \sup_{t\in[a,b]}|f(t)|} dt \leq \|f\|_\infty \int_a^b dt = (b-a)\|f\|_\infty.$$

□

Théorème 4. *Interversion limite/intégrale*

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues et f une fonction de $[a,b]$ dans $\mathbb{K}$. Si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[a,b]$ alors $f \in C([a,b],\mathbb{K})$ et

$$\lim_{n\to+\infty}\int_a^b f_n(t)dt = \int_a^b \lim_{n\to+\infty} f_n(t)dt = \int_a^b f(t)dt.$$

Démonstration.

On suppose que $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur $[a,b]$. Alors, comme les (f_n) sont continues sur $[a,b]$, alors f est continue sur $[a,b]$. Ainsi, pour tout $n\in\mathbb{N}$, $f_n - f \in C([a,b],\mathbb{K})$ et on a, d'après le lemme précédent :

$$\|f_n - f\|_1 \leq (b-a)\|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0.$$

Donc (f_n) converge en moyenne vers f. Par suite, d'après la proposition précédente, on a :

$$\lim_{n\to+\infty}\int_a^b f_n(t)dt = \int_a^b f(t)dt.$$

□

Exercice 15.

Déterminer $\lim_{n\to+\infty}\int_0^1 n\sin^n(x)dx$.

b. Intégration des séries de fonctions

Théorème 5. *Interversion intégrale/somme*

Soit $\sum f_n$ une série de fonctions continues de $[a,b]$ dans $\mathbb{K}$. Si $\sum f_n$ converge uniformément sur $[a,b]$ vers sa somme S, alors on a :

$$\sum_{n=0}^{+\infty}\left(\int_a^b f_n(t)dt\right) = \int_a^b S(t)dt = \int_a^b \left(\sum_{n=0}^{+\infty} f_n(t)\right)dt.$$

Exercice 16.

Montrer que $f : x \mapsto \sum_{n=1}^{+\infty}\dfrac{\sin(nx)}{n^2}$ est définie et continue sur $[0,\pi]$, puis démontrer que $\int_0^\pi f(t)dt = \frac{7}{4}\zeta(3)$.

c. Primitives

Théorème 6.

Soit $a \in I$, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues et f une fonction de I dans $\mathbb{K}$. Si $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur tout segment de I, alors pour tout $x \in I$:

$$\lim_{n\to+\infty} \int_a^x f_n(t)dt = \int_a^x \lim_{n\to+\infty} f_n(t)dt = \int_a^x f(t)dt.$$

Soit $a \in I$ et $\sum f_n$ une série de fonctions continues de I dans $\mathbb{K}$. Si $\sum f_n$ converge uniformément sur tout segment de I vers sa somme S, alors on a :

$$\sum_{n=0}^{+\infty} \left(\int_a^x f_n(t)dt \right) = \int_a^x S(t)dt = \int_a^x \left(\sum_{n=0}^{+\infty} f_n(t) \right) dt.$$

Exercice 17.

Montrer que pour tout $x \in]-1, 1[$, $\sum_{n=1}^{+\infty} \dfrac{x^n}{n} = -\ln(1-x)$.

4. Dérivation

Dans ce paragraphe, les fonctions considérées sont définies sur un intervalle d'intérieur non vide I de $\mathbb{R}$ et sont à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$; de plus, a, b sont des éléments de I tels que $a < b$.

a. Dérivation des suites de fonctions

Théorème 7. *Interversion dérivation/limite*

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de classe C^1 sur I et f, g des fonctions de I dans $\mathbb{K}$. Si

i) la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f sur I et ;

ii) la suite $(f_n')_{n\in\mathbb{N}}$ converge uniformément vers g sur tout segment de I.

Alors la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers f sur tout segment de I et la fonction f est de classe C^1 sur I où on a :

$$f' = \left(\lim_{n\to+\infty} f_n \right)' = g = \lim_{n\to+\infty} f_n'.$$

Corollaire 3. *Cas des fonctions de classe C^k*

Soit $k \in \mathbb{N}^*$, $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de I dans $\mathbb{K}$ de classe C^k sur I. Si

i) pour tout $i \in [\![0, k-1]\!]$, la suite $(f_n^{(i)})_{n\in\mathbb{N}}$ converge simplement vers f sur I et ;

ii) la suite $(f_n^{(k)})_{n\in\mathbb{N}}$ converge uniformément sur tout segment de I.

Alors la limite simple f de la suite $(f_n)_{n\in\mathbb{N}}$ est de classe C^k sur I et, pour tout $k \in [\![0, k]\!]$, on a $f^{(k)} = \lim_{n\to+\infty} f_n^{(k)}$.

b. Dérivation des séries de fonctions

Théorème 8. *Interversion dérivation/somme*

Soit $\sum f_n$ une série de fonctions de I dans $\mathbb{K}$ de classe C^1 sur I. Si

i) la série $\sum f_n$ converge simplement vers sa somme S sur I et ;

ii) la série $\sum f_n'$ converge uniformément sur tout segment de I.

Alors la série $\sum f_n$ converge uniformément vers S sur tout segment de I et la fonction S est de classe C^1 sur I où on a :

$$\left(\sum_{n=0}^{+\infty} f_n\right)' = S' = \sum_{n=0}^{+\infty} f_n'.$$

Exemple 9.

La série de fonctions $\sum f_n$ de terme général $f_n : x \mapsto \frac{x^n}{n!}$ est C^1 de somme S sur $\mathbb{R}$ et on a $S' = S$ et ainsi $S = \exp$ car $S(0) = 1$.

On applique le théorème précédent :

- les f_n sont de classe C^1 sur $\mathbb{R}$ car ce sont des fonctions polynomiales et on a, pour tout $x \in \mathbb{R}$:
$$f_n'(x) = \begin{cases} 0 & \text{si } n = 0 \\ \dfrac{x^{n-1}}{(n-1)!} & \text{si } n > 0. \end{cases}$$
- CVS sur $\mathbb{R}$: Soit $x \in \mathbb{R}$. En appliquant la règle de D'Alembert, on montre que $\dfrac{x^n}{n!}$ est le terme général d'une série numérique absolument convergente et donc convergente.
Ainsi, $\sum f_n$ converge simplement sur $\mathbb{R}$.
- CVU sur tout segment de $\mathbb{R}$ de la série des dérivées :
Soit $a > 0$. On établit la Convergence Normale sur $[-a, a]$.
Soit $n \in \mathbb{N}$. On a $f_0' = 0$ et pour tout $x \in [-a, a]$,
$$|f_n'(x)| = \frac{|x|^{n-1}}{(n-1)!} \leq \frac{a^{n-1}}{(n-1)!}$$
donc $\|f_n'\|_\infty \leq \dfrac{a^{n-1}}{(n-1)!}$ (c'est même une égalité) qui est le terme général d'une série convergente d'après, par exemple, la règle de d'Alembert.
Ainsi, par comparaison, $\sum \|f_n'\|_\infty$ converge. Il en résulte que, pour tout $a > 0$, $\sum f_n'$ converge normalement sur $[-a, a]$, et donc que $\sum f_n'$ converge uniformément sur tout segment de $\mathbb{R}$.

Par suite, d'après le théorème d'interversion dérivation/somme, S est de classe C^1 sur $\mathbb{R}$ et on a pour tout $x \in \mathbb{R}$:
$$S'(x) = \frac{d}{dx} \sum_{n=0}^{+\infty} \frac{x^n}{n!} = \sum_{n=0}^{+\infty} \frac{d}{dx} \frac{x^n}{n!} = \sum_{n=1}^{+\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{+\infty} \frac{x^n}{n!} = S(x).$$

Remarque : Ainsi, comme $S(0) = 1$, S est solution du problème de Cauchy :
$$\begin{cases} y' = y \\ y(0) = 1 \end{cases}$$
dont l'unique solution est la fonction $x \mapsto e^x$.
On a donc montré l'égalité qu'il faut *absolument* connaître :
$$\forall\, x \in \mathbb{R}, \quad e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

Corollaire 4. *Cas des séries de fonctions de classe C^k*

Soit $k \in \mathbb{N}^*$, $\sum f_n$ une série de fonctions de I dans $\mathbb{K}$ de classe C^k sur I. Si

i) pour tout $i \in [\![0, k-1]\!]$, la série $\sum f_n^{(i)}$ converge simplement vers f sur I et ;

ii) la série $\sum f^{(k)}$ converge uniformément sur tout segment de I.
Alors la somme S de la série $\sum f_n$ est de classe C^k sur I et, pour tout $i \in [\![0, k]\!]$, on a

$$S^{(i)} = \left(\sum_{n=0}^{+\infty} f_n\right)^{(i)} = \sum_{n=0}^{+\infty} f_n^{(i)}.$$

c. Suites et séries de fonctions de classe C^∞

Théorème 9. *Cas des fonctions de classe C^∞*

- **Théorème d'interversion dérivation /limite**
  Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de I dans $\mathbb{K}$ de classe C^∞ sur I. Si,
  i) la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur I, et
  ii) pour tout $k \in \mathbb{N}^*$, la suite $(f_n^{(k)})_{n\in\mathbb{N}}$ converge uniformément sur tout segment de I,
  alors la limite simple f de la suite $(f_n)_{n\in\mathbb{N}}$ est de classe C^∞ sur I et, pour tout $k \in \mathbb{N}$, on a $f^{(k)} = (\lim_{n\to+\infty} f_n)^{(k)} = \lim_{n\to+\infty} f_n^{(k)}$.
- **Théorème d'interversion dérivation /somme**
  Soit $\sum f_n$ une série de fonctions de I dans $\mathbb{K}$ de classe C^∞ sur I. Si,
  i) la série $\sum f_n^{(k)}$ converge simplement sur I, et
  ii) pour tout $k \in \mathbb{N}^*$, la série $\sum f_n^{(k)}$ converge uniformément sur tout segment de I,
  alors la somme S de la série $\sum f_n$ est de classe C^∞ sur I et, pour tout $k \in \mathbb{N}$, on a :

$$S^{(k)} = \left(\sum_{n=0}^{+\infty} f_n\right)^{(k)} = \sum_{n=0}^{+\infty} f_n^{(k)}.$$

Exemple 10.

La série de fonctions $\sum f_n$ de terme général $f_n : x \mapsto \frac{x^n}{n!}$ est C^∞ de somme S sur $\mathbb{R}$ et on pour tout $k \in \mathbb{N}$, $S^{(k)} = S$.

Exercice 18.

1. Montrer que la série de fonctions $\sum f_n$ de terme général $f_n : x \mapsto x^n$ est C^∞ sur $]-1, 1[$. Pour $k \in \mathbb{N}$, en déduire une expression, pour tout $x \in]-1, 1[$, de :
$$\sum_{n=0}^{+\infty} \binom{k+n}{k} x^n.$$
2. Montrer que la fonction ζ est C^∞ sur $]1, +\infty[$ et pour $k \in \mathbb{N}$ déterminer $\zeta^{(k)}$.
3. Montrer que la fonction $\varphi : x \mapsto \sum_{n=0}^{+\infty} \dfrac{e^{-nx}}{n^2+1}$ est continue sur $\mathbb{R}_+$ et C^∞ sur $\mathbb{R}_+^*$. Montrer que f est croissante et déterminer $\lim_{x\to 0^+} f(x)$. En déduire que f n'est pas dérivable en 0.

Correction.

2. Montrons que $\zeta : x \mapsto \sum_{n=1}^{+\infty} \dfrac{1}{n^x}$ est de classe C^∞ sur $]1,+\infty[$. Pour cela, on vérifie les hypothèses du théorème d'interversion dérivation/somme (version C^∞) :

   i) Pour tout $n \in \mathbb{N}^*$, $f_n : x \mapsto \frac{1}{n^x}$ est de classe C^∞ sur $]1,+\infty[$;

   ii) pour tout $k \in \mathbb{N}$, la série $\sum f_n^{(k)}$ converge uniformément sur tout segment de $]1,+\infty[$.

   i) Soit $n \in \mathbb{N}^*$. On a, pour $x \in]1,+\infty[$, $f_n(x) = \frac{1}{n^x} = e^{-x\ln(n)}$. Ainsi, f_n est de classe C^∞ sur $\mathbb{R}$ comme composée de la fonction exp et $x \mapsto -x\ln(n)$ qui sont C^∞ sur $\mathbb{R}$. Ainsi, en particulier, f_n est de classe C^∞ sur $]1,+\infty[$ et on a, pour $k \in \mathbb{N}$ et $x \in]1,+\infty[$:

$$f_n^{(k)} = (-1)^k \ln(n)^k e^{-x\ln(n)} = \frac{(-1)^k \ln(n)^k}{n^x}.$$

   ii) Soit $k \in \mathbb{N}$. Soit $a > 1$. On étudie la convergence normale de $\sum f_n^{(k)}$ sur $[a,+\infty[$. Soit $n \in \mathbb{N}$. On a :

$$\|f_n^{(k)}\|_\infty = \sup_{x\in[a,+\infty[} \left|\frac{(-1)^k \ln(n)^k}{n^x}\right| \leq \frac{\ln(n)^k}{n^a}.$$

   — Pour $k = 0$, $\dfrac{\ln(n)^k}{n^a} = \dfrac{1}{n^a}$ est le terme général d'une série convergente d'après le critère de Riemann car $a > 1$. Ainsi, $\sum f_n$ converge normalement sur $[a,+\infty[$.

   — Soit $k \in \mathbb{N}^*$. Pour tout $\beta > 0$, $\ln(n) = o(n^\beta)$; ainsi, pour $\beta = \dfrac{a-1}{2k} > 0$, on a :

$$\frac{\ln(n)^k}{n^a} = o\left(\frac{1}{n^{a-k\beta}}\right)$$

   Or $a - k\beta = a - \frac{a-1}{2} = \dfrac{a+1}{2} > 1$ car $a > 1$; par suite, d'après le critère de Riemann, $\dfrac{1}{n^{a-k\beta}}$ est le terme général d'une série convergente et donc $\dfrac{\ln(n)^k}{n^a}$ l'est aussi par comparaison. Ainsi, $\sum f_n^{(k)}$ converge normalement sur $[a,+\infty[$

   Il en résulte que, pour tout $k \in \mathbb{N}$, $\sum f_n^{(k)}$ converge uniformément sur $[a,+\infty$ pour tout $a > 1$ et donc sur tout segment de $]1,+\infty[$.
   Il en résulte que, par le théorème d'interversion dérivation/somme, ζ est de classe C^∞ sur $]1,+\infty[$, et on a, pour tout $k \in \mathbb{N}$ et tout $x \in]1,+\infty[$:

$$\zeta^{(k)}(x) = \left(\sum_{n=1}^{+\infty} f_n\right)^{(k)}(x) = \sum_{n=1}^{+\infty} f_n^{(k)}(x) = (-1)^k \sum_{n=1}^{+\infty} \frac{\ln(n)^k}{n^x}.$$

Partie C
Approximation uniforme

Dans cette partie, les fonctions considérées sont définies sur un intervalle I de $\mathbb{R}$ d'intérieur non vide et sont à valeurs dans un espace vectoriel F sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ de dimension finie. De plus, a, b désignent deux réels de I tels que $a < b$.

1. Définitions

Définition 8. *Subdivision d'un segment*

Soit $n \in \mathbb{N}^*$ et $\sigma = (a_i)_{i \in [\![0,n]\!]}$ une famille finie à valeurs dans $[a, b]$. On dit que σ est une **subdivision du segment** $[a, b]$ si

$$a = a_0 < a_1 < \ldots < a_{n-1} < a_n = b.$$

Définition 9. *Fonctions en escalier*

Soit $\varphi : [a, b] \to \mathbb{F}$. On dit que φ est une **fonction en escalier sur** $[a, b]$ s'il existe une subdivision $\sigma = (a_0, \ldots, a_n)$ de $[a, b]$ telle que pour tout $i \in [\![0, n-1]\!]$, $\varphi_{|]a_i, a_{i+1}[}$ est constante.

On note $\mathrm{Esc}([a, b], F)$ l'ensemble des fonctions en escalier sur $[a, b]$.

Définition 10. *Fonctions continues par morceaux*

Soit $f : I \to F$. On dit que f est **continue par morceaux sur le segment** $[a, b]$ s'il existe une subdivision $\sigma = (a_0, \ldots, a_n)$ de $[a, b]$ telle que pour tout $i \in [\![0, n-1]\!]$, $f_{|]a_i, a_{i+1}[}$ est continue et prolongeable par continuité sur $[a_i, a_{i+1}]$.

On dit que f est **continue par morceaux sur l'intervalle** I si f est continue par morceaux sur tout segment de I.

On note $C_{\mathrm{pm}}(I, F)$ l'ensemble des fonctions continues par morceaux sur I.

Exercice 19.

1. Dessiner quelques graphes de fonctions en escalier et continues par morceaux sur un segment.
2. Montrer que $C_{\mathrm{pm}}(I, F)$ est un sous-espace vectoriel de $\mathcal{F}_b(I, F)$.

2. Approximation uniforme par des fonctions en escalier

Théorème 10.

Soit $f \in C_{pm}([a,b],F)$. Alors, pour tout $\varepsilon > 0$, il existe $\varphi \in \mathrm{Esc}([a,b],F)$ telle que $\|f-\varphi\|_\infty \leq \varepsilon$.

Autrement dit, toute fonction continue par morceaux sur $[a,b]$ est limite uniforme d'une suite de fonctions en escalier sur $[a,b]$.
Autrement dit, l'ensemble $\mathrm{Esc}([a,b],F)$ est dense dans $(C_{pm}([a,b],F), \|\cdot\|_\infty)$.

Démonstration.

- On traite tout d'abord le cas f continue sur $[a,b]$. Alors, d'après le théorème de Heine, f est uniformément continue sur $[a,b]$.
  Soit $\varepsilon > 0$. Alors il existe $\delta > 0$ tel que, pour tout $x, y \in [a,b]$ avec $|x-y| \leq \delta$, $\|f(x)-f(y)\|_F \leq \varepsilon$. On construit alors une subdivision de $[a,b]$ de la façon suivante :
  — comme la suite $(\frac{b-a}{n})_{n\in\mathbb{N}^*}$ converge vers 0, il existe $n \in \mathbb{N}^*$ tel que $\frac{b-a}{n} \leq \delta$.
  — pour $i \in [\![0,n]\!]$, on pose $a_i = a + i\frac{b-a}{n}$.
  Ainsi, $\sigma = (a_0,\ldots,a_n)$ est une subdivision de $[a,b]$. On définit alors la fonction en escalier $\varphi : [a,b] \to F$, pour $x \in [a,b]$, par :
  $$\varphi(x) = \begin{cases} f(a) & \text{si } x = a \\ f(a_{i+1}) & \text{si } a_i < x \leq a_{i+1}, i \in [\![0,n-1]\!]. \end{cases}$$
  Pour $x \in [a,b]$, on a l'alternative :
  — $x = a$. Alors $\|f(a)-\varphi(a)\|_F = 0 \leq \varepsilon$.
  — il existe $i \in i \in [\![0,n-1]\!]$ tel que $x \in]a_i,a_{i+1}]$. Alors $|x-a_{i+1}| \leq \delta$, donc :
  $$\|f(x)-\varphi(x)\|_F = \|f(x)-f(a_i)\|_F \leq \varepsilon.$$
  Dans tous les cas, $\|f(x)-\varphi(x)\|_F \leq \varepsilon$. Par suite, $\|f-\varphi\|_\infty \leq \varepsilon$.
- Traitons maintenant le cas général $f \in C_{pm}([a,b],F)$. On considère une subdivision $\sigma = (a_0,\ldots,a_n)$ adaptée à f et on note, pour tout $i \in [\![0,n-1]\!]$, f_i le prolongement par continuité de $f_{|]a_i,a_{i+1}[}$ sur $[a_i,a_{i+1}]$.
  Soit $\varepsilon > 0$. Comme pour chaque $i \in [\![0,n-1]\!]$, f_i est continue sur $[a_i,a_{i+1}]$, en appliquant le point précédent à f_i, on construit $\varphi_i \in \mathrm{Esc}([a,b],F)$ telle que $\|f_i-\varphi_i\|_\infty \leq \varepsilon$.
  On définit alors la fonction en escalier $\varphi : [a,b] \to F$, pour $x \in [a,b]$, par :
  $$\varphi(x) = \begin{cases} f(a_i) & \text{si } x = a_i,\ i \in [\![0,n]\!] \\ \varphi_i(x) & \text{si } a_i < x < a_{i+1},\ i \in [\![0,n-1]\!]. \end{cases}$$
  Pour $x \in [a,b]$, on a l'alternative :
  — il existe $i \in i \in [\![0,n]\!]$ tel que $x = a_i$. Alors $\|f(a_i)-\varphi(a_i)\|_F = 0 \leq \varepsilon$.
  — il existe $i \in i \in [\![0,n-1]\!]$ tel que $x \in]a_i,a_{i+1}[$. Alors, on a :
  $$\|f(x)-\varphi(x)\|_F = \|f_i(x)-\varphi_i(x)\|_F \leq \|f_i-\varphi_i\|_\infty \leq \varepsilon.$$
  Dans tous les cas, $\|f(x)-\varphi(x)\|_F \leq \varepsilon$. Par suite, $\|f-\varphi\|_\infty \leq \varepsilon$.
  Pour démontrer ce point, on pouvait également remarquer que tout fonction continue par morceaux sur $[a,b]$ est la somme d'une fonction continue et d'une fonction en escalier. On conclut alors en approximant, grâce au premier point, cette fonction continue par une fonction en escalier.

□

Exercice 20.

Montrer que toute fonction continue sur le segment $[a,b]$ est limite uniforme d'une suite de fonctions affines par morceaux.

Exercice 21. *Cas particulier du Lemme intégrale de Riemann-Lebesgue*

Soit $f \in \mathbb{C}_{pm}([a,b],\mathbb{C})$. Montrer que $\int_a^b f(t)e^{int}dt \xrightarrow[n\to+\infty]{} 0$.

3. Approximation uniforme par des fonctions polynomiales

Notation 3. *Fonctions polynomiales*

On note $\mathcal{P}([a,b],\mathbb{K})$ l'ensemble des fonctions polynomiales de $\mathbb{R}$ dans $\mathbb{K}$ restreintes au segment $[a,b]$ i.e. si $p \in \mathcal{P}([a,b],\mathbb{K})$, il existe $P \in \mathbb{K}[X]$ tel que, pour tout $x \in [a,b]$, $p(x) = P(x)$.

Théorème 11. *Théorème de Weierstrass*

Soit $f \in C([a,b],\mathbb{K})$. Alors, pour tout $\varepsilon > 0$, il existe $p \in \mathcal{P}([a,b],\mathbb{K})$ telle que $\|f-\varphi\|_\infty \leq \varepsilon$.
Autrement dit, toute fonction continue sur $[a,b]$ est limite uniforme d'une suite de fonctions polynomiales.
Autrement dit, l'ensemble $\mathcal{P}([a,b],\mathbb{K})$ est dense dans $(C([a,b],\mathbb{K}), \|\cdot\|_\infty)$.

Exercice 22.

Soit $f \in C([a,b],\mathbb{R})$. Montrer que si, pour tout $n \in \mathbb{N}$, $\int_0^1 t^n f(t)dt = 0$, alors f est la fonction nulle sur $[a,b]$.

Chapitre VII
Séries entières

Table des matières

Dans ce chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou le corps $\mathbb{C}$ et $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ désignent, sauf mention contraire, des suites à valeurs dans $\mathbb{K}$.

Partie A

Définitions et généralités sur les séries entières

1. Séries entières

Définition 1. *Série entière*

Soit $(a_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$. On appelle **série entière** associée à la suite $(a_n)_{n\in\mathbb{N}}$, la série de fonction $\sum f_n$ où, pour $n \in \mathbb{N}$, la fonction $f_n : \mathbb{C} \to \mathbb{C}$ est définie par :

$$f_n : z \mapsto a_n z^n.$$

On notera (abusivement) $\sum a_n z^n$ la série entière associée à la suite $(a_n)_{n\in\mathbb{N}}$.

Exemple 1.

On connaît déjà plusieurs séries entières :
- la série géométrique $\sum z^n$;
- la série de somme exponentielle $\sum \frac{z^n}{n!}$.

Exercice 1.

Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres complexes. Montrer que $\sum a_n z^{2n}$ est une série entière.

Correction.

Attention ! Il y a un piège ! $\sum a_n z^{2n}$ est bien une série entière : il s'agit de la série entière $\sum b_n z^n$ où, pour $n \in \mathbb{N}$,

$$\begin{cases} b_n = a_{n/2} & \text{si } n \text{ est pair} \\ b_n = 0 & \text{si } n \text{ est impair} \end{cases}$$

Définition 2. *Somme et domaine de convergence*

Soit $\sum a_n z^n$ une série entière. On appelle **domaine de convergence** de la série entière $\sum a_n z^n$ l'ensemble

$$D = \{z \in \mathbb{C} \mid \sum a_n z^n \text{ converge }\}.$$

On appelle **somme de la série entière** $\sum a_n z^n$ la fonction somme $S : D \to \mathbb{C}$ de la série, i.e.

$$S : z \mapsto \sum_{n=0}^{+\infty} a_n z^n.$$

Remarque 1.

- Si $(a_n)_{n\in\mathbb{N}}$ est à valeurs réelles, on peut également considérer le domaine réelle de convergence $D_{\mathbb{R}} = \{x \in \mathbb{R} \mid \sum a_n x^n \text{ converge }\}$ de la série entière et sa somme $S : D_{\mathbb{R}} \to \mathbb{R}$ où $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$.
- Le domaine de convergence d'une série entière $\sum a_n z^n$ est le plus grand ensemble sur lequel la série de fonctions $\sum a_n z^n$ converge simplement.

Question 1.

Que dire de la somme d'une série entière associée à une suite stationnaire en 0 ?

Réponse : Soit $(a_n)_{n\in\mathbb{N}}$ est une suite stationnaire en 0 ; on note $N = \min_{n\in\mathbb{N}}(a_n = 0)$ et $P = \sum_{n=0}^{N-1} a_n X^n \in \mathbb{C}[X]$. Alors la série $\sum a_n z^n$ converge pour tout $z \in \mathbb{C}$. En effet, la suite des sommes partielles est stationnaire en $\sum_{n=0}^{N-1} a_n z^n = P(z)$. De plus, pour la même raison, la somme S de la série entière est :

$$S : z \mapsto P(z).$$

On peut donc conclure que la somme d'une série entière associée à une suite stationnaire en 0 est une fonction polynomiale !

2. Rayon de convergence

a. Lemme d'Abel

Théorème 1. *Lemme d'Abel*

Soit $(a_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$ et $z_0 \in \mathbb{C}^*$. Si la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ est bornée, alors, pour tout $z \in \mathbb{C}$ tel que $|z| < |z_0|$, la série $\sum a_n z^n$ est absolument convergente.

Démonstration.

On suppose que suite $(a_n z_0^n)_{n\in\mathbb{N}}$ est bornée. Alors il existe $M > 0$ tel que $|a_n z_0^n| \leq M$.
Soit $z \in \mathbb{C}$ tel que $|z| < |z_0|$. Alors, pour tout $n \in \mathbb{N}$, on a :

$$|a_n z^n| = |a_n z_0^n|.\left(\frac{z}{z_0}\right)^n \leq M\left(\frac{z}{z_0}\right)^n,$$

qui est le terme général d'une série géométrique de raison strictement inférieure à 1 car $|z| < |z_0|$.
Par suite, $\sum |a_n z^n|$ est convergente. $\square$

b. Définition et propriétés du rayon de convergence

Lemme 1.

Soit $(a_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{K}$. L'ensemble $\{r \in \mathbb{R}_+ \mid (|a_n|r^n)_{n\in\mathbb{N}} \text{ est bornée}\}$ est un intervalle non vide de $\mathbb{R}$.

Démonstration.

On note $I = \{r \in \mathbb{R}_+ \mid (|a_n|r^n)_{n\in\mathbb{N}} \text{ est bornée}\}$. Alors I contient 0 car $(|a_n|0^n)_{n\in\mathbb{N}}$ est bornée.
De plus, si $r \in I$, alors, pour tout $s \in [0, r]$, $s \in I$ car pour tout $n \in \mathbb{N}$, $|a_n|s^n \leq |a_n|r^n$ donc $(|a_n|s^n)_{n\in\mathbb{N}}$ est bornée.
Il en résulte que I est un intervalle de la forme $[0, a)$. $\square$

Ce lemme justifie la définition suivante :

Définition 3. *Rayon de convergence*

Soit $\sum a_n z^n$ une série entière.

i) On appelle **rayon de convergence** et on note R la borne supérieure de l'intervalle $I = \{r \in \mathbb{R}_+ \mid (|a_n|r^n)_{n\in\mathbb{N}} \text{ est bornée}\}$ i.e.

$$R = \sup\{r \in \mathbb{R}_+ \mid (|a_n|r^n)_{n\in\mathbb{N}} \text{ est bornée}\}.$$

on convient que $R = +\infty$ si l'intervalle I n'est pas majoré.

ii) On appelle **disque ouvert de convergence** de la série entière $\sum a_n z^n$ l'ensemble $\mathbb{D}(0, R) = \{z \in \mathbb{C} \mid |z| < R\}$.

iii) Si $(a_n)_{n\in\mathbb{N}}$ est à valeurs dans $\mathbb{R}$, On appelle **intervalle ouvert de convergence** de la série entière $\sum a_n x^n$ l'intervalle $]-R, R[$.

Proposition 1.

Soit $\sum a_n z^n$ une série entière, R son rayon de convergence et $z \in \mathbb{C}$.

— Si $|z| < R$, alors la série numérique $\sum a_n z^n$ converge absolument.

— Si $|z| > R$, alors la série numérique $\sum a_n z^n$ diverge grossièrement.

Démonstration.

— On suppose $|z| < R$. Comme $R = \sup\{r \in \mathbb{R}_+ \mid (|a_n|r^n)_{n\in\mathbb{N}} \text{ est bornée}\}$, alors il existe $r_0 \in \{r \in \mathbb{R}_+ \mid (|a_n|r^n)_{n\in\mathbb{N}} \text{ est bornée}\}$ tel que $|z| < r_0 < R$.
Par conséquent, la suite $(a_n r_0^n)_{n\in\mathbb{N}}$ étant bornée, d'après le lemme d'Abel, la série $\sum a_n z^n$

est absolument convergente.

— On suppose $|z| > R$. Alors la suite $(a_n z^n)_{n\in\mathbb{N}}$ n'est pas bornée et donc ne converge pas vers 0. Ainsi, la série $\sum a_n z^n$ diverge grossièrement.

□

Remarque 2.

Si $|z| = R$, on ne peut, a priori, rien dire ! Il faut étudier la série dans ce cas.

Proposition 2.

Soit $\sum a_n z^n$ une série entière, R son rayon de convergence et D son domaine de convergence. Alors on a :

$$\mathbb{D}(0, R) = \{z \in \mathbb{C} \mid |z| < R\} \subset D \subset \overline{\mathbb{D}}(0, R) = \{z \in \mathbb{C} \mid |z| \leq R\}.$$

Démonstration.

- Si $z \in \mathbb{D}(0, R)$ alors $|z| < R$. Par suite, d'après la proposition précédente, $\sum a_n z^n$ converge absolument et donc converge. D'où $z \in D$.
  Il en résulte que $\mathbb{D}(0, R) \subset D$.
- Si $z \notin \overline{\mathbb{D}}(0, R)$ alors $|z| > R$. Par suite, d'après la proposition précédente, $\sum a_n z^n$ diverge grossièrement. D'où $z \notin D$.
  Ainsi $\overline{D}(0, R)^c \subset D^c$ et donc $D \subset \overline{D}(0, R)$.

□

Exemple 2.

— Pour la série entière $\sum z^n$, le rayon de convergence est 1 et son domaine de convergence est $D = \mathbb{D}(0, 1)$.

> On a
> $$\{r \in \mathbb{R}_+ \mid (r^n)_{n\in\mathbb{N}} \text{ est bornée}\} = [0, 1].$$
> Donc le rayon de convergence R de $\sum z^n$ est :
> $$R = \sup[0, 1] = 1.$$
> De plus, si $|z| = 1$, $|z^n| = |z|^n = 1 \underset{n\to+\infty}{\not\to} 0$, donc $\sum z^n$ diverge grossièrement.
>
> Il en résulte que $D = \mathbb{D}(0, 1)$.

— Pour la série entière $\sum_{n\geq 1} \dfrac{1}{n^2} z^n$, le rayon de convergence est 1 et son domaine de convergence est $D = \overline{\mathbb{D}}(0, 1)$.

On a
$$\{r \in \mathbb{R}_+ \mid (\frac{1}{n^2}r^n)_{n\in\mathbb{N}} \text{ est bornée}\} = [0,1].$$
Donc le rayon de convergence R de $\sum_{n\geq 1} \frac{1}{n^2}z^n$ est :
$$R = \sup[0,1] = 1.$$
De plus, si $|z| = 1$, $|\frac{1}{n^2}z^n| = \frac{1}{n^2}$ donc, d'après le critère de Riemann, $\sum \frac{z^n}{n^2}$ converge absolument.
Il en résulte que $D = \overline{\mathbb{D}}(0,1)$.

Exercice 2.

Déterminer le rayon de convergence et le domaine de convergence des séries entières $\sum \frac{z^n}{n!}$ et $\sum n!z^n$.

Correction.

1. On a
$$\{r \in \mathbb{R}_+ \mid (\frac{1}{n!}r^n)_{n\in\mathbb{N}} \text{ est bornée}\} = [0,+\infty[.$$
car, pour tout $r \in \mathbb{R}$, $\frac{1}{n!}r^n$ est le terme général d'une série convergente - donc converge vers 0 et donc est une suite bornée.
Ainsi le rayon de convergence R de $\sum \frac{1}{n!}z^n$ est :
$$R = +\infty.$$
Il en résulte que $D = \mathbb{C}$.
2. On a
$$\{r \in \mathbb{R}_+ \mid (n!r^n)_{n\in\mathbb{N}} \text{ est bornée}\} = \{0\}.$$
En effet, pour $0 < r < 1$, à partir du rang $N = E(r) + 1$, il existe $C > 0$ tel que pour tout $n \geq N$, $n!r^n \geq Cn \xrightarrow[n\to+\infty]{} +\infty$ (on peut prendre $C = (N-1)!r^N$)) donc pour tout $r > 0$, la suite $(n!r^n)_{n\in\mathbb{N}}$ n'est pas bornée (le cas $r \geq 1$ est immédiat - *le faire quand même pour vérifier que c'est bien immédiat !*).
Donc le rayon de convergence R de $\sum n!z^n$ est :
$$R = \sup\{0\} = 0.$$
Il en résulte que $D = \{0\}$.

3. Calcul du rayon de convergence d'une série entière

a. Caractérisation du rayon de convergence

Proposition 3.

Soit $\sum a_n z^n$ une série entière et R son rayon de convergence. Alors on a les égalités suivantes :
— $R = \sup\{\,|z| \mid (a_n z^n)_{n\in\mathbb{N}} \text{ est bornée}\}$;
— $R = \sup\{\,|z| \mid (a_n z^n)_{n\in\mathbb{N}} \text{ converge}\}$;
— $R = \sup\left\{\,|z| \mid \sum a_n z^n \text{ converge}\right\}$;
— $R = \sup\left\{\,|z| \mid \sum a_n z^n \text{ converge absolument}\right\}$.

Méthode : Minoration et majoration du rayon de convergence

Étant donné une série entière $\sum a_n z^n$ de rayon de convergence R et $z_0 \in \mathbb{C}$, on a :

- la minoration $R \geq |z_0|$, si on est dans l'un des cas suivants :
  i) la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ est bornée ;
  ii) la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ converge ;
  iii) la série $\sum a_n z_0^n$ converge ;
  iv) la série $\sum a_n z_0^n$ converge absolument ;
- la majoration $R \leq |z_0|$, si on est dans l'un des cas suivants :
  i) la suite $(a_n z_0^n)_{n\in\mathbb{N}}$ n'est pas bornée ;
  ii) la série $\sum a_n z_0^n$ diverge ;
  iii) la série $\sum |a_n z_0^n|$ diverge.

Exercice 3.

1. Déterminer le rayon de convergence de $\sum n z^n$.
2. Déterminer le rayon de convergence de la série entière $\sum a_n z^{2n}$ en fonction de celui de $\sum a_n z^n$.

Correction.

1. On remarque tout d'abord que la suite $(n1^n)_{n\in\mathbb{N}}$ n'est pas bornée. Donc, comme $R = \sup\{\,|z| \mid (nz^n)_{n\in\mathbb{N}} \text{ est bornée}\}$, on a $R \leq 1$.
   Soit $z \in \mathbb{C}^*$. Si $|z| < 1$, la suite $(n|z|^n)_{n\in\mathbb{N}}$ converge vers 0 par croissances comparées donc comme $R = \sup\{\,|z'| \mid (nz'^n)_{n\in\mathbb{N}} \text{ converge}\}$, on a $R \geq |z|$.
   Ceci étant vrai pour tout z tel que $|z| < 1$, on peut faire tendre $|z|$ vers 1 dans l'inégalité précédente, ce qui donne $R \geq 1$.
   Il en résulte que $R = 1$.
2. Notons R le rayon de convergence de la série entière $\sum a_n z^{2n}$ et R' celui de $\sum a_n z^n$.

Soit $z \in \mathbb{C}^*$. On suppose $|z| < R$. Alors la suite $(a_n z^n)_{n\in\mathbb{N}}$ est bornée et donc la suite $(|a_n|(\sqrt{|z|})^{2n})_{n\in\mathbb{N}}$ l'est aussi. Or, on a $R' = \sup\left\{ |z'| \mid (a_n z'^{2n})_{n\in\mathbb{N}} \text{ est bornée}\right\}$, donc $R' \geq \sqrt{|z|}$. Ceci étant vrai pour tout z tel que $|z| < R$, on fait tendre $|z|$ vers R et ainsi, par continuité de la fonction racine :

$$R' \geq \sqrt{R}.$$

Soit $z \in \mathbb{C}^*$. On suppose $|z| < R'$. Alors la suite $(a_n z^{2n})_{n\in\mathbb{N}}$ est bornée et donc la suite $(a_n (z^2)^n)_{n\in\mathbb{N}}$ l'est aussi. Or, on a $R = \sup\left\{ |z'| \mid (a_n z'^n)_{n\in\mathbb{N}} \text{ est bornée}\right\}$, donc $R \geq |z^2| = |z|^2$. Ceci étant vrai pour tout z tel que $|z| < R'$, on fait tendre $|z|$ vers R' et ainsi, par continuité de la fonction carrée :

$$R \geq R'^2.$$

Il en résulte que $R' = \sqrt{R}$.

b. Comparaison

Proposition 4. *Comparaison des rayons de convergence*

Soit $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières et R_a, R_b leurs rayons de convergence respectifs. Alors si, à partir d'un certain rang $N \in \mathbb{N}$, on a, pour tout $n \geq N$:

i) $|a_n| \leq |b_n|$, alors $R_a \geq R_b$;
ii) $a_n = O(b_n)$, alors $R_a \geq R_b$;
iii) $a_n = o(b_n)$, alors $R_a \geq R_b$;
iv) $|a_n| \underset{n\to+\infty}{\sim} |b_n|$ alors $R_a = R_b$.

Démonstration.

i) Soit $z \in \mathbb{C}^*$. On suppose $|z| < R_b$. Alors la suite $(b_n z^n)_{n\in\mathbb{N}}$ est bornée. Comme pour tout $n \geq N$, $|a_n| \leq |b_n|$, on a $|a_n z^n| \leq |b_n z^n|$ donc la suite $(a_n z^n)_{n\in\mathbb{N}}$ est bornée. Or, on a $R_a = \sup\left\{ |z'| \mid (a_n z'^n)_{n\in\mathbb{N}} \text{ est bornée}\right\}$, donc $R_a \geq |z|$. Ceci étant vrai pour tout z tel que $|z| < R_b$, on fait tendre $|z|$ vers R_b et ainsi :

$$R_a \geq R_b.$$

ii) On suppose $a_n = O(b_n)$. Alors il existe $M \geq 0$ tel que pour tout $n \in \mathbb{N}$, $|a_n| \leq M|b_n|$. On adpate alors la preuve précédente en remarquant que, pour un certain $z \in \mathbb{C}$, si $(b_n z^n)_{n\in\mathbb{N}}$ est bornée, alors $(M b_n z^n)_{n\in\mathbb{N}}$ l'est aussi.

iii) Si $a_n = o(b_n)$, alors $a_n = O(b_n)$ d'où $R_a \geq R_b$;

iv) On remarque que $|a_n| \underset{n\to+\infty}{\sim} |b_n|$ implique $a_n = O(b_n)$ et $b_n = O(a_n)$. En effet, par définition, $|a_n| \underset{n\to+\infty}{\sim} |b_n| \Leftrightarrow a_n = b_n + o(b_n) = O(b_n) + O(b_n) = O(b_n)$.

□

Exercice 4.

1. Déterminer les rayons de convergence de $\sum \frac{2^n(1+5^n n^2)}{10^n(n+\sqrt{3n+1})} z^n$ et de $\sum \frac{\sin(\frac{n}{3^n})}{n+1} z^n$.
2. Déterminer le rayon de convergence de $\sum_{n\geq 1} d(n) z^n$ où, pour $n \in \mathbb{N}^*$, $d(n) = \#\{d \in [\![1,n]\!] \mid d|n\}$.

Correction.

1. On a :
$$\frac{2^n(1+5^n n^2)}{10^n(n+\sqrt{3n+1})} \underset{n\to+\infty}{\sim} n$$
Or on a prouvé précédemment que $\sum nz^n$ a pour rayon de convergence 1 donc, par comparaison, le rayon de convergence de $\sum \frac{2^n(1+5^n n^2)}{10^n(n+\sqrt{3n+1})} z^n$ est égal à 1.
Comme $\sin(x) \underset{x\to 0}{\sim} x$, on a :
$$\frac{\sin(\frac{n}{3^n})}{n+1} \underset{n\to+\infty}{\sim} \frac{1}{3^n}$$
Or le rayon de convergence de $\sum \frac{1}{3^n} z^n$ est égal à 3 : en effet, pour $z \in \mathbb{C}$, la suite $((\frac{z}{3})^n)_{n\in\mathbb{N}}$ est bornée si, et seulement si $|z| \leq 3$. Ainsi, par comparaison, le rayon de convergence de $\sum \frac{\sin(\frac{n}{3^n})}{n+1} z^n$ est égal à 3.
2. Pour $n \in \mathbb{N}^*$, on remarque que $1 \leq d(n) \leq n$. Or les rayons de convergence de $\sum z^n$ et de $\sum nz^n$ sont tous deux égaux à 1, d'où, si on note R le rayon de convergence de $\sum_{n\geq 1} d(n) z^n$, on obtient $1 \geq R \geq 1$ et ainsi $R = 1$

c. Utilisation de la règle de D'Alembert

Théorème 2. *Règle de D'Alembert pour les séries entières*

Soit $\sum a_n z^n$ une série entière de rayon de convergence R telle que, à partir d'un certain rang $N \in \mathbb{N}$, pour tout $n \geq N$, $a_n \neq 0$. S'il existe $\ell \in [0,+\infty[\cup\{+\infty\}$ tel que :
$$\left|\frac{a_{n+1}}{a_n}\right| \xrightarrow[n\to+\infty]{} \ell,$$
alors on a :
$$R = \frac{1}{\ell} = \begin{cases} +\infty & \text{si } \ell = 0; \\ \frac{1}{\ell} & \text{si } \ell \in]0,+\infty[; \\ 0 & \text{si } \ell = +\infty. \end{cases}$$

Démonstration.

Soit $z \in \mathbb{C}$. On applique le critère de D'Alembert à la série de terme général $u_n = |a_n z^n|$. Alors on a, pour tout $n \geq N$,

$$\frac{u_{n+1}}{u_n} = \left|\frac{a_{n+1}z^{n+1}}{a_n z^n}\right| = \left|\frac{a_{n+1}}{a_n}\right|.|z|.$$

Par suite, si $\left|\frac{a_{n+1}}{a_n}\right| \xrightarrow[n\to+\infty]{} \ell$ où :

- $\ell \in \mathbb{R}_+^*$, alors $\frac{u_{n+1}}{u_n} \xrightarrow[n\to+\infty]{} \ell|z|$. Ainsi, d'après le critère de D'Alembert, si $|z| < \frac{1}{\ell}$, $\sum u_n$ converge et si $|z| > \frac{1}{\ell}$, $\sum u_n$ diverge. Par suite, $R = \frac{1}{\ell}$.
- $\ell = +\infty$, alors $\frac{u_{n+1}}{u_n} \xrightarrow[n\to+\infty]{} +\infty$. Ainsi, d'après le critère de D'Alembert, pour tout $z \in \mathbb{C}^*$, $\sum u_n$ diverge donc $R = 0$.
- $\ell = 0$, alors $\frac{u_{n+1}}{u_n} \xrightarrow[n\to+\infty]{} 0$. Ainsi, d'après le critère de D'Alembert, pour tout $z \in \mathbb{C}^*$, $\sum u_n$ converge. Par suite, $R = +\infty$.

$\square$

Remarque 3.

Attention le critère précédent n'est valable que si $(a_n)_{n\in\mathbb{N}}$ est différente de 0 à partir d'un certain rang !

Ainsi, pour une série entière du type $\sum a_n z^{\varphi(n)}$ avec $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante, on appliquera directement la règle de D'Alembert sur la série (tout court) $\sum a_n z^{\varphi(n)}$ i.e. on étudie la limite de

$$\left|\frac{a_{n+1}z^{\varphi(n+1)}}{a_n z^{\varphi(n)}}\right| = \left|\frac{a_{n+1}}{a_n}\right|.|z^{\varphi(n+1)-\varphi(n)}|,$$

en fonction des valeurs de $z \in \mathbb{C}^*$ afin de majorer et minorer le rayon de convergence de la série entière.

Exercice 5.

1. Déterminer les rayons de convergence des séries entières

$$\sum_{n\geq 1} n^\alpha z^n \text{ où } \alpha \in \mathbb{R}; \quad \sum \frac{n^n}{n!} z^n; \quad \sum \binom{4n}{2n+1} z^n;$$

$$\sum_{n\geq n_0} \frac{P(n)}{Q(n)} z^n \text{ où } P, Q \in \mathbb{K}[X].$$

2. Déterminer les rayons de convergence des séries entières :

$$\sum n! z^{2n} \qquad \sum n! z^{n^2} \qquad \sum n^n z^{\binom{3n}{n}}.$$

Correction.

1. Pour cette question, on remarque que les séries entières ne sont pas lacunaires et que les suites $(a_n)_{n\in\mathbb{N}}$ associées sont non nuls (à partir d'un certain rang). On peut donc appliquer le critère de D'Alembert pour les séries entières :
   - Ici, $a_n = n^\alpha$ pour $n \geq 1$ et $a_0 = 0$. Ainsi, à partir du rang 1, on a, par continuité de la fonction $x \mapsto x^\alpha$ en 1 :
   $$\left|\frac{a_{n+1}}{a_n}\right| = \left(\frac{n+1}{n}\right)^\alpha \xrightarrow[n\to+\infty]{} 1^\alpha = 1$$
   Ainsi, le rayon de convergence R de $\sum_{n\geq1} n^\alpha z^n$ est $R = \frac{1}{1} = 1$.
   - Ici, $a_n = \frac{n^n}{n!}$ pour $n \geq 0$. Ainsi, comme pour tout $x \in \mathbb{R}$, $(1+\frac{x}{n})^n \xrightarrow[n\to+\infty]{} e^x$, on a :
   $$\left|\frac{a_{n+1}}{a_n}\right| = \left(\frac{n+1}{n}\right)^n \xrightarrow[n\to+\infty]{} e^1 = e$$
   Ainsi, le rayon de convergence R de $\sum \frac{n^n}{n!} z^n$ est $R = \frac{1}{e}$.
   - Ici, $a_n = \begin{pmatrix} 4n \\ 2n+1 \end{pmatrix}$ pour $n \geq 0$. Ainsi, on a :
   $$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(4n+4)(4n+3)(4n+2)(4n+1)}{(2n+3)(2n+2)(2n+1)(2n)} \underset{n\to+\infty}{\sim} \frac{4^4n^4n^4}{2^4n^4} = 2^4 \xrightarrow[n\to+\infty]{} 2^4 = 16$$
   Ainsi, le rayon de convergence R de $\sum \begin{pmatrix} 4n \\ 2n+1 \end{pmatrix} z^n$ est $R = \frac{1}{16}$.
   - On suppose que P, Q sont des polynômes non nuls. Ici, $a_n = \frac{P(n)}{Q(n)}$ pour $n \geq n_0$ où $n_0 = E(\max\{x \in \mathbb{R}_+ \mid Q(x) = 0\}) + 1$ si Q admet des racines réelles positives et $n_0 = 0$ sinon *(pour s'assurer qu'on ne divise par par 0 ; dans le cas où Q possède des racines positives, ce "max" existe bien car Q étant un polynôme non nul, l'ensemble de ses racines est fini)* et $a_n = 0$ pour tout $n < n_0$.
   On va cette fois utiliser une comparaison avec la première série entière de la question pour déterminer le raton de convergence :
   Comme P, Q sont non nuls, il existe $p, q \in \mathbb{N}$ et des coefficients $\alpha_i, \beta_i \in \mathbb{K}$ tels que $P = \sum_{i=0}^p \alpha_i X^i$ et $Q = \sum_{i=0}^p \beta_i X^i$ avec $\alpha_p \neq 0$ et $\beta_q \neq 0$. Par suite, on a, pour tout $n \geq n_0$:
   $$\frac{P(n)}{Q(n)} \underset{n\to+\infty}{\sim} \frac{\alpha_p n^p}{\beta_q n^q} = \frac{\alpha_p}{\beta_q} n^{p-q}$$
   Or, pour $\alpha = p - q \in \mathbb{R}$, la série entière $\sum_{n\geq1} n^\alpha z^n$ possède un rayon de convergence égal à 1, donc, par comparaison, le rayon de convergence de $\sum_{n\geq n_0} \frac{P(n)}{Q(n)} z^n$ est égal à 1.
2. Les séries entières de cette question sont lacunaires, on ne peut donc pas aplliquer les critère de D'Alembert pour les séries entières. On se rabat donc sur le critère de D'Alembert... tout court !
   - Soit $z \in \mathbb{C}^*$. On pose, pour $n \in \mathbb{N}$, $u_n = |n!z^{2n}| = n!|z|^{2n} > 0$. On peut donc appliquer la règle de D'Alembert à la suite $(u_n)_{n\in\mathbb{N}}$. On a :
   $$\frac{u_{n+1}}{u_n} = (n+1)|z|^2 \xrightarrow[n\to+\infty]{} +\infty" > 1"$$
   Ainsi, d'après la règle de D'Alembert, $\sum u_n$ diverge.
   Par suite, pour tout $z \in \mathbb{C}^*$, $\sum n!z^n$ ne converge pas absolument. Or la rayon R de la série entière vérifie $R = \sup\{|z| \mid \sum n!z^n \text{ converge absolument}\}$, donc $R = 0$.

— Soit $z \in \mathbb{C}^*$. On pose, pour $n \in \mathbb{N}$, $u_n = |n!z^{n^2}| = n!|z|^{n^2} > 0$. On peut donc appliquer la règle de D'Alembert à la suite $(u_n)_{n\in\mathbb{N}}$. On a :

$$\frac{u_{n+1}}{u_n} = (n+1)|z|^{2n+1} \xrightarrow[n\to+\infty]{} \begin{cases} +\infty\text{"} > 1\text{"} & \text{si } |z| \geq 1 \\ 0 < 1 & \text{si } |z| < 1 \end{cases}$$

Donc, d'après la règle de D'Alembert, la série numérique $\sum n!z^{n^2}$ converge absolument si, et seulement si, $|z| < 1$.
Il en résulte que $R = 1$ car $R = \sup\{|z| \mid \sum n!z^{n^2} \text{ converge absolument}\}$.

— Soit $z \in \mathbb{C}^*$. On pose, pour $n \in \mathbb{N}$, $u_n = |n^n z^{\binom{3n}{n}}| = n^n|z|^{\binom{3n}{n}} > 0$. On peut donc appliquer la règle de D'Alembert à la suite $(u_n)_{n\in\mathbb{N}}$. On remarque que, comme $n \leq E(\frac{3n}{2})$, $\binom{3n}{1} \leq \binom{3n}{n}$ et donc :

$$\binom{3n+3}{n+1} - \binom{3n}{n} = \binom{3n}{n}\left(3\underbrace{\frac{(3n+2)(3n+1)}{(2n+2)(2n+1)}}_{\geq 1} - 1\right) \geq \binom{3n}{1} \times 2 = 6n \xrightarrow[n\to+\infty]{} +\infty$$

et ainsi,

$$\frac{u_{n+1}}{u_n} = (n+1)\left(\frac{n+1}{n}\right)^n |z|^{\binom{3n+3}{n+1}-\binom{3n}{n}} \xrightarrow[n\to+\infty]{} \begin{cases} +\infty\text{"} > 1\text{"} & \text{si } |z| \geq 1 \\ 0 < 1 & \text{si } |z| < 1 \end{cases}$$

Donc, d'après la règle de D'Alembert, la série numérique $\sum n^n z^{\binom{3n}{n}}$ converge absolument si, et seulement si, $|z| < 1$.
Il en résulte que $R = 1$ car $R = \sup\{|z| \mid \sum n^n z^{\binom{3n}{n}} \text{ converge absolument}\}$.

Exercice 6. *Apparté : Transformée d'Abel*

Soit $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites à valeurs dans K. On considère les séries $\sum a_n b_n$ et $\sum a_n$. On note $(S_n)_{n\in\mathbb{N}}$ la suite de ses sommes partielles de $\sum a_n b_n$ et $(A_n)_{n\in\mathbb{N}}$ celle de $\sum a_n$.

1. Montrer que pour tout $N \in \mathbb{N}$,

$$S_N = A_N b_N - \sum_{n=0}^{N-1} A_n(b_{n+1} - b_n).$$

2. En déduire que si $(A_n)_{n\in\mathbb{N}}$ est bornée ; $b_n \xrightarrow[n\to+\infty]{} 0$ et $\sum(b_{n+1} - b_n)$ est absolument convergente, alors $\sum a_n b_n$ converge.

Exercice 7. *Étude d'une série entière sur la frontière du disque*

On considère la série entière $\displaystyle\sum_{n\geq 1} \frac{z^n}{n}$.

1. Montrer que son rayon de convergence est 1. Que dire de la convergence en $z = 1$?

2. Soit $z_0 \in \mathbb{U} \setminus \{1\}$. En utilisant la transformée d'Abel, montrer que $\sum_{n\geq 1} \dfrac{z_0^n}{n}$ converge.
3. En déduire le domaine de convergence de $\displaystyle\sum_{n\geq 1} \frac{z^n}{n}$.

Partie B

Propriétés des séries entières

1. Opérations sur les séries entières

a. Combinaisons linéaires

Proposition 5. *Produit par un scalaire*

Soit $\sum a_n z^n$ une série entière et $\lambda \in \mathbb{C}^*$. Alors $\sum \lambda a_n z^n$ et $\sum a_n z^n$ ont le même rayon de convergence.

Démonstration.

Soit $z \in \mathbb{C}$. Comme $\lambda \neq 0$, la suite $(\lambda a_n z^n)_{n\in\mathbb{N}}$ est bornée si, et seulement si, $(a_n z^n)_{n\in\mathbb{N}}$ est bornée.
Il en résulte que $\sum \lambda a_n z^n$ et $\sum a_n z^n$ ont même rayon de convergence. □

Proposition 6. *Somme*

Soit $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières et R_a, R_b leurs rayons de convergence respectifs. Alors le rayon de convergence R de la série entière $\sum (a_n + b_n) z^n$ vérifie :

— si $R_a \neq R_b$, $R = \min(R_a, R_b)$

— si $R_a = R_b$, $R \geq R_a (= R_b)$.

Démonstration.

Soit $z \in \mathbb{C}$. Si $(a_n z^n)_{n\in\mathbb{N}}$ et $(b_n z^n)_{n\in\mathbb{N}}$ sont des suites bornées, alors la suite $((a_n + b_n) z^n)_{n\in\mathbb{N}}$ est bornée, donc $|z| \leq R$. Ceci étant vrai pour tout $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$, on obtient :

$$R \geq \min(R_a, R_b).$$

Supposons que $R_a \neq R_b$. Quitte à échanger R_a et R_b, on suppose que $R_a < R_b$.
Soit $z \in \mathbb{C}$ tel que $R_a < |z| < R_b$. Alors la suite $((a_n + b_n) z^n)_{n\in\mathbb{N}}$ n'est pas bornée car $(a_n z^n)_{n\in\mathbb{N}}$ n'est pas bornée et $(b_n z^n)_{n\in\mathbb{N}}$ est bornée.

> *Remarque :* pour démontrer le fait précédent, on peut utiliser la contraposée de l'assertion : "si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont bornées, alors $(u_n + v_n)_{n\in\mathbb{N}}$ est bornée.

Ainsi, on a $|z| \geq R$. Ceci étant vrai pour tout $z \in \mathbb{C}$ avec $\mathbf{R_a} < |\mathbf{z}| < R_b$, on obtient $\min(R_a, R_b) = R_a \geq R$.

Il en résulte que, si $R_a \neq R_b$,
$$R = \min(R_a, R_b).$$
□

Proposition 7. *Somme d'une combinaison linéaire de séries entières*

Soit $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières de rayon de convergence respectifs R_a, R_b et $\lambda, \mu \in \mathbb{C}$. On pose $R = \min(R_a, R_b)$. Alors, pour tout $z \in \mathbb{C}$ tel que $|z| < R$, on a :
$$\sum_{n=0}^{+\infty} (\lambda a_n + \mu b_n) z^n = \lambda \sum_{n=0}^{+\infty} a_n z^n + \mu \sum_{n=0}^{+\infty} b_n z^n.$$

Exercice 8.

Déterminer les rayons de convergence et la somme dans le disque ouvert de convergence des séries entières suivantes :
$$\sum \mathrm{ch}(n) z^n \qquad \sum \sin(n\theta) z^n \quad (\text{où } \theta \in \mathbb{R}).$$

Correction.

1. On a, par définition, pour tout $x \in \mathbb{R}$, $\mathrm{ch}(x) = \frac{e^x + e^{-x}}{2}$.
   La série entière $\sum e^n z^n$ a pour rayon de convergence $\frac{1}{e}$ et la série entière $\sum e^{-n} z^n$ a pour rayon de convergence e donc la série entière $\sum \mathrm{ch}(n) z^n$ a pour rayon de convergence $R = \min(\frac{1}{e}, e) = \frac{1}{e}$ et on a, pour tout $z \in \mathbb{C}$ tel que $|z| < \frac{1}{e}$:
   $$\sum_{n=0}^{+\infty} \mathrm{ch}(n) z^n = \frac{1}{2} \sum_{n=0}^{+\infty} e^n z^n + \frac{1}{2} \sum_{n=0}^{+\infty} e^{-n} z^n = \frac{1}{2} \frac{1}{1 - ez} + \frac{1}{2} \frac{1}{1 - \frac{z}{e}}.$$
2. On a, par définition, pour tout $x \in \mathbb{R}$, $\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}$.
   Les séries entières $\sum e^{in\theta} z^n$ et $\sum e^{-n} z^n$ ont pour rayons de convergence 1 donc la série entière $\sum \sin(n\theta) z^n$ a pour rayon de convergence $R \geq 1$ et on a, pour tout $z \in \mathbb{C}$ tel que $|z| < 1$:
   $$\begin{aligned} \sum_{n=0}^{+\infty} \sin(n\theta) z^n &= \frac{1}{2i} \sum_{n=0}^{+\infty} e^{in\theta} z^n - \frac{1}{2i} \sum_{n=0}^{+\infty} e^{-in\theta} z^n \\ &= \frac{1}{2i} \frac{1}{1 - e^{i\theta} z} - \frac{1}{2i} \frac{1}{1 - e^{-i\theta} z} \\ \sum_{n=0}^{+\infty} \sin(n\theta) z^n &= \frac{\sin(\theta) z}{1 - 2\cos(\theta) z + z^2}. \end{aligned}$$
   Maintenant, déterminons exactement le rayon de convergence R de $\sum \sin(n\theta) z^n$.
   Si $\theta \in 2\pi\mathbb{Z}$, alors $\sin(n\theta) = 0$ pout tout $n \in \mathbb{N}$. Donc dans ce cas, $R = +\infty$.
   Supposons $\theta \notin 2\pi\mathbb{Z}$. Comme la suite $(\sin(n\theta))_{n\in\mathbb{N}}$ ne tend pas vers 0 alors la suite

$(\sin(n\theta)1^n)_{n\in\mathbb{N}}$ n'est pas le terme général d'une série convergente et donc $R \leq 1$.
Il en résulte que $R = 1$.

b. Produits de Cauchy

On rappelle ici la notion de produit de Cauchy dans le cas de séries entières :

Définition 4. *Produit de Cauchy de deux séries entières*

Soit $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières. On appelle **produit de Cauchy** des séries entières $\sum a_n z^n$ et $\sum b_n z^n$, la série entière $\sum c_n z^n$ où, pour $n \in \mathbb{N}$,

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Proposition 8. *Produit de Cauchy*

Soit $\sum a_n z^n$ et $\sum b_n z^n$ des séries entières et R_a, R_b leurs rayons de convergence respectifs. Alors le rayon de convergence R du produit de Cauchy $\sum c_n z^n$ des séries entières $\sum a_n z^n$ et $\sum b_n z^n$ vérifie $R \geq \min(R_a, R_b)$ et on a, pour tout $z \in \mathbb{C}$ tel que $|z| < \min(R_a, R_b)$:

$$\sum_{n=0}^{+\infty} c_n z^n = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) z^n = \left(\sum_{n=0}^{+\infty} a_n z^n \right) \left(\sum_{n=0}^{+\infty} b_n z^n \right).$$

Exercice 9.

Déterminer le rayon de convergence et la somme du produit de Cauchy de $\sum z^n$ et $1 - z$. Qu'en conclure ?

2. Régularité d'une série entière

a. Convergence normale

Théorème 3.

Soit $\sum a_n z^n$ une série entière de rayon de convergence $R > 0$. Alors :

- la série entière $\sum a_n z^n$ converge normalement sur tout compact de $\mathbb{D}(0, R)$ ($= \mathbb{C}$ si $R = +\infty$) ;
- si $(a_n)_{n\in\mathbb{N}}$ est à valeurs réelles, la série entière $\sum a_n x^n$ converge normalement sur tout segment de $]-R, R[$ ($= \mathbb{R}$ si $R = +\infty$).

Démonstration.

— Soit $a > 0$ avec $a < R$. Montrons la série entière $\sum a_n z^n$ converge normalement $\overline{\mathbb{D}}(0, a)$. On note $f_n : z \mapsto a_n z^n$.

On a, pour $n \in \mathbb{N}$:

$$\|f_n\|_\infty = \sup_{z \in \overline{\mathbb{D}}(0,a)} (|a_n||z|^n) \leq |a_n|a^n.$$

Or $R = \sup\{r \in \mathbb{R}_+ \mid \sum |a_n|r^n \text{ converge}\}$, donc comme $a < R$, $|a_n|a^n$ est le terme général d'une série convergente.
Par suite, pour tout $0 < a < R$, la série entière $\sum a_n z^n$ converge normalement sur $\overline{\mathbb{D}}(0, a)$.

Ainsi, comme tout compact de $\mathbb{D}(0, R)$ est inclus dans un disque $\overline{\mathbb{D}}(0, a)$ avec $a < R$, on a convergence normale de $\sum a_n z^n$ sur tout compact de $\mathbb{D}(0, R)$.

— Soit $a > 0$ avec $a < R$. Montrons la série entière $\sum a_n x^n$ converge normalement $[-a, a]$. On note $f_n : x \mapsto a_n x^n$.

On a, pour $n \in \mathbb{N}$:

$$\|f_n\|_\infty = \sup_{x \in [-a,a]} (|a_n||x|^n) \leq |a_n|a^n.$$

Or $R = \sup\{r \in \mathbb{R}_+ \mid \sum |a_n|r^n \text{ converge}\}$, donc comme $a < R$, $|a_n|a^n$ est le terme général d'une série convergente.
Par suite, pour tout $0 < a < R$, la série entière $\sum a_n x^n$ converge normalement sur $[-a, a]$.

Ainsi, comme tout segment de $]-R, R[$ est inclus dans un intervalle $[-a, a]$ avec $0 < a < R$, on a convergence normale de $\sum a_n x^n$ sur tout segment $[-R, R]$.

Remarque : on aurait bien-sur pu utiliser le point précédent pour démontrer le cas réel. □

Remarque 4.

Sur le disque $\mathbb{D}(0, R)$, la convergence n'est pas normale en général : par exemple, les séries entières $\sum z^n$ ou $\sum_{n \geq 1} \dfrac{z^n}{n}$ ne convergent pas normalement sur $\mathbb{D}(0, 1)$.

b. Continuité

Théorème 4.

Soit $\sum a_n z^n$ une série entière de rayon de convergence $R > 0$. Alors sa somme

$$S : z \mapsto \sum_{n=0}^{+\infty} a_n z^n$$

est continue sur le disque ouvert de convergence $\mathbb{D}(0, R)$.
En particulier, S est continue sur l'intervalle $] - R, R[$.

Démonstration.

— D'après le théorème précédent, la série de fonctions $\sum a_n z^n$ converge normalement sur tout compact de $\mathbb{D}(0, R)$ et les fonctions $z \mapsto a_n z^n$ sont continues sur $\mathbb{C}$ et donc sur $\mathbb{D}(0, R)$ car polynomiales.
Ainsi, d'après le théorème de continuité des séries de fonctions, $S : z \mapsto \sum_{n=0}^{+\infty} a_n z^n$ est continue sur $\mathbb{D}(0, \mathbb{R})$.

— *Cas réel :* D'après le théorème précédent, la série de fonctions $\sum a_n x^n$ converge normalement sur tout segment de $]-R, R[$ et les fonctions $z \mapsto a_n x^n$ sont continues sur $\mathbb{R}$ et donc sur $]-R, R[$ car polynomiales.
Ainsi, d'après le théorème de continuité des séries de fonctions, $S : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est continue sur $]-R, R[$.
Remarque : on aurait bien-sur pu utiliser le point précédent pour démontrer le cas réel.

□

On se pose alors la question de la continuité de la somme d'une série entière de rayon de convergence R aux bornes de l'intervalle $]-R, R[$. Le théorème suivant nous fournit la réponse :

Théorème 5. *Théorème d'Abel radial*

Soit $\sum a_n x^n$ une série entière de la variable réelle de rayon de convergence $R > 0$ et de somme S. Si $\sum a_n R^n$ converge, alors

$$S(x) = \sum_{n=0}^{+\infty} a_n x^n \xrightarrow[x \to R^-]{} \sum_{n=0}^{+\infty} a_n R^n,$$

i.e. S est définie et continue à gauche en R.

Corollaire 1.

Si $\sum a_n x^n$ est une série entière de la variable réelle, alors sa somme est continue sur l'intervalle de convergence de la série entière.

Exercice 10.

1. Montrer que $f : x \mapsto \sum_{n=1}^{+\infty} \frac{x^n}{n^2}$ est continue sur $[-1, 1]$.
2. Montrer que $g : x \mapsto \sum_{n=2}^{+\infty} \frac{x^n}{\ln(n)}$ est continue sur $[-1, 1[$

c. Série entière dérivée

Définition 5. *Série entière dérivée*

Soit $\sum a_n z^n$ une série entière. On appelle **série entière dérivée** de $\sum a_n z^n$, la série entière

$$\sum (n+1) a_{n+1} z^n.$$

Lemme 2.

Soit $\sum a_n z^n$ une série entière. Pour tout $k \in \mathbb{N}$, $\sum a_n z^n$ et la série entière $\sum a_{n+k} z^n$ ont même rayon de convergence.

Démonstration.

Soit $z \in \mathbb{C}^*$. On a, pour tout $n \in \mathbb{N}$:

$$a_{n+k} z^n = a_{n+k} z^{n+k} . \frac{1}{z^k}$$

Donc la suite $(a_{n+k} z^n)_{n\in\mathbb{N}}$ est bornée si, et seulement si, la suite $(a_{n+k} z^{n+k})_{n\in\mathbb{N}}$ est bornée. Or cette dernière est bornée si, et seulement si, la suite $(a_n z^n)_{n\in\mathbb{N}}$ (car il s'agit de la "même" suite à laquelle on a "ajouté" un nombre fini de terme).
Par suite, $\sum a_n z^n$ et $\sum a_{n+k} z^n$ ont même rayon de convergence. □

Proposition 9.

Soit $\sum a_n z^n$ une série entière. Alors $\sum a_n z^n$ et sa série entière dérivée $\sum (n+1) a_{n+1} z^n$ ont même rayon de convergence.

Démonstration.

On note R le rayon de convergence de $\sum a_n z^n$ et R' celui de $\sum (n+1) a_{n+1} z^n$. D'après le lemme précédent, $\sum (n+1) a_{n+1} z^n$ et $\sum n a_n z^n$ ont même rayon de convergence R'.
Soit $z \in \mathbb{C}^*$ tel que $|z| < R'$. Alors la suite $(n a_n z^n)_{n\in\mathbb{N}}$ est bornée et donc $(a_n z^n)_{n\in\mathbb{N}}$ l'est aussi ; d'où $|z| \leq R$. Ceci étant vrai pour tout z tel que $|z| < R'$, on a $R' \leq R$. Montrons l'inégalité dans l'autre sens :

Soit $z \in \mathbb{C}^*$ tel que $|z| < R$ et $\rho \in \mathbb{R}$ tel que $|z| < \rho < R$. Alors la suite $(a_n \rho^n)_{n\in\mathbb{N}}$ est bornée, disons par une constante $M \geq 0$. Ainsi, pour tout $n \in \mathbb{N}$, on a, comme $0 \leq \frac{|z|}{\rho} < 1$, par croissances comparées :

$$|n a_n z^n| = n \left(\frac{|z|}{\rho}\right)^n \underbrace{|a_n \rho^n|}_{\leq M} \leq M n \left(\frac{|z|}{\rho}\right)^n \xrightarrow[n\to+\infty]{} 0$$

Il en résulte que $|z| \leq R'$. Ceci étant vrai pour tout z tel que $|z| < R$, on a $R \leq R'$.

Ainsi, $R' = R$. □

Corollaire 2.

Soit $\sum a_n z^n$ une série entière. Les séries entières $\sum a_n z^n$ et $\sum \frac{a_n}{n+1} z^{n+1}$ ont même rayon de convergence.

Corollaire 3.

Soit $\sum a_n z^n$ une série entière et $k \in \mathbb{Z}$. Les séries entières $\sum a_n z^n$ et $\sum n^k a_n z^n$ ont même rayon de convergence.

Exercice 11.

Montrer que pour tout $\alpha \in \mathbb{R}$, Les séries entières $\sum a_n z^n$ et la série entière $\sum n^\alpha a_n z^n$ ont même rayon de convergence.

3. Dérivée de la somme d'une série entière réelle

Théorème 6. *Dérivation d'une série entière réelle*

Soit $\sum a_n x^n$ une série entière de la variable réelle de rayon de convergence $R > 0$. Alors sa somme f est de classe C^1 sur $]-R, R[$ et on a, pour tout $x \in]-R, R[$:

$$f'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}.$$

Démonstration.

D'après la proposition 9, $\sum_{n\geq 1} n a_n x^{n-1} = \sum (n+1) a_{n+1} x^n$ a pour rayon de convergence R. Par suite, les séries entières $\sum a_n x^n$ et $\sum_{n\geq 1} n a_n x^{n-1}$ convergent normalement sur tout segment de $]-R, R[$. On peut alors vérifier les hypothèses du théorème d'interversion dérivation/somme :

- pour tout $n \in \mathbb{N}$, $f_n : x \mapsto a_n x^n$ est une fonction polynomiale et donc est de classe C^1 sur $]-R, R[$ et on a, pour tout $x \in]-R, R[$:
$$f_n'(x) = \begin{cases} 0 & \text{si } n = 0 \\ n a_n x^{n-1} & \text{si } n \geq 1 \end{cases}$$
- $\sum f_n$ converge simplement sur $]-R, R[$ car $\sum f_n$ converge normalement sur tout segment de $]-R, R[$.
- $\sum f_n'$ converge uniformément sur tout segment de $]-R, R[$ car $\sum f_n' = \sum_{n\geq 1} n a_n x^{n-1}$ converge normalement sur tout segment de $]-R, R[$.

Ainsi, d'après le théorème d'interversion dérivation/somme, $f : x \mapsto \sum_{n=0}^{+\infty} a_n x^n$ est de classe C^1 sur $]-R, R[$ et, pour tout $x \in]-R, R[$:

$$f'(x) = \left(\sum_{n=0}^{+\infty} f_n \right)'(x) = \sum_{n=0}^{+\infty} f_n'(x) = \sum_{n=1}^{+\infty} n a_n x^{n-1}.$$

□

On obtient alors, par récurrence :

Corollaire 4.

Soit $\sum a_n x^n$ une série entière de la variable réelle de rayon de convergence $R > 0$. Alors sa somme f est de classe C^∞ sur $]-R, R[$ et on a, pour tout $n \in \mathbb{N}$,

$$a_n = \frac{f^{(n)}(0)}{n!}$$

4. Primitive de la somme d'une série entière réelle

Théorème 7. *Primitive d'une série entière réelle*

Soit $\sum a_n x^n$ une série entière de la variable réelle de rayon de convergence $R > 0$, F une primitive sur $]-R, R[$ de sa somme f. Alors, pour tout $t \in]-R, R[$, on a :

$$F(x) = F(0) + \sum_{n=0}^{+\infty} a_n \frac{x^{n+1}}{n+1}.$$

Exemple 3.

On a, pour tout $x \in]-1, 1[$,

$$\ln(1-x) = -\sum_{n=1}^{+\infty} \frac{x^n}{n}.$$

En effet, $\ln(1-x) = -\int_0^x \frac{dt}{1-t}$, d'où

$$\ln(1-x) = -\sum_{n=0}^{+\infty} \frac{x^{n+1}}{n+1} = -\sum_{n=1}^{+\infty} \frac{x^n}{n}.$$

Exercice 12.

Déterminer une suite $(a_n)_{n\in\mathbb{N}}$ telle que, pour tout $x \in]-1, 1[$,

$$\arctan(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Partie C

Développements en série entière

Dans cette partie, U désigne un voisinage de 0 dans $\mathbb{R}$.

1. Développement en série entière d'une fonction d'une variable réelle

a. Définition et premier exemple

Définition 6. *Fonction développable en série entière*

Soit $r > 0$ et $f :]-r, r[\to \mathbb{C}$. On dit que f est **développable en série entière sur** $]-r, r[$ s'il existe une série entière $\sum a_n x^n$ telle que, pour tout $x \in]-r, r[$:

$$f(x) = \sum_{n=0}^{+\infty} a_n x^n.$$

Soit $f : U \to \mathbb{C}$. On dit que f est **développable en série entière** s'il existe $r > 0$ tel que f est développable en série entière sur $]-r, r[$.

Remarque 5.

Attention, le développement en série entière d'une fonction n'est pas, en général, valable sur tout l'ensemble de définition de la fonction !

Exemple 4.

Les fonctions suivantes sont développables en série entière :
- $x \mapsto \dfrac{1}{1-x}$
- $x \mapsto \ln(1-x)$

- On a pour tout $x \in]-1, 1[$,
$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n$$
donc $x \mapsto \dfrac{1}{1-x}$ est développable en série entière sur $]-1, 1[$.
- On remarque tout d'abord, que pour tout $x \in]-1, +\infty[$,
$$\ln(1-x) = -\int_0^x \frac{1}{1-t} dt.$$

Considérons alors la série entière $\sum x^n$ de rayon de convergence égal à 1 et notons f sa somme.
Alors f admet une primitive sur $]-1,1[$. On considère la primitive F de f qui s'annule en 0. Par le théorème d'interversion intégrale/somme, on a, pour tout $x \in]-1,1[$:

$$F(x) = \int_0^x f(t)dt = \int_0^x \left(\sum_{n=0}^{+\infty} t^n\right) dt = \sum_{n=0}^{+\infty} \left(\int_0^x t^n dt\right) = \sum_{n=0}^{+\infty} \frac{x^{n+1}}{n+1} = \sum_{n=1}^{+\infty} \frac{x^n}{n}.$$

Or, pour $t \in]-1,1[$, $f(t) = \dfrac{1}{1-t}$, donc pour $x \in]-1,1[$

$$\ln(1-x) = -\int_0^x \frac{1}{1-t}dt = -F(x) = -\sum_{n=1}^{+\infty} \frac{x^n}{n}.$$

Par suite, $x \mapsto \ln(1-x)$ est développable en série entière sur $]-1,1[$.

Exercice 13.

Montrer que $\dfrac{1}{(1-x)^2}$ est développable en série entière sur $]-1,1[$.

Proposition 10. *Unicité du développement en série entière*

Soit $\sum a_n x^n$ et $\sum b_n x^n$ des séries entière de la variable réelle. S'il existe $r > 0$ tel que pour tout $x \in]-r,r[$:

$$\sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} b_n x^n,$$

alors, pour tout $n \in \mathbb{N}$, $a_n = b_n$.

b. Séries de Taylor

Définition 7. *Série de Taylor*

Soit $f : U \to \mathbb{R}$ une fonction C^∞. On appelle **série de Taylor** (ou plus précisément, série de Taylor-MacLaurin) de f, la série entière :

$$\sum \frac{f^{(n)}(0)}{n!} x^n.$$

Théorème 8.

Soit $r > 0$ et $f :]-r,r[\to \mathbb{C}$. Si f est développable en série entière, alors f est C^∞ sur $]-r,r[$

et f coïncide avec sa série de Taylor-MacLaurin sur $]-r,r[$ i.e., pour tout $x \in]-r,r[$:

$$f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Remarque 6.

Attention ! La réciproque du théorème précédent est fausse ! Par exemple, la fonction $f : x \mapsto e^{-\frac{1}{x^2}}$ (prolongée par continuité en 0 par $f(0) = 0$) est C^∞ sur $\mathbb{R}$ mais n'est pas égale à sa série de Taylor-MacLaurin au voisinage de 0.

c. Opérations sur les développements en série entière

A FINIR !

2. Développements en série entière usuels

a. L'exponentielle, les fonctions trigonométriques et hyperboliques

Définition 8. *Exponentielle complexe*

Pour $z \in \mathbb{C}$, on pose :

$$\exp(z) = \sum_{n=0}^{+\infty} \frac{z^n}{n!}$$

Théorème 9.

La fonction $x \mapsto e^x$ est développable en série entière sur $\mathbb{R}$ et pour tout $x \in \mathbb{R}$, $e^x = \exp(x)$.

Exercice 14.

1. Montrer que pour tout $z \in \mathbb{C}$,
$$\exp(z + z') = \exp(z)\exp(z').$$
2. En déduire que :
   a) $\exp(z) \neq 0$
   b) $\exp(-z) = \dfrac{1}{\exp(z)}$;
   c) $\overline{\exp(z)} = \exp(\overline{z})$;
   d) $|\exp(z)| = e^{\mathrm{Re}(z)}$;
   e) $|\exp(z)| = 1 \ \Leftrightarrow \ z \in i\mathbb{R}$.

Définition 9. *Les fonctions cosinus et sinus*

On définit les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ notées cos et sin, pour $x \in \mathbb{R}$, par :

$$\cos(x) = \mathrm{Re}(\exp(ix)) \qquad \text{et} \qquad \sin(x) = \mathrm{Im}(\exp(ix)).$$

Exercice 15.

1. Montrer que pour tout $x \in \mathbb{R}$, on a :

$$\cos(x) = \frac{\exp(ix) + \exp(-ix)}{2}, \quad \sin(x) = \frac{\exp(ix) - \exp(-ix)}{2i} \quad \text{et} \quad \cos^2(x) + \sin^2(x) = 1.$$

2. Montrer que pour $z = x + iy \in \mathbb{C}$, $\exp(z) = e^x\,(\cos(y) + i\sin(y))$.

À toujours savoir retrouver ! Les développements en série entières sur $\mathbb{R}$ des fonctions trigonométriques et hyperboliques :

$$\cos(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n}}{(2n)!} \qquad \sin(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$$

$$\mathrm{ch}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n}}{(2n)!} \qquad \mathrm{sh}(x) = \sum_{n=0}^{+\infty} \frac{x^{2n+1}}{(2n+1)!}.$$

Correction.

Montrons que sin est développable en série entière sur $\mathbb{R}$ et déterminons ce développement. Pour tout $x \in \mathbb{R}$, on a :

$$\sin(x) = \frac{\exp(ix) - \exp(-ix)}{2i}.$$

Or, les fonctions $x \mapsto e^{\pm ix}$ sont développables en série entière sur $\mathbb{R}$, et on a, pour tout $x \in \mathbb{R}$:

$$e^{\pm ix} = \sum_{n=0}^{+\infty} \frac{(\pm ix)^n}{n!}$$

La série entière $\sum a_n x^n$ avec $a_n = \dfrac{(ix)^n - (-ix)^n}{n!}$ est de rayon de convergence égal à $+\infty$ car il s'agit de la somme des séries entières $\sum \frac{(ix)^n}{n!}$ et $\sum -\frac{(-ix)^n}{n!}$ qui ont un rayon de convergence égal à $+\infty$. Et de plus, on a, pour tout $x \in \mathbb{R}$:

$$\sum_{n=0}^{+\infty} a_n x^n = \sum_{n=0}^{+\infty} \frac{(ix)^n}{n!} + \sum_{n=0}^{+\infty} -\frac{(-ix)^n}{n!} = \exp(ix) - \exp(-ix) = 2i\sin(x).$$

Ainsi, sin est développable en série entière sur $\mathbb{R}$ et on a, pour tout $x \in \mathbb{R}$:

$$\begin{aligned}
2i\sin(x) &= \sum_{n=0}^{+\infty} a_n x^n \\
&= \sum_{n=0}^{+\infty} \frac{(ix)^n - (-ix)^n}{n!} \\
&= \sum_{n=0}^{+\infty} i^n(1-(-1)^n)\frac{x^n}{n!} \\
&= \sum_{n \text{ impair}} 2i^n \frac{x^n}{n!} \\
&= \sum_{k=0}^{+\infty} 2 \underbrace{i^{2k+1}}_{=i((i)^2)^k = i(-1)^k} \frac{x^{2k+1}}{(2k+1)!} \\
&= \sum_{k=0}^{+\infty} 2i(-1)^k \frac{x^{2k+1}}{(2k+1)!},
\end{aligned}$$

d'où

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

b. Les fonctions $x \mapsto (1+x)^\alpha$

Théorème 10.

Soit $\alpha \in \mathbb{R}$. Alors la fonction $x \mapsto (1+x)^\alpha$ est développable en série entière sur $]-1,1[$ et on a, pour tout $x \in]-1,1[$:

$$(1+x)^\alpha = \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} x^n.$$

À toujours savoir retrouver ! Les développements en série entières sur $]-1,1[$ de :

$$\frac{1}{(1-x)^{k+1}} = \sum_{n=0}^{+\infty} \binom{n+k}{k} x^n \qquad \ln(1+x) = \sum_{n=0}^{+\infty} \frac{(-1)^{n+1} x^n}{n}$$

$$\arctan(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n x^{2n+1}}{2n+1} \qquad \arcsin(x) = \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} \frac{x^{2n+1}}{2n+1}.$$

Correction Pour le arcsinus !.

On se rappelle que pour tout $x \in]-1,1[$, on a :

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-\frac{1}{2}}.$$

> *Remarque :* il se serait bien de savoir démontrer ce résultat... donc : Exercice, montrer l'égalité précédente !
> Indication : on sait que arcsin est la fonction réciproque de sin restreinte à $[-\pi/2, \pi/2]$ (la fonction sin est bien une bijection de $[-\pi/2, \pi/2]$ dans $[-1,1]$ car elle est strictement croissante sur $[-\pi/2, \pi/2]$). De plus, on connaît (ou on sait retrouver) la formule qui relie la dérivée de la réciproque d'une fonction et la dérivée de la fonction. Et on conclut grâce à la formule de trigonométrie la plus connue du monde qui relie le cosinus et le sinus !

La fonction $y \mapsto (1+y)^{-\frac{1}{2}}$ est développable en série entière sur $]-1,1[$ et on a, pour tout $y \in]-1,1[$:

$$\begin{aligned}
(1+y)^{-\frac{1}{2}} &= \sum_{n=0}^{+\infty} \frac{-\frac{1}{2}(-\frac{1}{2}-1)...(-\frac{1}{2}-n+1)}{n!} y^n \\
&= \sum_{n=0}^{+\infty} \frac{(-1)^n \times 1 \times 3 \times 5 \times ... \times (2n-1)}{2^n n!} y^n \\
&= \sum_{n=0}^{+\infty} \frac{(-1)^n (2n)!}{(2 \times 4 \times ... \times 2n) 2^n n!} y^n \\
&= \sum_{n=0}^{+\infty} \frac{(-1)^n (2n)!}{4^n (n!)^2} y^n
\end{aligned}$$

Ainsi, pour tout $x \in]-1,1[$, on a $y = -x^2 \in]-1,0[\subset]-1,1[$, donc :

$$\begin{aligned}
(1-x^2)^{-\frac{1}{2}} &= \sum_{n=0}^{+\infty} \frac{(-1)^n (2n)!}{4^n (n!)^2} (-x^2)^n \\
&= \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} x^{2n}
\end{aligned}$$

Il en résulte que $x \mapsto (1-x^2)^{-\frac{1}{2}}$ est développable en série entière sur $]-1,1[$. De plus, le rayon de convergence R de la série entière associée à ce développement est $R = 1$ (on peut obtenir ce résultat en utilisant la règle de D'Alembert pour les séries **tout court** puisqu'il s'agit d'une série entière lacunaire).

Ainsi, on peut primitiver sur $]-1,1[$ la somme $S : x \mapsto \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} x^{2n}$, et on a, pour tout $x \in]-1,1[$:

$$\begin{aligned}
\int_0^x \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} t^{2n} dt &= \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} \int_0^x t^{2n} dt \\
&= \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} \frac{x^{2n+1}}{2n+1}
\end{aligned}$$

Ainsi, en remarquant que deux primitives sont égales à une constante près et que la primitive précédente s'annule en 0 tout comme arcsin(0), on obtient, pour tout $x \in]-1,1[$:

$$\arcsin = \sum_{n=0}^{+\infty} \frac{(2n)!}{4^n (n!)^2} \frac{x^{2n+1}}{2n+1}.$$

et donc que arcsin est développable en série entière sur $]-1,1[$.

Chapitre VIII

Fonctions vectorielles

Table des matières

Dans ce chapitre, I désigne un intervalle de R ; $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou le corps $\mathbb{C}$ et $(E, \|\cdot\|)$ désigne un espace vectoriel normé sur $\mathbb{K}$ de dimension finie $n \in \mathbb{N}^*$.

Partie A
Dérivation des fonctions vectorielles

1. Dérivée en un point

a. Définition et premières propriétés

Définition 1. *Dérivabilité et vecteur dérivé*

Soit $f : I \to E$ une fonction et $t_0 \in I$. On dit que f est **dérivable** en t_0 si l'application $\tau_{t_0} : I \setminus \{t_0\} \to E$, appelée **taux d'accroissement de f en t_0** et définie par :

$$t \mapsto \frac{f(t) - f(t_0)}{t - t_0},$$

admet une limite $\ell \in E$ en t_0.
Cette limite est appelée **vecteur dérivé** de f en t_0 et est notée usuellement

$$f'(t_0) \quad \text{ou} \quad \frac{df}{dt}(t_0).$$

Proposition 1. *Développement limité à l'ordre 1*

Soit $f : I \to E$ et $t_0 \in I$. La fonction f est dérivable en t_0 si, et seulement si, elle admet une développement limité à l'ordre 1 en t_0 i.e. s'il existe un vecteur $\ell \in E$ et une fonction $\varepsilon : I \to E$ avec $\lim\limits_{t \to t_0} \varepsilon(t) = 0_E$ telle que, pour tout $t \in I$:

$$f(t) = f(t_0) + (t - t_0)\ell + (t - t_0)\varepsilon(t).$$

Dans ce cas, on a $f'(t_0) = \ell$.

Proposition 2.

Soit $f : I \to E$ et $t_0 \in I$. Si f est dérivable en t_0, alors f est continue en t_0.

On rappelle la définition suivante :

Définition-Proposition 2. *Fonctions composantes*

Soit $f : I \to E$ et $\mathcal{B} = (e_1, ..., e_n)$ une base de E. Il existe un unique n-uplet $(f_1, ..., f_n)$ de fonctions de I dans $\mathbb{R}$ appelées **fonctions composantes** de f, qui vérifie, pour tout $t \in I$:

$$f(t) = \sum_{i=1}^{n} f_i(t)e_i.$$

Proposition 3.

Soit $f : I \to E$, $t_0 \in I$ et $\mathcal{B} = (e_1, ..., e_n)$ une base de E. La fonction f est dérivable en t_0 si, et seulement si, pour tout $i \in [\![1, n]\!]$, la fonction composante f_i de f est dérivable en t_0.

Dans ce cas, on a :

$$f'(t_0) = \sum_{i=1}^{n} f_i'(t_0)e_i.$$

Exemple 1.

Soit $f : [-1, 1] \to \mathbb{R}^2$ définie par $f : t \mapsto (\arctan(t), \arcsin(t))$. Alors f est dérivable en tout point de $]-1, 1[$ et on a, pour $t_0 \in]-1, 1[$,

$$f'(t_0) = (\frac{1}{1+t^2}, \frac{1}{\sqrt{1-t^2}}).$$

Exercice 1.

On considère la fonction $f : t \mapsto \begin{pmatrix} \cos(t^2) & \sin(\frac{1}{1-t}) \\ 2^{t^3+1} & \frac{1}{\ln(t^2)^2+1} \end{pmatrix}$.

Quelle est le domaine de définition de f ? En quels points cette fonction est-elle dérivable et quel est le vecteur dérivée dans ce cas ?

b. Dérivée à gauche et à droite en un point

Définition 3. *Vecteur dérivée à gauche/à droite*

Soit $f : I \to E$, $t_0 \in I$. On dit que f est **dérivable à gauche (resp. à droite) en** t_0 si le taux d'accroissement τ_{t_0} de f admet une limite à gauche (resp. à droite) en t_0.

Dans cette limite est appelée **vecteur dérivé à gauche (resp. à droite)** de f en t_0 et est notée

$$f_g'(t_0) \quad (\text{resp. } f_d'(t_0))$$

c. Fonction dérivée

Définition 4. *Fonction dérivée*

Soit $f : I \to E$. On dit que f est **dérivable sur** I si elle est dérivable en tout point de I. On note alors f' (ou encore $\frac{df}{dt}$) la fonction de I dans E définie par :

$$f' : t \mapsto f'(t).$$

Exercice 2.

1. En utilisant le théorème des accroissements finis, montrer qu'une fonction $f : I \to \mathbb{K}$ est constante si, et seulement si, elle est dérivable sur $\mathring{I}$ et sa fonction dérivée f' est nulle sur $\mathring{I}$.
2. En déduire qu'une fonction $f : I \to E$ est constante si, et seulement si, elle est dérivable sur $\mathring{I}$ et sa fonction dérivée f' est nulle sur $\mathring{I}$.

2. Opérations sur les fonctions dérivables

a. Combinaisons linéaires

Proposition 4.

Soit $f, g : I \to E$, $\lambda, \mu \in \mathbb{K}$ et $t_0 \in I$. Si f et g sont dérivables en t_0, alors $\lambda f + \mu g$ est dérivable en t_0 et on a :

$$(\lambda f + \mu g)'(t_0) = \lambda f'(t_0) + \mu g'(t_0).$$

De plus, si f et g sont dérivables sur I, on a $(\lambda f + \mu g)' = \lambda f' + \mu g'$.

Remarque 1.

On en déduit que l'ensemble des fonctions de I dans E dérivables sur I est un espace vectoriel sur K et que l'application $f \mapsto f'$ est une application linéaire de l'ensemble des fonctions de I dans E dérivables sur I vers l'ensemble des fonctions de I dans E.

b. Compositions

Proposition 5. *Composition de fonctions*

Soit $f : I \to E$, $\varphi : J \to \mathbb{R}$ et $t_0 \in J$ avec $\varphi(t_0) \in I$. Si φ est dérivable en t_0 et f est dérivable en $\varphi(t_0)$, alors $f \circ \varphi$ est dérivable en t_0 et on a :

$$(f \circ \varphi)'(t_0) = \varphi'(t_0).f'(\varphi(t_0)).$$

De plus, si φ est dérivable sur J avec $\varphi(J) \subset I$ et f est dérivable sur I, alors on a $(f \circ \varphi)' = \varphi'.f' \circ \varphi$.

Proposition 6. *Composition par une application linéaire*

Soit F un espace vectoriel normé sur $\mathbb{K}$, $f : I \to E$, $L : E \to F$ une application linéaire et $t_0 \in I$. Si f est dérivable en t_0, alors $L \circ f$ est dérivable en t_0 et on a :

$$(L \circ f)'(t_0) = L(f'(t_0)).$$

De plus, si f est dérivable sur I, alors on a $(L \circ f)' = L \circ f'$.

Démonstration.

On suppose f dérivable en t_0. Comme E est de dimension finie, L est continue sur E. Notons τ_{t_0} le taux d'accroissement de f. On a, pour tout $t \in I$ avec $t \neq t_0$, par linéarité et continuité de L :

$$\frac{L \circ f(t) - L \circ f(t_0)}{t - t_0} = L\left(\frac{f(t) - f(t_0)}{t - t_0}\right) \xrightarrow[t \to t_0]{} L(f'(t_0)).$$

D'où le résultat. $\square$

c. Produit de fonctions

Proposition 7.

Soit E, F des espaces vectoriels normés de dimension finie, G un espace vectoriel normé, $B : E \times F \to G$ une application bilinéaire, $f : I \to E$, $g : I \to F$ des fonctions et $t_0 \in I$. Si f et g sont dérivables en t_0, alors $B(f, g) : t \mapsto B(f(t), g(t))$ est dérivable en t_0 et on a :

$$(B(f, g))'(t_0) = B(f'(t_0), g(t_0)) + B(f(t_0), g'(t_0)).$$

De plus, si f et g sont dérivables sur I, alors on a $(B(f, g))' = B(f', g) + B(f, g')$.

Démonstration.

On suppose que f et g sont dérivables en t_0. Alors on a, par bilinéarité de B, pour tout $t \in I$ avec $t \neq t_0$:

$$\begin{aligned} B(f,g)(t) - B(f,g)(t_0) &= B(f(t), g(t)) - B(f(t_0), g(t)) + B(f(t_0), g(t)) - B(f(t_0), g(t_0)) \\ &= B(f(t) - f(t_0), g(t)) + B(f(t_0), g(t) - g(t_0)). \end{aligned}$$

Comme E et F sont de dimension finie, l'application bilinéaire B est continue. Par suite, on a :

$$\begin{aligned} \frac{B(f,g)(t) - B(f,g)(t_0)}{t - t_0} &= \underbrace{B\left(\frac{f(t) - f(t_0)}{t - t_0}, g(t)\right)}_{\xrightarrow[n \to +\infty]{} B(f'(t_0), g(t_0))} + \underbrace{B\left(f(t_0), \frac{g(t) - g(t_0)}{t - t_0}\right)}_{\xrightarrow[n \to +\infty]{} B(f(t_0), g'(t_0))} \\ &\xrightarrow[n \to +\infty]{} B(f'(t_0), g(t_0)) + B(f(t_0), g'(t_0)). \end{aligned}$$

D'où le résultat. $\square$

Exemple 2. *Produit scalaire de fonctions*

Soit E un espace euclidien sur $\mathbb{K}$ de produit scalaire noté $(\cdot|\cdot)$ et $f, g : I \to E$ des fonctions dérivables sur I.
Alors l'application $(f|g) : t \mapsto (f(t)|g(t))$ est dérivable sur I et

$$(f|g)' = (f'|g) + (f|g').$$

Exercice 3.

Soit E un espace euclidien de produit scalaire noté $(\cdot|\cdot)$, de norme associée notée $\|\cdot\|$ et $f : I \to E$ une fonction dérivable sur I.

1. Discuter de la dérivabilité de la fonction $\|f\| : t \mapsto \|f(t)\|$ et déterminer sa dérivée aux points de dérivabilité.
2. On suppose que pour tout $t \in I$, $f(t) \in S(0, 1)$. Montrer que, pour tout $t \in I$, $f(t)$ et $f'(t)$ sont orthogonaux.

3. Fonctions de classe C^k

a. Définitions

Définition 5. *Fonctions de classe C^1*

Soit $f : I \to E$. On dit que f est de **classe** C^1 sur I si f est dérivable sur I et sa fonction dérivée f' est continue sur I. On note $C^1(I, E)$ l'ensemble des fonctions de classe C^1 sur I.

On peut généraliser cette notion par récurrence :

Définition 6. *Fonctions de classe C^k*

Soit $f : I \to E$. On dit que f est :

- **de classe** C^0 **sur** I si f est continue sur I ;
- **de classe** C^k **sur** I, pour $k \in \mathbb{N}^*$, si f est dérivable sur I et que f' est de classe C^{k-1} sur I ;
- **de classe** C^∞ si f est de classe C^k pour tout $k \in \mathbb{N}$.

Soit $k \in \mathbb{N} \cup \{\infty\}$. On note $C^k(I, E)$ l'ensemble des fonctions de classe C^k sur I.

Proposition 8.

Soit $k \in \mathbb{N} \cup \{\infty\}$. L'ensemble $C^k(I, E)$ est un espace vectoriel sur $\mathbb{K}$. De plus, si $k \geq 1$, l'application $D : C^k(I, E) \to C^{k-1}(I, E)$ telle que $D : f \mapsto f'$ est une application linéaire.

Remarque 2.

On a $C^\infty(I,E) = \bigcap_{k\in\mathbb{N}} C^k(I,E)$.

b. Opérations sur les fonctions de classe C^k

Proposition 9. *Opérations usuelles*

Soit $k \in \mathbb{N} \cup \{\infty\}$.
Soit F un espace vectoriel normé de dimension finie.
Soit $f, g : I \to E$ et $\varphi : J \to \mathbb{R}$ avec $\varphi(J) \subset I$.
Soit $L : E \to F$ une application linéaire.

- Si f, g sont de classe C^k sur I, alors, pour tous $\lambda, \mu \in \mathbb{K}$, $\lambda f + \mu g$ est de classe C^k sur I ;
- Si f, φ sont de classe C^k sur I, J respectivement, alors, $f \circ \varphi$ est de classe C^k sur J ;
- Si f est de classe C^k sur I, alors, $L \circ f$ est de classe C^k sur I.

Proposition 10. *Formule de Leibniz*

Soit $k \in \mathbb{N} \cup \{\infty\}$.
Soit E, F des espaces vectoriels normés de dimension finie, G un espace vectoriel normé, $B : E \times F \to G$ une application bilinéaire et $f : I \to E$, $g : I \to F$. Si f et g sont de classe C^k sur I, alors $B(f,g)$ est de classe C^k sur I et on a, pour tout $p \leq k$:

$$(B(f,g))^{(p)} = \sum_{i=0}^{p} \binom{p}{i} B(f^{(i)}, g^{(p-i)}).$$

Partie B

Intégration des fonctions vectorielles sur un segment

Dans cette partie, a, b désignent des réels tels que $a < b$.

a. Définition

Lemme 1.

Soit $f \in C_{pm}([a,b], E)$, $\mathcal{B} = (e_1, ..., e_n)$ une base de E et $(f_1, ..., f_n)$ les fonctions composantes de f dans $\mathcal{B}$. Le vecteur $\displaystyle\sum_{i=1}^{n} \left(\int_a^b f_i(t)dt \right) e_i$ ne dépend pas de la base $\mathcal{B}$.

Grâce au lemme précédent, on peut donner le sens suivant à l'intégrale sur un segment d'une fonction continue par morceaux à valeurs vectorielles :

Définition 7. *Intégrale d'une fonction continue par morceaux*

Soit $f \in C_{pm}([a,b], E)$, $\mathcal{B} = (e_1, ..., e_n)$ une base de E et $(f_1, ..., f_n)$ les fonctions composantes de f dans $\mathcal{B}$. On note $\displaystyle\int_a^b f(t)dt$, $\displaystyle\int_a^b f$ ou encore $\displaystyle\int_{[a,b]} f$ le vecteur :

$$\int_a^b f(t)dt = \sum_{i=1}^{n} \left(\int_a^b f_i(t)dt \right) e_i.$$

b. Propriétés de l'intégrale

Proposition 11.

On a les propriétés suivantes :

— $f \mapsto \displaystyle\int_a^b f$ est linéaire (de quel espace dans quel espace ?). Soit $f \in C_{pm}([a,b], E)$.

— pour $u : E \to F$ une application linéaire, $u(\displaystyle\int_a^b f) = \int_a^b u(f)$.

— pour $a < b < c$, $\displaystyle\int_a^c f = \int_a^b f + \int_b^c f$.

Proposition 12.

Soit $f \in C_{pm}([a,b],E)$. On a :

$$\| \int_a^b f(t)dt \| \leq \int_a^b \|f(t)\|dt \leq (b-a)\|f\|_\infty.$$

c. Sommes de Riemann

On rappelle et on généralise les résultats sur les sommes de Riemann (ici dans le cas particulier d'une subdivision régulière)

Définition 8. *Somme de Riemann*

Soit $f \in C_{pm}([a,b],E)$ et $n \in \mathbb{N}^*$. On appelle **somme de Riemann d'ordre n de f** le vecteur

$$\frac{b-a}{n}\sum_{k=0}^{n-1} f(a+k\frac{b-a}{n})$$

Théorème 1.

Soit $f \in C_{pm}([a,b],E)$. Alors les sommes de Riemann d'ordre n convergent vers $\int_a^b f$ quand n tend vers l'infini i.e.

$$\frac{b-a}{n}\sum_{k=0}^{n-1} f(a+k\frac{b-a}{n}) \xrightarrow[n\to+\infty]{} \int_a^b f.$$

Exercice 4.

Rappels de sup. Déterminer les limites des suite de termes généraux suivants :

$$u_n = \sum_{k=0}^{n-1} \frac{n}{n^2+k^2} \qquad v_n = \sum_{k=0}^{n-1} \frac{k}{n^2+k^2} \qquad w_n = \sum_{k=0}^{n-1} \frac{1}{\sqrt{n^2+2kn}}.$$

1. Primitives et intégrales

a. Extension de la définition de l'intégrale

On rappelle qu'une fonction est continue par morceaux sur I (autrement dit, $f \in C_{pm}(I,E)$) si elle est continue par morceaux sur tout segment de I.

Définition 9. *Extension de l'intégrale*

Soit $f \in C_{pm}(I,E)$ et $a,b \in I$. On pose

$$\int_a^b f(t)dt = \begin{cases} \int_a^b f(t)dt & \text{si } a < b \\ 0 & \text{si } a = b \\ -\int_b^a f(t)dt & \text{si } a > b. \end{cases}$$

Proposition 13. *Relation de Chasles*

Soit $f \in C_{pm}(I,E)$ et $a,b,c \in I$ avec $a \leq b \leq c$. Alors on a :

$$\int_a^c f(t)dt = \int_a^b f(t)dt + \int_b^c f(t)dt.$$

b. Primitives des fonctions continues

Définition 10. *Primitive*

Soit $f : I \to E$ une fonction continue. On dit qu'une fonction $g : I \to E$ est **une primitive** de f sur I si g est dérivable sur I et $g' = f$.

Proposition 14.

Soit $f,g,h \in C(I,E)$. Si g et sont des primitives de f sur I, alors $g-h$ est une fonction constante sur I.

Théorème 2. *Théorème fondamental de l'Analyse*

Soit $f \in C(I,E)$ et $t_0 \in I$.

i) L'application $g : I \to E$ définie par :

$$g : t \mapsto \int_{t_0}^t g(x)dx$$

est l'unique primitive de f qui s'annule en t_0.

ii) Pour toute primitive F de f sur I, on a, pour $a,b \in I$:

$$\int_a^b f(t)dt = F(b) - F(a).$$

c. Inégalité des accroissements finis

Théorème 3. *Inégalités des accroissements finis*

Soit $f \in C(I, E)$ une fonction de classe C^1 sur $\mathring{I}$. S'il existe $M \geq 0$ tel que, pour tout $t \in \mathring{I}$,

$$\|f'(t)\| \leq M$$

alors, pour tout $a, b \in I$, on a :

$$\|f(b) - f(a)\| \leq M(b - a).$$

Démonstration.

On suppose qu'il existe $M \geq 0$ tel que, pour tout $t \in \mathring{I}$, $\|f'(t)\| \leq M$.

— *Cas particulier :* $a, b \in \mathring{I}$. Alors $f \in C^1([a, b], E)$ donc, d'après le théorème fondamental de l'Analyse

$$\|f(b) - f(a)\| = \|\int_a^b f'(t)dt\| \leq \int_a^b \|f'(t)\|dt \leq M(b - a).$$

— *Cas général :* $a, b \in I$. Soit $x, y \in \mathring{I}$. D'après le cas précédent, on a :

$$\|f(y) - f(x)\| =\leq M(y - x).$$

Et donc en passant à la limite quand $x \to a$ puis quand $y \to b$, on obtient le résultat. □

d. Exercices de révision

Exercice 5.

Déterminer une primitive pour chacune des fonctions suivantes :

$$t \mapsto \frac{t^3 + t^2 - 1}{t^2 + 2t + 2} \quad t \mapsto \frac{1}{1 + t^3} \quad t \mapsto \left(\frac{2t}{\sqrt{1 - t^4}}, e^t \cos(t)\right).$$

Correction.

1. La fonction $t \mapsto t^2 + 2t + 2$ ne s'annule pas sur $\mathbb{R}$ et on a, par division euclidienne :

$$t^3 + t^2 - 1 = (t^2 + 2t + 2)(t - 1) + 1$$

donc

$$\frac{t^3 + t^2 - 1}{t^2 + 2t + 2} = \frac{(t^2 + 2t + 2)(t - 1) + 1}{t^2 + 2t + 2} = t - 1 + \frac{1}{t^2 + 2t + 2}$$

et de plus, $t^2 + 2t + 2 = 1 + (t + 1)^2$, donc :

$$\int \frac{1}{t^2 + 2t + 2} dt = \int \frac{1}{1 + (t + 1)^2} dt = \arctan(t + 1)$$

Par suite,

$$\int \frac{t^3+t^2-1}{t^2+2t+2}dt = \frac{t^2}{2} - t + \arctan(t+1) + \text{ constante.}$$

Partie C
Formules de Taylor

Dans cette partie, n désigne un entier naturel.

1. Formule de Taylor avec reste intégrale

Théorème 4. *Formule de Taylor avec reste intégrale*

Soit $f \in C^{n+1}(I, E)$. Pour $a, b \in I$, on a :

$$f(b) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(b-a)^i + R_n,$$

où

$$R_n = \int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n dt.$$

Exercice 6.

Montrer que le reste R_n peut s'écrire :

$$R_n = (b-a)^{n+1} \int_0^1 \frac{f^{(n+1)}(a+t(b-a))}{n!}(1-t)^n dt.$$

2. Inégalité de Taylor-Lagrange

Théorème 5. *Inégalité de Taylor-Lagrange*

Soit $f \in C^{n+1}(I, E)$. S'il existe $M \geq 0$ tel que, pour tout $t \in I$, $\|f^{(n+1)}(t)\| \leq M$, alors, pour tous $a, b \in I$, on a :

$$\|f(b) - \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(b-a)^i\| \leq M \frac{|b-a|^{n+1}}{(n+1)!}.$$

3. Formule de Taylor-Young

Théorème 6. *Formule de Taylor-Young*

Soit $f \in C^n(I, E)$ et $x_0 \in I$. Alors on a :

$$f(x) = \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i + o((x - x_0)^n).$$

Chapitre IX
Endomorphismes d'un espace euclidien

Table des matières

Dans ce chapitre, E désigne un espace vectoriel réel (pas forcément de dimension finie).

Partie $*$

Rappels sur les espaces préhilbertiens réels

1. Rappels sur le produit scalaire

a. Produit scalaire

Définition $*$1. Produit scalaire

Une application $\varphi : E \times E \to \mathbb{R}$ est appelée **produit scalaire** sur E si φ est une forme bilinéaire symétrique définie positive de $E \times E$ dans $\mathbb{R}$.

Remarque $*$1.

Usuellement, on note un produit scalaire de deux vecteurs $(u|v)$, $\langle u|v \rangle$, $\langle u, v\rangle$ ou même $u \cdot v$. Dans la suite, on utilisera principalement la notation $(\cdot|\cdot)$.

Exercice $*$1.

Donner des exemples de produits scalaires sur les espaces : $\mathbb{R}^n$, $\ell^2(\mathbb{R})$, $C(I)$ où I est un intervalle fermé borné de $\mathbb{R}$ et $M_n(\mathbb{R})$.

Correction.

— pour $x = (x_1, ..., x_n)$ et $y = (y_1, ..., y_n)$ des vecteurs de $\mathbb{R}^n$, le produit scalaire canonique est donné par :

$$(x|y) = \sum_{i=1}^{n} x_i y_i$$

— pour $u = (u_n)_{n\in\mathbb{N}}$ et $v = (v_n)_{n\in\mathbb{N}}$ des vecteurs de $\ell^2(\mathbb{R})$, le produit scalaire canonique est donné par :

$$(u|v) = \sum_{i=0}^{\infty} u_i v_i$$

— pour f et g des vecteurs de $C(I)$, le produit scalaire canonique est donné par :

$$(f|g) = \int_I f(t)g(t)dt$$

— pour A et B des vecteurs de $M_n(\mathbb{R})$, le produit scalaire canonique est donné par :

$$(A|B) = \mathrm{Tr}({}^tAB) = \sum_{1\leq i,j\leq n} a_{i,j}b_{i,j}$$

Définition ∗2. Espace préhilbertien

Soit $(\cdot|\cdot)$ un produit scalaire sur E. Le couple $(E,(\cdot|\cdot))$, ou simplement E s'il n'y a pas d'ambiguïté, est appelé **espace préhilbertien**.
Si de plus E est de dimension finie, on dit que $(E,(\cdot|\cdot))$ est un **espace euclidien**.

Exemple ∗1. Important

L'espace vectoriel $M_{n,1}(\mathbb{R})$ muni du produit scalaire $\langle\cdot,\cdot\rangle$ où, pour tous $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$,

$$\langle X,Y\rangle = \sum_{i=1}^n x_iy_i = {}^tXY,$$

est un espace euclidien. Le produit scalaire précédent est souvent désigné comme le produit scalaire **canonique** de $M_{n,1}$.

b. Norme associée à un produit scalaire

Définition ∗3. Norme associé au produit scalaire

Soit E un espace préhilbertien réel. On note $\|\cdot\|$ et on appelle **norme associée au produit scalaire** $(\cdot|\cdot)$, l'application :

$$\|\cdot\| : \left| \begin{array}{ccc} E & \to & \mathbb{R}_+ \\ x & \mapsto & \|x\| = \sqrt{(x|x)}. \end{array}\right.$$

De plus, on note $d : E\times E \to \mathbb{R}_+$ et on appelle distance associée à $\|\cdot\|$, l'application définie, pour $x,y\in E$, par :

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

Théorème ∗1. Inégalité de Cauchy-Schwarz

Soit E un espace préhilbertien réel. On a l'inégalité suivante, pour tous $x,y\in E$:

$$|(x|y)| \leq \|x\|\|y\|$$

Et de plus, on a $|(x|y)| = \|x\|\|y\|$ si, et seulement si, x et y sont colinéaires.

Démonstration.

Soit $x, y \in E$ et $t \in \mathbb{R}$. On considère la fonction $f : \mathbb{R} \to \mathbb{R}$ définie, pour $t \in \mathbb{R}$, par :

$$f(t) = \|tx + y\|^2 = (tx + y|tx + y).$$

Alors on a, pour tout $t \in \mathbb{R}$,

$$f(t) = t^2(x|x) + 2(x|y) + (y|y) = t^2\|x\|^2 + 2(x|y) + \|y\|^2.$$

Pour $x \neq 0_E$, f est une fonction polynomiale de degré 2 et on a $f(t) \geq 0$ pour tout $t \in \mathbb{R}$, donc son discriminant Δ est négatif i.e.

$$\Delta = 4(x|y)^2 - 4\|x\|^2\|y\|^2 \leq 0.$$

Par suite,

$$(x|y)^2 \leq (\|x\|\|y\|)^2$$

Et cette inégalité est trivialement vraie pour $x = 0_E$ - et c'est même une égalité dans ce cas. La fonction racine étant croissante sur $\mathbb{R}_+$, il en résulte que pour tous $x, y \in E$:

$$|(x|y)| \leq \|x\|\|y\|.$$

De plus, si $x \neq 0_E$ et y colinéaire à x, alors il existe $\lambda \in \mathbb{R}$ tel que $y = \lambda x$ et

$$|(x|y)| = |\lambda|\|x\|^2 = \pm\|x\|.|\lambda|\|x\| = \|x\|\|y\|.$$

Réciproquement, pour $x \neq 0_E$, si l'égalité est vérifiée, alors $\Delta = 4(x|y)^2 - 4\|x\|^2\|y\|^2 = 0$ donc f possède une racine λ i.e.

$$(\lambda x + y|\lambda x + y) = f(\lambda) = 0.$$

Donc par définie positivité du produit scalaire, $\lambda x + y = 0_E$; d'où x et y sont colinéaires. $\square$

Proposition ∗1.

Soit E un espace préhilbertien réel. La norme $\|\cdot\|$ associée au produit scalaire est une norme sur E.

Démonstration.

- $\|\cdot\|$ est positive par positivité du produit scalaire).
- Soit $x \in E$. Si $\|x\| = 0$, alors $(x|x) = 0$ donc $x = 0_E$ car le produit scalaire est défini positif.
- Soit $x \in E$ et $\lambda \in \mathbb{R}$. Par bilinéarité de $(\cdot|\cdot)$, on a :

$$\|\lambda x\|^2 = (\lambda x|\lambda x) = \lambda^2(x|x) = \lambda^2\|x\|^2,$$

  d'où l'homogénéité.
- Soit $x, y \in E$. On a :

$$\|x + y\|^2 = (x + y|x + y) = \|x\|^2 + 2(x|y) + \|y\|^2,$$

donc, d'après l'inégalité de Cauchy-Schwarz :

$$\|x+y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 .$$

d'où l'inégalité triangulaire.

□

Exercice ∗2.

Soit $x, y \in E$. Démontrer les égalités suivantes :

1. $d(x,y) = \sqrt{\|x\|^2 + \|y\|^2 - 2(x|y)}$.
2. $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$ *(Identité du parallélogramme).*
3. $(x|y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$ *(Identité de polarisation).*

Correction.

Tout cet exerice repose sur les égalités :

$$\|x \pm y\|^2 = (x \pm y | x \pm y) = \|x\|^2 \pm 2(x|y) + \|y\|^2.$$

1. On a :
$$\begin{array}{rcl} d(x,y)^2 & = & \|x-y\|^2 \\ & = & \|x\|^2 - 2(x|y) + \|y\|^2 \end{array}$$
d'où le résultat.
2. On a :
$$\begin{array}{rcl} \|x+y\|^2 + \|x-y\|^2 & = \|x\|^2 + 2(x|y) + \|y\|^2 + (\|x\|^2 - 2(x|y) + \|y\|^2) & \\ & = & 2\|x\|^2 + 2\|y\|^2 \\ & = & 2\left(\|x\|^2 + \|y\|^2\right) \end{array}$$
3. On a :
$$\begin{array}{rcl} \|x+y\|^2 - \|x-y\|^2 & = \|x\|^2 + 2(x|y) + \|y\|^2 - (\|x\|^2 - 2(x|y) + \|y\|^2) & \\ & = & 4(x|y) \end{array}$$
d'où le résultat.

2. Rappels et compléments sur l'orthogonalité

Définition ∗4. Orthogonalité

Soit E un espace préhilbertien réel.

— Soit $x, y \in E$. On dit que x et y sont **orthogonaux** et on note $x \perp y$ si $(x|y) = 0$.

— Soit $A \subset E$. On note $A^\perp$ et on appelle **orthogonal** de A, l'ensemble :

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

— Soit F, G des sous-espaces vectoriels de E. On dit que F et G sont **orthogonaux** et on note $F \perp G$ si, pour tout $x \in F$ et tout $y \in G$, $(x|y) = 0$. Autrement dit si $F \subset G^\perp$ (ou de manière équivalente, $G \subset F^\perp$).

Exercice ∗3.

Soit E un espace préhilbertien réel et $A \subset E$.

— Montrer que $A^\perp$ est un sous-espace vectoriel de E de deux façons.

— Montrer que $A^\perp$ est un fermé de E pour la norme $\|\cdot\|$.

Correction.

1. Pour la première façon, il s'agit d'utiliser la définition de sous-espace vectoriel.
   Pour la deuxième, on remarque que $A^\perp = \bigcap_{a \in A} \mathrm{Ker}(\varphi_a)$ où, pour $a \in A$, $\varphi_a : x \mapsto (a|x)$, donc $A^\perp$ est un sous-espace vectoriel de E comme intersection de sous-espaces vectoriels de E.

2. On reprend l'égalité $A^\perp = \bigcap_{a \in A} \mathrm{Ker}(\varphi_a)$ et on remarque que :
   — pour $a \in A$ et $x \in X$, d'après l'inégalité de Cauchy-Schwarz,
   $$|f_a(x)| = |(a|x)| \leq \|a\|\|x\|.$$
   Donc f_a est $\|a\|$-lipschitzienne et donc continue pour $\|\cdot\|$.
   — $\mathrm{Ker}(\varphi_a) = \varphi_a^{-1}(\{0\})$ donc $\mathrm{Ker}(\varphi_a)$ est un fermé pour $\|\cdot\|$ comme image réciproque d'un fermé par une application continue.
   Par suite, $A^\perp$ est fermé pour $\|\cdot\|$ comme intersection de fermé.

Exemple ∗2.

— Dans un espace préhilbertien réel E, on a : $E^\perp = \{0_E\}$ et $\{0_E\}^\perp = E$.

— Dans $\mathbb{R}^2$ muni de son produit scalaire canonique, $\{(1,-1)\}^\perp = \{(x,x) \mid x \in \mathbb{R}\}$.

— Dans $\mathbb{R}^3$ muni de son produit scalaire canonique,

$$\mathrm{Vect}((1,1,1),(1,2,3)) \perp \mathrm{Vect}(1,-2,1).$$

Définition ∗5. Famille orthogonale/orthonormale

Soit E un espace préhilbertien réel et $\mathcal{F} = (x_i)_{i \in I}$ une famille de vecteurs de E. On dit que $\mathcal{F}$ est :

— **orthogonale** si, pour tous $i, j \in I$ avec $i \neq j$,

$$(x_i|x_j) = 0.$$

— **orthonormale** si, pour tous $i, j \in I$,

$$(x_i|x_j) = \delta_{ij} = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$$

On dit qu'une famille de E est une **base orthonormale** si c'est une base et une famille orthonormale.

Exercice $*$4.

Soit E un espace préhilbertien réel et $n \in \mathbb{N}^*$. Montrer les proposition suivantes :

1. Toute famille orthogonale de vecteurs non nuls de E est libre.
2. On suppose que E admet une base orthonormale $\mathcal{B} = (e_i)_{i\in I}$. Montrer que, pour tous $x, y \in E$ de décompositions dans $\mathcal{B}$, $x = \sum_{i\in I} x_i e_i$ et $y = \sum_{i\in I} y_i e_i$ (sommes finies), on a :
$$(x|y) = \sum_{i\in I} x_i y_i \text{ et } x = \sum_{i\in I} (x|e_i)e_i.$$
3. (Théorème de Pythagore) Soit $(x_1, ...x_n)$ une famille finie orthogonale de vecteurs de E. Alors
$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2.$$

Proposition 1.

Soit $(E, (\cdot|\cdot))$ un espace euclidien de dimension n, $\mathcal{B}$ une base orthonormale de E. Pour tous $x, y \in E$, on a :

$$(x|y) = {}^tXY \ \ (= \langle X, Y\rangle) \text{ où } X = \mathrm{Mat}_{\mathcal{B}}(x), Y = \mathrm{Mat}_{\mathcal{B}}(y).$$

Démonstration.

Soit $x, y \in E$. On note $e_1, ..., e_n$ les vecteurs de $\mathcal{B}$. Alors, si on note $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{i=1}^n y_i e_i$ les décompositions de x, y dans $\mathcal{B}$, par bilinéarité du produit scalaire $(\cdot|\cdot)$ et par orthonormalité de $\mathcal{B}$, on a, $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ et :

$$(x|y) = \sum_{i=1}^{n} x_i y_i = {}^tXY$$

□

Proposition ∗2. Matrice d'un endormrphisme dans une base orthonormale

Soit E un espace euclidien, $u \in \mathcal{L}(E)$ et $\mathcal{B} = (e_1, ..., e_n)$ une base **orthonormale** de E. Si on note $A = \mathrm{Mat}_{\mathcal{B}}(u)$, alors :

$$A = ((u(e_j)|e_i))_{1 \leq i,j \leq n},$$

et, pour tous $x, y \in E$,

$$(u(x)|y) = {}^tX{}^tAY = {}^tYAX \text{ où } X = \mathrm{Mat}_{\mathcal{B}}(x), Y = \mathrm{Mat}_{\mathcal{B}}(y).$$

Démonstration.

On note $A = \mathrm{Mat}_{\mathcal{B}}(u) = (a_{ij})_{1 \leq i,j \leq n}$.

— Soit $j \in [\![1, n]\!]$. D'une part, comme $\mathcal{B}$ est orthonormale, on a $u(e_j) = \sum_{i=1}^{n} (u(e_j)|e_i)e_i$ et d'autre part, par définition de A, $u(e_j) = \sum_{i=1}^{n} a_{ij} e_i$; ainsi, par unicité de la décomposition d'un vecteur dans une base, pour tout $i \in [\![1, n]\!]$, $a_{ij} = (u(e_j)|e_i)$.

— Soit $x, y \in E$. On note $X = \mathrm{Mat}_{\mathcal{B}}(x), Y = \mathrm{Mat}_{\mathcal{B}}(y)$. Alors $\mathrm{Mat}_{\mathcal{B}}(u(x)) = AX$ et donc :

$$(u(x)|y) = {}^tAXY = {}^tX{}^tAY.$$

□

Proposition ∗3. Orthonormalisation de Gram-Schmidt

Soit E un espace préhilbertien réel et $(x_1, ..., x_n)$ une famille libre de E. Alors il existe une famille $(e_1, ..., e_n)$ orthonormale de E telle que, pour tout $k \in [\![1, n]\!]$:

$$\mathrm{Vect}(e_1, ..., e_k) = \mathrm{Vect}(x_1, ..., x_k).$$

Plus précisément, on peut construire un telle famille $(e_1, ..., e_n)$ par le procédé de Gram-Schmidt : pour $k = 1, ..., n$

$$e_k = \frac{\varepsilon_k}{\|\varepsilon_k\|} \text{ où } \varepsilon_k = x_k - \sum_{i=1}^{k-1} (x_k|e_i)e_i.$$

Démonstration.

On raisonne par récurrence finie sur $k \in [\![1, n]\!]$ pour montrer que $\mathrm{Vect}(e_1, ..., e_k) = \mathrm{Vect}(x_1, ..., x_k)$ où les e_i sont donnés par le procédé de Gram-Schmidt :

- *Initialisation.* Pour $k = 1$, la propriété est vraie car

$$e_1 = \frac{\varepsilon_1}{\|\varepsilon_1\|} = \frac{x_1}{\|x_1\|}.$$

- *Hérédité.* Soit $k \in [\![1, n-1]\!]$. On suppose la propriété vraie pour k. On a, par hypothèse de récurrence :

$$\varepsilon_{k+1} = x_{k+1} + \underbrace{\sum_{i=1}^{k} (x_{k+1}|e_i)e_i}_{\in \mathrm{Vect}(x_1, ..., x_k)} \in \mathrm{Vect}(x_1, ..., x_{k+1}).$$

Par suite, $e_{k+1} = \dfrac{\varepsilon_{k+1}}{\|\varepsilon_{k+1}\|} \in \mathrm{Vect}(x_1, ..., x_{k+1})$. De plus, on a, pour $l \in [\![1, k]\!]$,

$$\begin{aligned}
(\varepsilon_{k+1}|\varepsilon_l) &= (x_{k+1} - \sum_{i=1}^{k}(x_{k+1}|e_i)e_i | x_l - \sum_{i=1}^{l-1}(x_l|e_i)e_i) \\
&= (x_{k+1}|x_l) - \sum_{i=1}^{k}(x_{k+1}|e_i)(x_l|e_i) - \sum_{i=1}^{l-1}(x_l|e_i)(x_{k+1}|e_i) + \sum_{i=1}^{l-1}\sum_{j=1}^{k}(x_l|e_i)(x_{k+1}|e_j)\underbrace{(e_i|e_j)}_{=0 \text{ si } j \neq i} \\
&= (x_{k+1}|x_l) - \sum_{i=1}^{k}(x_{k+1}|e_i)(x_l|e_i) - \sum_{i=1}^{l-1}(x_l|e_i)(x_{k+1}|e_i) + \sum_{i=1}^{l-1}(x_l|e_i)(x_{k+1}|e_i) \\
&= (x_{k+1}|x_l) - \sum_{i=1}^{k}(x_{k+1}|e_i)(x_l|e_i) - \sum_{i=1}^{l-1}(x_l|e_i)(x_{k+1}|e_i) + \sum_{i=1}^{l-1}(x_l|e_i)(x_{k+1}|e_i)
\end{aligned}$$

□

Corollaire ∗1.

Soit E un espace préhilbertien réel et F un sous-espace vectoriel de dimension finie. Alors F admet une base orthonormale.

Démonstration.

Comme F est de dimension finie (disons p), il existe une base $\mathcal{B} = (x_1, ..., x_p)$ de F. Alors on orthonormalise cette base grâce au procédé de Gram-Schmidt pour obtenir une base $\mathcal{B}'$ orthonormale de F. □

Exercice ∗5.

On munit $\mathbb{R}^3$ de son produit scalaire canonique. Déterminer, grâce au procédé de Gram-Schmidt, la base orthonormale de $\mathbb{R}^3$ obtenue à partir de la base formée des vecteurs :

$$(1, 1, 1), \quad (1, 2, 3), \quad (1, -2, 1).$$

Remarque ∗2.

Pour $A, B \subset E$ et F un sous-espace vectoriel de E, on a les propriétés suivantes :

— Si $A \subset B$ alors $B^\perp \subset A^\perp$,

— $A \subset (A^\perp)^\perp$.

Définition ∗6. Somme directe orthogonale

Soit E un espace préhilbertien réel et F, G des sous-espaces vectoriels de E. On dit que F et G sont en **somme directe orthogonale** si F et G sont en somme directe et $F \perp G$.

Dans ce cas, on note alors $F \overset{\perp}{\oplus} G$ la somme $F \oplus G$.

Proposition ∗4.

Soit E un espace préhilbertien réel et F un sous-espace vectoriel de E. Les sous-espaces vectoriels F et $F^\perp$ sont en somme directe orthogonale.

Démonstration.

Soit $x \in F \cap F^\perp$. Alors $(\underbrace{x}_{\in F} | \underbrace{x}_{\in F^\perp}) = 0$ donc par définie positivité de $(\cdot|\cdot)$, $x = 0_E$.
De plus, par définition, $F \perp F^\perp$.
Par suite, F et $F^\perp$ sont en somme directe orthogonal. □

Définition ∗7. Supplémentaire orthogonal

Soit E un espace préhilbertien réel et F un sous-espace vectoriel de E. Si $F \overset{\perp}{\oplus} F^\perp = E$, on dit que $F^\perp$ est le supplémentaire orthogonal de E .

Proposition ∗5.

Soit E un espace préhilbertien réel et F, G des sous-espaces vectoriels de E.
- On suppose que F et G sont supplémentaires. Alors $F \perp G$ si, et seulement si, $G = F^\perp$.
- Si $E = F \overset{\perp}{\oplus} F^\perp$, alors $(F^\perp)^\perp = F$.

Démonstration.

- On suppose $F \perp G$. Alors $G \subset F^\perp$. Montrons l'inclusion réciproque.
  Soit $x \in F^\perp = E = F \oplus G$. Alors $x = \underbrace{x_F}_{\in F} + \underbrace{x_G}_{\in G}$ et on a :
  $$\begin{aligned}(x - x_G | x - x_G) &= (x - x_G | x_F)\\ &= \underbrace{(x|x_F)}_{=0} - \underbrace{(x|x_G)}_{=0}\\ &= 0.\end{aligned}$$
  Donc, par définie positivité de $(\cdot|\cdot)$, $x - x_G = 0_E$, d'où $x = x_G \in G$.
  Par suite, $G = F^\perp$.
  La réciproque est immédiate car $F^\perp \perp F$.
- On applique le point précédent à "F" $= F^\perp$ et "G" $= F$. Comme $F^\perp \perp F$, on obtient alors $F = (F^\perp)^\perp$.

□

Exemple ∗3.

On considère les espaces vectoriels suivants munis de leurs produits scalaires canoniques respectifs.
- Dans $\mathbb{R}^3$, on considère le plan $\mathcal{P}$ d'équation $\mathcal{P} : x+y=0$. Alors la droite $\mathcal{P}^\perp = \text{Vect}(1,1,0)$ est le supplémentaire orthogonal de $\mathcal{P}$.
- Dans $\ell^2(\mathbb{N})$, on considère le sous-espace vectoriel F des suites stationnaires en 0. Alors $F^\perp = \{0\}$ et donc F n'admet pas de supplémentaire orthogonal.

Proposition ∗6.

Soit E un espace préhilbertien réel et F un sous-espace vectoriel de E de dimension finie.

Alors $E = F \overset{\perp}{\oplus} F^\perp$ et $(F^\perp)^\perp = F$.

Démonstration.

D'après la proposition 4, F et $F^\perp$ sont en somme directe orthogonale. Montrons alors que $E = F + F^\perp$.

Soit $x \in E$. Le sous-espace vectoriel F de E est de dimension finie alors on peut considérer une base orthonormale $\mathcal{B} = (e_1, ..., e_p)$ de F. On pose $x_F = \sum_{i=1}^p (x|e_i)e_i \in F$. Alors on a, pour tout $j \in [\![1,p]\!]$:

$$\begin{aligned}(x - x_F|e_j) &= (x|e_j) - (x_F|e_j)\\ &= (x|e_j) - \sum_{i=1}^{p}(x|e_i)(e_i|e_j)\\ &= (x|e_j) - (x|e_j) = 0.\end{aligned}$$

Par suite, $x - x_F$ est orthogonal avec chacun des éléments d'une base de F, donc $x - x_F \in F^\perp$. Ainsi, $x = x_F + (x - x_F) \in F + F^\perp$.

Il en résulte que $E = F \overset{\perp}{\oplus} F^\perp$.

De plus, d'après ce qui précède, on a $E = F \overset{\perp}{\oplus} F^\perp$ et on applique alors la proposition 5. □

Corollaire ∗2.

Soit E un espace **euclidien** et F un sous-espace vectoriel de E. Alors

$$\dim(F) + \dim(F^\perp) = \dim(E).$$

Démonstration.

Comme E est de dimension finie, alors F et $F^\perp$ sont de dimension finie et donc d'après la proposition précédente, on a $E = F \overset{\perp}{\oplus} F^\perp$, d'où le résultat. □

Partie $**$

Projection orthogonale

Dans cette partie, E désigne un espace préhilbertien réel de produit scalaire noté $(\cdot|\cdot)$, de norme associée notée $\|\cdot\|$ et de distance associée notée d.

1. Projection orthogonale sur un sous-espace vectoriel de dimension finie

Définition $ $1.** Projection orthogonale

Soit F un sous-espace vectoriel de dimension finie de E. On appelle **projection orthogonale** sur F et on note p_F, la projection sur F parallèlement à son supplémentaire orthogonal $F^{\perp}$. L'image $p_F(x)$ d'un vecteur $x \in E$ par la projection orthogonale sur F est appelée **projeté orthogonal** de x sur F.

Remarque $ $1.**

Pour F un sous-espace vectoriel de E de dimension finie, on a $\operatorname{Im}(p_F) = F$ et $\operatorname{Ker}(p_F) = F^{\perp}$.

Proposition $ $1.**

Soit F un sous-espace vectoriel de E de dimension finie p, $\mathcal{B} = (e_1, ..., e_p)$ une base orthonormale de F et $x \in E$. Alors la projection orthogonale de x sur F vérifie :

$$p_F(x) = \sum_{i=1}^{p} (x|e_i)e_i.$$

Démonstration.

Comme dans la démonstration de 6, on décompose $x = y + z$ avec $y = \sum_{i=1}^{p}(x|e_i)e_i \in F$ et $z \in F^{\perp}$. Par suite,

$$p_F(x) = \underbrace{p_F(y)}_{=y} + \underbrace{p_F(z)}_{=0_E} = \sum_{i=1}^{p} (x|e_i)e_i.$$

$\square$

Remarque $ $2.**

La famille orthonormale $(e_1, ..., e_n)$ obtenue à partir d'une famille $(x_1, ...x_n)$ libre de E grâce au procédé de Gram-Schmidt peut alors s'exprimer de la façon suivante : pour $k = 1, ..., n-1$, on note

$$F_k = \text{Vect}(x_1, ...x_k) \ \ (= \text{Vect}(e_1, ...e_k)) \, .$$

et on a :

$$e_{k+1} = \frac{1}{\|\varepsilon_{k+1}\|}\varepsilon_{k+1} \text{ où } \varepsilon_{k+1} = x_{k+1} - p_{F_k}(x_{k+1});$$

de plus,

$$\|\varepsilon_{k+1}\| = \sqrt{\|x_{k+1}\|^2 - \|p_{F_k}(x_{k+1})\|^2}.$$

Proposition ∗∗2.

Soit F un sous-espace vectoriel de dimension finie de E et $(x_1, ..., x_k)$ une famille génératrice de F. Pour $x, y \in E$, on a :

$$y = p_F(x) \ \Leftrightarrow \ \begin{cases} y \in F \\ (x - y|x_i) = 0 \ \forall i \in [\![1, k]\!]. \end{cases}$$

Exercice ∗∗1.

Soit $n \in \mathbb{N}^*$. On munit $\mathbb{R}[X]$ du produit scalaire

$$(P|Q) = \int_0^1 P(t)Q(t)dt.$$

Déterminer la projection orthogonale de X^n sur $F = \mathbb{R}_1[X]$.

Correction.

On a $\deg(p_F(X^n)) = 1$ donc $p_F(X^n) = aX + b$. De plus, on a :

$$\begin{cases} (X^n - p_F(X^n)|1) = 0 \\ (X^n - p_F(X^n)|X) = 0 \end{cases} \Leftrightarrow \begin{cases} \dfrac{1}{n+1} = \dfrac{a+2b}{2} \\ \dfrac{1}{n+2} = \dfrac{2a+3b}{6} \end{cases} \Leftrightarrow \begin{cases} a = \dfrac{6n}{(n+1)(n+2)} \\ b = \dfrac{2-2n}{(n+1)(n+2)} \end{cases}$$

d'où $p_F(X^n) = \frac{1}{(n+1)(n+2)}(6nX + 2 - 2n)$.

2. Distance à un sous-espace de dimension finie

On rappelle ici la définition de distance à une partie de E :

Définition ∗∗2. Distance à une partie

Soit $x \in E$ et $A \subset E$. On appelle **distance** de x à A et on note $d(x, A)$ la quantité

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

Proposition $ $3.**

Soit F un sous-espace vectoriel de dimension finie de E et $x \in E$. Alors la distance $d(x,F)$ de x à F est atteinte en un unique point de F : le projeté orthogonale $p_F(x)$ de x sur F. Autrement dit :

- $d(x,F) = \|x - p_F(x)\|$ et ;
- pour tout $y \in F$, $d(x,F) = \|x-y\|$ implique $y = p_F(x)$.

Démonstration.

Soit $y \in F$. Alors on a $x - y = \underbrace{x - p_F(x)}_{\in F^\perp} + \underbrace{p_F(x) - y}_{\in F}$.
D'après le théorème de Pythagore, on a

$$\|x-y\|^2 = \|x - p_F(x)\|^2 + \|p_F(x) - y\|^2 \geq \|x - p_F(x)\|^2.$$

Ceci étant vrai pour tout $y \in F$, on a : $d(x,F) \geq \|x - p_F(x)\|$. Or, $p_F(x) \in F$ donc $d(x,F) \leq \|x - p_F(x)\|$. Il en résulte que $d(x,F) = \|x - p_F(x)\|$.
De plus, pour $y \in F$, si $d(x,F) = \|x-y\|$, alors, d'après le théorème de Pythagore,

$$\|y - p_F(x)\|^2 = \|x-y\|^2 - \|x - p_F(x)\|^2 = \|x-y\|^2 - d(x,F)^2 = 0.$$

□

Exercice $ $2.**

Déterminer la quantité $A = \inf_{(a,b)\in\mathbb{R}^2} \int_0^1 (t^2 - at - b)^2 dt$.

Correction.

On remarque, en considérant $\mathbb{R}[X]$ muni du produit scalaire $(P|Q) = \int_0^1 P(t)Q(t)dt$ et en notant $\mathbb{R}_1[X]$, que

$$A = \inf_{(a,b)\in\mathbb{R}^2} \int_0^1 (t^2 - at - b)^2 dt = d(X^2, F)^2 = d(X^2, p_F(X^2))^2$$

En reprenant le résultat de l'exercice $**$1, on obtient $p_F(X^2) = X - \frac{1}{6}$, d'où

$$A = \int_0^1 (t^2 - t + \frac{1}{6})^2 dt = \frac{1}{180}.$$

Corollaire $ $1.**

Soit F un sous-espace de dimension finie k, $\mathcal{B} = (e_1, ..., e_k)$ une base orthonormale de F et $x \in E$. Alors on a :

$$d(x,F)^2 = \|x\|^2 - \sum_{i=1}^{k} (e_i|x)^2.$$

Démonstration.

On a $p_F(x) = \sum_{i=1}^{k}(e_i|x)e_i$, et $x - p_F(x) \perp p_F(x)$, donc, d'après le théorème de Pythagore :

$$d(x,F)^2 = \|x - p_F(x)\|^2 = \|x\|^2 - \|p_F(x)\|^2 = \|x\|^2 - \sum_{i=1}^{k}(e_i|x)^2.$$

$\square$

Théorème **1.

Soit $(e_1, ..., e_k)$ une famille orthonormale de E. Alors on a :

$$\sum_{i=1}^{k}(e_i|x)^2 \leq \|x\|^2.$$

Démonstration.

On applique le corollaire précédent à $F = \mathrm{Vect}(e_1, ..., e_k)$. Alors on a :

$$\|x\|^2 = d(x,F)^2 + \sum_{i=1}^{k}(e_i|x)^2 \geq \sum_{i=1}^{k}(e_i|x)^2.$$

$\square$

Partie A
Adjoint d'un endomorphisme

Dans cette partie, n désigne un entier naturel et E désigne un espace **euclidien** de dimension n de produit scalaire noté $(\cdot|\cdot)$, de norme associée notée $\|\cdot\|$ et de distance associée notée d.

1. Représentation des formes linéaires

Théorème 1. *Théorème de représentation de Riesz*

Soit φ une forme linéaire sur l'espace euclidien E. Alors il existe un unique vecteur $a \in E$ tel que, pour tout $x \in E$,
$$\varphi(x) = (a|x).$$

Démonstration.

On considère une base orthonormale $\mathcal{B} = (e_1, ..., e_n)$ de E.

- **Existence :** On pose $a = \sum_{i=1}^n \varphi(e_i)e_i$. Alors, pour tout $x = \sum_{i=1}^n x_ie_i \in E$, on a, par linéarité de f :
$$(a|x) = \sum_{i=1}^n x_i\varphi(e_i) = \varphi(\sum_{i=1}^n x_ie_i) = \varphi(x).$$
- **Unicité :** Soit $a, b \in E$ tels que, pour tous $x \in E$, $(a|x) = \varphi(x) = (b|x)$. Alors, pour tous $x \in E$:
$$(a - b|x) = (a|x) - (b|x) = \varphi(x) - \varphi(x) = 0.$$
Ainsi, $a - b$ appartient à l'orthogonal de E d'où $a - b = 0_E$ i.e. $a = b$.

□

Corollaire 1.

Soit H un hyperplan de E. Il existe un vecteur non nul $n \in E$ tel que $H = \{n\}^\perp$.

Démonstration.

Soit H un hyperplan de E. Alors il existe une forme linéaire non nulle φ telle que $H = \mathrm{Ker}(\varphi)$. D'après le théorème de représentation de Riesz, il existe (un unique) $n \in E$ tel que $\varphi : x \mapsto (n|x)$; de plus, comme $\varphi \neq \mathbf{0}$, $n \neq 0_E$ et on a :
$$H = \mathrm{Ker}(\varphi) = \{x \in E \mid (n|x) = \varphi(x) = 0\} = \{n\}^\perp.$$

□

Remarque 1.

Un tel vecteur n est appelé *vecteur normal* à H ; il n'y a pas unicité d'un vecteur normal : en effet, si n est normal à H, pour $\lambda \in \mathbb{R}^*$, λn est normal à H.

Exercice 1.

Soit φ une forme linéaire sur $M_n(\mathbb{R})$. Montrer qu'il existe une unique matrice $A \in M_n(\mathbb{R})$ telle que pour tout $M \in M_n(\mathbb{R})$,

$$\varphi(M) = \mathrm{Tr}(AM).$$

Démonstration.

On munit $M_n(\mathbb{R})$ du produit scalaire $(M|N) = \mathrm{Tr}({}^tMN)$. Comme φ est une forme linéaire, d'après le théorème de représentation de Riesz, il existe un unique $B \in M_n(\mathbb{R})$ tel que pour tout $M \in M_n(\mathbb{R})$, $\varphi(M) = (B|M) = \mathrm{Tr}({}^tBM)$. Ainsi, en posant $A = {}^tB$, on obtient :

$$\varphi(M) = \mathrm{Tr}(AM),$$

et de plus, A est unique par unicité de B. □

2. Adjoint d'un endomorphisme

Lemme 1.

Soit $u \in \mathcal{L}(E)$. Il existe un unique $v \in \mathcal{L}(E)$ tel que, pour tous $x, y \in E$:

$$(u(x)|y) = (x|v(y)).$$

Démonstration.

- **Existence :** Soit $y \in E$. On note $\varphi : E \to \mathbb{R}$ l'application définie, pour $x \in E$ par $\varphi_y(x) = (u(x)|y)$. Par linéarité de u et du produit scalaire par rapport à sa première variable, φ est une forme linéaire sur l'espace euclidien E. Ainsi, d'après le théorème de représentation de Riesz, il existe un unique vecteur $a_y \in E$ tel que $\varphi_y(x) = (a_y|x)$ pour tous $x \in E$.

  Par suite, l'application $v : y \mapsto a_y$ est bien définie de E dans lui-même et on a alors, pour tous $x, y \in E$:
  $$(u(x)|y) = \varphi_y(x) = (a_y|x) = (x|a_y) = (x|v(y)).$$
  De plus, pour $y, z \in E$ et $\lambda, \mu \in \mathbb{R}$, on a, pour tout $x \in E$, par linéarité du produit scalaire par rapport à la seconde variable :
  $$\begin{array}{rcl} (x|v(\lambda y + \mu z)) & = & (u(x)|\lambda y + \mu z) \\ & = & \lambda(u(x)|y) + \mu(u(x)|z) \\ & = & \lambda(x|v(y)) + \mu(x|v(z)) \\ (x|v(\lambda y + \mu z)) & = & (x|\lambda v(y) + \mu v(z)); \end{array}$$

donc $(x|v(\lambda y+\mu z)-(\lambda v(y)+\mu v(z)))=0$, pour tout $x\in E$, d'où :

$$v(\lambda y+\mu z)=\lambda v(y)+\mu v(z).$$

Ainsi, v est un endomorphisme de E.
Ce qui prouve l'existence.

- **Unicité :** Soit $v,w\in\mathcal{L}(E)$ tels que, pour tous $x,y\in E$:

$$(x|v(y))=(u(x)|y)=(x|w(y)).$$

Soit $y\in E$. alors pour tout $x\in E$:

$$(x|v(y)-w(y))=(x|v(y))-(x|w(y))=(u(x)|y)-(u(x)|y)=0.$$

Par suite, $v(y)=w(y)$.
Donc $v=w$. Ce qui prouve l'unicité.

□

Le lemme précédent légitime la définition suivante :

Définition 1.

Soit $u\in\mathcal{L}(E)$. On appelle **endomorphisme adjoint** - ou simplement **adjoint** - de u l'unique endomorphisme de E noté u^* tel que, pour tous $x,y\in E$:

$$(u(x)|y)=(x|u^*(y))$$

Exemple 1.

— Si u est une homothéthie de E, alors $u^*=u$.

> En effet, pour $u=\lambda\mathrm{Id}_E$ avec $\lambda\in\mathbb{R}$, on a, pour tous $x,y\in E$:
>
> $$(u(x)|y)=(\lambda x|y)=\lambda(x|y)=(x|\lambda y)=(x|u(y)).$$
>
> Par suite, $u^*=u$.

— Pour $E=\mathbb{R}^2$ muni de son produit scalaire canonique et $f\in\mathcal{L}(E)$ tel que $f:(x,y)\mapsto(x+2y,x)$, on a $f^*=g:(x,y)\mapsto(x+y,2x)$.

> En effet, pour tous $(x,y),(a,b)\in E$:
>
> $$(f(x,y)|(a,b))=(x+2y)a+xb=x(a+b)+y.2a=((x,y)|g(a,b))$$
>
> Par suite, $f^*=g$.

3. Propriétés de l'adjoint

Proposition 2. *Matrice de l'adjoint dans une base orthonormale*

Soit $u, v \in \mathcal{L}(E)$ et $\mathcal{B}$ une base **orthonormale** de E. On note $A = \mathrm{Mat}_{\mathcal{B}}(u)$ et $B = \mathrm{Mat}_{\mathcal{B}}(v)$.
L'endomorphisme v est l'adjoint de u i.e. $v = u^*$ si, et seulement si, $B = {}^t\!A$.
En particulier, $\mathrm{Mat}_{\mathcal{B}}(u^*) = {}^t\!A$.

Démonstration.

Soit $u, v \in \mathcal{L}(E)$, u^* son adjoint, $\mathcal{B}$ une base **orthonormale** de E. On note $A = (a_{ij})_{1 \leq i,j \leq n} = \mathrm{Mat}_{\mathcal{B}}(u)$ et $B = (b_{ij})_{1 \leq i,j \leq n} = \mathrm{Mat}_{\mathcal{B}}(v)$.

($\Rightarrow$) On suppose $v = u^*$. Comme $\mathcal{B}$ est orthonormale, on a, pour tous $i, j \in [\![1, n]\!]$:

$$b_{ij} = (u^*(e_j)|e_i) = (e_j|u(e_i)) = (u(e_i)|e_j) = a_{ji}.$$

car u^* est l'adjoint de u.
Par suite, on a $B = {}^t\!A$.

($\Leftarrow$) On suppose $B = {}^t\!A$. Soit $x, y \in E$. On note $X = \mathrm{Mat}_{\mathcal{B}}(x)$ et $Y = \mathrm{Mat}_{\mathcal{B}}(y)$. On a :

$$(u(x)|y) = {}^t(AX)Y = {}^tX \underbrace{{}^t\!A}_{=B} Y = {}^tX(BY) = (x|v(y))$$

Par suite, par unicité de l'adjoint de u, on a $v = u^*$.

$\square$

Proposition 3.

L'application $u \mapsto u^*$ est un automorphisme involutif de $\mathcal{L}(E)$.

Démonstration.

Notons $\mathrm{Ad} : u \mapsto u^*$.

— *1ère façon : avec la définition.*
L'application Ad va bien de $\mathcal{L}(E)$ dans lui-même. Soit $u, v \in \mathcal{L}(E)$ et $\lambda, \mu \in \mathbb{R}$. On a, pour tous $x, y \in E$:

$$\begin{aligned} ((\lambda u + \mu v)(x)|y) &= \lambda(u(x)|y) + \mu(v(x)|y) \\ &= \lambda(x|u^*(y)) + \mu(x|v^*(y)) \\ ((\lambda u + \mu v)(x)|y) &= (x|(\lambda u^* + \mu v^*)(y)). \end{aligned}$$

Donc, par unicité de l'adjoint :

$$\mathrm{Ad}\,(\lambda u + \mu v) = (\lambda u + \mu v)^* = \lambda u^* + \mu v^* = \lambda \mathrm{Ad}(u) + \mu \mathrm{Ad}(v).$$

Par suite Ad est linéaire.
Montrons que Ad est involutive. Soit $u \in \mathcal{L}(E)$. On note $v = u^*$. Pour tous $x, y \in E$:

$$(v(x)|y) = (u^*(x)|y) = (x|u(y)),$$

donc $v^* = u$ par unicité de l'adjoint. Ainsi, $\mathrm{Ad}^2(u) = \mathrm{Ad}(v) = v^* = u$ i.e. Ad est une involution.
Ainsi, Ad est un élément inversible de l'anneau $(\mathcal{L}\left((\mathcal{L}(E)\right), +, \circ)$ car, étant une involution, il est sa propre inverse ; d'où Ad est bijective et donc un automorphisme de $\mathcal{L}(E)$.

Il en résulte que Ad est un automorphisme involutif de $\mathcal{L}(E)$.

— *2nde façon : avec la proposition précédente.* Soit $\mathcal{B}$ une base orthonormale de E. L'application $M = \mathrm{Mat}_{\mathcal{B}} : \mathcal{L}(E) \to M_n(\mathbb{R})$ est un isomorphisme d'espaces vectoriels (et même d'algèbres) et l'application $T : A \mapsto {}^tA$ est un automorphisme involutif de $M_n(\mathbb{R})$. Or, on a :

$$\mathrm{Ad} = M^{-1} \circ T \circ M,$$

en effet, pour tout $u \in \mathcal{L}(E)$ avec $A = \mathrm{Mat}_{\mathcal{B}}(u)$, on a, d'après la proposition précédente :

$$M^{-1} \circ T \circ M(u) = M^{-1}(T(A)) = M^{-1}\left({}^tA\right) = u^* = \mathrm{Ad}(u).$$

Donc Ad : $\mathcal{L}(E) \to \mathcal{L}(E)$ est un isomorphisme d'espaces vectoriels comme composée d'isomorphismes d'espces vectoriels.
De plus, comme T est involutif, on a :

$$\mathrm{Ad}^2 = \left(M^{-1} \circ T \circ M\right) \circ \left(M^{-1} \circ T \circ M\right) = M^{-1} \circ T^2 \circ M = M^{-1} \circ T \circ M = \mathrm{Ad}.$$

Il en résulte que Ad est un automorphisme involutif de $\mathcal{L}(E)$.

□

Proposition 4.

Pour tous $u, v \in \mathcal{L}(E)$, $(u \circ v)^* = v^* \circ u^*$.

Correction.

Soit $u, v \in \mathcal{L}(E)$. Pour tous $x, y \in E$, on a :

$$\begin{array}{rcl} (u \circ v(x)|y) & = & (u(v(x))|y) \\ & = & (v(x)|u^*(y)) \\ & = & (x|v^*(u^*(y))) \\ (u \circ v(x)|y) & = & (x|v^* \circ u^*(y)). \end{array}$$

Donc, par unicité de l'adjoint $(u \circ v)^* = v^* \circ u^*$.

Remarque : on aurait pu également utiliser une base orthonormale et utiliser la caractérisation matricielle de l'adjoint en remarquant que, pour toutes matrices A, B, ${}^tAB = {}^tB{}^tA$.

Proposition 5.

Soit $u \in \mathcal{L}(E)$. On a :

$$\mathrm{Ker}(u^*) = \mathrm{Im}(u)^{\perp} \text{ et } \mathrm{Im}(u^*) = \mathrm{Ker}(u)^{\perp}$$

Démonstration.

Soit $u \in \mathcal{L}(E)$. Montrons tout d'abord les inclusions suivantes :

— Montrons $\mathrm{Ker}(u^*) \subset \mathrm{Im}(u)^{\perp}$. Soit $x \in \mathrm{Ker}(u^*)$. Alors, pour tout $y \in \mathrm{Im}(u)$, il existe

$x' \in E$ tel que $y = u(x')$ et on a :

$$(x|y) = (x|u(x')) = (\underbrace{u^*(x)}_{=0_E}|x') = 0$$

D'où $x \in \mathrm{Im}(u)^\perp$. Par suite, $\mathrm{Ker}(u^*) \subset \mathrm{Im}(u)^\perp$.

— Montrons que $\mathrm{Im}(u^*) \subset \mathrm{Ker}(u)^\perp$. Soit $y \in \mathrm{Im}(u^*)$. Alors il existe $x \in E$ tel que $y = u^*(x)$ et on a, pour tout $x' \in \mathrm{Ker}(u)$:

$$(y|x') = (u^*(x)|x') = (x|\underbrace{u(x)}_{=0_E}) = 0$$

D'où $y \in \mathrm{Ker}(u)^\perp$. Par suite, $\mathrm{Im}(u^*) \subset \mathrm{Ker}(u)^\perp$.

Montrons les inclusions réciproques. On rappelle que si F est un sous-espace vectoriel de dimension finie de E (comme E est euclidien ici, tout sous-espace vectoriel de E est de dimension finie), $(F^\perp)^\perp = F$ et si $A, B \subset E$ tels que $A \subset B$ alors $B^\perp \subset A^\perp$.
Ainsi, on a, d'après les inclusions précédentes appliquées à $v = (u^*)^* = u$ (car la passage à l'adjoint est involutif) :

—
$$\mathrm{Ker}(u) = \mathrm{Ker}((u^*)^*) \subset \mathrm{Im}(u^*)^\perp$$

d'où :

$$\mathrm{Ker}(u)^\perp \supset \mathrm{Im}(u^*).$$

Par suite, $\mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp$.

—
$$\mathrm{Im}(u) = \mathrm{Im}((u^*)^*) \subset \mathrm{Ker}(u^*)^\perp$$

d'où :

$$\mathrm{Im}(u)^\perp \supset \mathrm{Ker}(u^*).$$

Par suite, $\mathrm{Ker}(u^*) = \mathrm{Im}(u)^\perp$.

□

Exercice 2.

Soit $u \in \mathcal{L}(E)$. Montrer (par forcément dans l'ordre indiqué) que :

$$\mathrm{rg}(u^*) = \mathrm{rg}(u) \quad \mathrm{Tr}(u^*) = \mathrm{Tr}(u) \quad \det(u^*) = \det(u)$$

et en terme de réduction :

$$\chi_{u^*} = \chi_u \quad \mathrm{Sp}(u^*) = \mathrm{Sp}(u) \quad \pi_{u^*} = \pi_u.$$

En déduire les liens potentiels entre diagonalisation/trigonalisation de u et u^*.

Proposition 6.

Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel de E. Si F est stable par u, alors $F^{\perp}$ est stable par u^*.

Démonstration.

On suppose F stable par u. Soit $x \in F^{\perp}$. Alors, pour tout $y \in F$, $u(y) \in F$ et :

$$(u^*(x)|y) = (\underbrace{x}_{\in F^{\perp}}|\underbrace{u(y)}_{\in F}) = 0$$

D'où $u^*(x) \in F^{\perp}$.
Il en résulte que $F^{\perp}$ est stable par u^*. $\square$

Partie B

Isométries vectorielles et matrices orthogonales

Dans cette partie, n désigne un entier naturel et E désigne un espace **euclidien** de dimension n de produit scalaire noté $(\cdot|\cdot)$, de norme associée notée $\|\cdot\|$ et de distance associée notée d.

1. Matrices orthogonales

a. Définitions

Définition 2. *Matrice orthogonale*

Soit $M \in M_n(\mathbb{R})$. On dit M est une **matrice orthogonale** si ${}^tMM = I_n$.
On appelle **groupe orthogonal** d'ordre n et on note $O_n(\mathbb{R})$ (ou encore $O(n)$) l'ensemble des matrices orthogonales de $M_n(\mathbb{R})$ i.e.

$$O_n(\mathbb{R}) = \{M \in M_n(\mathbb{R}) \mid {}^tMM = I_n\}.$$

Définition-Proposition 3.

Soit $M \in M_n(\mathbb{R})$. Si $M \in O_n(\mathbb{R})$ alors $\det(M) = \pm 1$;

- si $\det(M) = 1$, on dit M est une matrice orthogonale **directe** (ou **positive**) ;
- si $\det(M) = -1$, on dit M est une matrice orthogonale **indirecte** (ou **négative**) ;

On appelle **groupe spécial orthogonal** d'ordre n et on note $SO_n(\mathbb{R})$ (ou encore $SO(n)$) l'ensemble des matrices orthogonales directes i.e.

$$SO_n(\mathbb{R}) = \{M \in O_n(\mathbb{R}) \mid \det(M) = 1\}.$$

Définition 4. *Matrices orthogonalement semblables*

Soit $A, B \in M_n(\mathbb{R})$. On dit que A et B sont **orthogonalement semblables** s'il existe $P \in O_n(\mathbb{R})$ telle que $B = PA{}^tP$.

b. Propriétés des matrices orthogonales

On justifie ici la terminologie de "groupe" (spécial) orthogonal :

Proposition 7.

Le groupe orthogonal $O_n(\mathbb{R})$ est un sous-groupe de $GL_n(\mathbb{R})$ et le groupe spécial orthogonal $SO_n(\mathbb{R})$ est un sous-groupe de $O_n(\mathbb{R})$.
En particulier, si $M \in O_n(\mathbb{R})$, on a $M^{-1} = {}^tM$.

Proposition 8.

Soit $M \in M_n(\mathbb{R})$. On note $C_1, ..., C_n$ ses colonnes.
La matrice M est une matrice orthogonale si, et seulement si, la famille $(C_1, ..., C_n)$ est une base orthonormale de $M_{n,1}(\mathbb{R})$ muni de son produit scalaire canonique.
Le même résultat est valable pour les lignes de M.

2. Orientation d'un espace vectoriel réel de dimension finie

On rappelle que dans cette partie, E est en particulier un espace vectoriel réel de dimension finie.

Définition 5. *Orientation*

Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. On dit que $\mathcal{B}$ et $\mathcal{B}'$ **définissent la même orientation de** E si $\det(P) > 0$ où P est la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$.

Orienter l'espace E revient à se fixer une base $\mathcal{B}$ de référence. Ce choix étant fait, on appelle **bases directes**, les bases qui définissent la même orientation que $\mathcal{B}$ et **bases indirectes**, les autres.

Exemple 2.

On oriente $\mathbb{R}^3$ grâce à sa base canonique $\mathcal{B} = (e_1, e_2, e_3)$. Alors la base (e_2, e_3, e_1) est directe et la base (e_1, e_3, e_2) est indirecte.

Exercice 3.

Montrer que le procédé d'orthonormalisation de Gram-Schmidt ne change pas l'orientation d'une base.

Proposition 9.

Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormales de E et P la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$. Alors $P \in O_n(\mathbb{R})$ et de plus, $P \in SO_n(\mathbb{R})$ si, et seulement si, $\mathcal{B}, \mathcal{B}'$ définissent la même orientation de E.

Proposition 10.

On suppose que E est orienté. Si $\mathcal{B}$ et $\mathcal{B}'$ deux bases orthonormales directes de E, alors $\det_{\mathcal{B}} = \det_{\mathcal{B}'}$.

3. Isométries vectorielles

a. Définitions et premières propriétés

Définition 6. *Isométrie vectorielle*

Soit $u \in \mathcal{L}(E)$. On dit que u est une **isométrie vectorielle** (ou également un **automorphisme orthogonal**) si, pour tout $x \in E$,
$$\|u(x)\| = \|x\|.$$
On note $O(E)$ l'ensemble des isométries vectorielles de E.

Proposition 11.

Soit $u \in \mathcal{L}(E)$. Si $u \in O(E)$ alors u est un automorphisme de E.

Démonstration.

On suppose $u \in O(E)$. Alors $\mathrm{Ker}(u) = \{0_E\}$; en effet, si $x \in \mathrm{Ker}(u)$, alors $\|x\| = \|u(x)\| = \|0_E\| = 0$, donc par séparation de la norme, $x = 0_E$. Par suite, u est un endomorphisme injectif en dimension finie : il est donc bijectif ; ainsi, u est un automorphisme de E. $\square$

Exercice 4.

Soit $u \in O(E)$. Quelles sont les seules valeurs propres possibles pour u ?

Exemple 3.

— Toute symétrie orthogonale est une isométrie vectorielle.

> En effet, soit s une symétrie orthogonale de E. Alors s est la symétrie sur $F = \mathrm{Ker}(s - \mathrm{Id}_E)$ parallèlement à $F^\perp$, donc, pour tout $x \in E$, on a $x = y + z$ où $y \in F$ et $z \in F^\perp$ et donc $s(x) = y - z$; ainsi, d'après le theorème de Pythagore :
> $$\|s(x)\|^2 = \|y - z\|^2 = \|y\|^2 + \|z\|^2 = \|y + z\|^2 = \|x\|^2.$$
> Donc s est une isométrie vectorielle.

— En particulier, toute réflexion i.e. symétrie orthogonale par rapport à un hyperplan, est une isométrie vectorielle.

b. Caractérisations des isométries vectorielles

Proposition 12.

Soit $u \in \mathcal{L}(E)$. Alors $u \in O(E)$ si, et seulement si, pour tous $x, y \in E$, on a :
$$(u(x)|u(y)) = (x|y).$$

Démonstration.

($\Leftarrow$) On suppose que pour tous $x, y \in E$, $(u(x)|u(y)) = (x|y)$. Alors, pour tout $x \in E$, on a :

$$\|u(x)\| = \sqrt{(u(x)|u(x))} = \sqrt{(x|x)} = \|x\|.$$

D'où $u \in O(E)$.

($\Rightarrow$) Soit $x, y \in E$. On a :

$$(u(x)|u(y)) = \frac{1}{2}\left(\|u(x)\|^2 + \|u(y)\|^2 - \|u(x-y)\|^2\right) = \frac{1}{2}\left(\|x\|^2 + \|y\|^2 - \|x-y\|^2\right) = (x|y).$$

□

Proposition 13.

Soit $u \in \mathcal{L}(E)$. Alors $u \in O(E)$ si, et seulement si, $u^* = u^{-1}$.

Démonstration.

($\Rightarrow$) On suppose $u \in O(E)$. Soit $x, y \in E$. Comme u est bijective, il existe $z \in E$ tel que $y = u(z)$. Par suite, comme $z = u^{-1}(y)$ et d'après la proposition précédente :

$$(u(x)|y) = (u(x)|u(z)) \underbrace{=}_{u \in O(E)} (x|z) = (x|u^{-1}(y))$$

Ainsi, par unicité de l'adjoint de u, $u^* = u^{-1}$.

($\Leftarrow$) On suppose $u^* = u^{-1}$. Soit $x \in E$. On a :

$$\|u(x)\|^2 = (u(x)|u(x)) = (x|\underbrace{u^*(u(x))}_{=u^{-1}(u(x))}) = (x|x) = \|x\|^2.$$

Par suite, u est une isométrie vectorielle de E.

□

Proposition 14.

Soit $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base **orthonormale** de E. Alors $u \in O(E)$ si, et seulement si, l'image de la base $\mathcal{B}$ par u est une base orthonormale de E.

Démonstration.

On considère une base orthonormale $\mathcal{B} = (e_1, ..., e_n)$ de E.

- ($\Rightarrow$). On suppose $u \in O(E)$. Comme u est un isomorphisme, $(u(e_1), ..., u(e_n))$ est une base de E. De plus, pour tous $i, j \in [\![1, n]\!]$, on a :

$$(u(e_i)|u(e_j)) = (e_i|e_j) = \delta_{ij}.$$

Donc $(u(e_1), ..., u(e_n))$ est une base orthonormale.

- $(\Leftarrow)$. On suppose que $\mathcal{B}' = (u(e_1), ..., u(e_n))$ est une base orthonormale. Soit $x \in E$. On a :
$$u(x) = u(\sum_{i=1}^{n}(x|e_i)e_i) = \sum_{i=1}^{n}(x|e_i)u(e_i),$$
Par suite, comme $\mathcal{B}$ et $\mathcal{B}'$ sont des bases orthonormales, d'après le théorème de Pythagore :
$$\|u(x)\|^2 = \sum_{i=1}^{n}(x|e_i)^2 = \sum_{i=1}^{n}(x|e_i)^2 = \|x\|^2.$$
Il en résulte que $u \in O(E)$.

□

Corollaire 2.

Soit $u \in \mathcal{L}(E)$, $\mathcal{B}$ une base **orthonormale** de E et $A = \mathrm{Mat}_{\mathcal{B}}(u)$. Alors $u \in O(E)$ si, et seulement si, $A \in O_n(\mathbb{R})$.

Démonstration.

Cela découle du fait que $A \in O_n(\mathbb{R})$ si, et seulement si, les vecteurs colonnes de A forment une base orthonormale de $M_{n,1}(\mathbb{R})$ muni du produit scalaire $(X|Y) = {}^tXY$. □

Remarque 2.

On a alors un isomorphisme de groupes entre $O(E)$ et $O_n(\mathbb{R})$.

c. Déterminant d'une isométrie vectorielle

Proposition 15.

Soit $u \in \mathcal{L}(E)$. Si $u \in O(E)$, alors $\det(u) = \pm 1$.

Démonstration.

On suppose $u \in O(E)$. Alors, pour $\mathcal{B}$ une base orthonormale de E, on a, d'après le corollaire précédent, $M = \mathrm{Mat}_{\mathcal{B}}(u) \in O_n(\mathbb{R})$ donc $\det(u) = \det(M) = \pm 1$. □

Définition 7.

Soit $u \in O(E)$. On dit que u est une isométrie vectorielle **directe** si $\det(u) = 1$ et **indirecte** si $\det(u) = -1$.
On appelle **groupe spécial orthogonal** et on note $SO(E)$ l'ensemble des isométries vectorielles

directes de E i.e.

$$SO(E) = \{u \in O(E) \mid \det(u) = 1\}.$$

Remarque 3.

On considère que E est orienté. Si $u \in SO(E)$ et $\mathcal{B}$ est une base orthonormale directe, alors la base orthonormale formée par l'image de $\mathcal{B}$ par u est directe.

Proposition 16.

Le groupe orthogonal est un sous-groupe de $GL(E)$ et le groupe spécial orthogonal $SO(E)$ est un sous-groupe de $O(E)$.

4. isométries vectorielles en dimension 2

Dans ce paragraphe E est un espace euclidien **de dimension 2**.

Notation 1.

Soit $\theta \in \mathbb{R}$. On note :

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \text{et} \quad S_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$$

Proposition 17.

Soit $\theta, \theta' \in \mathbb{R}$. On a :
- $R_\theta R_{\theta'} = R_{\theta+\theta'} = R_{\theta'} R_\theta$.
- R_θ est inversible et $R_\theta^{-1} = R_{-\theta}$.

Démonstration.

- On a, d'après les formules d'additions de cos et sin :

$$\begin{aligned} R_\theta R_{\theta'} &= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta) & -\sin(\theta)\cos(\theta') - \cos(\theta)\sin(\theta') \\ \sin(\theta)\cos(\theta') + \cos(\theta)\sin(\theta') & \cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta+\theta') & -\sin(\theta+\theta') \\ \sin(\theta+\theta') & \cos\theta + (\theta') \end{pmatrix} \\ R_\theta R_{\theta'} &= R_{\theta+\theta'} \end{aligned}$$

De plus, $R_{\theta'} R_\theta = R_{\theta'+\theta} = R_{\theta+\theta'} = R_\theta R_{\theta'}$.

— D'après le résultat précédent, on a :

$$R_\theta R_{-\theta} = R_{\theta-\theta} = R_0 = I_2.$$

donc R_θ est inversible et son inverse est $R_{-\theta}$.

$\square$

Proposition 18.

Soit E un espace euclidien de dimension 2, $\mathcal{B}$ une base orthonormale de E et $u \in O(E)$.

— Si u est indirecte il existe un unique (à 2π près) $\theta \in \mathbb{R}$ tel que :

$$\mathrm{Mat}_{\mathcal{B}}(u) = S_\theta.$$

— Si u est directe i.e. $u \in SO(E)$, il existe un unique (à 2π près) $\theta \in \mathbb{R}$ tel que :

$$\mathrm{Mat}_{\mathcal{B}}(u) = R_\theta.$$

Démonstration.

Soit E un espace euclidien de dimension 2, $u \in O(E)$ et $\mathcal{B} = (e_1, e_2)$ une base orthonormale de E.
Alors, comme u est linéaire, il existe $a, b, c, d \in \mathbb{R}$ tels que : $u(e_1) = ae_1 + be_2$ et $u(e_2) = ce_1 + de_2$.
Comme $u \in O(E)$, d'après la proposition 14, $(u(e_1), u(e_2))$ est une base orthonormale de E et donc

$$\begin{cases} \text{i)} & a^2 + b^2 = \|u(e_1)\|^2 = 1 \\ \text{ii)} & c^2 + d^2 = \|u(e_2)\|^2 = 1 \\ \text{iii)} & ac + bd = (u(e_1)|u(e_2)) = 0 \end{cases}$$

Par suite, d'après les égalités i) et ii), il existe des uniques (à 2π près) $\theta, \theta' \in \mathbb{R}$ tels que $a = \cos(\theta)$, $b = \sin(\theta)$ et $c = \sin(\theta')$, $d = \cos(\theta')$.
De plus, on a, d'après l'égalité iii) :

$$\sin(\theta + \theta') = \sin(\theta)\cos(\theta') + \cos(\theta)\sin(\theta') = bd + ac = 0$$

Ainsi $\theta + \theta' = 0\ [\pi]$ d'où $\theta' = -\theta + k\pi$ avec $k \in \mathbb{Z}$ et donc :

$$d = (-1)^k \cos(\theta) \text{ et } c = (-1)^{k+1} \sin(\theta)$$

i.e.

$$M = \mathrm{Mat}_{\mathcal{B}}(u) = R_\theta \text{ ou } S_\theta$$

On remarque alors que $\det(R_\theta) = 1$ et $\det(S_\theta) = -1$. Ainsi :

— Si u est indirecte, $\det(u) = -1$ d'où $M = S_\theta$.

— Si u est directe, $\det(u) = 1$ d'où $M = R_\theta$.

$\square$

Théorème 2. *Classification des matrices orthogonales 2×2*

Soit $M \in O_2(\mathbb{R})$.
- Si M est indirecte i.e. $\det(M) = -1$, il existe $\theta \in \mathbb{R}$ tel que $M = S_\theta$.
- Si M est directe i.e. $\det(M) = 1$, il existe $\theta \in \mathbb{R}$ tel que $M = R_\theta$.

Démonstration.

Soit $M \in O_2(\mathbb{R})$. On considère alors l'endomorphisme u de $\mathbb{R}^2$ canoniquement associé à M. Alors $M = \mathrm{Mat}_{\mathcal{B}}(u)$ où $\mathcal{B}$ est la base canonique de $\mathbb{R}^2$ qui est une base orthonormale pour le produit scalaire canonique de $\mathbb{R}^2$.
Par suite, d'après le corollaire 2, $u \in O(\mathbb{R}^2)$. On remarque que $\det(M) = \det(u)$ et donc, d'après la proposition 18 :
- Si $\det(M) = -1$, alors $\det(u) = -1$ d'où, il existe $\theta \in \mathbb{R}$ tel que $M = S_\theta$.
- Si $\det(M) = 1$, alors $\det(u) = 1$ d'où, il existe $\theta \in \mathbb{R}$ tel que $M = R_\theta$.

□

Proposition 19.

Le groupe $SO_2(\mathbb{R})$ est commutatif.

Démonstration.

Soit $M, N \in SO_2(\mathbb{R})$. D'après le théorème précédent, il existe $\theta, \theta' \in \mathbb{R}$ tels que $M = R_\theta$ et $N = R_{\theta'}$. Ainsi, d'après la proposition 17, on a :

$$MN = R_\theta R_{\theta'} = R_{\theta'} R_\theta = NM.$$

Donc $SO_2(\mathbb{R})$ est commutatif.

□

Théorème 3. *Classification des isométries vectorielles en dimension 2*

Soit E un espace euclidien de dimension 2 orienté et $u \in O(E)$.
- Si u est indirecte, alors u est une réflexion i.e. une symétrie othogonale par rapport à une droite vectorielle (car $\dim(E) = 2$).
- Si u est directe i.e. $u \in SO(E)$, il existe un unique $\theta \in \mathbb{R}$ à 2π près tel que, **pour toute base orthonormale <u>directe</u>** $\mathcal{B}$:

$$\mathrm{Mat}_{\mathcal{B}}(u) = R_\theta$$

Démonstration.

Soit $\mathcal{B}$ une base orthonormale directe.
- On suppose u est indirecte. D'après la proposition 18, il existe une unique, à 2π près, $\theta \in \mathbb{R}$ tel que $M = \mathrm{Mat}_{\mathcal{B}}(u) = S_\theta$. De plus, on a $M^2 = I_2$, donc $u^2 = \mathrm{Id}_E$ d'où u est une

symétrie. De plus, pour $x, y \in E$ tels que $u(x) = x$ et $u(y) = -y$, on a, comme $u \in O(E)$:

$$(x|y) = (u(x)|u(y)) = (x| - y) = -(x|y),$$

d'où $(x|y) = 0$. Ainsi, $\mathrm{Ker}(u - \mathrm{Id}_E)$ et $\mathrm{Ker}(u + \mathrm{Id}_E)$ sont orthogonaux, d'où u est une symétrie orthogonale.
Comme $\det(u) = -1$, $u \neq \pm I_2$, donc u étant une symétrie de E de dimension 2, on a $\dim(\mathrm{Ker}(u - \mathrm{Id}_E)) = 1$ et ainsi, u est une symétrie orthogonale par rapport à la droite vectorielle $\mathrm{Ker}(u - \mathrm{Id}_E)$.
Remarque : on aurait également pu considérer la famille $\mathcal{B}' = (\varepsilon_1, \varepsilon_2)$ tel que :

$$\varepsilon_1 = \cos(\frac{\theta}{2})e_1 + \sin(\frac{\theta}{2})e_2 \text{ et } \varepsilon_2 = -\sin(\frac{\theta}{2})e_1 + \cos(\frac{\theta}{2})e_2$$

et montrer que $\mathcal{B}'$ est une base orthonormale de E et que $\mathrm{Mat}_{\mathcal{B}'}(u) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et donc que u est la symétrie orthogonale par rapport à $\mathrm{Vect}(\varepsilon_1)$.

— On suppose u est directe. D'après la proposition 18, il existe une unique, à 2π près, $\theta \in \mathbb{R}$ tel que $M = \mathrm{Mat}_{\mathcal{B}}(u) = R_\theta \in SO_2(\mathbb{R})$.
Soit $\mathcal{B}'$ une base orthonormale directe et donc de même orientation que $\mathcal{B}$. Notons $N = \mathrm{Mat}_{\mathcal{B}'}(u)$.
D'après la proposition 9, la matrice de passage P de $\mathcal{B}$ vers $\mathcal{B}'$ appartient à $SO_2(\mathbb{R})$. Ainsi, par commutativité de $SO_2(\mathbb{R})$, on a :

$$N = P^{-1}MP = P^{-1}PM = M.$$

Il en résulte que la forme matricielle de u ne dépend pas de la base orthonormale directe choisie et donc θ est le même (toujours à 2π près) quelque soit la base orthonormale directe choisie.

$\square$

Proposition 20.

Les groupes $(\mathbb{U}, \times)$ et $(SO(E), \circ)$ sont isomorphes.

Démonstration.

On suppose que E est orienté. Soit $\mathcal{B}$ une base orthonormale directe de E. Pour $r \in SO(E)$, d'après le théorème précédent, il existe un unique $\theta_r \in \mathbb{R}$ tel que $\mathrm{Mat}_{\mathcal{B}}(u) = R_{\theta_r}$.
On construit alors l'application $\varphi : SO(E) \to \mathbb{U}$ telle que, pour $r \in SO(E)$:

$$\varphi(r) = e^{i\theta_r}.$$

Comme pour tout $k \in \mathbb{Z}$, $e^{ix} = e^{i(x+2k\pi}$, l'unicité à 2π près de θ_r pour $r \in SO(E)$ garantit que $\varphi(r)$ est bien défini et donc que φ est bien définie sur $SO(E)$.

— Montrons que φ est un morphisme. On a, pour tout $r, r' \in SO(E)$,

$$\mathrm{Mat}_{BB}(r \circ r') = \mathrm{Mat}_{BB}(r)\mathrm{Mat}_{BB}(r') = R_{\theta_r}R_{\theta_{r'}} = R_{\theta_r + \theta_{r'}},$$

donc
$$\varphi(r \circ r') = e^{i(\theta_r + \theta_{r'})} = e^{i\theta_r} e^{i\theta_{r'}} = \varphi(r)\varphi(r')$$
Par suite, φ est un morphisme de groupes.

— Montrons que φ est injective. Si $r \in \mathrm{Ker}(\varphi)$, alors $\varphi(r) = 1 = e^{i0}$, donc $\theta_r = 0\ [2\pi]$. Par suite, $r = \mathrm{Id}_E$. Ainsi $\mathrm{Ker}(\varphi) = \{\mathrm{Id}_E\}$ et donc φ est injective.

— Montrons que φ est surjective. Soit $e^{i\theta} \in \mathbb{U}$ où $\theta \in \mathbb{R}$. Soit $\mathcal{B}$ une base orthonormale de E. On pose $r = \mathrm{Mat}_{\mathcal{B}}^{-1}(R_\theta)$. Alors , par construction de φ, on a $\varphi(r) = e^{i\theta}$.
Par suite, φ est surjective.

Il en résulte que φ est un ismorphisme et donc $\mathbb{U}$ et $SO(E)$ sont isomorphes. □

Lemme 2.

On note $U = S(0_E, 1)$ l'ensemble des vecteurs unitaires de E.

— pour tous $u, v \in U$, il existe un unique $r \in SO(E)$ tel que $v = r(u)$.

— La relation binaire $\sim$ sur U^2 définie, pour $(u, v), (u', v') \in U^2$, par :
$$(u, v) \sim (u', v') \text{ si } \exists\, r \in SO(E), v = r(u) \text{ et } v' = r(u')$$
est une relation d'équivalence sur U^2.

— Les classes d'équivalence de $\sim$ sont en bijection avec $SO(E)$.

Démonstration.

— Soit $u, v \in U$. On pose $e_1 = u$ et on complète la famille (e_1) en une base orthonormale $\mathcal{B} = (e_1, e_2)$ de E. Alors $v = xe_1 + ye_2$ et d'après le théorème de Pythagore, on a :
$$x^2 + y^2 = \|v\|^2 = 1$$
Considérons l'application linéaire $r : E \to E$ telle que $M = \mathrm{Mat}_{\mathcal{B}}(r) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$. Comme $\mathcal{B}$ est une base orthonormale, ${}^t\!MM = I_2$ et $\det(M) = 1$ (faire les calculs), alors $r \in SO(E)$. De plus, on a :
$$r(u) = r(e_1) = xe_1 + ye_2 = v$$
D'où l'existence de $r \in SO(E)$ tel que $v = r(u)$.
Montrons l'unicité : soit $r \in SO(E)$ tel que $v = r(u)$. On reprend les notations utilisée dans l'existence, et on pose $N = \mathrm{Mat}_{\mathcal{B}}(r)$. Comme $\mathcal{B}$ est une base orthonormale de E, $N \in O_2(\mathbb{R})$ et :
$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(v) = \mathrm{Mat}_{\mathcal{B}}(r(u)) = \mathrm{Mat}_{\mathcal{B}}(r).\mathrm{Mat}_{\mathcal{B}}(u) = N \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
Donc la première colonne de B est $N_1 = \begin{pmatrix} x \\ y \end{pmatrix}$; or comme $N \in O_2(\mathbb{R})$, sa deuxième colonne est $N_2 = \begin{pmatrix} -y \\ x \end{pmatrix}$ car (N_1, N_2) est une base orthonormale de $M_{2,1}(\mathbb{R})$. On aurait le même résultat pour $r' \in SO(E)$ tel que $v = r'(u)$, d'où, pour tous $r, r' \in SO(E)$ tels que

$r(u) = v = r'(u)$, on a :

$$\mathrm{Mat}_{\mathcal{B}}(r) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = \mathrm{Mat}_{\mathcal{B}}(r')$$

Et donc $r = r'$ par injectivité de l'application $\mathrm{Mat}_{\mathcal{B}}$.
D'où l'unicité.

— Montrons que $\sim$ est une relation d'équivalence sur U^2.
  - (Réflexivité)
    Soit $(u, v) \in U^2$. D'après le point précédent, il existe $r \in SO(E)$ tel que $v = r(u)$ (et $v = r(u)$), d'où $(u, v) \sim (u, v)$.
  - (Symétrie)
    Soit $(u, v), (u', v') \in U^2$. On suppose $(u, v) \sim (u', v')$. Alors il existe $r \in SO(E)$ tel que $v = r(u)$ et $v' = r(u')$; donc il existe $r \in SO(E)$ tel que $v' = r(u')$ et $v = r(u)$. D'où $(u', v') \sim (u, v)$.
  - (Transitivité)
    Soit $(u, v), (u', v'), (u'', v'') \in U^2$. On suppose $(u, v) \sim (u', v')$ et $(u', v') \sim (u'', v'')$. Alors il existe $r, r' \in SO(E)$ tels que $v = r(u)$ et $v' = r(u')$, et $v' = r'(u')$ et $v'' = r'(u'')$; par unicité du $r \in SO(E)$ tel que $v' = r(u')$ (point précédent), on a $r = r'$ donc $v = r(u)$ et $v'' = r(u'')$ i.e. $(u, v) \sim (u'', v'')$.

— On considère l'application $f : U^2 \to SO(E)$ qui à un couple $(u, v) \in U^2$ associe $r \in SO(E)$ telle que $v = r(u)$. L'application f est bien définie par unicité d'un tel r. Voyons quelques propriétés de f qui nous aiderons à construire notre bijection :

  $*$ pour $(u, v) \in U^2$ et pour tout couple $(u', v') \in \mathrm{Cl}(u, v)$ (classe d'équivalence de (u, v)), on a, par définition de $\sim$, $f(u, v) = f(u', v')$.
  Par suite, f est constante sur chaque classe d'équivalence.

  $**$ Réciproquement, si pour $(u, v), (u', v') \in U^2$, $f(u, v) = r = f(u', v')$, alors $v = r(u)$ et $v' = r(u')$, donc $(u, v) \sim (u', v')$.

  $***$ L'application f est surjective : en effet, pour $r \in SO(E)$, on a, pour $u \in U$, $f(u, r(u)) = r$.

  Notons $\mathcal{C}$ l'ensemble des classes d'équivalence de $\sim$. On construit alors l'application $\tilde{f} : \mathcal{C} \to SO(E)$ de la manière suivante :
  pour $C \in \mathcal{C}$, on pose $\tilde{f}(C) = f(u, v)$ pour un certain (u, v) quelconque dans C.
  D'après $*$, pour tout $C \in \mathcal{C}$, $\tilde{f}(C)$ est bien défini, et donc l'application $\tilde{f}$ est bien définie.
  Remarque : cette construction est classique quand on a une application constante sur des classes d'équivalence, on l'appelle la factorisation de f
  Montrons que $\tilde{f}$ est une bijection.
  - (injectivité) Soit $C, C' \in \mathcal{C}$. On suppose $\tilde{f}(C) = \tilde{f}(C')$. Alors pour $(u, v) \in C$ et $(u'v') \in C'$, on a :
  $$f(u, v) = \tilde{f}(C) = \tilde{f}(C') = f(u', v')$$
  donc d'après $**$, $(u, v) \sim (u', v')$ et donc $C = C'$.
  D'où $\tilde{f}$ est injective.
  - (surjectivité) Pour $r \in SO(E)$, d'après $***$, il existe $(u, v) \in U^2$ tel que $f(u, v) = r$. On pose $C = \mathrm{Cl}(u, v) \in \mathcal{C}$. Alors,
  $$\tilde{f}(C) = f(u, v) = r$$
  D'où $\tilde{f}$ est surjective.

  Il en résulte que $\tilde{f}$ est une bijection.

$\square$

Dans cette définition, on reprend les notations du lemme précédent.

Définition 8. *Angle orienté*

Soit $(u,v) \in (E \setminus \{0_E\})^2$. On appelle **angle orienté** du couple (u,v) et on note $\widehat{(u,v)}$ la classe d'équivalence du couple $\left(\frac{u}{\|u\|}, \frac{v}{\|v\|}\right)$ par la relation d'équivalence $\sim$.

En utilisant le lemme 2 et le théorème 3, on justifie la bonne définition de la notion de *mesure d'un angle orienté* :

Définition 9. *Mesure d'un angle orienté*

On suppose E orienté.
Soit $\widehat{(u,v)}$ l'angle orienté d'un couple de vecteurs non nuls (u,v). On appelle **mesure de l'angle orienté** $\widehat{(u,v)}$ l'unique (à 2π près) $\theta \in \mathbb{R}$ tel que $r = R_\theta$ où r est l'unique élément de $SO(E)$ qui envoie $\frac{u}{\|u\|}$ sur $\frac{v}{\|v\|}$.

Remarque 4.

L'orientation de E est importante : si on change l'orientation de E i.e. on prend une base orthonormale indirecte par rapport à l'orientation initiale et qu'on oriente E avec cette nouvelle base, la mesure d'un angle qui valait θ vaudra $-\theta$.

Définition 10. *Rotation d'un espace de dimension 2*

Soit $\widehat{(u,v)}$ l'angle orienté de mesure $\theta \in \mathbb{R}$.
On appelle **rotation d'angle orienté** $\widehat{(u,v)}$ l'unique élément r de $SO(E)$ qui envoie $\frac{u}{\|u\|}$ sur $\frac{v}{\|v\|}$.
De plus, si E est orienté, on dira que r est **la rotation d'angle de mesure** θ.

Proposition 21.

Soit $r \in SO(E)$ et $u \in E \setminus \{0_E\}$.
- l'application r est la rotation d'angle orienté $\widehat{(u, r(u))}$.
- si E est orienté et si on note θ est la mesure de l'angle orienté $\widehat{(u, r(u))}$, alors r est la rotation d'angle de mesure θ.
  En particulier, dans toute base $\mathcal{B}$ orthonormale directe de E,
  $$\mathrm{Mat}_{\mathcal{B}}(r) = R_\theta.$$

5. Réduction des isométries vectorielles

a. Cas général

Proposition 22.

Soit $u \in O(E)$ et F un sous-espace vectoriel de E. Si F est stable par u alors l'endomorphisme u_F induit par u sur F est une isométrie vectorielle de F i.e. $u_F \in O(F)$.

Démonstration.

Pour tout $x \in F$, on a :
$$\|u_F(x)\| = \|u(x)\| = \|x\|.$$
Donc $u_F \in O(F)$. □

Proposition 23.

Soit $u \in O(E)$ et F un sous-espace vectoriel de E. Si F est stable par u, alors $F^\perp$ est stable par u.

Démonstration.

Soit $y \in F^\perp$. Comme $u_F \in O(F)$, u_F est une bijection, pour tout $x \in F$, il existe $x' \in F$ tel que $x = u_F(x') = u(x')$ et donc on a :
$$(u(y)|x) = (u(y)|u(x')) = (\underbrace{y}_{\in F^\perp} | \underbrace{x'}_{\in F}) = 0.$$
Ainsi, $u(y) \in F^\perp$ donc $F^\perp$ est stable par u. □

Lemme 3.

Soit $u \in \mathcal{L}(E)$. Alors E possède un sous-espace vectoriel stable par u de dimension 1 ou 2 i.e. une droite ou un plan.

Démonstration.

Le polynôme minimal π_u de u est un polynôme de degré supérieur ou égal à 1 et comme $\pi_u \in \mathbb{R}[X]$ alors π_u se décompose en produit $P_1...P_k$ de facteurs irréductibles de degré 1 ou 2. De plus, comme π_u est annulateur, on a : $\pi_u(u) = P_1(u) \circ ... \circ P_k(u) = 0$, donc il existe $i \in [\![1, k]\!]$ tel que $P_i(u)$ n'est pas injectif. Alors, pour $x \in \text{Ker}(P_i(u))$, $F = \text{Vect}(x, u(x))$ est stable par u. En effet, comme $P_i(u)(x) = 0$ et $\deg(P_i) = 1$ ou 2, $u^2(x)$ peut s'écrire comme combinaison linéaire de $u(x)$ et de x.
Il en résulte que F est une droite ou un plan stable par u. □

Théorème 4.

Soit $u \in O(E)$. Alors il existe une base orthonormale $\mathcal{B}$ de E dans laquelle la matrice de u est diagonale par blocs où les blocs sont de la forme :

$$I_p \quad ; \quad -I_q \quad ; \quad R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

où $p, q \in [\![0, n]\!]$ et $\theta \in \mathbb{R}$.

Démonstration.

On démontre le résultat par récurrence double sur la dimension de E, c'est-à-dire, on montre que pour tout $n \in \mathbb{N}^*$, la propriété :

$\mathcal{P}_n$ ="pour tout espace euclidien de dimension n, pour tout $u \in O(E)$, il existe une base orthonormale $\mathcal{B}$ de E telle la matrice de u possède la forme annoncée."

est vraie.

- **Initialisation :**
  — Montrons que $\mathcal{P}_1$ est vraie. Soit E un espace euclidien de dimension 1 et $u \in O(E)$. Comme u est linéaire et E de dimension 1 sur $\mathbb{R}$, il existe $a \in \mathbb{R}$ tel que $u : x \mapsto ax$. Or, comme $u \in O(E)$, on a, pour $x \in E \setminus \{0_E\}$ (un tel x existe car $\dim(E) > 0$) :
  $$\|x\| = \|u(x)\| = \|ax\| = |a|\|x\|$$
  d'où $a = \pm 1$. Il en résulte que toute matrice de u (et donc dans une base orthonormée de E) possède la forme annoncée - ici (± 1).
  — Montrons que $\mathcal{P}_2$ est vraie. D'après le théorème 3, selon le déterminant de u, s est une réflexion ou une rotation. Or, pour une réflexion s, il existe une base orthonormale $\mathcal{B}$ tel que $\mathrm{Mat}_{\mathcal{B}}(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ et pour une rotation r, dans n'importe quelle base orthonormale $\mathcal{B}$, $r = R_\theta$ pour un certain θ dans $\mathbb{R}$.
  Donc $\mathcal{P}_2$ est vraie.
- **Hérédité :** soit $n \in \mathbb{N}^*$. On suppose $\mathcal{P}_n$ et $\mathcal{P}_{n+1}$ vraies. Montrons que $\mathcal{P}_{n+2}$ est vraie.
  Soit E un espace euclidien de dimension $n + 2$ et $u \in O(E)$. D'après le lemme précédent, il existe F de dimension 1 ou 2 tel que F est stable par u. Comme $u \in O(E)$ et F est stable par u, d'après la proposition 23, $F^\perp$ est stable par u.
  — Comme $\dim(F) = 1$ ou 2 et $u_F \in O(F)$ (d'après la proposition 22), d'après l'initialisation, il existe une base orthonormale $\mathcal{B}_F$ de F tel que sa matrice possède la forme annoncée.
  — On a $E = F \oplus F^\perp$ donc $\dim(F^\perp) = n$ ou $n+1$ et $u_{F^\perp} \in O(F^\perp)$ (d'après la proposition 22). Ainsi, d'après l'hypothèse de récurrence, il existe une base orthornormale $\mathcal{B}_{F^\perp}$ de $F^\perp$ telle que la matrice de $u_{F^\perp}$ possède la forme annoncée.
  La base $\mathcal{B}$ de E obtenue en concaténant les bases précédentes de F et $F^\perp$ est orthonormale (c'est bien une base car F et $F^\perp$ sont supplémentaires, et elle est orthonormale car F et $F^\perp$ sont orthogonaux et les deux bases sont orthornormales). De plus, la matrice de u dans la base $\mathcal{B}$ est diagonale par blocs dont le premier bloc est la matrice de u_F dans $\mathcal{B}_F$ et le second bloc est la mtrice de $u_{F^\perp}$ dans $\mathcal{B}_{F^\perp}$: ainsi la matrice de u a bien la forme annoncée.

Ce qui achève le raisonnement par récurrence. Il en résulte que pour tout $n \in \mathbb{N}^*$, $\mathcal{P}_n$ est vraie. □

Remarque 5.

Ainsi, si u est une isométrie vectorielle, il existe une base orthonormale telle que sa matrice A dans cette base est :

$$A = \left(\begin{array}{c|c|c|c|c} I_p & 0 & \dots & \dots & 0 \\ \hline 0 & -I_q & \ddots & & \vdots \\ \hline \vdots & \ddots & R_{\theta_1} & \ddots & \vdots \\ \hline \vdots & & \ddots & \ddots & 0 \\ \hline 0 & \dots & \dots & 0 & R_{\theta_k} \end{array}\right)$$

où $p, q, k \in [\![0, n]\!]$ (plus précisément $k \leq n/2$) et pour tout $i \in [\![1, k]\!]$, $\theta_i \in \mathbb{R}$.

b. Cas d'une isométrie vectorielle directe en dimension 3

Proposition 24.

Soit $u \in \mathcal{L}(\mathbb{R}^3)$ une isométrie vectorielle directe i.e. $u \in O(\mathbb{R}^3)$ et $\det(u) = 1$. Alors il existe une base orthonormale $\mathcal{B}$ et $\theta \in [0, 2\pi[$ telle que :

$$\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Démonstration.

On applique le théorème précédent et en utilisant le fait que $\det(u) = 1$. □

Définition-Proposition 11. *Rotation de l'espace*

Soit $u \in \mathcal{L}(\mathbb{R}^3)$ une isométrie vectorielle directe. Alors il existe une droite $\mathcal{D}$ stable par u et $\theta \in [0, 2\pi]$ tel que l'endomorphisme $u_{\mathcal{P}}$ induit par u sur $\mathcal{P}$ est une rotation d'angle θ où $\mathcal{P}$ est le plan orthogonal à $\mathcal{D}$.
Un tel endomorphisme u de $\mathbb{R}^3$ est appelé **rotation d'axe** $\mathcal{D}$.

Démonstration.

D'après la proposition précédente, il existe une base $\mathcal{B} = (e_1, e_2, e_3)$ orthonormale telle que :

$$A = \mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Alors la droite $\mathcal{D} = \mathrm{Vect}(e_1)$ est stable par u et le plan $\mathcal{P} = \mathrm{Vect}(e_2, e_3)$ vérifie $\mathcal{P} \perp \mathcal{D}$ et

$$u_{\mathcal{P}} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

□

Partie C

Endomorphismes autoadjoints

Dans cette partie, n désigne un entier naturel et E désigne un espace **euclidien** de dimension n de produit scalaire noté $(\cdot|\cdot)$, de norme associée notée $\|\cdot\|$ et de distance associée notée d.
On note $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices carrées symétriques d'ordre n.

1. Définition et propriétés

Définition 12. *Endomorphisme autoadjoint*

Soit $u \in \mathcal{L}(E)$. On dit que u est **autoadjoint** (ou encore **symétrique**) si $u^* = u$ i.e si, pour tous $x, y \in E$:

$$(u(x)|y) = (x|u(y)).$$

On note $\mathcal{S}(E)$ l'ensemble des endomorphismes autoadjoints de E.

Proposition 25.

L'ensemble $\mathcal{S}(E)$ des endomorphismes autoadjoints de E est un sous-espace vectoriel de $\mathcal{L}(E)$.

Démonstration.

L'endomorphisme nul vérifie pour tous $x, y \in E$:

$$(x|0(y)) = (x|0) = 0 = (0|y) = (0(x)|y).$$

donc $0 \in \mathcal{S}(E)$.
Soit $\lambda, \mu \in \mathbb{R}$ et $u, v \in \mathcal{S}(E)$. Pour tous $x, y \in E$, on a :

$$\begin{aligned}((\lambda u + \mu v)(x)|y) &= \lambda(u(x)|y) + (\mu v(x)|y) \\ &= \lambda(x|u(y)) + (\mu x|v(y)) \\ &(x|(\lambda u + \mu v)(y))\end{aligned}$$

D'où $\mathcal{S}(E)$ est un sous-espace vectoriel de $\mathcal{L}(E)$.
On aurait pu également remarquer que $\mathcal{S}(E)$ est un sous-espace vectoriel de $\mathcal{L}(E)$ comme noyau de l'endomorphisme f de $\mathcal{L}(E)$ défini par $f : u \mapsto u^* - u$; et f est bien linéaire car $f = Ad - Id_{\mathcal{L}(E)}$. □

Proposition 26.

Soit $u \in \mathcal{L}(E)$ et $\mathcal{B}$ une base **orthonormale** de E. Alors u est autoadjoint i.e. $u \in \mathcal{S}(E)$ si, et seulement si, la matrice $\mathrm{Mat}_{\mathcal{B}}(u)$ de u dans $\mathcal{B}$ est symétrique i.e. $M \in \mathcal{S}_n(\mathbb{R})$.

Démonstration.

Soit $u \in \mathcal{L}(E)$ et $\mathcal{B} = (e_1, ..., e_n)$ une base orthonormale de E. On note $A = \mathrm{Mat}_{\mathcal{B}}(u)$. Comme $\mathcal{B}$ est orthonormale, d'après la proposition 2, u est autoadjoint si, et seulement si, ${}^tA = A$. D'où le résultat.

□

Remarque 6.

La proposition précédente fournit un isomorphisme entre $\mathcal{S}(E)$ et $\mathcal{S}_n(\mathbb{R})$. On obtient alors que :

$$\dim(\mathcal{S}(E)) = \dim(\mathcal{S}_n(\mathbb{R})) = \frac{n(n+1)}{2}.$$

Proposition 27.

Soit $u \in \mathcal{S}(E)$ et F est un sous-espace vectoriel de E. Si F est stable par u alors l'endomorphisme induit u_F de u sur F est autoadjoint i.e. $u_F \in \mathcal{S}(F)$.

Démonstration.

Soit $x, y \in F$. Alors on a

$$(x|u_F(y)) = (x|u(y)) = (u(x)|y) = (u_F(x)|y).$$

Donc u_F est autoadjoint. □

Proposition 28.

Soit p un projecteur et $F = \mathrm{Im}(p)$. Alors p est la projection orthogonale sur F si, et seulement si, p est autoadjoint.

Démonstration.

- $(\Rightarrow)$. On suppose que p est la projection orthogonale sur F. On a $F^\perp = \mathrm{Ker}(p)$, donc pour tout $x, y \in E$, on a :

  $$(p(x)|y) = (p(x)|\underbrace{y - p(y)}_{\in F^\perp} + \underbrace{p(y)}_{\in F}) = (p(x)|p(y))$$

  et par un calcul similaire,

  $$(x|p(y)) = (p(x)|p(y))$$

  d'où $(p(x)|y) = (x|p(y))$. Ainsi, p est autoadjoint.
- $(\Leftarrow)$. On suppose p autoadjoint. Il suffit de montrer que $\mathrm{Im}(p) \perp \mathrm{Ker}(p)$. Soit $x \in \mathrm{Im}(p)$ et $y \in \mathrm{Ker}(p)$. Alors :

  $$(x|y) = (p(x)|y) = (x|p(y)) = (x|0_E) = 0.$$

  Donc p est la projection orthogonale sur F.

□

Exercice 5.

Soit $s \in \mathcal{L}(E)$ une symétrie de E i.e. $s^2 = \mathrm{Id}_E$. Donner une condition nécessaire et suffisante pour que $s \in \mathcal{S}(E)$.

Correction.

On pose $p = \frac{1}{2}(s + \mathrm{Id}_E)$. Alors p est un projecteur de E et on a :

$$s \in \mathcal{S}(E)$$

si, et seulement si,

$$p = \tfrac{1}{2}(s + \mathrm{Id}_E) \in \mathcal{S}(E) \text{ (car } \mathcal{S}(E) \text{ est un espace vectoriel et } \mathrm{Id}_E \in \mathcal{S}(E))$$

si, et seulement si,

$$p \text{ est un projecteur orthogonal}$$

si, et seulement si,

$$s = 2p - \mathrm{Id}_E \text{ est une symétrie orthogonale.}$$

Proposition 29.

Soit $u \in \mathcal{S}(E)$ et $\lambda, \mu \in \mathbb{R}$. Si $\lambda, \mu \in \mathrm{Sp}(u)$ et $\lambda \neq \mu$ alors

$$E_\lambda(u) \perp E_\lambda(u).$$

Démonstration.

On suppose $\lambda, \mu \in \mathrm{Sp}(u)$ et $\lambda \neq \mu$. Soit $x \in E_\lambda(u)$ et $y \in E_\mu(u)$. Alors on a :

$$\lambda(x|y) = (\lambda x|y) = (u(x)|y) = (x|u(y)) = (x|\mu y) = \mu(x|y),$$

et donc $(\lambda - \mu)(x|y) = 0$ d'où $(x|y) = 0$ car $\lambda - \mu \neq 0$. □

Remarque 7.

Compte tenu de la proposition précédente, on sait ainsi que les sous-espaces propres d'un endomorphisme autoadjoint sont en somme directe orthogonale.

2. Réduction des endomorphismes autoadjoints

Proposition 30.

Soit $A \in \mathcal{S}_n(\mathbb{R})$. Alors A admet au moins une valeur propre réelle et de plus, toutes ses valeurs propres de A sont réelles i.e. $\mathrm{Sp}(A) \subset \mathbb{R}$.

Démonstration.

On considère A comme une matrice à coefficients complexes. Alors A possède au moins une valeur propre $\lambda \in \mathbb{C}$. Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(\mathbb{C})$ un vecteur propre de A associé à λ. Alors on a, d'une part :

$$ {}^t\!XA\overline{X} = {}^t\!AX\overline{X} = \lambda {}^t\!X\overline{X} = \lambda \sum_{i=1}^{n} |x_i|^2; $$

et d'autre part

$$ {}^t\!XA\overline{X} = {}^t\!X\overline{A\overline{X}} = {}^t\!X\overline{\lambda X} = \overline{\lambda}{}^t\!X\overline{X} = \overline{\lambda} \sum_{i=1}^{n} |x_i|^2; $$

Or, comme $X \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, on a $\sum_{i=1}^{n} |x_i|^2 \neq 0$. Par suite, $\lambda = \overline{\lambda}$ d'où $\lambda \in \mathbb{R}$. $\square$

Proposition 31.

Soit $u \in \mathcal{S}(E)$. Alors u possède au moins une valeur propre.

Démonstration.

Soit $\mathcal{B} = (e_1, ..., e_n)$ une base orthonormale de E. Alors $A = \mathrm{Mat}_{\mathcal{B}}(u)$ est une matrice symétrique réelle. De plus, le polynôme caractéristique χ_u de u possède au moins une racine λ dans $\mathbb{C}$ et donc $\lambda \in \mathrm{Sp}(A)$. Ainsi, d'après la proposition précédente, $\lambda \in \mathbb{R}$. Donc λ est une valeur propre réelle de u. $\square$

Proposition 32.

Soit $u \in \mathcal{S}(E)$ et F un sous-espace vectoriel de E. Si F est stable par u, alors $F^\perp$ est stable par u.

Démonstration.

On applique la proposition 6 avec l'hypothèse $u^* = u$. $\square$

Théorème 5. *Théorème spectral*

Soit $u \in \mathcal{S}(E)$. Alors u est diagonalisable et de plus :

— E est la somme directe orthogonale des sous-espaces propres de u ; ou de manière équivalente,

— E possède une base orthonormale de vecteurs propres.

Démonstration.

D'après le proposition 31, u possède au moins une valeur propre. Soit $\lambda_1, ..., \lambda_k \in \mathbb{R}$ la liste de toutes les valeurs propres deux à deux distinctes de u. On note F le sous-espace vectoriel de E :

$$F = E_{\lambda_1}(u) \overset{\perp}{\oplus} ... \overset{\perp}{\oplus} E_{\lambda_k}(u).$$

Alors F est stable par u donc $G = F^\perp$ l'est aussi.
Comme u est autoadjoint, l'endomorphisme u_G par u sur G l'est aussi. Supposons par l'absurde que $G \neq \{0_E\}$, u_G est autoadjoint donc il possède une valeur propre réelle λ et ainsi un vecteur propre $x \in G \setminus \{0_E\}$ associé à λ. Or tout vecteur propre appartient à F, donc $x \in F \cap G = F \cap F^\perp = \{0_E\}$. contradiction !
Par suite $F^\perp = \{0_E\}$ et donc

$$E = F \oplus F^\perp = F = E_{\lambda_1}(u) \overset{\perp}{\oplus} ... \overset{\perp}{\oplus} E_{\lambda_k}(u).$$

$\square$

Remarque 8.

Si $u \in \mathcal{S}(E)$, on dit alors que u est **orthogonalement diagonalisable**.

Corollaire 3.

Soit $A \in \mathcal{S}_n(\mathbb{R})$. Alors A est diagonalisable et de plus, il existe $P \in O_n(\mathbb{R})$ i.e. $P^{-1} = {}^tP$ et une matrice $D \in M_n(\mathbb{R})$ diagonale tel que

$${}^tPAP(= P^{-1}AP) = D.$$

Démonstration.

On munit $\mathbb{R}^n$ de son produit scalaire canonique, de sa base canonique $\mathcal{B}$ qui est orthonormale pour ce produit scalaire et on considère l'endomorphisme u de $\mathbb{R}^n$ canoniquement associée à A. Alors $A = \mathrm{Mat}_{\mathcal{B}}(u)$ donc, comme $A \in \mathcal{S}_n(\mathbb{R})$ et $\mathcal{B}$ est orthonormale, $u \in \mathcal{S}(\mathbb{R}^n)$. Ainsi, d'après le théorème spectral, u est diagonalisable dans une base orthonormale $\mathcal{B}'$. Par suite, la matrice P formée des vecteurs de la base $\mathcal{B}'$ vérifie que $A = PDP^{-1}$ où D est une matrice diagonale et de plus, comme ses vecteurs colonnes forment une base orthonormale de $M_{n,1}(\mathbb{R})$ muni de son produit scalaire canonique, $P \in O_n(\mathbb{R})$. $\square$

Remarque 9.

Attention, ce dernier résultat est faux dans $\mathbb{C}$: une matrice symétrique complexe n'est pas forcément diagonalisable (cf. exercice suivant) .

Exercice 6.

1. Diagonaliser dans une base orthonormale la matrice suivante :

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

2. Que dire de la diagonalisabilité de la matrice symétrique complexe suivante :

$$\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$$

Correction.

1. La matrice A est diagonalisable dans une base orthonormale de $M_{n,1}(\mathbb{R})$. De plus, on a $A = PD^tP$ où

$$D = \begin{pmatrix} -5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad P = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

2. On a $\begin{vmatrix} X-i & -1 \\ -1 & X+i \end{vmatrix} = X^2$. Par suite la matrice est nilpotente non nulle donc elle n'est pas diagonalisable.

Exercice 7. *norme d'opérateur d'un endomorphisme autoadjoint positif*

On suppose $\dim(E) > 0$. Soit $u \in \mathcal{L}(E)$. On note $|||u|||$ la norme d'opérateur de u de $(E, \|\cdot\|)$ dans lui-même.

1. Montrer :

$$|||u||| = \sup_{\|x\|=1, \|y\|=1} |(u(x)|y)| \,,$$

puis, si $u \in \mathcal{S}(E)$:

$$|||u||| = \sup_{\|x\|=1} |(u(x)|x)| \,.$$

2. Montrer que si $u \in \mathcal{S}(E)$, $|||u|||$ ou $-|||u|||$ est valeur propre de u et qu'il s'agit de la plus grande en valeur absolue.

Correction.

1. On note S la sphère unité de E et $A = \{|(u(x)|y)| \mid x, y \in S\}$ et $B = \{|(u(x)|x)| \mid x \in S\}$.
On a, pour tout $x, y \in S$, d'après l'inégalité de Cauchy-Schwarz :
$$|(u(x)|y)| \leq \underbrace{\|u(x)\|}_{\leq |||u|||.\|x\|}.\|y\| = |||u|||$$
Ainsi A est non vide (car S est non vide - ici $\dim(E) > 0$) et majoré par $|||u|||$ donc $\sup(A)$ existe et vérifie $\sup(A) \leq |||u|||$.
Notons $a = \sup(A)$. On a, pour tous $x, y \in E$:
$$|(u(x)|y)| \leq a\|x\|.\|y\|. \quad (*)$$
En effet, si $x = 0_E$ ou $y = 0_E$, l'inégalité est vérifié (c'est même une égalité) et sinon, on pose $x' = \frac{x}{\|x\|} \in S$ et $y' = \frac{y}{\|y\|} \in S$, et ainsi :
$$|(u(x)|y)| = \|x\|.\|y\|.|(u(x')|y')| \leq a\|x\|.\|y\|$$
Par suite, on a, pour $x \in E$, en posant $y = u(x) \in E$:
$$\|u(x)\|^2 = (u(x)|y) \leq a\|x\|.\|y\| = a\|u(x)\|.\|x\|$$
D'où, pour tout $x \in E \setminus \{0_E\}$:
$$\frac{\|u(x)\|}{\|x\|} \leq a$$
(le cas $u(x) = 0_E$ étant immédiatement vérifié). Ainsi, on a $|||u||| \leq a$.
Il en résulte que $|||u||| = \sup(A)$.
On a $B \subset A$ et B non vide car S est non vide, donc $b = \sup(B)$ existe et $b \leq a$.
De plus, on remarque que pour tous $x, y \in E$, comme $u^* = u$:
$$(u(x \pm y)|x \pm y) = (u(x)|x) \pm 2(u(x)|y) + (u(y)|y).$$
et, par un raisonnement similaire à celui décrit pour établir $(*)$, on a, pour tout $z \in E$,
$$(u(z)|z) \leq b\|z\|^2,$$
Ainsi, pour tous $x, y \in S$:
$$\begin{array}{lcl} |(u(x)|y)| & = & \frac{1}{4}\,|(u(x+y)|x+y) - (u(x-y)|x-y)| \\ & \leq & \frac{1}{4}(|(u(x+y)|x+y)| + |(u(x-y)|x-y)|) \\ & \leq & \frac{b}{4}\underbrace{(\|x+y\|^2 + \|x-y\|^2)}_{=2\|x\|^2+2\|y\|^2=4} \\ |(u(x)|y)| & \leq b. & \end{array}$$
Par suite, $a \leq b$. Et donc $b = a$.
Il en résulte que $|||u||| = b = a$.

2. Comme u est autoadjoint, u est diagonalisable dans une base orthonormale $\mathcal{B} = (e_1, ..., e_n)$.
Pour chaque $i \in [\![1, n]\!]$, on note λ_i la valeur propre à laquelle e_i est associée.

On peut reformuler l'assertion à démontrer par :

$$|||u||| = \max\{|\lambda| \mid \lambda \in \mathrm{Sp}(u)\}$$

Comme $\mathrm{Sp}(u) = \{\lambda_1, ..., \lambda_n\}$ est un ensemble fini non vide, il existe $i_0 \in [\![1, n]\!]$ tel que $|\lambda_{i_0}| = \max\{|\lambda| \mid \lambda \in \mathrm{Sp}(u)\}$.
Soit $x \in S$. Comme $\mathcal{B}$ est orthonormale, on a $x = \sum_{i=1}^n x_i e_i$ où $x_i = (x|e_i)$ et $\sum_{i=1}^n |x_i| = \|x\| = 1$.De plus, on a $u(x) = \sum_{i=1}^n x_i \lambda_i e_i$ car chaque e_i est vecteur propre associé à λ_i.
Par suite, $\mathcal{B}$ étant orthonormale, on a :

$$|(u(x)|x)| = \left|\sum_{i=1}^n \lambda_i.x_i\right| \leq \sum_{i=1}^n |x_i|.\underbrace{|\lambda_i|}_{\leq |\lambda_{i_0}|} \leq |\lambda_{i_0}|.$$

Ainsi, d'après la question précédente, on a $|||u||| \leq |\lambda_{i_0}|$.
De plus, on a, $e_{i_0} \in S$ et :

$$|(u(e_{i_0})|e_{i_0})| = |\lambda_{i_0}|.|(e_{i_0}|e_{i_0})| = |\lambda_{i_0}|$$

Donc $|||u||| = |\lambda_{i_0}| = \max\{|\lambda| \mid \lambda \in \mathrm{Sp}(u)\}$.
Il en résulte que $\lambda_{i_0} = \pm|||u|||$ est valeur propre de u et il s'agit de la plus grande en valeur absolue.

3. Endomorphismes autoadjoints positifs et définis positifs

a. Définitions

Définition 13.

Soit $u \in O(E)$. On dit que u est un endomorphisme autoadjoint :
- **positif** si, pour tout $x \in E$,
$$(u(x)|x) \geq 0;$$
- **défini positif** si, pour tout $x \in E \setminus \{0_E\}$,
$$(u(x)|x) > 0.$$

On note $\mathcal{S}^+(E)$ l'ensemble des endomorphismes autoadjoints positifs et $\mathcal{S}^{++}(E)$ l'ensemble des endomorphismes autoadjoints définis positifs.

Définition 14.

Soit $A \in \mathcal{S}_n(\mathbb{R})$. On dit que A est matrice symétrique :
- **positive** si, pour tout $X \in M_{n,1}(\mathbb{R})$,
$${}^tXAX \geq 0;$$
- **définie positive** si, pour tout $X \in M_{n,1}(\mathbb{R}) \setminus \{0_{n,1}\}$,
$${}^tXAX > 0.$$

On note $\mathcal{S}_n^+(\mathbb{R})$ l'ensemble des matrices symétriques positives et $\mathcal{S}_n^{++}(\mathbb{R})$ l'ensemble des matrices symétriques positives définies positives.

b. Propriétés

Proposition 33.

Soit $u \in \mathcal{S}(E)$.
- On a $u \in \mathcal{S}^+(E)$ si, et seulement si, $\mathrm{Sp}(u) \subset \mathbb{R}_+$.
- On a $u \in \mathcal{S}^{++}(E)$ si, et seulement si, $\mathrm{Sp}(u) \subset \mathbb{R}_+^*$.

De même dans le cas matriciel.

Démonstration.

Soit $u \in \mathcal{S}(E)$. Alors, pour tout $\lambda \in \mathrm{Sp}(u)$ et x vecteur propre associé à λ, on a, comme $x \neq 0$:

$$(u(x)|x) = (\lambda x|x) = \lambda \underbrace{(x|x)}_{>0}$$

donc $(u(x)|x)$ et λ sont de même signe.
On obtient alors les deux implications directes de la proposition.
a □

c. Exercices

Exercice 8. *Caractérisation des produits scalaires sur $M_{n,1}(\mathbb{R})$*

On pose $E = M_{n,1}(\mathbb{R})$. Soit $\varphi : E \times E \to \mathbb{R}$. Montrer que φ est un produit scalaire sur E si, et seulement si, il existe $A \in \mathcal{S}_n^{++}(\mathbb{R})$ tel que, pour tous $X, Y \in E$:

$$\varphi(X,Y) = {}^tXAY$$

Correction.

($\Rightarrow$) On suppose qu'il existe $A \in \mathcal{S}_n^{++}(\mathbb{R})$ tel que, pour tous $X, Y \in E$, $\varphi(X,Y) = {}^tXAY$.

1ère façon : sans rien remarquer ;)

Pour tous $X, Y, Z \in E$ et $\lambda \in \mathbb{R}$, on a :

- On remarque que pour tout $m \in \mathbb{R} = M_{1,1}(\mathbb{R})$, ${}^tm = m$ d'où, par symétrie de A :

$$\varphi(X,Y) = {}^tXAY = {}^t{}^tXAY = {}^tY{}^tAX = {}^tYAX = \varphi(Y,X)$$

  donc φ est symétrique.
- On a, par linéarité du passage à la transposée et par distributivité du produit matriciel :

$$\varphi(\lambda X + Y, Z) = {}^t\lambda X + YAZ = \lambda {}^tXAZ + {}^YAZ = \lambda\varphi(X,Z) + \varphi(Y,Z)$$

donc φ est linéaire par rapport à sa première variable, et ainsi, pas symétrie de φ, φ est bilinéaire.

- On remarque que $\varphi(X,X) = {}^tXAX$ donc, par positivité de A, $\varphi(X,X) \geq 0$ et par définie positivité de A, si $X \neq 0$, $\varphi(X,X) > 0$. Or on a $\varphi(0,0) = 0$, donc si $\varphi(X,X) = 0$ alors $X = 0$.

2ème façon : en remarquant que $\varphi(X,Y) =< X, AY >$ où $< \cdot, \cdot >$ est le produit scalaire canonique et en utilisant le fait que $A \in \mathcal{S}_n^{++}(\mathbb{R})$ (ce qui revient quasiment à faire la 1ère façon !).

De toutes ces façons, il en résulte que φ est une forme bilinéaire symétrique définie positive i.e. un produit scalaire.

($\Leftarrow$) On suppose que φ est un produit scalaire. On montre tout d'abord le résultat suivant :

> Soit $M = (m_{ij}) \in M_n(\mathbb{R})$. On note $(E_i)_{1 \leq i \leq n}$ la base canonique de $M_{n,1}(\mathbb{R})$.
> Pour tous $i, j \in [\![1, n]\!]$, on a :
> $$m_{ij} = {}^tE_i M E_j$$
> *Démonstration :* On a, pour tous $i, j \in [\![1, n]\!]$:
> $${}^tE_i M E_j = {}^tE_i \begin{pmatrix} m_{1j} \\ \vdots \\ m_{nj} \end{pmatrix} = m_{ij}.$$

Montrons qu'il existe $A \in \mathcal{S}_n^{++}(\mathbb{R})$ tel que, pour tous $X, Y \in E$, $\varphi(X,Y) = {}^tXAY$.

On considère $(E_i)_{1 \leq i \leq n}$ la base canonique de E et on pose, pour $i, j \in [\![1, n]\!]$, $a_{ij} = \varphi(E_i, E_j)$ puis $A = (a_{ij})_{1 \leq i,j \leq n}$.

Par suite, on a, pour tous $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in E$, $X = \sum_{i=1}^n x_i E_i$ et $Y = \sum_{i=1}^n y_i E_i$ et donc, par bilinéarité de φ :

$$\varphi(X,Y) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \underbrace{\varphi(E_i, E_j)}_{a_{ij}} = \sum_{i=1}^n x_i \underbrace{\left(\sum_{j=1}^n a_{i,j} y_j \right)}_{=(AY)_i} = {}^tXAY.$$

Il reste à montrer que $A \in \mathcal{S}_n^{++}(\mathbb{R})$.

- On a, pour tous $i, j \in [\![1, n]\!]$, par symétrie de φ :
$$a_{ji} = \varphi(E_j, E_i) = \varphi(E_i, E_j) = a_{ij}$$
donc ${}^tA = A$ i.e. $A \in \mathcal{S}_n(\mathbb{R})$.
- On a, pour tout $X \in E$ tel que $X \neq 0$, par définie positivité de φ :
$${}^tXAX = \varphi(X,X) > 0$$

Par suite, $A \in \mathcal{S}_n^{++}(\mathbb{R})$. D'où l'existence recherchée.

Exercice 9. *Racine carrée d'une matrice symétrique positive*

Soit $A \in \mathcal{S}_n^+(\mathbb{R})$.

1. Montrer qu'il existe $R \in \mathcal{S}_n(\mathbb{R})$ tel que $R^2 = A$.
2. On suppose $A \in \mathcal{S}_n^+ + (\mathbb{R})$. Montrer que $\{R \in M_n(\mathbb{R}) \mid R^2 = A\}$ est fini et déterminer son cardinal.
3. Si on suppose seulement $A \in \mathcal{S}_n^+(\mathbb{R})$, A a-t-elle un nombre fini de racines ?

Correction.

1. Comme $A \in \mathcal{S}_n(\mathbb{R})$, A est orthogonalement diagonalisable et $\mathrm{Sp}(A) \subset \mathbb{R}_+$. Ainsi, il existe $P \in O_n(\mathbb{R})$ et $D = \mathrm{diag}(\alpha_1, ..., \alpha_n)$ où, pour tout $i \in [\![1, n]\!]$, $\alpha_i \in \mathbb{R}_+$.
On note $D' = \mathrm{diag}(\sqrt{\alpha_1}, ..., \sqrt{\alpha_n})$ et $R = PD'^tP$. Alors on a :
$$R^2 = PD'^tPPD'^tP = PD'^{2t}P = PD^tP = A.$$
2. Reamrquons tout d'abord deux faits :
— On note $\alpha_1, ..., \alpha_k$ les valeurs propres distinctes de A. Alors, comme $A \in \mathcal{S}_n^+ + (\mathbb{R})$, pour tout $i \in [\![1, n]\!]$, $\alpha_k \in \mathbb{R}_+^*$.
— Considérons le polynôme minimal π_A de A. Alors $\pi_A = \prod_{i=1}^n (X - \alpha_k)$ car A est diagonalisable d'après le théorème spectral.
Soit R tel que $R^2 = A$. On pose $Q = \pi_A(X^2)$. Alors :
$$Q = \prod_{i=1}^n (X^2 - \alpha_k) = \prod_{i=1}^n (X - \sqrt{\alpha_k})(X + \sqrt{\alpha_k})$$
est un polynôme scindé à racines simples car les a_k étant tous différents, leurs racines le sont et comme les a_k sont non nuls, $\sqrt{\alpha_k} \neq -\sqrt{\alpha_k}$.
De plus on a :
$$Q(R) = \pi_A(R^2) = \pi_A(A) = 0,$$
donc R possède un polynôme annulateur scindé à racines simples, donc R est diagonalisable.
On remarque alors que R et A sont diagonalisables et commutent (on a $AR = R^3 = RA$) donc elle sont simultanément diagonalisables i.e. il existe $P \in GL_n(\mathbb{R})$ et D, D' diagonales telles que :
$$A = PDP^{-1} \text{ et } R = PD'P^{-1}$$

> Prouvons ce résultat classique : pour $M, N \in M_n(\mathbb{R})$, si M, N sont diagonalisables et commutent, alors elles sont simultanément diagonalisables.
>
> *Démonstration :* On suppose que $u, v \in \mathcal{L}(E)$ sont diagonalisables et commutent. Soit λ une valeur propre de u (qui en possède au moins une car u est diagonalisable). Comme v et $u - \lambda \mathrm{Id}_E$ commutent, $F = E_\lambda(u)$ est stable par v. Ainsi, l'endomorphisme induit v_F par v sur F est diagonalisable car v l'est. Par suite, il existe $\mu \in \mathbb{R}$ tel que, pour tout $x \in E$,
> $$v(x) = v_F(x) = \mu x$$
> Il en résulte que tout vecteur propre de u est une vecteur propre de v et vice-versa par un raisonnement analogue.
> Ainsi, si $\mathcal{B}$ est une base de E qui diagonalise u, alors $\mathcal{B}$ diagonalise v et réciproquement i.e. u et v sont diagonalisables dans une même base.
>
> Maintenant, pour $M, N \in M_n(\mathbb{R})$ qui sont diagonalisables et qui commutent, on considère leurs endomorphismes de $\mathbb{R}^n$ canoniquement associés ; ceux-ci sont donc diagonalisables et commutent donc d'après ce qui précède, ils sont diagonalisables dans une même base $\mathcal{B}$ de $\mathbb{R}^n$. Il en résulte que M et N sont toutes deux semblables à des matrices diagonales avec la même matrice de passage : la matrice de passage de la base canonique de $\mathbb{R}^n$ vers $\mathcal{B}$. Par suite, M et N sont simultanément diagonalisables.

On a donc :
$$D'^2 = D$$

Chapitre X

Suites, séries d'intégrales et intégrales à paramètres

Table des matières

Dans ce chapitre, $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$; les intervalles considérés sont supposés d'intérieur non vide et a désigne un nombre réel.

Partie A

Suites et séries d'intégrales

Dans cette partie, I désigne un intervalle d'intérieur non vide.

1. Théorème de convergence dominée

Théorème 1. *Théorème de convergence dominée*

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de I dans $\mathbb{K}$. Si les propriétés suivantes sont vérifiées :

- pour tout $n \in \mathbb{N}$, f_n est continue par morceaux sur I.
- la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur I vers une fonction f et f est continue par morceaux ;
- il existe une fonction g intégrable sur I qui domine tous les $|f_n|$, *i.e.* pour tout $n \in \mathbb{N}$, pour tout $t \in I$,

$$|f_n(t)| \leq g(t) \qquad \textbf{hypothèse de domination}$$

alors, les f_n et f sont intégrables sur I et on a :

$$\lim_{n\to+\infty} \int_I f_n(t)\,\mathrm{d}t = \int_I \left(\lim_{n\to+\infty} f_n(t)\right)\,\mathrm{d}t = \int_I f(t)\,\mathrm{d}t.$$

Exemple 1.

Pour $n \in \mathbb{N}^*$, l'intégrale $I_n = \displaystyle\int_0^{+\infty} \frac{\cos(nx)}{1+n^2x^2}\,\mathrm{d}x$ est bien définie et $I_n \xrightarrow[n\to+\infty]{} 0$.

Pour $n \in \mathbb{N}^*$, on pose $f_n : x \mapsto \frac{\cos(nx)}{1+n^2x^2}$.
Pour tout $n \in \mathbb{N}^*$, f_n est continue sur $[0, +\infty[$ comme quotient de fonctions continues sur $[0, +\infty[$ dont le dénominateur ne s'annule pas. De plus, on a, pour tout $n \in \mathbb{N}^*$:

$$|f_n(x)| \leq \frac{1}{1+n^2x^2} \underset{n\to+\infty}{\sim} \frac{1}{n^2x^2}$$

Or l'intégrale de Riemann $\int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x$ converge car $2 > 1$, donc par comparaison f_n est intégrable sur $[1, +\infty[$ et donc sur $[0, +\infty[$ par continuité des f_n sur le segment $[0, 1]$.

Par suite, pour tout $n \in \mathbb{N}^*$, $\int_0^{+\infty} \dfrac{\cos(nx)}{1+n^2x^2}\,\mathrm{d}x$ converge et donc I_n est bien définie.

Déterminons cette limite grâce au théorème de convergence dominée :

- pour tout $n \in \mathbb{N}^*$, f_n est continue sur $[0, +\infty[$.
- CVS sur $[0, +\infty[$:
  Soit $x \in \mathbb{R}_+$. On a, si $x \neq 0$:

  $$|f_n(x)| \leq \frac{1}{1+n^2x^2} \xrightarrow[n\to+\infty]{} 0$$

  et $f_n(0) = 1 \xrightarrow[n\to+\infty]{} 1$. Ainsi, $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $[0, +\infty[$ vers $f : x \mapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{si } x > 0 \end{cases}$ qui est continue par morceaux sur $[0, +\infty[$.
- Pour tout $n \in \mathbb{N}^*$, pour tout $x \in [0, +\infty[$,

  $$|f_n(x)| \frac{1}{1+n^2x^2} \leq \frac{1}{1+x^2} = g(t) \qquad \textbf{hypothèse de domination}$$

  et $g : x \mapsto \frac{1}{1+x^2}$ est intégrable sur $[0, +\infty[$ car continue sur $[0, +\infty[$ et $g(x) \underset{n\to+\infty}{\sim} \frac{1}{x^2}$.

Par suite, les f_n et f sont intégrables sur $[0, +\infty[$ (on le savait déjà !) et on a :

$$\lim_{n\to+\infty} I_n = \lim_{n\to+\infty} \int_0^{+\infty} f_n(t)\,\mathrm{d}t = \int_0^{+\infty} \left(\lim_{n\to+\infty} f_n(t)\right) \mathrm{d}t = \int_0^{+\infty} f(t)\,\mathrm{d}t = 0.$$

Remarque 1.

Attention, l'hypothèse de domination est **essentielle** ! Par exemple, la suite de fonctions $(f_n)_{n\in\mathbb{N}^*}$ telle que, pour $n \in \mathbb{N}^*$

$$\text{sur } [0, 1[, \quad f_n : t \mapsto n^2 t^{n-1}.$$

converge simplement vers 0, mais la limite des intégrales n'est pas nulle !

Exercice 1.

Justifier l'existence des intégrales I_n et calculer la limite de la suite (I_n) pour :

1. $I_n = \displaystyle\int_0^{+\infty} \frac{1}{1+t^2+t^n e^{-t}} \,\mathrm{d}t$.
2. $I_n = \displaystyle\int_0^{1} \frac{1}{(1-t^n)^{\frac{1}{n}}} \,\mathrm{d}t$.

Correction.

1. Pour $n \in \mathbb{N}$, on pose $f_n : t \mapsto \frac{1}{1+t^2+t^n e^{-t}}$.
   — Montrons que, pour tout $n \in \mathbb{N}$, $\int_0^{+\infty} \frac{1}{1+t^2+t^n e^{-t}} \,\mathrm{d}t$ converge.
   Soit $n \in \mathbb{N}$. La fonction f_n est continue sur $[0, +\infty[$ et on a, pour tout $t \in [0, +\infty[$, $|f_n(t)| \leq \frac{1}{1+t^2}$; donc f_n est intégrable sur $[0, +\infty[$ par comparaison à la fonction $t \mapsto \frac{1}{1+t^2}$ intégrable sur $[0, +\infty[$ car continue sur $[0, +\infty[$ et $\frac{1}{1+t^2} \underset{t\to+\infty}{\sim} \frac{1}{t^2}$ dont l'intégrale converge sur $[1, +\infty[$ (d'après le critère de Riemann en $+\infty$ - $2 > 1$).
   Par suite, $\int_0^{+\infty} \frac{1}{1+t^2+t^n e^{-t}} \,\mathrm{d}t$ converge et donc I_n est bien définie.
   — Déterminons la limite potentielle de la suite $(I_n)_{n\in\mathbb{N}}$. On vérifie les hypothèses du théorème de convergence dominée :
   - pour tout $n \in \mathbb{N}$, f_n est continue sur $[0, +\infty[$.
   - la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement sur $[0, +\infty[$ vers $f : t \mapsto \begin{cases} \frac{1}{1+t^2} & \text{si } t \in [0,1[\\ \frac{1}{2+e^{-1}} & \text{si } t = 1 \\ 0 & \text{si } t > 1 \end{cases}$

     qui est continue par morceaux sur $[0, +\infty[$;
   - pour tout $n \in \mathbb{N}$, pour tout $t \in [0, +\infty[$,
     $$|f_n(t)| \leq \frac{1}{1+t^2} = g(t) \qquad \textbf{hypothèse de domination}$$
     et $g : t \mapsto \frac{1}{1+t^2}$ est intégrable sur $[0, +\infty[$.

   Ainsi, d'après le théroème de convergence dominée les f_n et f sont intégrables sur $[0, +\infty[$ et on a :
   $$\lim_{n\to+\infty} I_n = \int_0^{+\infty} f(t) \,\mathrm{d}t = \int_0^1 \frac{1}{1+t^2} \,\mathrm{d}t = [\arctan(t)]_0^1 = \frac{\pi}{4}.$$
2. Pour $n \in \mathbb{N}^*$, on pose $f_n : t \mapsto \frac{1}{(1-t^n)^{\frac{1}{n}}}$.
   — Pour tout $n \in \mathbb{N}^*$, déterminons la nature de $\int_0^1 f_n(t) \,\mathrm{d}t$ converge.
   Soit $n \in \mathbb{N}^*$. La fonction f_n est continue sur $[0, 1[$ et on a, pour tout $t \in [0, 1[$,
   $$1 - t^n = (1-t) \sum_{k=0}^{n-1} t^k \underset{t\to1}{\sim} n(1-t)$$
   d'où :
   $$|f_n(t)| = \frac{1}{(1-t^n)^{\frac{1}{n}}} \underset{t\to1}{\sim} \frac{1}{\sqrt[n]{n}(1-t)^{\frac{1}{n}}}$$

Or, d'après le critère de Riemann en 1, $\int_0^1 \frac{1}{(1-t)^{\frac{1}{n}}}\,dt$ converge si, et seulement si, $\frac{1}{n} < 1$ *i.e.* $n > 1$.
Par suite, par comparaison, f_n est intégrable sur $[0, 1[$ si, et seulement si, $n \geq 2$.
La fonction f_n étant positive, son intégrale sur $[0, 1[$ converge si, et seulement si, elle est intégrable sur $[0, 1[$ et donc I_n est bien définie pour tout entier $n \geq 2$ (et $I_1 = +\infty$).

— Déterminons la limite potentielle de la suite $(I_n)_{n\geq 2}$. On vérifie les hypothèses du théorème de convergence dominée :
- pour tout entier $n \geq 2$, f_n est continue sur $[0, 1[$.
- la suite $(f_n)_{n\geq 2}$ converge simplement sur $[0, +\infty[$ vers $f : t \mapsto 1$ qui est continue sur $[0, 1[$;
- pour tout entier $n \geq 2$, pour tout $t \in [0, 1[$,

$$|f_n(t)| \leq \frac{1}{\sqrt{1-t}} = g(t) \qquad \textbf{hypothèse de domination}$$

et $g : t \mapsto \frac{1}{\sqrt{1-t}}$ est intégrable sur $[0, +1[$ d'après le critère de Riemann en 1 ($\frac{1}{2} < 1$).

Ainsi, d'après le théroème de convergence dominée les f_n et f sont intégrables sur $[0, 1[$ et on a :

$$\lim_{n\to+\infty} I_n = \int_0^1 f(t)\,\mathrm{d}t = \int_0^1 1\,\mathrm{d}t = 1.$$

Problème 1. *Calcul de l'intégrale de Gauss*

Pour $n \in \mathbb{N}^*$, on considère la fonction $f_n : \mathbb{R}_+ \to \mathbb{R}$ définie, pour $x \in \mathbb{R}_+$, par :

$$f_n(t) = \begin{cases} \left(1 - \frac{t^2}{n}\right)^n & \text{si } t \leq \sqrt{n} \\ 0 & \text{si } t > \sqrt{n}. \end{cases}$$

1. Montrer que la suite $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathbb{R}_+$ vers une fonction f qu'on explicitera.
2. Montrer que f est intégrable sur $\mathbb{R}_+$ et que :

$$\int_0^{+\infty} f(t)\,\mathrm{d}t = \lim_{n\to+\infty} \int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t.$$

3. pour $n \in \mathbb{N}$, on note $W_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n(t)\,\mathrm{d}t$.
   (a) En utilisant le changement de variable $x = \arccos\left(\frac{t}{\sqrt{n}}\right)$, montrer que, pour tout $n \in \mathbb{N}^*$:

$$\int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t = \sqrt{n}W_{2n+1}$$

   (b) On rappelle le résultat *classique* $W_n \underset{n\to+\infty}{\sim} \sqrt{\frac{\pi}{2n}}$. En déduire la limite de $\displaystyle\int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t$ quand n tend vers $+\infty$.
4. En déduire que la valeur de l'intégrale de Gauss $\displaystyle\int_0^{+\infty} e^{-t^2}\,\mathrm{d}t$.

Correction.

1. CVS sur $\mathbb{R}_+$:
Soit $t \in \mathbb{R}_+$. Alors, à partir d'un certain rang $N \in \mathbb{N}^*$ (par exemple : $N = \lfloor t^2 \rfloor + 1$), pour tout entier $n \geq N$, $t \leq \sqrt{n}$ et donc, comme $\ln(1-x) = -x + o(x)$ quand $x \to 0$, on a :
$$f_n(t) = \left(1 - \frac{t^2}{n}\right)^n = e^{n\ln(1-\frac{t^2}{n})} = e^{-t^2} e^{o(1)} \xrightarrow[n\to+\infty]{} e^{-t^2}$$
Par suite, $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathbb{R}_+$ vers une fonction $f : t \mapsto e^{-t^2}$.
2. On applique le théorème de convergence dominée :
   - pour tout $n \in \mathbb{N}^*$, f_n est continue sur $\mathbb{R}_+$ car continue sur $[0, \sqrt{n}[$ et $]\sqrt{n}, +\infty[$ et les limites à gauche et à droite de f_n en $\sqrt{n}$ valent toutes deux 0.
   - la suite $(f_n)_{n\in\mathbb{N}^*}$ converge simplement sur $\mathbb{R}_+$ vers une fonction $f : t \mapsto e^{-t^2}$ (question 1) et f est continue (composée de fonctions continues) ;
   - Pour tout $n \in \mathbb{N}^*$, pour tout $t \in \mathbb{R}_+$,
   $$|f_n(t)| \leq f(t) \qquad \textbf{hypothèse de domination}$$
   Montrons cette inégalité. Comme f est positive, il suffit de la montrer sur $[0, \sqrt{n}]$. On a, pour tout $x \in \mathbb{R}$, $e^x \geq 1 + x$ (par convexité de l'exponentielle). De plus, pour tout $t \in [0, \sqrt{n}]$, en prenant $x = -\frac{t^2}{n}$, on a $1 + x \geq 0$ et donc, par croissance de la fonction $u \mapsto u^n$ sur $\mathbb{R}_+$, on obtient $\left(1 - \frac{t^2}{n}\right)^n \leq (e^{-\frac{t^2}{n}})^n = f(t)$.
   La fonction f est continue sur $\mathbb{R}_+$ et $f(t) = o(\frac{1}{x^2})$ en $+\infty$ donc, par comparaison à une intégrale de Riemann convergente $(2 > 1)$, f est intégrable sur $[0, +\infty[$.

   Ainsi, d'après le théorème de convergence dominée, (les f_n et f sont intégrables sur $\mathbb{R}_+$ - on le sait déjà !) et on a :
$$\int_0^{+\infty} f(t)\,\mathrm{d}t = \lim_{n\to+\infty} \int_0^{+\infty} f_n(t)\,\mathrm{d}t = \lim_{n\to+\infty} \int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t.$$
3. (a) On applique le changement de variable annoncé - qui est bien valable car la fonction arcos est strictement monotone et C^1 sur $[0, 1]$ (et $\frac{t}{\sqrt{n}} \in [0, 1]$ pour $t \in [0, \sqrt{n}]$) ; ce qui donne :
$$\begin{aligned}
\int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t &= \int_{\frac{\pi}{2}}^{0} \left(1 - \cos^2(x)\right)^n \left(-\sqrt{n}\sin(x)\right)\mathrm{d}x \\
&= \sqrt{n} \int_0^{\frac{\pi}{2}} \left(\sin^2(x)\right)^n \sin(x)\,\mathrm{d}x \\
&= \sqrt{n} \int_0^{\frac{\pi}{2}} \sin^{2n+1}\,\mathrm{d}x \\
\int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t &= \sqrt{n} W_{2n+1}.
\end{aligned}$$
   (b) On a :
$$\int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t = \sqrt{n} W_{2n+1} \underset{n\to+\infty}{\sim} \sqrt{n}\sqrt{\frac{\pi}{4n+2}} \underset{n\to+\infty}{\sim} \frac{\sqrt{\pi}}{2}.$$
4. Il en résulte, d'après la question 2 et la question 3b, que :
$$\int_0^{+\infty} e^{-t^2}\,\mathrm{d}t = \int_0^{+\infty} f(t)\,\mathrm{d}t = \lim_{n\to+\infty} \int_0^{\sqrt{n}} f_n(t)\,\mathrm{d}t = \frac{\sqrt{\pi}}{2}.$$

Théorème 2. *Théorème de convergence dominée (cas continu)*

Soit J un intervalle de $\mathbb{R}$, $(f_\lambda)_{\lambda\in J}$ une suite de fonctions de I dans $\mathbb{K}$ et $\lambda_0 \in \bar{J}$. Si les propriétés suivantes sont vérifiées :

- pour tout $\lambda \in J$, f_λ est continue par morceaux sur I.
- pour tout $t \in I$, $f_\lambda(t) \xrightarrow[\lambda\to\lambda_0]{} f(t)$ où $f : I \to \mathbb{K}$ est continue par morceaux sur I ;
- il existe une fonction g **intégrable** sur I qui **domine** tous les $|f_\lambda|$, *i.e.* pour tout $\lambda \in J$, pour tout $t \in I$,
$$|f_\lambda(t)| \leq g(t) \qquad \textbf{hypothèse de domination}$$

alors, les f_λ et f sont intégrables sur I et on a :

$$\lim_{\lambda\to\lambda_0} \int_I f_\lambda(t)\,\mathrm{d}t = \int_I \left(\lim_{\lambda\to\lambda_0} f_\lambda(t)\right)\,\mathrm{d}t = \int_I f(t)\,\mathrm{d}t.$$

Exercice 2.

1. Soit $\lambda \in \mathbb{R}_+$. Montrer que :
$$\int_0^{+\infty} \frac{e^{-\lambda t}}{1+t^2}\,\mathrm{d}t$$
converge. On note $I(\lambda)$ la valeur de cette intégrale.
2. Montrer que la fonction I admet une limite en $+\infty$ et la déterminer. *Remarque : on déterminera la valeur de deux façons ; en utilisant le théorème pour se faire la main ; puis en étant malin !*

Correction.

Pour $\lambda \in \mathbb{R}_+$, on pose $f_\lambda : t \mapsto \frac{e^{-\lambda t}}{1+t^2}$.

1. Soit $\lambda \in \mathbb{R}_+$. Alors f_λ est continue sur $[0,+\infty[$ et, pour tout $t \in [0,+\infty[$, $|f_\lambda(t)| \leq \frac{1}{1+t^2} \underset{n\to+\infty}{\sim} \frac{1}{t^2}$ donc f_λ est intégrable sur $[0,+\infty[$ par continuité sur le segment $[0,1]$ et par comparaison à une intégrale de Riemann convergente $(2>1)$ sur $[1,+\infty[$.
Par suite, $\int_0^{+\infty} \frac{e^{-\lambda t}}{1+t^2}\,\mathrm{d}t$ converge.
2. On remarque que par positivité des f_λ sur $[0,+\infty[$ et par croissance l'intégrale, $I(\lambda) \geq 0$ pour tout $\lambda \in \mathbb{R}_+$. De plus, pour tous $\lambda, \lambda' \in \mathbb{R}_+$ avec $\lambda \leq \lambda'$, par croissance de l'exponentielle, on a $f_\lambda \geq f_{\lambda'}$ sur $[0,+\infty[$; donc, par croissance de l'intégrale, $I(\lambda) \geq I(\lambda')$.
La fonction $I : \mathbb{R}_+ \to \mathbb{R}$ est donc décroissante, minorée par 0, d'où, d'après le théorème de la limite monotone, I admet une limite finie en $+\infty$.
Déterminons cette limite :
1ère façon : avec le théorème de convergence dominée (cas continu).
On vérifie les hypothèses du théorème :
   - pour tout $\lambda \in \mathbb{R}_+$, f_λ est continue sur $[0,+\infty[$.
   - pour tout $t \in [0,+\infty[$, on a
$$f_\lambda(t) = \frac{e^{-\lambda t}}{1+t^2} \xrightarrow[\lambda\to+\infty]{} 0$$
et $f : t \mapsto 0$ est continue sur $[0,+\infty[$;

- On a, pour tout $\lambda \in \mathbb{R}_+$, pour tout $t \in [0, +\infty[$,

$$|f_\lambda(t)| \leq \frac{1}{1+t^2} = g(t) \qquad \textbf{hypothèse de domination}$$

et $g : t \mapsto \frac{1}{1+t^2}$ est intégrable sur $[0, +\infty[$ car continue sur $[0, +\infty[$ et intégrable au voisinage de $+\infty$ par comparaison à une intégrable de Riemann convergente.

Ainsi, d'après le théorème de convergence dominée (cas continu), les f_λ et f sont intégrables sur $[0, +\infty[$ (on le savait déjà !) et on a :

$$\lim_{\lambda \to +\infty} I(\lambda) = \int_0^{+\infty} \lim_{\lambda \to +\infty} f_\lambda(t)\,\mathrm{d}t = \int_0^{+\infty} 0\,\mathrm{d}t = 0.$$

2nde façon : par majoration de $I(\lambda)$ par une intégrale connue.
Pour tout $\lambda > 0$, on a, par croissance de l'intégrale :

$$I(\lambda) = \int_0^{+\infty} \frac{e^{-\lambda t}}{1+t^2}\,\mathrm{d}t \leq \int_0^{+\infty} e^{-\lambda t}\,\mathrm{d}t = \left[-\frac{e^{-\lambda t}}{\lambda}\right]_0^{+\infty} = \frac{1}{\lambda} \xrightarrow[\lambda \to +\infty]{} 0.$$

Ainsi, comme $I(\lambda) \geq 0$, d'après le théorème des gendarmes, $I(\lambda) \xrightarrow[\lambda \to +\infty]{} 0$.

2. Intégration terme à terme d'une série de fonctions

Théorème 3. *Intégration terme à terme (cas à valeurs positives)*

Soit $\sum f_n$ une série de fonctions de I dans $\mathbb{R}$. Si les propriétés suivantes sont vérifiées :

- pour tout $n \in \mathbb{N}$, f_n est à valeurs dans $\mathbb{R}_+$ et continue par morceaux sur I ;
- pour tout $n \in \mathbb{N}$, f_n est intégrable sur I ;
- la série de fonctions $\sum f_n$ converge simplement sur I et sa somme $S : t \mapsto \sum_{n=0}^{+\infty} f_n(t)$ est continue par morceaux sur I,

alors, l'égalité suivante est valable dans $[0, +\infty]$:

$$\int_I S(t)\,\mathrm{d}t = \int_I \left(\sum_{n=0}^{+\infty} f_n\right)\mathrm{d}t = \sum_{n=0}^{+\infty}\left(\int_I f_n(t)\,\mathrm{d}t\right).$$

En particulier, S est intégrable sur I si, et seulement si,

$$\sum \left(\int_I f_n(t)\,\mathrm{d}t\right) \text{ converge.}$$

Remarque 2.

Le théorème précédent nous dit donc que si (f_n) est une suite de fonctions **positives** sur I tel que $\sum f_n$ **converge simplement** sur I - dans les faits, on vérifiera succinctement et sans écrire toutes les justifications, les hypothèses de continuité par morceaux, alors on peut intervertir toujours la somme et l'intégrale.

Et ce, <u>même si un des deux termes vaut $+\infty$</u> : en fait, dans ce cas, l'intégrale finale diverge si, et seulement si, la série finale diverge - c'est ce que dit le "en particulier" du théorème d'ailleurs !

Très souvent on utilisera ce théorème (et le suivant) pour calculer l'intégrale d'une fonction f que l'on peut, en partie, développer en série entière :

Exemple 2.

On a $\displaystyle\int_0^1 \frac{\ln(t)}{t-1}\,\mathrm{d}t = \zeta(2)$.

On a remarque que, pour $t \in]0,1[$,

$$\frac{\ln(t)}{t-1} = \sum_{n=0}^{+\infty} (-t^n \ln(t)).$$

Ainsi, on pose, pour $n \in \mathbb{N}$, $f_n : t \mapsto -t^n \ln(t)$.

- pour tout $n \in \mathbb{N}$, f_n est positive sur $]0,1[$ car $-\ln$ l'est ;
- On a, pour tout $n \in \mathbb{N}$ et tout $t \in]0,1[$, $|f_n(t)| \leq |\ln(t)|$ donc f_n est intégrable sur $]0,1[$.
- $\sum f_n$ converge simplement sur $]0,1[$ vers $t \mapsto \frac{\ln(t)}{t-1}$ qui est continue sur $]0,1[$.

Les hypothèses du théorème d'intégration terme à terme (cas positif) sont donc vérifiées, d'où, dans $[0,+\infty]$:

$$\begin{aligned}
\int_0^1 \frac{\ln(t)}{t-1}\,\mathrm{d}t &= -\sum_{n=0}^{+\infty} \int_0^1 t^n \ln(t)\,\mathrm{d}t \\
&= \sum_{n=0}^{+\infty} \frac{1}{(n+1)^2} \quad \text{via une IPP} \\
&= \sum_{n=1}^{+\infty} \frac{1}{n^2}
\end{aligned}$$

Or la série $\sum_{n\geq 1} \frac{1}{n^2}$ converge vers $\zeta(2)$ donc, d'après le théorème d'intégration terme à terme, $S : t \mapsto \dfrac{\ln(t)}{t-1}$ est intégrable sur $]0,1[$ (*on aurait d'ailleurs pu le montrer dès le début !*) et on a :

$$\int_0^1 \frac{\ln(t)}{t-1}\,\mathrm{d}t = \sum_{n=1}^{+\infty} \frac{1}{n^2} = \zeta(2).$$

Exercice 3.

Calculer les intégrales suivantes en fonction de $\zeta(3)$:

1. $\displaystyle\int_0^{+\infty} \frac{\ln(t)\ln(1-t)}{t}\,\mathrm{d}t$;
2. $\displaystyle\int_0^{+\infty} \frac{t^2}{e^t-1}\,\mathrm{d}t$.

Correction.

1. On note $f : t \mapsto \frac{\ln(t)\ln(1-t)}{t}$. On a remarque que, pour $t \in]-1,1[$,

$$\ln(1-t) = \sum_{n=0}^{+\infty}(-\frac{t^{n+1}}{n+1}).$$

d'où, pour tout $t \in]0,1[$:

$$f(t) = \sum_{n=0}^{+\infty}(-\frac{\ln(t)t^n}{n+1}).$$

Ainsi, on pose, pour $n \in \mathbb{N}$, $f_n : t \mapsto -\frac{\ln(t)t^n}{n+1}$.

- pour tout $n \in \mathbb{N}$, f_n est positive sur $]0,1[$ car $-\ln$ l'est ;
- On a, pour tout $n \in \mathbb{N}$ et tout $t \in]0,1[$, $|f_n(t)| \leq \frac{|\ln(t)|}{n+1}$ donc f_n est intégrable sur $]0,1[$.
- $\sum f_n$ converge simplement sur $]0,1[$ vers f qui est continue sur $]0,1[$.

Les hypothèses du théorème d'intégration terme à terme (cas positif) sont donc vérifiées, d'où, dans $[0,+\infty]$:

$$\begin{aligned}\int_0^1 \frac{\ln(t)\ln(1-t)}{t}\,\mathrm{d}t &= -\sum_{n=0}^{+\infty}\int_0^1 \frac{t^n\ln(t)}{n+1}\,\mathrm{d}t \\ &= \sum_{n=0}^{+\infty}\frac{1}{(n+1)^3} \quad \text{via une IPP} \\ &= \sum_{n=1}^{+\infty}\frac{1}{n^3}\end{aligned}$$

Or la série $\sum_{n\geq 1}\frac{1}{n^3}$ converge vers $\zeta(3)$ donc, d'après le théorème d'intégration terme à terme, f est intégrable sur $]0,1[$ (*on aurait encore une fois pu le montrer dès le début !*) et on a :

$$\int_0^1 \frac{\ln(t)\ln(1-t)}{t}\,\mathrm{d}t = \sum_{n=1}^{+\infty}\frac{1}{n^3} = \zeta(3).$$

2. On note $f : t \mapsto \dfrac{t^2}{e^t-1}$. Pour $t > 0$, on a $e^{-t} \in]0,1[$, d'où :

$$\begin{aligned} f(t) &= t^2e^{-t}\frac{1}{1-e^{-t}} \\ &= t^2e^{-t}\sum_{n=0}^{+\infty}\left(e^{-t}\right)^n \\ &= \sum_{n=0}^{+\infty} t^2e^{-(n+1)t} \end{aligned}$$

On pose alors $f_n : x \mapsto t^2e^{-(n+1)t}$.

- pour $n \in \mathbb{N}$, f_n est positive et continue sur $[0,+\infty[$;
- pour $n \in \mathbb{N}$, f_n est continue sur $[0,+\infty[$ et on a $f_n(t) = \underset{x\to+\infty}{o}(\frac{1}{t^2})$, donc f_n est intégrable sur $[0,+\infty[$.
- $\sum f_n$ converge simplement sur $]0,+\infty[$ vers f qui est continue sur $]0,+\infty[$.

Ainsi, d'après le théorème d'intégration terme à terme sont donc vérifiées, d'où, dans $[0,+\infty]$:

$$\int_0^{+\infty}\frac{t^2}{e^t-1}\,\mathrm{d}t = \sum_{n=0}^{+\infty}\left(\int_0^{+\infty} t^2e^{-(n+1)t}\,\mathrm{d}t\right).$$

De plus, la série numérique $\sum\int_0^{+\infty} f_n(t)\,\mathrm{d}t$ converge car, en effectuant deux IPP, on obtient l'égalité, pour $n \in \mathbb{N}$:

$$\int_0^{+\infty} t^2e^{-(n+1)t}\,\mathrm{d}t = \left[-\frac{t^2e^{-(n+1)t}}{n+1}\right]_0^{+\infty} + \left[-\frac{2te^{-(n+1)t}}{(n+1)^2}\right]_0^{+\infty} + \left[-\frac{2e^{-(n+1)t}}{(n+1)^3}\right]_0^{+\infty} = \frac{2}{(n+1)^3}.$$

et $\sum\frac{2}{(n+1)^3}$ converge (série de Riemann convergente).

Il en résulte, d'après que le théorème d'intégration terme à terme, que f est intégrable sur $]0,+\infty[$ et on a :

$$\int_0^{+\infty}\frac{t^2}{e^t-1}\,\mathrm{d}t = \sum_{n=0}^{+\infty}\frac{2}{(n+1)^3} = 2\sum_{n=1}^{+\infty}\frac{1}{n^3} = 2\zeta(3).$$

Remarque : on aurait également pu effectuer au départ un changement de variable pour montrer :

$$\int_0^{+\infty}\frac{t^2}{e^t-1}\,\mathrm{d}t = \int_0^1\frac{\ln(u)^2}{1-u}\,\mathrm{d}u,$$

puis effectuer un raisonnement similaire à celui de l'exemple précédent pour trouver le résultat.

Problème 2. *Calcul de $\zeta(2)$*

1. (a) Rappeler (ou retrouver) le développement en série entière de arcsin sur $]-1,1[$; ou plus précisément, la suite $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ telle que, pour tout $x \in]-1,1[$:

$$\arcsin(x) = \sum_{n=0}^{+\infty} a_nx^{2n+1}.$$

(b) En déduire que, pour tout $t \in [0, \frac{\pi}{2}[$:

$$\sum_{n=0}^{+\infty} a_n \sin^{2n+1}(t) = t.$$

2. On pose, pour $n \in \mathbb{N}$, $W_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n(t)\,\mathrm{d}t$. Montrer que :

$$\sum_{n=0}^{+\infty} a_n W_{2n+1} = \frac{\pi^2}{8}.$$

3. Montrer que, pour tout $n \in \mathbb{N}$, $a_n W_{2n+1} = \frac{1}{(2n+1)^2}$.

4. En déduire la valeur de $\displaystyle\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2}$ puis de $\zeta(2) = \displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^2}$.

Correction.

1. (a) On rappelle (voire chapitre "Séries entières") que, pour $n \in \mathbb{N}$:

$$a_n = \frac{1 \times 3 \times \ldots \times (2n-1)}{2 \times 4 \times \ldots \times (2n)} \frac{1}{2n+1} = \frac{(2n)!}{4^n (n!)^2} \frac{1}{2n+1} = \frac{\binom{2n}{n}}{4^n} \frac{1}{2n+1}.$$

   (b) Soit $t \in [0, \frac{\pi}{2}[$. On pose $x = \sin(t)$. Alors $x \in [0, 1[$, donc, d'après la question précédente :

$$t = \arcsin(x) = \sum_{n=0}^{+\infty} a_n x^{2n+1} = \sum_{n=0}^{+\infty} a_n \sin^{2n+1}(t).$$

2. On pose, pour $n \in \mathbb{N}$, $f_n : t \mapsto a_n \sin^{2n+1}(t)$.
   - pour tout $n \in \mathbb{N}$, f_n est positive sur $[0, \frac{\pi}{2}[$ car a_n et sin le sont ;
   - On a, pour tout $n \in \mathbb{N}$, f_n est continue sur $[0, \frac{\pi}{2}]$ donc f_n est intégrable sur $[0, \frac{\pi}{2}[$.
   - D'après la question précédente, $\sum f_n$ converge simplement sur $[0, \frac{\pi}{2}[$ vers $S : t \mapsto t$ qui est continue sur $[0, \frac{\pi}{2}[$ (*ça ne semble pas très satisfaisant de justifier la CVS de cette façon même si c'est tout à fait correct... mais on peut tout de même, pour être "sûr", montrer que $\sum f_n$ CVS sur $[0, \frac{\pi}{2}[$ directement - cela revient quasiment au calcul du rayon de convergence de $\sum a_n x^n$!*).

   Par suite, d'après le théorème d'intégration terme à terme, S étant continue sur $[0, \frac{\pi}{2}]$ et donc intégrable sur $[0, \frac{\pi}{2}[$, la série $\sum a_n (\int_0^{\frac{\pi}{2}} \sin^{2n+1}(t)\,\mathrm{d}t)$ converge et on a :

$$\sum_{n=0}^{+\infty} a_n W_{2n+1} = \sum_{n=0}^{+\infty} a_n \int_0^{\frac{\pi}{2}} \sin^{2n+1}(t)\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} S(t)\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} t\,\mathrm{d}t = \frac{\pi^2}{8}.$$

Remarque : On pouvait se passer de l'utilisation du théorème d'intégration terme à terme ici :

En effet, supposons que $\sum a_n$ converge et posons $g_n : x \mapsto a_n x^n$.

— Comme (a_n) est à valeurs positives, on a, sur $[-1,1]$, $\|g_n\|_\infty = a_n$ donc $\sum \|g_n\|_\infty$ converge ; d'où la convergence normale sur $[-1,1]$.

  Les g_n sont continues sur $[-1,1]$ car polynomiaux et $\sum g_n$ converge uniformément sur $[-1,1]$ vers sa somme g donc g est continue. Par suite, comme arcsin est continue sur $[-1,1]$ et $\arcsin_{|]-1,1[} = g_{|]-1,1[}$, on a $g = \arcsin$ sur $[-1,1]$.

  On aurait pu avoir directement cette conclusion grâce au théorème d'Abel radial.

  Il en résulte que la formule de la question 1.b) est valable sur $[0, \frac{\pi}{2}]$.

— Par des raisonnement et calcul similaires à ceux effectués pour $\sum g_n$, $\sum f_n$ converge uniformément sur $[0, \frac{\pi}{2}]$ vers $S : t \mapsto t$ et de plus, les f_n étant continues sur $[0, \frac{\pi}{2}]$, on a, par le théorème d'interversion somme/intégrale :

$$\sum_{n=0}^{+\infty} a_n \int_0^{\frac{\pi}{2}} \sin^{2n+1}(t)\,\mathrm{d}t = \int_0^{\frac{\pi}{2}} S(t)\,\mathrm{d}t$$

Il nous reste à démontrer que notre hypothèse "$\sum a_n$ converge" est bien vérifiée !

— *1ère façon : rapide mais pas très économique.*
  D'après la formule de Stirling $n! \underset{n\to+\infty}{\sim} \sqrt{2\pi n}\left(\dfrac{n}{e}\right)^n$, on a :

$$a_n = \frac{(2n)!}{4^n (n!)^2}\frac{1}{2n+1} \underset{n\to+\infty}{\sim} \frac{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}}{4^n \times 2\pi n\left(\frac{n}{e}\right)^{2n}}\frac{1}{2n} = \frac{1}{2\sqrt{\pi}n^{\frac{3}{2}}}$$

  d'où, par comparaison avec le terme général d'une série de Riemann convergente ($\frac{3}{2} > 1$), $\sum a_n$ converge.

— *2nde façon : plus astucieuse mais bien moins chère.*
  On remarque que, pour tout $n \in \mathbb{N}^*$:

$$\sqrt{2n+1} = \sqrt{\prod_{i=1}^{n} \frac{2i+1}{2i-1}} \text{ et } \forall\, i \in \mathbb{N}^*, \sqrt{\frac{2i+1}{2i-1}} \times \frac{2i-1}{2i} = \frac{\sqrt{(2i)^2-1}}{2i} \leq 1.$$

  Par suite, on a :

$$(2n+1)\sqrt{2n+1}a_n = \prod_{i=1}^{n}\left(\sqrt{\frac{2i+1}{2i-1}} \times \frac{2i-1}{2i}\right) \leq 1.$$

  D'où $0 \leq a_n \leq \dfrac{1}{(2n+1)\sqrt{2n+1}} = \underset{n\to+\infty}{O}\left(\dfrac{1}{n^{\frac{3}{2}}}\right)$; et donc $\sum a_n$ converge par comparaison avec le terme général d'une série de Riemann convergente.

3. On rappelle que, pour tout $n \in \mathbb{N}$, $W_{2n+1} = \frac{4^n}{\binom{2n}{n}} \cdot \frac{1}{2n+1}$ (voire problème 5) ; donc :
$$W_{2n+1} = \frac{1}{a_n} \times \frac{1}{(2n+1)^2}.$$
4. Par suite, on a :
$$\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \sum_{n=0}^{+\infty} a_n W_{2n+1} = \frac{\pi^2}{8}.$$

Passons à la valeur de $\zeta(2)$. On a besoin des résultats suivants sur les familles sommables :
— une suite à valeurs numériques est une famille sommable si, et seulement si, la série associée est absolument convergente ; et dans ce cas, la somme de la famille et la somme de la série sont égales.
— pour I, J des ensembles, si $(a_i)_{i \in I}$ est une famille sommable et $\sigma : J \to I$ est en bijection, alors $(a_{\sigma(j)})_{j \in J}$ est sommable et les sommes de ces deux familles sont égales (c'est ce qu'on appelle le *changement d'indice*).
— le théorème de sommation par paquets - dans notre cas, il y aura seulement deux paquets.

La suite $(\frac{1}{n^2})_{n \in \mathbb{N}^*}$ est le terme générale d'une série absolument convergente, donc c'est une famille sommable et les sommes famille/série coïncident. Par sommation par paquets sur la partition $\mathbb{N}^* = 2\mathbb{N}^* \sqcup 2\mathbb{N}+1$, on a :
$$\begin{aligned}
\zeta(2) &= \sum_{n=1}^{+\infty} \frac{1}{n^2} \\
&= \sum_{n \in \mathbb{N}^*} \frac{1}{n^2} \\
&= \sum_{n \in 2\mathbb{N}^*} \frac{1}{n^2} + \sum_{n \in 2\mathbb{N}+1} \frac{1}{n^2} \\
&= \sum_{p \in \mathbb{N}^*} \frac{1}{(2p)^2} + \sum_{p \in \mathbb{N}} \frac{1}{(2p+1)^2} \\
&= \frac{1}{4} \sum_{p=1}^{+\infty} \frac{1}{p^2} + \sum_{p=0}^{+\infty} \frac{1}{(2p+1)^2} \\
\zeta(2) &= \frac{1}{4}\zeta(2) + \frac{\pi^2}{8}
\end{aligned}$$

On obtient donc :
$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \zeta(2) = \frac{4}{3}\frac{\pi^2}{8} = \frac{\pi^2}{6}.$$

Théorème 4. *Intégration terme à terme (cas à valeurs dans $\mathbb{K}$)*

Soit $\sum f_n$ une série de fonctions de I dans $\mathbb{K}$. Si les propriétés suivantes sont vérifiées :
- pour tout $n \in \mathbb{N}$, f_n est continue par morceaux sur I ;

- pour tout $n \in \mathbb{N}$, f_n est intégrable sur I ;
- la série de fonctions $\sum f_n$ converge simplement sur I et sa somme $S : t \mapsto \sum_{n=0}^{+\infty} f_n(t)$ est continue par morceaux sur I ;
- La série numérique :

$$\sum \left(\int_I |f_n(t)| \, \mathrm{d}t \right) \text{ converge,}$$

alors, S est intégrable sur I et :

$$\int_I S(t) \, \mathrm{d}t = \int_I \left(\sum_{n=0}^{+\infty} f_n \right) \mathrm{d}t = \sum_{n=0}^{+\infty} \left(\int_I f_n(t) \, \mathrm{d}t \right).$$

Exemple 3.

On a $\displaystyle\int_0^1 \frac{\ln(t)}{1+t} \, \mathrm{d}t = \frac{-\pi^2}{12}$.

En effet, cette intégrale existe car, $f : t \mapsto \dfrac{\ln(t)}{1+t}$ est continue sur $]0,1]$ et pour tout $t \in]0,1]$,

$$|\frac{\ln(t)}{1+t}| \leq |\ln(t)|.$$

Or ln est intégrable sur $]0,1]$; donc f l'est aussi.
On a, pour $t \in]0,1[$,

$$\frac{\ln(t)}{1+t} = \sum_{n=0}^{+\infty} (-1)^n t^n \ln(t).$$

Ainsi, on pose, pour $n \in \mathbb{N}$, $f_n : t \mapsto (-1)^n t^n \ln(t)$.

- On a, pour tout $n \in \mathbb{N}$ et tout $t \in]0,1]$, $|f_n(t)| \leq |\ln(t)|$ donc f_n est intégrable sur $]0,1]$.
- $\sum f_n$ converge simplement sur I vers $t \mapsto \dfrac{\ln(t)}{1+t}$ qui est continue sur $]0,1]$.
- Soit $n \in \mathbb{N}$. On a :

$$\begin{aligned} \int_0^1 |f_n(t)|\,\mathrm{d}t &= -\int_0^1 t^n \ln(t)\,\mathrm{d}t \\ &\overset{\text{IPP}}{=} -\left[\frac{t^{n+1}}{n+1}\ln(t)\right]_0^1 + \int_0^1 \frac{t^n}{n+1}\,\mathrm{d}t \\ &= \frac{1}{(n+1)^2}. \end{aligned}$$

  Donc, $\sum \int_0^1 |f_n| = \sum \dfrac{1}{(n+1)^2}$ est une série numérique convergente (série de Riemann avec $2 > 1$).

Les hypothèses du théorème d'intégration terme à terme sont donc vérifiées, d'où

$$\begin{aligned} \int_0^1 \frac{\ln(t)}{1+t}\,\mathrm{d}t &= \sum_{n=0}^{+\infty} \int_0^1 (-1)^n t^n \ln(t)\,\mathrm{d}t \\ &= \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}}{(n+1)^2} \\ &= \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} \\ &= -\frac{\pi^2}{12}. \end{aligned}$$

Cette dernière égalité peut être obtenue en effectuant une sommation par paquets sur les nombres pairs et impairs pour faire apparaître une relation avec la somme $\sum_{n=1}^{+\infty} \dfrac{1}{n^2} = \dfrac{\pi^2}{6}$.

Exercice 4.

Dans chaque question, on s'assurera de la convergence des intégrales en jeu.

1. Montrer que $\displaystyle\int_0^1 \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)^2}$.
2. Montrer que $\displaystyle\int_0^{+\infty} \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = 0$ et déterminer une expression de $\displaystyle\int_1^{+\infty} \frac{\ln(t)}{1+t^2}\,\mathrm{d}t$ sous forme de somme de série.

Correction.

1. On applique la même technique que pour l'exemple précédent en developpant en série entière $x \mapsto \frac{1}{1+x^2}$ et on trouve :
$$\int_0^1 \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)^2}.$$
les intégrale et somme en jeu étant bien convergentes.
2. On a, via le changement de variable *licite* (à vérifier !) $u = \frac{1}{t}$:
$$\int_0^1 \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = -\int_1^{+\infty} \frac{\ln(u)}{1+u^2}\,\mathrm{d}u \quad (*)$$
donc, comme somme d'intégrales convergente, $\int_0^{+\infty} \frac{\ln(t)}{1+t^2}\,\mathrm{d}t$ converge et on a, d'après $(*)$:
$$\int_0^{+\infty} \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = \int_0^1 \frac{\ln(t)}{1+t^2}\,\mathrm{d}t + \int_1^{+\infty} \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = 0$$
et de plus, toujours d'après $(*)$:
$$\int_1^{+\infty} \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = -\int_0^1 \frac{\ln(t)}{1+t^2}\,\mathrm{d}t = -\sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)^2}.$$

Remarque 3.

Parfois, les hypothèses du théorème d'intégration terme à terme ne sont pas satisfaites (plus précisément, le dernier point) mais on peut montrer que sa conclusion est tout de même valide en appliquant le théorème de convergence dominée à la suite des sommes partielles de la série $\sum f_n$.
Voici un exemple typique d'un tel cas dans l'exercice suivant :

Exercice 5.

1. Montrer que, pour tout $\alpha \in \mathbb{R}_+^*$,
$$\int_0^1 \frac{1}{1+t^\alpha}\,\mathrm{d}t = \sum_{n=0}^{+\infty} \frac{(-1)^n}{\alpha n+1}.$$
2. En déduire la valeur de chacune des sommes suivantes :
$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} \qquad \sum_{n=0}^{+\infty} \frac{(-1)^n}{3n+1} \qquad \sum_{n=0}^{+\infty} \frac{(-1)^n}{4n+1}$$

Correction.

1. Soit $\alpha \in \mathbb{R}_+^*$. On remarque que, pour tout $t \in [0,1[$, $y = t^\alpha \in [0,1[\subset]-1,1[$, donc :
$$\frac{1}{1+t^\alpha} = \frac{1}{1+y} = \sum_{n=0}^{+\infty}(-1^n)y^n = \sum_{n=0}^{+\infty}(-1)^n t^{\alpha n}.$$
On pose donc, pour $n \in \mathbb{N}$, $f_n : t \mapsto (-1)^n t^{\alpha n}$.
Les f_n sont intégrables sur $[0,1[$ car continues sur $[0,1]$ et on a $\int_0^1 |f_n(t)|\,\mathrm{d}t = \frac{1}{\alpha n+1}$ qui est le terme général d'une série divergente. Donc le théorème d'intégration terme à terme est inutilisable ! Mais sans la valeur absolue, on obtiendrait $\frac{(-1)^n}{\alpha n+1}$ qui est cette fois le terme général d'une série convergente (d'après le critère des séries alternées) ; on se dit alors qu'on peut peut-être contourner le problème en revenant au théorème de convergence dominée sur la suite des sommes partielles de la série $\sum f_n$.
C'est le même principe que dans le cas de certaines séries alternées de fonctions, lorsqu'il n'y a pas convergence normale, mais qu'on a tout de même convergence uniforme !
En parlant de convergence uniforme, on pourra vérifier que le théorème d'interversion somme/intégrale est inutilisable ici !.

On pose, pour $n \in \mathbb{N}$, $S_n = \displaystyle\sum_{k=0}^{n} f_k$.
   - Comme les f_k sont continues sur $[0,1[$, les S_n le sont ;
   - D'après la remarque initiale, on a $(S_n)_{n\in\mathbb{N}} = \sum f_n$ converge simplement sur $[0,1[$ vers $f : t \mapsto \frac{1}{1+t^\alpha}$ qui est bien continue sur $[0,1[$;
   - Pour tout $n \in \mathbb{N}$, pour tout $t \in [0,1[$, on a, d'après le critère des séries alternées :
   $$|S_n(t)| = |\sum_{n=0}^{+\infty}(-1)^n t^{\alpha n}| \leq |(-1)^0 t^{\alpha 0}| = 1 = g(t)$$
   Et $g : t \mapsto 1$ est bien intégrable sur $[0,1]$ et donc sur $[0,1[$ car continue sur le segment $[0,1]$.

Par suite, d'après le théorème de convergence dominée, les S_n et f sont intégrables sur $[0,1[$

(mais on le savait déjà !), et on a :

$$\begin{aligned}
\int_0^1 \frac{1}{1+t^\alpha}\,\mathrm{d}t &= \int_0^1 f(t)\,\mathrm{d}t \\
&= \lim_{N\to+\infty} \int_0^1 S_N(t)\,\mathrm{d}t \\
&= \lim_{N\to+\infty} \sum_{n=0}^{N} (-1)^n \int_0^1 t^{\alpha n}\,\mathrm{d}t \\
&= \lim_{N\to+\infty} \sum_{n=0}^{N} \frac{(-1)^n}{\alpha n+1} \\
\int_0^1 \frac{1}{1+t^\alpha}\,\mathrm{d}t &= \sum_{n=0}^{+\infty} \frac{(-1)^n}{\alpha n+1}.
\end{aligned}$$

2. D'après la question précédente appliquée à $\alpha = 2, 3$ et 4, on obtient :
   — pour $\alpha = 2$:
   $$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = \int_0^1 \frac{1}{1+t^2}\,\mathrm{d}t = [\arctan(t)]_0^1 = \frac{\pi}{4}.$$
   — pour $\alpha = 3$ et 4, après une décomposition en éléments simples de la fraction rationnelle $1/(1+X^\alpha)$ pour déterminer une primitive de $t \mapsto \frac{1}{1+t^\alpha}$, on obtient - après beaucoup de calculs ! - les valeurs suivantes :
   $$\sum_{n=0}^{+\infty} \frac{(-1)^n}{3n+1} = \frac{1}{3}\left(\frac{\sqrt{3}}{3}\pi + \ln(2)\right) \quad \sum_{n=0}^{+\infty} \frac{(-1)^n}{4n+1} = \frac{1}{4\sqrt{2}}\left(\pi + 2\ln(1+\sqrt{2})\right).$$

Partie B

Intégrale dépendant d'un paramètre

Dans cette partie, A désigne une partie d'un espace vectoriel normé E de dimension finie et I désigne un intervalle d'intérieur non vide de $\mathbb{R}$.

1. Limite d'une intégrale à paramètre

On suppose que A est un intervalle de $\mathbb{R}$. Le théorème suivant n'est qu'une reformulation de la version continue du théorème de convergence dominée :

Théorème 5. *Limite d'une intégrale à paramètre*

Soit $f : A \times I \to \mathbb{K}$ une fonction et a une extrémité de l'intervalle A. Si f vérifie les propriétés suivantes :

- pour tout $x \in A$, $f(x, \cdot) : t \mapsto f(x,t)$ est continue par morceaux sur I ;
- pour tout $t \in I$, $f(\cdot, t) : x \mapsto f(x,t)$ admet une limite en a et l'application $h : t \mapsto \lim\limits_{x \to a} f(x,t)$ est continue par morceaux sur I ;
- il existe une fonction $g : I \to \mathbb{R}$ positive et **intégrable** sur I telle que, pour tout $(x,t) \in A \times I$:
$$|f(x,t)| \leq g(t), \qquad \textbf{hypothèse de domination}$$

alors la fonction h est intégrable sur I et on a :

$$\lim_{x \to a} \left(\int_I f(x,t)\,\mathrm{d}t \right) = \int_I \left(\lim_{x \to a} f(x,t) \right) \mathrm{d}t.$$

Exercice 6.

Justifier, pour tout $x \in \mathbb{R}_+$, l'existence de

$$I(x) = \int_0^{+\infty} \frac{1}{(1+t^2)(1+tx)}\,\mathrm{d}t$$

et calculer $\lim_{x \to +\infty} I(x)$.

Correction.

On pose, pour $x \in \mathbb{R}_+$ et $t \in [0, +\infty[$, $f(x,t) = \frac{1}{(1+t^2)(1+tx)}$.

— Soit $x \in \mathbb{R}_+$. Montrons la convergence de l'intégrale $\int_0^{+\infty} \frac{1}{(1+t^2)(1+tx)}\,\mathrm{d}t$.

La fonction $t \mapsto f(x,t)$ est continue sur $[0, +\infty[$ et on a :

$$|f(x,t)| \leq \frac{1}{1+t^2}$$

Or, $t \mapsto \frac{1}{1+t^2}$ est intégrable sur $[0, +\infty[$ donc, par comparaison, $t \mapsto f(x,t)$ l'est aussi. Par suite, $\int_0^{+\infty} \frac{1}{(1+t^2)(1+tx)}\,\mathrm{d}t$ converge.
Il en résulte que la fonction I est bien définie sur $\mathbb{R}_+$.

— On vérifie les hypothèses du théorème de la limite d'une intégrale à paramètres :
  - pour tout $x \in \mathbb{R}_+$, $f(x,\cdot) : t \mapsto f(x,t)$ est continue sur $[0, +\infty[$;
  - pour tout $t \in [0, +\infty[$, on a :

$$\frac{1}{(1+t^2)(1+tx)} \xrightarrow[x\to+\infty]{} 0 = h(t)$$

  et $h : t \mapsto 0$ est continue sur $[0, +\infty[$;
  - Soit $t \in [0, +\infty[$. Pour tout $x \in \mathbb{R}_+$, on a :

$$|f(x,t)| \leq \frac{1}{1+t^2} = g(t), \qquad \textbf{hypothèse de domination}$$

  et $g : t \mapsto \frac{1}{1+t^2}$ est intégrable sur $[0, +\infty[$ *(comme d'habitude !)*.

Par suite, la fonction $h : t \mapsto 0$ est intégrable sur $[0, +\infty[$ *(mais on ne le savait que trop bien !)* et on a :

$$\lim_{x\to+\infty} I(x) = \int_0^{+\infty} \left(\lim_{x\to+\infty} \frac{1}{(1+t^2)(1+tx)} \right) \mathrm{d}t = \int_0^{+\infty} 0 \,\mathrm{d}t = 0.$$

2. Continuité d'une intégrale à paramètre

Théorème 6. *Théorème de continuité d'une intégrale à paramètre*

Soit $f : A \times I \to \mathbb{K}$ une fonction. Si f vérifie les propriétés suivantes :
- pour tout $t \in I$, $f(\cdot, t) : x \mapsto f(x,t)$ est **continue** sur A ;
- pour tout $x \in A$, $f(x, \cdot) : t \mapsto f(x,t)$ est continue par morceaux sur I ;
- il existe une fonction $g : I \to \mathbb{R}$ positive et **intégrable** sur I telle que, pour tout $(x,t) \in A \times I$:

$$|f(x,t)| \leq g(t), \qquad \textbf{hypothèse de domination}$$

alors la fonction $F : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et continue sur A.

Exemple 4. *Transformée de Fourier*

Soit $\varphi : \mathbb{R} \to \mathbb{C}$ une fonction continue par morceaux et intégrable sur $\mathbb{R}$. On appelle **transformée de Fourier de** φ et on note $\mathcal{F}(\varphi)$ la fonction de $\mathbb{R} \to \mathbb{C}$:

$$\mathcal{F}(\varphi) : x \mapsto \int_{\mathbb{R}} \varphi(t) e^{-ixt}\,\mathrm{d}t.$$

La transformée de Fourier de φ est bien définie et continue sur $\mathbb{R}$.

Pour $x, t \in \mathbb{R}$, on pose $f(x,t) = \varphi(t)e^{-ixt}$. Montrons que $F(\varphi)$ est bien définie sur $\mathbb{R}$.

Soit $x \in \mathbb{R}$. Alors, la fonction $t \mapsto f(x,t)$ est continue par morceaux sur $]-\infty, +\infty[$. En effet, φ est continue par morceaux sur $]-\infty, +\infty[$ par hypothèse et la fonction $t \mapsto e^{-ixt}$ est continue sur $\mathbb{R}$ comme composée des applications continues $t \mapsto -ixt$ (continue sur $\mathbb{R}$ car polynomiale) et $z \mapsto \exp(z)$ (continue sur $\mathbb{C}$ car somme d'une série entière de rayon de convergence $+\infty$) ; donc $t \mapsto f(x,t)$ est continue par morceaux sur $]-\infty, +\infty[$ comme produit de fonctions continues par morceaux sur $]-\infty, +\infty[$.
De plus, on a, pour tout $t \in]-\infty, +\infty[$,

$$|f(x,t)| \leq |f(t)|$$

donc, φ étant intégrable sur $]-\infty, +\infty[$, $t \mapsto f(x,t)$ l'est aussi.
Par suite, l'intégrale $\int_{\mathbb{R}} \varphi(t)e^{-ixt}\,\mathrm{d}t$ converge et donc $F(\varphi)(x)$ est bien définie.
Il en résulte que la fonction $F(\varphi)$ est définie sur $\mathbb{R}$.

Montrons la continuité sur $\mathbb{R}$ de $F(\varphi)$.

- pour tout $t \in]-\infty, +\infty[$, $f(\cdot, t) : x \mapsto \varphi(t)e^{-ixt}$ est continue sur $\mathbb{R}$ car la fonction $x \mapsto e^{-ixt}$ est continue sur $\mathbb{R}$ (même raisonnement que précédemment) ;
- pour tout $x \in \mathbb{R}$, $f(x, \cdot) : t \mapsto f(x,t)$ est continue par morceaux sur $]-\infty, +\infty[$ (déjà fait) ;
- On a, pour tout $x \in \mathbb{R}$ et pour tout $t \in]-\infty, +\infty[$:

$$|f(x,t)| \leq |\varphi(t)| = g(t), \qquad \textbf{hypothèse de domination}$$

  Et comme φ est intégrable sur $]-\infty, +\infty[$, g l'est aussi.

Ainsi, d'après le théorème de continuité des intégrales à paramètres, $F(\varphi)$ est continue sur $\mathbb{R}$.

Exercice 7.

On pose $F : x \mapsto \displaystyle\int_0^{+\infty} \frac{e^{-xt}}{1+t^2}\,\mathrm{d}t$.

1. Déterminer le domaine de définition de F.
2. Montrer que F est continue sur son domaine de définition.
3. Établir le tableau des variations de F en précisant les valeurs/limites aux bornes de son domaine. *Indication : on justifiera la monotonie de F en utilisant la définition et pas par la dérivation !*

Correction.

On pose $f : (x,t) \mapsto \frac{e^{-xt}}{1+t^2}$.

1. Soit $x \in \mathbb{R}$. La fonction $t \mapsto f(x,t)$ est continue et positive sur $\mathbb{R}$ et donc sur $[0, +\infty[$ et on a :
   — si $x \geq 0$, $t^2|f(x,t)| \xrightarrow[t \to +\infty]{} 0$ ou 1, donc $f(x,t) \underset{t \to +\infty}{=} O\left(\frac{1}{t^2}\right)$; d'où, par comparaison à

une fonction intégrable sur $[1,+\infty$ (critère de Riemann en $+\infty$ avec $2>1$), $t \mapsto f(x,t)$ est intégrable sur $[1,+\infty[$. La fonction $t \mapsto f(x,t)$ étant de plus continue sur $[0,+\infty[$, elle est donc intégrable sur $[0,+\infty[$.
Par suite, $\int_0^{+\infty} \frac{e^{-xt}}{1+t^2}\,dt$ converge et donc $F(x)$ est bien défini.

— si $x<0$, on a $f(x,t) \xrightarrow[t\to+\infty]{} +\infty$ donc il existe un $t_0 \in [0,+\infty[$ tel que, pour tout $t \geq t_0$, $f(x,t) \geq 1$. Or $\int_0^{+\infty} 1\,\mathrm{d}t$ diverge, donc $\int_0^{+\infty} f(x,t)\,\mathrm{d}t$ diverge.
Par suite, $F(x) = +\infty$!

Ainsi, $F(x)$ est défini si, et seulement si, $x \in \mathbb{R}_+$. Donc $\mathcal{D}_F = \mathbb{R}_+$.

2. On vérifie les hypothèses du théorème de continuité des intégrales à paramètres :
   - pour tout $t \in [0,+\infty[$, $f(\cdot,t) : x \mapsto f(x,t)$ est continue sur $\mathbb{R}_+$;
   - pour tout $x \in \mathbb{R}_+$, $f(x,\cdot) : t \mapsto f(x,t)$ est continue sur $[0,+\infty[$;
   - Soit $t \in [0,+\infty[$. Pour tout $x \in \mathbb{R}_+$:

   $$|f(x,t)| \leq \frac{1}{1+t^2} = g(t), \qquad \textbf{hypothèse de domination}$$

   et $g : t \mapsto \frac{1}{1+t^2}$ est intégrable sur $[0,+\infty[$.

   Ainsi, la fonction $F : x \mapsto \displaystyle\int_0^{+\infty} \frac{e^{-xt}}{1+t^2}\,\mathrm{d}t$ est continue sur $\mathbb{R}_+$.

3. Pour tout $x, x' \in \mathbb{R}_+$ avec $x < x'$ et $t \in]0,+\infty[$, on a $f(x,t) > f(x',t)$ par stricte croissance de la fonction exponentielle, donc, par croissance de l'intégrale, F est strictement décroissante sur $\mathbb{R}_+$.
   De plus, on a :
   $$F(0) = \int_0^{+\infty} \frac{1}{1+t^2} = [\arctan(t)]_0^{+\infty} = \frac{\pi}{2}.$$
   et, d'après l'exercice 2 ou en appliquant le théorème de limite d'une intégrale à paramètres, on a :
   $$\lim_{x\to+\infty} F(x) = \int_0^{+\infty} \lim_{x\to+\infty}\left(\frac{e^{-xt}}{1+t^2}\right)\mathrm{d}t = \int_0^{+\infty} 0\,\mathrm{d}t = 0.$$
   d'où le tableau de variations de F :

x	0	$+\infty$
F	$\frac{\pi}{2}$ $\searrow$	0

Corollaire 1. *Extension du théorème de continuité*

On suppose ici que A est un intervalle de $\mathbb{R}$. Si on a les mêmes hypothèses que le théorème précédent en modifiant l'hypothèse de domination par :
- **Pour tout segment** $[a,b]$ **de** A, il existe une fonction $g : I \to \mathbb{R}$ positive et **intégrable**

sur I telle que, pour tout $(x,t) \in [a,b] \times I$:

$$|f(x,t)| \leq g(t), \qquad \textbf{hypothèse de domination sur tout segment}$$

alors la fonction $F : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et continue sur A.

Exercice 8.

Déterminer le domaine de définition de F et montrer que F est continue sur ce domaine pour :

1. $F : x \mapsto \displaystyle\int_1^{+\infty} \frac{e^{-xt}}{t}\,\mathrm{d}t$.
2. $F : x \mapsto \displaystyle\int_0^1 \frac{\cos(xt)}{x^2+t^2}\,\mathrm{d}t$.

Correction.

1. On pose $f(x,t) = \frac{e^{-xt}}{t}$.
   — Déterminons le domaine de F.
   Soit $x \in \mathbb{R}$. La fonction $t \mapsto f(x,t)$ est continue sur $[1,+\infty[$ et on a :
   - si $x > 0$: par croissances comparées,

   $$t^2|f(x,t)| = te^{-xt} \xrightarrow[t\to+\infty]{} 0 \text{ d'où } |f(x,t)| \underset{t\to+\infty}{=} o\left(\frac{1}{t^2}\right).$$

   Par comparaison à une fonction intégrable sur $[1,+\infty[$ (critère de Riemann en $+\infty$ avec $2 > 1$), $t \mapsto f(x,t)$ est intégrable sur $[1,+\infty[$.
   Par suite, $\int_1^{+\infty} \frac{e^{-xt}}{t}\,\mathrm{d}t$ converge.
   - si $x \leq 0$:

   $$f(x,t) = \frac{e^{-xt}}{t} \geq \frac{1}{t}$$

   Or $\int_1^{+\infty} \frac{1}{t}\,\mathrm{d}t$ est une intégrale de Riemann divergente $(1 \leq 1)$ donc, par comparaison, $\int_1^{+\infty} \frac{e^{-xt}}{t}\,\mathrm{d}t$ diverge.
   Il en résulte que $F(x)$ existe si, et seulement si, $x \in \mathbb{R}_+^*$ donc $\mathcal{D}_F = \mathbb{R}_+^*$.
   — Montrons la continuité de F sur $\mathbb{R}_+^*$. On vérifie les hypothèses du théorème de continuité des intégrales à paramètres.
   - pour tout $t \in [1,+\infty[$, $f(\cdot,t) : x \mapsto f(x,t)$ est continue sur $\mathbb{R}_+^*$;
   - pour tout $x \in \mathbb{R}_+^*$, $f(x,\cdot) : t \mapsto f(x,t)$ est continue sur $[1,+\infty[$;
   - Soit $a > 0$. Soit $t \in [1,+\infty[$. Pour tout $x \in [a,+\infty[$:

   $$|f(x,t)| \leq \frac{e^{-at}}{t} = g(t), \qquad \textbf{hypothèse de domination}$$

   et $g : t \mapsto \frac{e^{-at}}{t}$ est intégrable sur $[1,+\infty[$ (même raisonnement que précédemment).

Ainsi, la fonction $F : x \mapsto \displaystyle\int_1^{+\infty} \frac{e^{-xt}}{t}\,\mathrm{d}t$ est continue sur $\mathbb{R}_+^*$.

2. On pose $f(x,t) = \frac{\cos(xt)}{x^2+t^2}$.
   — Déterminons le domaine de F.
   Soit $x \in \mathbb{R}$.
   - si $x \neq 0$: la fonction $t \mapsto f(x,t)$ est continue sur le segment $[0,1]$, donc elle est intégrable sur $[0,1]$.
     Par suite, $\int_0^1 \frac{\cos(xt)}{x^2+t^2}\,\mathrm{d}t$ converge.
   - si $x = 0$: la fonction $t \mapsto f(x,t)$ est continue sur $]0,1]$ et

$$f(x,t) = \frac{1}{t^2}$$

     Or $\int_0^1 \frac{1}{t^2}\,\mathrm{d}t$ est une intégrale de Riemann divergente $(2 \geq 1)$ donc, par comparaison, $\int_0^1 f(0,t)\,\mathrm{d}t$ diverge.

   Il en résulte que $F(x)$ existe si, et seulement si, $x \in \mathbb{R}^*$ donc $\mathcal{D}_F = \mathbb{R}^*$.
   — Montrons la continuité de F sur $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$. On remarque que F est paire, donc il suffit de le montrer sur $\mathbb{R}_+^*$. On vérifie les hypothèses du théorème de continuité des intégrales à paramètres.
   - pour tout $t \in [0,1]$, $f(\cdot,t) : x \mapsto f(x,t)$ est continue sur $\mathbb{R}_+^*$;
   - pour tout $x \in \mathbb{R}_+^*$, $f(x,\cdot) : t \mapsto f(x,t)$ est continue sur $[0,1]$;
   - Soit $a > 0$. Soit $t \in [0,1]$. Pour tout $x \in [a,+\infty[$:

$$|f(x,t)| \leq \frac{1}{x^2+t^2} \leq \frac{1}{a^2+x^2}, \qquad \textbf{hypothèse de domination}$$

     et $g : t \mapsto \frac{1}{a^2+x^2}$ est intégrable sur $[0,1]$ car continue sur le segment $[0,1]$.

   Ainsi, la fonction $F : x \mapsto \displaystyle\int_0^1 \frac{\cos(xt)}{x^2+t^2}\,\mathrm{d}t$ est continue sur $\mathbb{R}_+^*$ et donc également sur $\mathbb{R}_-^*$ par parité.

Problème 3. *Transformée de Laplace*

Dans ce problème, $I = \mathbb{R}_+$ ou $\mathbb{R}_+^*$.
Pour $f : I \to \mathbb{C}$ une fonction continue par morceaux sur I, on appelle transformée de Laplace de f et on note $\mathcal{L}(f)$ la fonction de la variable p réelle :

$$\mathcal{L}(f) : p \mapsto \int_0^{+\infty} f(x)e^{-px}\,\mathrm{d}x.$$

1. *Premiers exemples.*
   Déterminer le domaine et une expression explicite de $\mathcal{L}(f)$ pour :

$$f \text{ constante}; \quad f : x \mapsto x^n \text{ où } n \in \mathbb{N}; \quad f : x \mapsto e^{ax} \text{ où } a \in \mathbb{R}.$$

2. *Domaine de définition de $\mathcal{L}(f)$ pour une famille de fonctions spécifiques.*
   Soit $f \in C_{pm}(I,\mathbb{C})$ intégrable sur $]0,1]$ vérifiant qu'il existe $n \in \mathbb{N}$ et $a \in \mathbb{R}$ tels que

$f(x) \underset{x\to+\infty}{=} O(x^n e^{ax})$.
Montrer que $\mathcal{L}(f)$ est bien définie et continue sur $]a, +\infty[$.

3. Déterminer le domaine et une expression de la transformée de Laplace de :
   (a) $x \mapsto e^{iax}$ puis en déduire les transformées de Laplace de $x \mapsto \sin(ax)$ et $x \mapsto \cos(ax)$ pour $a \in \mathbb{R}$.
   (b) $x \mapsto \sqrt{x}$ puis $x \mapsto \frac{1}{\sqrt{x}}$.

4. *Continuité en 0 de la transformée de Laplace.*
   Soit f une fonction continue par morceaux sur I.
   (a) Montrer que si f est intégrable sur I, alors $\mathcal{L}(f)$ est continue sur $\mathbb{R}_+$.
   (b) *(Difficile)* Montrer que si f est continue sur $\mathbb{R}_+$ et $\displaystyle\int_0^{+\infty} f(x)\,\mathrm{d}x$ converge, alors $\mathcal{L}(f)$ est continue sur $\mathbb{R}_+$.
   En particulier, sous ces hypothèses, on a donc :

$$\mathcal{L}(f)(p) \xrightarrow[p\to 0^+]{} \int_0^{+\infty} f(x)\,\mathrm{d}x.$$

Indications :
i) Commencer par montrer, en utilisant la deuxième formule de la moyenne, que pour tout $\varepsilon > 0$, il existe $a_\varepsilon \in \mathbb{R}_+$ tel que pour tout $p \in \mathbb{R}_+$ et tout $a \geq a_\varepsilon$, $|\int_a^{+\infty} f(x)e^{-px}\,dx| \leq \varepsilon$.
ii) Soit $a \in \mathbb{R}_+$. Montrer la continuité de la fonction $H : p \mapsto \int_0^a f(x)e^{-px}\,dx$.
iii) Déduire de i) et ii) la continuité de $\mathcal{L}(f)$ avec la définition de la continuité et en s'inspirant de la démonstration de "convergence uniforme implique continuité" pour les limites de suites et séries de fonctions continues.

Correction.

Dans la suite, on note $\varphi(p, x) = f(x)e^{-px}$.

1. *Premiers exemples.*
   - On suppose qu'il existe $c \in \mathbb{R}$ tel que $f :\mapsto c$. Alors, pour $p \in \mathbb{R}$:

$$\mathcal{L}(f)(p) = \int_0^{+\infty} ce^{-px}\,\mathrm{d}x.$$

   Or, $\int_0^{+\infty} e^{-px}\,\mathrm{d}x$ converge si, et seulement si, $p \in \mathbb{R}_+^*$ (en utilisant par exemple une primitive de $p \mapsto e^{-px}$), d'où :
   — si $c \neq 0$, le domaine de $\mathcal{L}(f)$ est $\mathbb{R}_+^*$ et on a :

$$\mathcal{L}(f)(p) = \int_0^{+\infty} ce^{-px}\,\mathrm{d}x = c\int_0^{+\infty} e^{-px}\,\mathrm{d}x = c\left[\frac{e^{-px}}{x}\right]_0^{+\infty} = \frac{c}{x}.$$

   — si $c = 0$, le domaine de $\mathcal{L}(f)$ est $\mathbb{R}$ et $\mathcal{L}(f) : p \mapsto 0$.
   - On suppose qu'il existe $n \in \mathbb{N}$ tel que $f : x \mapsto x^n$.
   — Si $p \leq 0$, on a, pour tout $x \in [0, +\infty[$, $x^n e^{-px} \geq x^n = \frac{1}{x^{-n}}$ qui est d'intégrale divergente sur $[0, +\infty[$ d'après le critère de Riemann en $+\infty$ avec $-n \leq 1$ (ou on

aurait pu remarquer également que $x^n \geq 1$ pour tout $x \geq 1$), donc $\int_0^{+\infty} x^n e^{-px}\,\mathrm{d}x$ diverge.

— Soit $p \in \mathbb{R}_+^*$. On effectue le changement de variable *licite* $t = px$:

$$\mathcal{L}(f)(p) = \frac{1}{p^{n+1}} \int_0^{+\infty} t^n e^{-t}\,\mathrm{d}t \quad (*)$$

Ainsi, les intégrales $\mathcal{L}(f)(p)$ et $\int_0^{+\infty} t^n e^{-t}\,\mathrm{d}t$ sont de même nature. La fonction $t \mapsto t^n e^{-t}$ est continue sur $[0, +\infty[$ et $t^n e^{-t} \underset{t \to +\infty}{=} o\left(\frac{1}{t^2}\right)$ donc l'intégrale $\int_0^{+\infty} t^n e^{-t}\,\mathrm{d}t$ converge.
Par suite, $\int_0^{+\infty} x^n e^{-px}\,\mathrm{d}x$ converge.

Il en résulte que le domaine de $\mathcal{L}(f)$ est $\mathbb{R}_+^*$.

Notons, pour $k \in \mathbb{N}$, $I_k = \displaystyle\int_0^{+\infty} t^k e^{-t}\,\mathrm{d}t$. Alors, pour tout $k \in \mathbb{N}$, $I_k > 0$ car $t \mapsto t^k e^{-t}$ est continue et strictement positive sur $]0, +\infty[$ et par une intégration par partie *licite*, on obtient $I_{k+1} = (k+1)I_k$. Par suite, comme $I_0 = 1$, on a, pour tout $k \in \mathbb{N}$, $I_k = k!$.
Ainsi, pour $p \in \mathbb{R}_+^*$, d'après l'égalité $(*)$, on a :

$$\mathcal{L}(f)(p) = \frac{1}{p^{n+1}} I_n = \frac{n!}{p^{n+1}}.$$

Par continuité de $p \mapsto \frac{1}{p^{n+1}}$ sur $\mathbb{R}_+^*$, $\mathcal{L}(f)$ est donc continue sur son domaine de définition.

- On suppose qu'il existe $a \in \mathbb{R}$ tel que $f : x \mapsto e^{ax}$.
  On a alors, pour $p \in \mathbb{R}$:

$$. \int_0^{+\infty} e^{-(p-a)x}\,\mathrm{d}x$$

  Or $\int_0^{+\infty} e^{-\alpha x}\,\mathrm{d}x$ converge si, et seulement si, $\alpha > 0$ (on le prouve, par exemple, en déterminant une primitive de $x \mapsto e^{-\alpha x}$).
  Par suite, le domaine de définition de $\mathcal{L}(f)$ est $]a, +\infty[$.
  De plus, on a, pour tout $p > a$:

$$\mathcal{L}(f)(p) = \left[-\frac{e^{-(p-a)x}}{p-a}\right]_0^{+\infty} = \frac{1}{p-a}.$$

  Par continuité de $p \mapsto \frac{1}{p-a}$ sur $]a, +\infty[$, $\mathcal{L}(f)$ est donc continue sur son domaine de définition.

2. On suppose que f est une fonction continue par morceaux sur I, intégrable sur $]0, 1]$ et telle que $f(x) \underset{x \to +\infty}{=} O(x^n e^{ax})$.
   Alors il existe $M \geq 0$ et $x_0 > 0$ tel que pour tout $x \geq x_0$, $|f(x)| \leq M x^n e^{ax}$.
   De plus, on remarque pour tout $y \in I$, f est intégrable sur $]0, y]$: en effet, f est intégrable sur $]0, 1]$ par hypothèse et :
   — si $y \leq 1$, $]0, y] \subset]0, 1]$ donc f est intégrable sur $]0, y]$;
   — si $y > 1$, f est continue par morceaux sur I et donc sur le segment $[1, y]$ d'où f est intégrable sur $[1, y]$; par suite, d'après la relation de Chasles, f est intégrable sur $]0, y]$.
   En particulier, f est donc intégrable sur $]0, x_0]$.

   On vérifie les hypothèses du théorème de continuité des intégrales à paramètres.

- pour tout $x \in I$, $\varphi(\cdot, x) : p \mapsto \varphi(p, x)$ est continue sur $]a, +\infty[$;
- pour tout $p \in \mathbb{R}_+^*$, $\varphi(p, \cdot) : x \mapsto \varphi(p, x)$ est continue par morceaux sur I ;
- Soit $b > a$. Soit $x \in I$. Pour tout $p \in [b, +\infty[$:

$$|\varphi(p,x)| \leq \begin{cases} |f(x)| & \text{si } x \leq x_0 \\ Mx^n e^{-(b-a)x} & \text{si } x > x_0 \end{cases}, \qquad \textbf{hypothèse de domination}$$

et $g : x \mapsto \begin{cases} |f(x)| & \text{si } x \leq x_0 \\ Mx^n e^{-(b-a)x} & \text{si } x > x_0 \end{cases}$ est intégrable sur I.

En effet, g est continue par morceaux sur I ; d'après la remarque initiale, g est intégrable sur $]0, x_0]$ et on a, par croissances comparées, $x^2 g(x) \xrightarrow[x\to+\infty]{} 0$ d'où $g(x) = \underset{x\to+\infty}{o}(\frac{1}{x^2})$ et ainsi g est intégrable sur $[x_0, +\infty[$. D'où, par la relation de Chasles, g est intégrable sur I.

Par suite, d'après le théorème de continuité d'une intégrale à paramètres, $\mathcal{L}(f)$ est définie et continue sur $]a, +\infty[$.

3. *Autres exemples :*

(a) Soit $a \in \mathbb{R}$.

— On pose $f : x \mapsto e^{iax}$. Alors, pour tout $x \in \mathbb{R}$, on a $|f(x)| \leq 1$ donc f est continue sur $\mathbb{R}_+$, intégrable sur $[0, 1]$ et $f(x) = \underset{x\to+\infty}{O}(x^0 e^{0x})$. Ainsi, d'après la question précédente, $\mathcal{L}(f)$ est définie et continue sur $\mathbb{R}_+^*$. De plus, on a, pour tout $p \in \mathbb{R}_+^*$:

$$\mathcal{L}(f)(p) = \left[-\frac{e^{-(p-ia)x}}{p - ia} \right]_0^{+\infty} = \frac{1}{p - ia}.$$

— On a, pour tout $x \in \mathbb{R}$, $\cos(ax) = \mathrm{Re}(e^{iax})$ et $\sin(ax) = \mathrm{Im}(e^{iax})$. Or, pour $g : J \to C$ une fonction intégrable sur J, $\mathrm{Re}(g)$ et $\mathrm{Im}(g)$ sont intégrables sur J et :

$$\mathrm{Re}\left(\int_J g(t)\,\mathrm{d}t \right) = \int_J \mathrm{Re}(g(t))\,\mathrm{d}t \text{ et } \mathrm{Im}\left(\int_J g(t)\,\mathrm{d}t \right) = \int_J \mathrm{Im}(g(t))\,\mathrm{d}t,$$

donc, en posant $s_a : x \mapsto \sin(ax)$ et $c_a : x \mapsto \cos(ax)$, on obtient, pour $p \in \mathbb{R}_+^*$:

$$\mathcal{L}(s_a)(p) = \mathrm{Im}\left(\frac{1}{p - ia} \right) = \frac{a}{p^2 + a^2} \text{ et } \mathcal{L}(c_a)(p) = \mathrm{Re}\left(\frac{1}{p - ia} \right) = \frac{p}{p^2 + a^2}.$$

et $\mathcal{L}(s_a)$ et $\mathcal{L}(c_a)$ sont continues sur $\mathbb{R}_+^*$.

(b) Comme les fonctions $x \mapsto x^{\pm\frac{1}{2}}$ sont continues sur I, intégrables sur $]0, 1]$ (critère de Riemann en 0) et $\underset{x\to+\infty}{o}(x)$, d'après la question 2., leur transformées de Laplace sont définies et continues sur $\mathbb{R}_+^*$. De plus, en posant $f : x \mapsto \sqrt{x}$, on a, par des changement de variable et intégration par parties *licites*, pour tout $p \in \mathbb{R}_+^*$:

$$\begin{aligned} \mathcal{L}(f)(p) &\underbrace{=}_{t=\sqrt{px}} \frac{1}{p^{\frac{3}{2}}} \int_0^{+\infty} t \times 2te^{-t^2}\,\mathrm{d}t \\ &\underbrace{=}_{\text{IPP}} \frac{1}{p^{\frac{3}{2}}} \int_0^{+\infty} e^{-t^2}\,\mathrm{d}t \\ \mathcal{L}(f)(p) &= \frac{\sqrt{\pi}}{2p^{\frac{3}{2}}}. \end{aligned}$$

et pour $g : x \mapsto \frac{1}{\sqrt{x}}$, en effectuant une intégration par partie encore une fois licite, on obtient, pour tout $p \in \mathbb{R}_+^*$:

$$\mathcal{L}(g)(p) = 2p\mathcal{L}(f)(p) = \sqrt{\frac{\pi}{p}}.$$

4. *Continuité en* 0

   (a) On suppose f intégrable sur I.
   - pour tout $x \in I$, $\varphi(\cdot, x)$ est continue sur $\mathbb{R}$ par continuité de la fonction exponentielle ;
   - pour tout $p \in \mathbb{R}$, $\varphi(p, \cdot) : x \mapsto \varphi(p, x)$ est continue par morceaux sur I comme produit de fonctions continues par morceaux sur I ;
   - On a, pour tout $p \in \mathbb{R}_+$ et pour tout $x \in I$:

   $$|\varphi(p, x)| \leq |f(x)| = g(x), \qquad \textbf{hypothèse de domination}$$

   Et comme f est intégrable sur I, g l'est aussi.
   Ainsi, d'après le théorème de continuité des intégrales à paramètres, $\mathcal{L}(f)$ est bien définie et continue sur $\mathbb{R}_+$.

   (b) On suppose f est continue sur $\mathbb{R}_+$ et $\displaystyle\int_0^{+\infty} f(x)\,\mathrm{d}x$ converge.
   Remarque : Comme f n'est pas forcément intégrable sur $\mathbb{R}_+$, la majoration $|\varphi(p, x)| \leq |f(x)|$ de la question précédente ne permet plus d'appliquer directement le théorème de continuité des intégrales à paramètres !
   Procédons comme indiqué dans les indications :
   i) Soit $p \in \mathbb{R}_+$ et $a, b \in \mathbb{R}_+$ tels que $a \leq b$. On pose $g : x \mapsto e^{-px}$. Alors f est continue sur $[a, b]$ et g est de classe C^1 et décroissante sur $[a, b]$ donc, d'après la deuxième formule de la moyenne, il existe $c_p \in [a, b]$ tel que :

   $$\int_a^b f(x)e^{-px}\,\mathrm{d}x = e^{-pa}\int_a^{c_p} f(x)\,\mathrm{d}x$$

   Soit $\varepsilon > 0$. Comme $\int_0^{+\infty} f(x)\,\mathrm{d}x$ converge, pour tout $\varepsilon > 0$, il existe $a_\varepsilon \in \mathbb{R}_+$ tel que pour tout $A \geq a_\varepsilon$, $|\int_A^{+\infty} f(x)\,\mathrm{d}x| \leq \frac{\varepsilon}{2}$. Par suite, pour tous $a, b \in \mathbb{R}_+$ tels que $a_\varepsilon \leq a \leq b$, on obtient, comme $c_p \geq a \geq a_\varepsilon$:

   $$\left|\int_a^b f(x)e^{-px}\,\mathrm{d}x\right| = \underbrace{e^{-pa}}_{\leq 1}\left|\int_a^{c_p} f(x)\,\mathrm{d}x\right| \leq \left|\int_a^{+\infty} f(x)\,\mathrm{d}x\right| + \left|\int_{c_p}^{+\infty} f(x)\,\mathrm{d}x\right| \leq \varepsilon;$$

   puis en passant à la limite quand $b \to +\infty$, on obtient alors :

   $$\left|\int_a^{+\infty} f(x)e^{-px}\,\mathrm{d}x\right| \leq \varepsilon.$$

   ii) On note $H : p \mapsto \int_0^{a_\varepsilon} f(x)e^{-px}\,\mathrm{d}x$. On applique le théorème de continuité des intégrales à paramètres. Les hypothèses de continuité sont bien satisfaites, et on a, pour tout $p \in \mathbb{R}_+$ et pour tout $x \in [0, a_\varepsilon]$:

   $$|f(x)e^{-px}| \leq |f(x)| = g(x), \qquad \textbf{hypothèse de domination}$$

et g est bien intégrable sur $[0, a_\varepsilon]$ car continue sur un segment.
Par suite, d'après le théorème de continuité des intégrales à paramètres, H est continue sur $\mathbb{R}_+$.

iii) Montrons que $\mathcal{L}(f)$ est continue sur $\mathbb{R}_+$.
Soit $\varepsilon > 0$ et $p_0 \in \mathbb{R}_+$. Alors, d'après i) il existe a_ε tel que, pour tous $p, a \in \mathbb{R}_+$ avec $a \geq a_\varepsilon$, on a :

$$\left| \int_a^{+\infty} f(x) e^{-px} \, \mathrm{d}x \right| \leq \frac{\varepsilon}{3}.$$

De plus, comme H est continue sur $\mathbb{R}_+$, il existe $\delta_\varepsilon > 0$ tel que pour tout $p \in \mathbb{R}_+$, $|p - p_0| \leq \delta_\varepsilon$ implique $|H(p) - H(p_0)| \leq \frac{\varepsilon}{3}$.
Concluons : pour tout $p \in \mathbb{R}_+$ tel que $|p - p_0| \leq \delta_\varepsilon$, on a :

$$\begin{aligned}
|\mathcal{L}(f)(p) - \mathcal{L}(f)(p_0)| &\leq |\mathcal{L}(f)(p) - H(p)| + |H(p) - H(p_0)| + |H(p_0) - \mathcal{L}(f)(p_0)| \\
&\leq |\mathcal{L}(f)(p) - H(p)| + |H(p) - H(p_0)| + |H(p_0) - \mathcal{L}(f)(p_0)| \\
&= \left| \int_{a_\varepsilon}^{+\infty} f(x) e^{-px} \, \mathrm{d}x \right| + \frac{\varepsilon}{3} + \left| \int_{a_\varepsilon}^{+\infty} f(x) e^{-p_0 x} \, \mathrm{d}x \right| \\
|\mathcal{L}(f)(p) - \mathcal{L}(f)(p_0)| &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\end{aligned}$$

D'où $\mathcal{L}(f)$ est continue sur $\mathbb{R}_+$.

3. Dérivation d'une intégrale à paramètre

On suppose que A est un intervalle de $\mathbb{R}$.

Théorème 7. *Théorème de dérivation d'une intégrale à paramètre*

Soit $f : A \times I \to \mathbb{K}$ une fonction. Si f vérifie les propriétés suivantes :

- pour tout $t \in I$, $f(\cdot, t) : x \mapsto f(x, t)$ est C^1 sur A ;
- pour tout $x \in A$, $f(x, \cdot) : t \mapsto f(x, t)$ est intégrable sur I ;
- pour tout $x \in A$, $\dfrac{\partial f}{\partial x}(x, \cdot) : t \mapsto \dfrac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur I ;
- il existe une fonction $g : I \to \mathbb{R}$ positive et **intégrable** sur I telle que, pour tout $(x, t) \in A \times I$ *(ou $\in [a, b] \times I$ pour tout segment $[a, b]$ de A si $A \subset \mathbb{R}$)* :

$$\left| \frac{\partial f}{\partial x}(x, t) \right| \leq g(t), \qquad \textbf{hypothèse de domination } \textit{(sur tout segment)}$$

alors la fonction $F : x \mapsto \displaystyle\int_I f(x, t) \, \mathrm{d}t$ est définie et C^1 sur A. De plus, on a, pour tout $x \in A$:

$$F'(x) = \frac{\partial}{\partial x} \int_I f(x, t) \, \mathrm{d}t = \int_I \frac{\partial f}{\partial x}(x, t) \, \mathrm{d}t.$$

Théorème 8. *Théorème de dérivation C^∞ d'une intégrale à paramètre*

Soit $f : A \times I \to \mathbb{K}$ une fonction. Si f vérifie les propriétés suivantes :

- pour tout $t \in I$, $f(\cdot, t) : x \mapsto f(x,t)$ est C^∞ sur A ;
- pour tout $x \in A$, $f(x, \cdot) : t \mapsto f(x,t)$ est intégrable sur I ;
- pour tout $k \in \mathbb{N}^*$, pour tout $x \in A$, $\dfrac{\partial^k f}{\partial x^k}(x, \cdot) : t \mapsto \dfrac{\partial^k f}{\partial x^k}(x,t)$ est continue par morceaux sur I ;
- pour tout $k \in \mathbb{N}^*$, il existe une fonction $g_k : I \to \mathbb{R}$ positive et **intégrable** sur I telle que, pour tout $(x,t) \in A \times I$ *(ou sur tout segment si $A \subset \mathbb{R}$)* :

$$\left| \frac{\partial^k f}{\partial x^k}(x,t) \right| \leq g_k(t),$$ **hypothèse de domination** *(sur tout segment)*

alors la fonction $F : x \mapsto \displaystyle\int_I f(x,t)\,\mathrm{d}t$ est définie et C^∞ sur A. De plus, on a, pour tout $x \in A$:

$$F'(x) = \frac{\partial^k}{\partial x^k} \int_I f(x,t)\,\mathrm{d}t = \int_I \frac{\partial^k f}{\partial x^k}(x,t)\,\mathrm{d}t.$$

Remarque 4.

Pour montrer qu'une intégrale à paramètre est de classe C^n sur A pour un certain $n \in \mathbb{N}^*$, on vérifiera les mêmes hypothèses que le théorème précédent en remplaçant les "pour tout $k \in \mathbb{N}^*$" par "pour tout $k \in [\![1, n]\!]$".

Exemple 5.

La fonction $F : x \mapsto \displaystyle\int_0^{+\infty} e^{-xt^2}\,\mathrm{d}t$ est C^1 sur $\mathbb{R}_+^*$ et vérifie l'équation différentielle $y' = \dfrac{1}{2x}y$. On obtient alors l'expression explicite de F : pour tout $x \in \mathbb{R}_+^*$,

$$F(x) = \frac{1}{2}\sqrt{\frac{\pi}{x}}.$$

En effet, on a, en posant $f(x,t)=e^{-xt^2}$:

- pour tout $t\in[0,+\infty[$, $f(\cdot,t):x\mapsto e^{-xt^2}$ est C^1 sur $\mathbb{R}_+^*$ (et même C^∞) ;
- pour tout $x\in\mathbb{R}_+^*$, $f(x,\cdot):t\mapsto e^{-xt^2}$ est intégrable sur $[0,+\infty[$ car $t\mapsto e^{-xt^2}$ est continue sur $[0,+\infty[$ et en $+\infty$, comme $x>0$, $e^{-xt^2}=o(\frac{1}{t^2})$.
- pour tout $x\in\mathbb{R}_+^*$, $\dfrac{\partial f}{\partial x}(x,\cdot):t\mapsto -t^2e^{-xt^2}$ est continue sur $[0,+\infty[$;
- *(Domination sur tout segment de $\mathbb{R}_+^*$)* Soit $a>0$. Soit $t\in[0,+\infty[$. Pour tout $x\in[a,+\infty[$, on a :
$$|-t^2e^{-xt^2}|\leq t^2e^{-at^2}=g(t).$$
Or la fonction $g:t\mapsto t^2e^{-at^2}$ est intégrable sur $[0,+\infty[$ car elle est continue sur cet intervalle et en $+\infty$, comme $a>0$, $t^2e^{-at^2}=o(\frac{1}{t^2})$ (car $t^2g(t)\to 0$).

Ainsi, d'après le théorème de dérivation des intégrales à paramètre, pour tout $a>0$, F est C^1 sur $[a,+\infty[$. Il en résulte que F est C^1 sur $\mathbb{R}_+^*$ et on a, pour tout $x\in\mathbb{R}_+^*$:

$$F'(x)=\int_0^{+\infty}\underbrace{-t^2e^{-xt^2}}_{=(-\frac{1}{2}t).(2t)e^{-xt^2}}\,\mathrm{d}t\overset{\text{IPP}}{=}\left[\frac{-te^{-xt^2}}{2x}\right]_0^{+\infty}-\int_0^{+\infty}\frac{-e^{-xt^2}}{2x}\,\mathrm{d}t=\frac{1}{2x}F(x).$$

Ainsi, F est solution de l'équation différentielle $y'=\dfrac{1}{2x}y$ qui admet pour solutions $y=Ce^{\frac{1}{2}\ln(x)}=C\dfrac{1}{\sqrt{x}}$ pour $C\in\mathbb{R}$.
Or on a :
$$C=F(1)=\frac{\sqrt{\pi}}{2},$$
donc, pour tout $x\in\mathbb{R}_+^*$:
$$F(x)=\frac{1}{2}\sqrt{\frac{\pi}{x}}.$$

Exercice 9.

Montrer que les fonctions suivantes sont de classe C^n sur leur domaine de définition :

1. $F:t\mapsto\displaystyle\int_0^{+\infty}e^{-x^2}cos(xt)\,\mathrm{d}x$ avec $n=1$. En exprimant F' en fonction de F, déterminer une expression explicite de F.
2. $G:x\mapsto\displaystyle\int_0^{+\infty}\frac{e^{ixt}}{1+t^4}\,\mathrm{d}t$ avec $n=2$.
3. $H:x\mapsto\displaystyle\int_0^{+\infty}\frac{\sin(t)}{x^2+t^2}\,\mathrm{d}t$ avec $n=1$ puis déterminer un équivalent simple de H en $+\infty$.
4. pour $k\in\mathbb{N}$, $I_k:t\mapsto\displaystyle\int_1^{+\infty}\frac{e^{-xt}}{x^k}\,\mathrm{d}x$ avec $n=\infty$ puis déterminer une expression explicite

de $I_k^{(k)}$.

Correction.

1. On pose $f(t,x) = e^{-x^2}\cos(xt)$.
   — Déterminons le domaine $\mathcal{D}_F$ de F.
   Soit $t \in \mathbb{R}$. La fonction $x \mapsto f(t,x)$ est continue sur $[0,+\infty[$ et on a $|f(t,x)| \leq e^{-x^2}$, donc par comparaison à la fonction $x \mapsto e^{-x^2}$ qui intégrable sur $[0,+\infty[$, $x \mapsto f(t,x)$ est intégrable sur $[0,+\infty[$ et ainsi $\int_0^{+\infty} e^{-x^2}\cos(xt)\,\mathrm{d}x$ converge.
   Par suite, $F(t)$ existe et est fini pour pour tout $t \in \mathbb{R}$. Donc $\mathcal{D}_F = \mathbb{R}$.
   — Montrons que F est de classe C^1 sur $\mathbb{R}$. On vérifie les hypothèses du théorème de dérivation des intégrales à paramètres.
   - pour tout $x \in [0,+\infty[$, $f(\cdot,x) : t \mapsto e^{-x^2}\cos(xt)$ est de classe C^1 sur $\mathbb{R}$ (et même C^∞) et on a, pour tout $t \in \mathbb{R}$:
   $$\frac{\partial f}{\partial t}(t,x) = -xe^{-x^2}\sin(xt).$$
   - pour tout $t \in \mathbb{R}$, $f(t,\cdot) : x \mapsto e^{-x^2}\cos(xt)$ est intégrable sur $[0,+\infty[$ (voire l'étude du domaine).
   - pour tout $t \in \mathbb{R}$, $\dfrac{\partial f}{\partial t}(t,\cdot)$ est continue par morceaux sur $[0,+\infty[$ car continue sur $[0,+\infty[$;
   - *(Domination)* Soit $x \in [0,+\infty[$. Pour tout $t \in \mathbb{R}$, on a :
   $$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq xe^{-x^2} = g(x),$$
   Or la fonction $g : x \mapsto xe^{-x^2}$ est intégrable sur $[0,+\infty[$ car elle est continue sur cet intervalle et en $+\infty$, $xe^{-x^2} = o(\dfrac{1}{x^2})$ (car $x^2 g(x) \to 0$).

   Ainsi, d'après le théorème de dérivation des intégrales à paramètre, F est de classe C^1 sur $\mathbb{R}$ et on a, pour tout $t \in \mathbb{R}$:
   $$F'(t) = \int_0^{+\infty} -xe^{-x^2}\sin(xt)\,\mathrm{d}x \overset{\text{IPP}}{=} \left[\frac{e^{-x^2}}{2}.\sin(xt)\right]_0^{+\infty} - \frac{t}{2}\int_0^{+\infty} e^{-x^2}\cos(xt)\,\mathrm{d}x = -\frac{t}{2}F(t).$$
   Ainsi, F est solution de l'équation différentielle $y' + \dfrac{t}{2}y = 0$ qui admet pour solutions $y = Ce^{-\int \frac{t}{2}\,\mathrm{d}t} = Ce^{-\frac{t^2}{4}}$ pour $C \in \mathbb{R}$.
   Or on a :
   $$C = F(0) = \frac{\sqrt{\pi}}{2},$$
   donc, pour tout $t \in \mathbb{R}$:
   $$F(t) = \frac{\sqrt{\pi}}{2}e^{-\frac{t^2}{4}}.$$

2. On pose $g(x,t) = \frac{e^{-ixt}}{1+t^4}$.
   — Déterminons le domaine $\mathcal{D}_G$ de G.
   Soit $x \in \mathbb{R}$. La fonction $t \mapsto g(x,t)$ est continue sur $[0,+\infty[$ et pour tout $t \in [0,+\infty[$, $|g(x,t)| \leq \frac{1}{1+t^4}$, donc, par comparaison à la fonction $t \mapsto \frac{1}{1+t^4}$ qui est intégrable sur $[0,+\infty[$, $t \mapsto g(x,t)$ est intégrable sur $[0,+\infty[$ et ainsi $\int_0^{+\infty} g(x,t)\,\mathrm{d}t$ converge.
   Par suite, $G(x)$ existe et est fini pour tout $x \in \mathbb{R}$. Donc $\mathcal{D}_G = \mathbb{R}$.
   — Montrons que G est de classe C^2 sur $\mathbb{R}$. On vérifie les hypothèses du théorème de dérivation des intégrales à paramètres.
   - pour tout $t \in [0,+\infty[$, $g(\cdot,t) : x \mapsto g(x,t)$ est C^2 sur $\mathbb{R}$ (et même C^∞) et on a, pour tout $x \in \mathbb{R}$:
   $$\frac{\partial g}{\partial x}(x,t) = \frac{ite^{ixt}}{1+t^4} \text{ et } \frac{\partial^2 g}{\partial x^2}(x,t) = \frac{-t^2e^{ixt}}{1+t^4}.$$
   - pour tout $x \in \mathbb{R}$, $g(x,\cdot) : t \mapsto g(x,t)$ est intégrable sur $[0,+\infty[$ (fait avec le domaine de G).
   - pour tout $x \in \mathbb{R}$, $\dfrac{\partial g}{\partial x}(x,\cdot)$ est continue par morceaux sur $[0,+\infty[$ car continue sur $[0,+\infty[$;
   - *(Domination)* Soit $t \in [0,+\infty[$. Pour tout $x \in \mathbb{R}$, on a :
   $$\left|\frac{\partial g}{\partial x}(x,t)\right| \leq \frac{t}{1+t^4} = g_1(t)$$
   et
   $$\left|\frac{\partial^2 g}{\partial x^2}(x,t)\right| \leq \frac{t^2}{1+t^4} = g_2(t).$$
   Or les fonctions g_1, g_2 sont intégrables sur $[0,+\infty[$ car elles sont continues sur cet intervalle et en $+\infty$, $g_1(t)$ et $g_2(t)$ sont des $O(\dfrac{1}{t^2})$ (car $t \to t^2 g_i(t)$ sont bornées sur $[0,+\infty[$).

   Ainsi, d'après le théorème de dérivation des intégrales à paramètre, G est C^2 sur $\mathbb{R}$.
3. On pose $h(x,t) = \frac{sin(t)}{x^2+t^2}$.
   — Déterminons le domaine $\mathcal{D}_H$ de H.
   Soit $x \in \mathbb{R}$.
   ∗ Si $x = 0$, $\varphi : t \mapsto h(x,t) = \frac{sin(t)}{t^2}$ est continue sur $]0,+\infty[$. La fonction φ est positive sur $]0,\pi]$ et $\varphi(t) \underset{t\to 0}{\sim} \frac{1}{t}$, donc, par comparaison $\int_0^\pi \varphi(t)\,\mathrm{d}t$ diverge car l'intégrale de Riemann $\int_0^\pi \frac{1}{t}\,\mathrm{d}t$ diverge.
   Par suite, $\int_0^\pi h(0,t)\,\mathrm{d}t$ diverge, donc H n'est pas défine en 0.
   ∗ Si $x \neq 0$, la fonction $t \mapsto h(x,t)$ est continue sur $[0,+\infty[$ et pour tout $t \in [0,+\infty[$, $|h(x,t)| \leq \frac{1}{x^2+t^2} \underset{t\to+\infty}{\sim} \frac{1}{t^2}$, donc, par comparaison à la fonction $t \mapsto \frac{1}{t^2}$ qui est intégrable sur $[1,+\infty[$, $t \mapsto h(x,t)$ est intégrable sur $[1,+\infty[$ et donc sur $[0,+\infty[$. Ainsi $\int_0^{+\infty} h(x,t)\,\mathrm{d}t$ converge.
   Par suite, $H(x)$ existe et est fini si, et seulement si, $x \in \mathbb{R}^*$. Donc $\mathcal{D}_H = \mathbb{R}^*$.
   — La fonction H est paire, donc on restreint l'étude à $\mathbb{R}_+^*$. Montrons que H est de classe C^1 sur $\mathbb{R}_+^*$. On vérifie les hypothèses du théorème de dérivation des intégrales à paramètres.

- pour tout $t \in [0,+\infty[$, $h(\cdot,t) : x \mapsto h(x,t)$ est C^1 sur $\mathbb{R}_+^*$ (et même C^∞) et on a, pour tout $x \in \mathbb{R}_+^*$:
$$\frac{\partial h}{\partial x}(x,t) = \frac{-2x\sin(t)}{(x^2+t^2)^2}$$
- pour tout $x \in \mathbb{R}_+^*$, $h(x,\cdot) : t \mapsto h(x,t)$ est intégrable sur $[0,+\infty[$ (fait avec le domaine de H).
- pour tout $x \in \mathbb{R}_+^*$, $\dfrac{\partial h}{\partial x}(x,\cdot)$ est continue par morceaux sur $[0,+\infty[$ car continue sur $[0,+\infty[$;
- *(Domination sur tout segment de $\mathbb{R}_+^*$)* Soit $a,b \in \mathbb{R}_+^*$ avec $a<b$. Soit $t \in [0,+\infty[$. Pour tout $x \in [a,b]$, on a :
$$\left|\frac{\partial h}{\partial x}(x,t)\right| \leq \frac{2b}{(a^2+t^2)^2} = g(t)$$
Or la fonction g est intégrable sur $[0,+\infty[$ car elle est continue sur cet intervalle et en $+\infty$, $g(t) \underset{t\to+\infty}{\sim} \frac{2b}{t^4}$.

Ainsi, d'après le théorème de dérivation des intégrales à paramètre, H est C^1 sur $\mathbb{R}_+^*$ et donc également sur $\mathbb{R}_-^*$ par parité.

Déterminons un équivalent de H en $+\infty$.
Soit $x \in \mathbb{R}_+^*$. On effectue le changement de variable *licite* $t = ux$ dans l'intégrale convergente $H(x)$ et on obtient :
$$H(x) = \frac{1}{x}\int_0^{+\infty} \frac{\sin(ux)}{1+u^2}\,\mathrm{d}u$$
puis, par des Intégrations Par Parties *licites*, on a :
$$\begin{aligned}
H(x) &= \frac{1}{x}\left(\left[-\frac{\cos(ux)}{x}.\frac{1}{1+u^2}\right]_0^{+\infty} - \int_0^{+\infty} -\frac{\cos(ux)}{x}.\frac{-2u}{(1+u^2)^2}\,\mathrm{d}u\right)\\
&= \frac{1}{x^2} - \frac{2}{x^2}\int_0^{+\infty}\frac{u\cos(ux)}{(1+u^2)^2}\,\mathrm{d}u\\
&= \frac{1}{x^2} - \frac{2}{x^2}\left(\left[\frac{\sin(ux)}{x}.\frac{u}{(1+u^2)^2}\right]_0^{+\infty} - \int_0^{+\infty}\frac{\sin(ux)}{x}.\frac{1-3u^2}{(1+u^2)^3}\,\mathrm{d}u\right)\\
&= \frac{1}{x^2} - \frac{2}{x^3}\int_0^{+\infty}\underbrace{\frac{\sin(ux)(1-3u^2)}{(1+u^2)^3}}_{=f(x,u)}\,\mathrm{d}u.
\end{aligned}$$

La fonction $u \mapsto f(x,u)$ est continue sur $[0,+\infty[$ et $|f(x,u)| \leq \frac{1+3u^2}{(1+u^2)^3} \underset{u\to+\infty}{\sim} \frac{3}{u^4}$; donc $u \mapsto f(x,u)$ est intégrable sur $[0,+\infty[$. Par suite, par inégalité triangulaire, on a :
$$\left|\int_0^{+\infty} f(x,u)\,\mathrm{d}u\right| \leq \int_0^{+\infty}\frac{1+3u^2}{(1+u^2)^3}\,\mathrm{d}u = C$$

Ainsi, la fonction $x \mapsto \int_0^{+\infty} f(x,u)\,\mathrm{d}u$ est bornée sur $\mathbb{R}_+^*$ par C, donc on a :

$$\begin{aligned} H(x) &= \frac{1}{x^2} - \frac{2}{x^3}\int_0^{+\infty} f(x,u)\,\mathrm{d}u \\ &= \frac{1}{x^2} + \underset{x\to+\infty}{O}\left(\frac{1}{x^3}\right) \\ &= \frac{1}{x^2} + \underset{x\to+\infty}{o}\left(\frac{1}{x^2}\right) \\ H(x) &\underset{u\to+\infty}{\sim} \frac{1}{x^2}. \end{aligned}$$

4. Pour $k \in \mathbb{N}$, on pose $f_k(t,x) = \frac{e^{-xt}}{x^k}$.
   — Soit $k \in \mathbb{N}$. Déterminons le domaine $\mathcal{D}_{I_k}$ de I_k.
   Soit $t \in \mathbb{R}$. La fonction $x \mapsto f_k(t,x)$ est continue sur $[1,+\infty[$.
   (∗) Si $t < 0$, par croissances comparées, $\frac{e^{-xt}}{x^k} \xrightarrow[x\to+\infty]{} +\infty$, donc $\int_0^{+\infty} f_k(t,x)\,\mathrm{d}x$ diverge (car, par exemple, à partir d'un certain $x_0 \geq 1$, pour tout $x \geq x_0$, $|f_k(t,x)| \geq 1$).
   (∗) Si $t = 0$, on a $\int_0^{+\infty} f_k(0,x)\,\mathrm{d}x = \int_0^{+\infty} \frac{1}{x^k}\,\mathrm{d}x$ converge si, et seulement si, $k \geq 2$. Donc $I_k(0)$ existe et est fini si, et seulement si, $k \geq 2$.
   (∗) Si $t > 0$, on a $|f_k(t,x)| \leq e^{-xt}$ et $x \mapsto e^{-xt}$ est intégrable sur $[1,+\infty[$, donc, par comparaison, $x \mapsto f_k(t,x)$ l'est aussi et donc $\int_0^{+\infty} f_k(t,x)\,\mathrm{d}x$ converge.
   Il en résulte que, pour $k = 0, 1$, $\mathcal{D}_{I_k} = \mathbb{R}_+^*$ et pour $k \geq 2$, $\mathcal{D}_{I_k} = \mathbb{R}_+$.
   — Soit $k \in \mathbb{N}$. Montrons que I_k est de classe C^∞ sur $\mathbb{R}_+^*$ *(on pourrait même montrer que pour $k \geq 2$, I_k est de classe C^{k-2} sur $\mathbb{R}_+$)*. On vérifie les hypothèses du théorème de dérivation des intégrales à paramètres.
   - pour tout $x \in [1,+\infty[$, $f_k(\cdot,x)$ est de classe C^∞ sur $\mathbb{R}$ et donc sur $\mathbb{R}_+^*$ et on a, pour tout $n \in \mathbb{N}$ et pour tout $t \in \mathbb{R}$:
   $$\frac{\partial^n f_k}{\partial t^n}(t,x) = (-1)^n \frac{e^{-xt}}{x^{k-n}}.$$
   - pour tout $t \in \mathbb{R}$, $f_k(t,\cdot)$ est intégrable sur $[1,+\infty[$ (voire l'étude du domaine).
   - pour tout $n \in \mathbb{N}^*$ et pour tout $t \in \mathbb{R}$, $\dfrac{\partial^n f_k}{\partial t^n}(t,\cdot)$ est continue par morceaux sur $[1,+\infty[$ car continue sur $[1,+\infty[$;
   - *(Domination sur tout segment de $\mathbb{R}_+^*$)* Soit $a > 0$. Soit $x \in [1,+\infty[$. Pour tout $t \in [a,+\infty[$, on a :
   $$\left|\frac{\partial^n f_k}{\partial t^n}(t,x)\right| = \frac{e^{-xt}}{x^{k-n}} \leq \frac{e^{-ax}}{x^{k-n}} = g(x),$$
   Or la fonction $g : x \mapsto \frac{e^{-ax}}{x^{k-n}}$ est intégrable sur $[1,+\infty[$ car elle est continue sur cet intervalle et en $+\infty$, $g(x) = o(\frac{1}{x^2})$ (car $x^2 g(x) \to 0$ par croissances comparées).

   Ainsi, d'après le théorème de dérivation des intégrales à paramètre, I_k est de classe C^∞ sur $\mathbb{R}_+^*$ et on a, pour tout $n \in \mathbb{N}$ et pour tout $t \in \mathbb{R}_+^*$:
   $$I_k^{(n)}(t) = (-1)^n \int_1^{+\infty} \frac{e^{-xt}}{x^{k-n}}\,\mathrm{d}x;$$

En particulier, on a, pour $n = k$:

$$I_k^{(k)}(t) = (-1)^k I_0(t) = (-1)^k \left[-\frac{e^{-xt}}{t}\right]_1^{+\infty} = (-1)^k \frac{e^{-t}}{t}.$$

Exercice 10. *Fonction Gamma*

On appelle *fonction Gamma* et on note Γ la fonction de la variable réelle :

$$\Gamma : x \mapsto \int_0^{+\infty} t^{x-1}e^{-t}\,\mathrm{d}t.$$

1. Déterminer le domaine de définition $\mathcal{D}_\Gamma$ de Γ.
2. Pour $x \in \mathcal{D}_\Gamma$, trouver une relation entre $\Gamma(x+1)$ et $\Gamma(x)$. En déduire la valeur de $\Gamma(n+1)$ pour tout $n \in \mathbb{N}$.
3. Déterminer la valeur de $\Gamma(\frac{3}{2})$.
4. Montrer que la transformée de Laplace de $f : t \mapsto t^\alpha$ ($\alpha \in]-1,+\infty[$) est $\mathcal{L}(f) : p \mapsto \frac{\Gamma(\alpha+1)}{p^{\alpha+1}}$.
5. Montrer que Γ est de classe C^∞ sur $\mathcal{D}_\Gamma$.

Correction.

1. Déterminons le domaine de définition $\mathcal{D}_\Gamma$ de Γ. On pose $f(x,t) = t^{x-1}e^{-t}$. Soit $x \in \mathbb{R}$.
   - ⋆ La fonction $t \mapsto t^{x-1}e^{-t}$ est continue sur $]0,+\infty[$.
   - ⋆ en 0 :
   $$f(x,t) \underset{t\to 0^+}{\sim} t^{x-1}$$
   Or, d'après le critère de Riemann en 0, $t \mapsto t^{x-1} = \frac{1}{t^{1-x}}$ est intégrable sur $]0,1]$ si, et seulement si, $1-x < 1$ i.e. $x > 0$. Donc $t \mapsto f(x,t)$ est intégrable sur $]0,1]$ si, et seulement si, $x > 0$.
   - ⋆ en $+\infty$:
   On a $t^2t^{x-1}e^{-t} \xrightarrow[t\to+\infty]{} 0$, donc
   $$f(x,t) = t^{x-1}e^{-t} = t^{x+1}e^{-t} = \underset{t\to+\infty}{o}\left(\frac{1}{t^2}\right)$$
   et, d'après le critère de Riemann en $+\infty$, $t \mapsto \frac{1}{t^2}$ est intégrable sur $[1,+\infty[$. Ainsi, par comparaison, $t \mapsto f(x,t)$ est intégrable sur $[1,+\infty[$.

   Il en résulte que $t \mapsto f(x,t)$ est intégrable sur $]0,+\infty[$ si, et seulement si, $x > 0$.
   Par suite, $\mathcal{D}_\Gamma = \mathbb{R}_+^*$.

2. Soit $x \in \mathbb{R}_+^*$. Par une intégration par parties licite, on a :
   $$\Gamma(x+1) = \left[-t^x e^{-t}\right]_0^{+\infty} + x\int_0^{+\infty} t^{x-1}e^{-t}\,\mathrm{d}t = x\Gamma(x).$$
   Ainsi, pour $n \in \mathbb{N}$, on obtient (par récurrence),
   $$\Gamma(n+1) = n!\,\Gamma(0) = n!.$$

3. On a :
$$\Gamma\left(\frac{3}{2}\right)=\int_0^{+\infty}\sqrt{t}e^{-t}\,\mathrm{d}t$$
En effectuant le changement de variable $t=u^2$ puis une intégration par parties (tous deux licites), on obtient :
$$\begin{aligned}\Gamma\left(\frac{3}{2}\right) &= -\int_0^{+\infty} u\times(-2ue^{-u^2})\,\mathrm{d}u\\ &= -\left[ue^{-u^2}\right]_0^{+\infty}+\int_0^{+\infty}e^{-u^2}\,\mathrm{d}u\\ \Gamma\left(\frac{3}{2}\right) &= \frac{\sqrt{\pi}}{2}.\end{aligned}$$

4. Soit $\alpha\in]-1,+\infty[$. Comme $f:t\mapsto t^\alpha$ est continue sur $]0,+\infty[$, intégrable sur $]0,1]$ (critère de Riemann en 0) et $O(t^{\lfloor\alpha\rfloor+1})$ en $+\infty$, d'après le problème 3 question 2., $\mathcal{L}(f)$ est définie (et continue) sur $\mathbb{R}_+^*$.
Soit $p\in\mathbb{R}_+^*$. On effectue le changement de variable licite $u=pt$ et on obtient :
$$\mathcal{L}(f)(p)=\frac{1}{p^{\alpha+1}}\int_0^{+\infty}u^\alpha e^{-u}\,\mathrm{d}u=\frac{\Gamma(\alpha+1)}{p^{\alpha+1}}.$$

5. Montrons que Γ est de classe C^∞ sur $\mathbb{R}_+^*$.
Soit $k\in\mathbb{N}^*$.
   - pour tout $t\in]0,+\infty[$, $x\mapsto f(x,t)$ est C^∞ sur $\mathbb{R}_+^*$ et, pour tout $k\in\mathbb{N}$ et pour tout $x\in\mathbb{R}_+^*$:
   $$\frac{\partial^k}{\partial x^k}f(x,t)=\ln(t)^k t^{x-1}e^{-t}.$$
   - Soit $x\in\mathbb{R}_+^*$. La fonction $t\mapsto f(x,t)$ est intégrable sur $]0,+\infty[$ (on l'a montré lors de la détermination du domaine de définition).
   - Soit $k\in\mathbb{N}^*$. Pour tout $x\in\mathbb{R}_+^*$, $t\mapsto\dfrac{\partial^k}{\partial x^k}f(x,t)$ est continue par morceaux sur $]0,+\infty[$ car continue sur cet intervalle.
   - Soit $k\in\mathbb{N}^*$. Soit $a,b\in\mathbb{R}_+^*$ tels que $a<b$.
   <u>Domination sur le segment $[a,b]$</u>.
   Soit $t\in]0,+\infty[$. On effectue des dominations différentes selon que t soit dans $]0,1]$ ou dans $]1,+\infty[$.
   Pour tout $x\in[a,b]$, on a :
     * <u>si $t\in]0,1]$</u>, comme $x-1\geq a-1$, on a $t^{x-1}\leq t^{a-1}$ d'où :
     $$\left|\frac{\partial^k}{\partial x^k}f(x,t)\right|\leq|\ln(t)|^k t^{a-1}e^{-t}\quad(1)$$
     * <u>si $t\geq 1$</u>, comme $x-1\leq b-1$, $t^{x-1}\leq t^{b-1}$ d'où :
     $$\left|\frac{\partial^k}{\partial x^k}f(x,t)\right|\leq|\ln(t)|^k t^{b-1}e^{-t}\quad(2)$$

Soit $u \in \mathbb{R}_+^*$. On vérifie que la fonction $\varphi_u : t \mapsto |\ln(t)|^k t^{u-1} e^{-t}$ est intégrable sur $]0, +\infty[$. En effet, elle est continue sur $]0, +\infty[$ et :

⋆ en 0 :
On a :

$$\varphi_u(t) \underset{t \to 0^+}{\sim} |\ln(t)|^k t^{u-1}$$

Comme $u > 0$, on a $\frac{u}{2} > 0$ et donc, par croissances comparées :

$$t^{1-\frac{u}{2}} \varphi_u(t) = t^{\frac{u}{2}} |\ln(t)|^k \xrightarrow[t \to 0^+]{} 0.$$

Par suite, $\varphi_u(t) = \underset{t \to +\infty}{o}\left(\frac{1}{t^{1-\frac{u}{2}}}\right)$ et $t \mapsto \frac{1}{t^{1-\frac{u}{2}}}$ est intégrable sur $]0, 1]$ d'après le critère de Riemann en 0 car $1 - \frac{u}{2} < 1$.
Donc, par comparaison, φ_u est intégrable sur $]0, 1]$.

⋆ en $+\infty$: On a, par croissances comparées :

$$t^2 \varphi_u(t) = t^{u+1} |\ln(t)|^k e^{-t} \xrightarrow[t \to +\infty]{} 0.$$

Par suite, $\varphi_u(t) = \underset{t \to +\infty}{o}\left(\frac{1}{t^2}\right)$ et $t \mapsto \frac{1}{t^2}$ est intégrable sur $[1, +\infty[$ d'après le critère de Riemann en $+\infty$ car $2 > 1$.
Donc, par comparaison, φ_u est intégrable sur $[1, +\infty[$.

Il en résulte que, pour tout $u > 0$, φ_u est intégrable sur $]0, +\infty[$ et donc en particulier, φ_a et φ_b sont intégrables sur $]0, +\infty[$.

On pose $g_k = \varphi_a + \varphi_b$. Alors g_k est intégrable sur $]0, +\infty[$ comme somme de fonctions intégrables sur $]0, +\infty[$ et on a, pour tout $t \in]0, +\infty[$, d'après (1) et (2) :

$$\left| \frac{\partial^k}{\partial x^k} f(x, t) \right| \leq g_k(t).$$

Remarque : on aurait également pu prendre pour fonction de domination

$$g_k : t \mapsto \begin{cases} \varphi_a(t) & \text{si } t \in]0, 1] \\ \varphi_b(t) & \text{si } t \in]1, +\infty[\end{cases}$$

Par suite, d'après le théorème de dérivation des intégrales à paramètre, Γ est de classe C^∞ sur tout segment $[a, b]$ de $\mathbb{R}_+^*$ et donc Γ est de classe C^∞ sur $\mathbb{R}_+^*$.
De plus, on a, pour tout $k \in \mathbb{N}$ et pour tout $x \in \mathbb{R}_+^*$:

$$\Gamma^{(k)}(x) = \int_0^{+\infty} (\ln t)^k t^{x-1} e^{-t} \, \mathrm{d}t.$$

Problème 4. *Calcul de l'intégrale du sinus cardinal*

Le but de ce problème est de calculer la valeur de :

$$\int_0^{+\infty} \frac{\sin(x)}{x} \, \mathrm{d}x.$$

On rappelle que cette intégrale est convergente mais pas absolument convergente. *On pourra retrouver des démonstrations de ces deux résultats dans le chapitre "Intégration généralisée" ; la convergence est à savoir refaire !*

On reprend ici les notations et résultats du problème 3.

1. Soit f une fonction continue par morceaux sur $\mathbb{R}_+^*$. On suppose qu'il existe $\alpha \in \mathbb{R}$ et $\beta \in]-1,+\infty[$ tels que $f(x) = \underset{x\to+\infty}{O}(x^\alpha)$ et $f(x) = \underset{x\to 0^+}{O}(x^\beta)$.
   (a) Montrer que $\mathcal{L}(f)$ admet une limite en $+\infty$ et la déterminer.
   (b) Montrer que la transformée de Laplace $\mathcal{L}(f)$ est de classe C^∞ sur $\mathbb{R}_+^*$ et donner une expression de la dérivée k-ième de $\mathcal{L}(f)$ pour $k \in \mathbb{N}$.
   (c) On suppose $\beta > 0$. En déduire que, pour $g : x \mapsto \dfrac{f(x)}{x}$, $\mathcal{L}(g)$ est dérivable sur $\mathbb{R}_+^*$ et, pour tout $p \in \mathbb{R}_+^*$:
   $$\mathcal{L}(g)(p) = \int_p^{+\infty} \mathcal{L}(f)(q)\,\mathrm{d}q.$$
2. (a) Déterminer une expression explicite de la transformée de Laplace de sinc : $x \mapsto \frac{\sin(x)}{x}$.
   (b) Justifier la continuité de $\mathcal{L}(\text{sinc})$ sur $\mathbb{R}_+$ et donc en 0 (voire problème 3) puis déterminer la valeur de $\displaystyle\int_0^{+\infty} \text{sinc}(x)\,\mathrm{d}x$.

Correction.

1. Pour $p \in \mathbb{R}_+$ et $x \in]0,+\infty[$, on pose $\varphi(p,x) = f(x)e^{-px}$.
   Commençons par traduire les hypothèses sur la fonction f :
   — comme $f(x) = \underset{x\to+\infty}{O}(x^\alpha)$, il existe $M \in \mathbb{R}_+$ et $x_0 \in]0,+\infty[$ tel que, pour tout $x \geq x_0$:
   $$|f(x)| \leq Mx^\alpha$$
   — comme $f(x) = \underset{x\to 0^+}{O}(x^\beta)$, il existe $M' \in \mathbb{R}_+$ et $x_0' \in]0,+\infty[$ tel que, pour tout $0 < x \leq x_0'$:
   $$|f(x)| \leq M'x^\beta.$$
   De plus, comme f est continue par morceaux sur $\mathbb{R}_+^*$, pour $\gamma \in \mathbb{R}$, la fonction $x \mapsto x^{-\gamma}f(x)$ l'est aussi ; ainsi, elle est continue par morceaux sur le segment $[x_0, x_0']$ (ou $[x_0', x_0]$), et donc bornée sur ce segment par disons $M''(\gamma)$. Ainsi, quitte à prendre le maximum de $M, M', M''(\alpha)$ et $M''(\beta)$, on peut supposer que $M = M'$ et $x_0' = x_0$.
   (a) Vérifions les hypothèses du théorème de la limite d'une intégrale à paramètre. Comme on veut déterminer la limite en $+\infty$, on peut placer la variable p sur le voisinage de $+\infty$ de notre choix : prendre $\mathbb{R}_+^*$ n'est pas un bon choix car, lors de la domination, l'exponentielle bien pratique en $+\infty$ "disparaitra" par la proximité de 0 dans $\mathbb{R}_+^*$. On se placera alors, par exemple, sur $[33,+\infty[$ qui a l'avantage de laisser 0 loin de l'affaire !
   - pour tout $p \in [33,+\infty[$, $\varphi(p,\cdot) : x \mapsto \varphi(p,x)$ est continue par morceaux sur $]0,+\infty[$;

- pour tout $x \in]0, +\infty[$, on a :

$$\varphi(p, x) = f(x)e^{-px} \xrightarrow[p \to +\infty]{} 0 = h(x)$$

et $h : x \mapsto 0$ est continue par morceaux sur $]0, +\infty[$ car continue sur $]0, +\infty[$;
- Soit $x \in]0, +\infty[$. Pour tout $p \in [33, +\infty[$, on a :

$$|\varphi(p, x)| = |f(x)|e^{-px} \leq |f(x)|e^{-33x} = g(x), \qquad \textbf{hypothèse de domination}$$

et $g : x \mapsto |f(x)|e^{-33x}$ est intégrable sur $]0, +\infty[$ car, g est continue par morceaux sur $]0, +\infty[$ et :
  - ⋆ sur $]0, x_0]$, $g(x) \leq Mx^{\beta}$ et $x \mapsto x^{\beta} = \frac{1}{x^{-\beta}}$ est intégrable sur $]0, x_0]$ d'après le critère de Riemann en 0 car $-\beta < 1$, donc par comparaison g est intégrable sur $]0, x_0]$;
  - ⋆ sur $[x_0, +\infty[$, $g(x) \leq Mx^{\alpha}e^{-33x} \underset{x \to +\infty}{=} o\left(\dfrac{1}{x^2}\right)$ et $x \mapsto \frac{1}{x^2}$ est intégrable sur $[x_0, +\infty[$ d'après le critère de Riemann en $+\infty$ car $2 > 1$, donc par comparaison g est intégrable sur $[x_0, +\infty[$.

Ainsi, d'après le théorème de la limite d'une intégrale à paramètre, on a :

$$\lim_{p \to +\infty} \left(\int_0^{+\infty} f(x)e^{-px} \, \mathrm{d}x \right) = \int_0^{+\infty} \left(\lim_{x \to +\infty} f(x)e^{-px} \right) \mathrm{d}x = \int_0^{+\infty} 0 \, \mathrm{d}x = 0.$$

(b) Vérifions les hypothèses du théorème de dérivation des intégrales à paramètre :
- pour tout $x \in]0, +\infty[$, $\varphi(\cdot, x)$ est C^{∞} sur $\mathbb{R}_+^*$ car exp est de classe C^{∞} sur $\mathbb{R}$ et on a, pour tout $k \in \mathbb{N}$ et pour tout $p \in \mathbb{R}_+^*$:

$$\frac{\partial^k \varphi}{\partial p^k}(p, x) = (-1)^k x^k f(x)e^{-px};$$

- pour tout $p \in \mathbb{R}_+^*$, $\varphi(p, \cdot)$ est intégrable sur $]0, +\infty[$ (même démonstration que pour la fonction g de la question précédente en remplaçant 33 par $p > 0$) ;
- pour tout $k \in \mathbb{N}$, pour tout $p \in \mathbb{R}_+^*$, $\dfrac{\partial^k \varphi}{\partial p^k}(p, \cdot)$ est continue par morceaux sur $]0, +\infty[$ car f l'est et $x \mapsto x^k e^{-px}$ est continue sur $]0, +\infty[$;
- Soit $k \in \mathbb{N}^*$. Soit $a > 0$. Pour tout $p \in [a, +\infty[$, on a :

$$\left| \frac{\partial^k \varphi}{\partial p^k}(x, t) \right| \leq x^k |f(x)|e^{-ax} = g_k(x), \quad \textbf{hypothèse de domination sur tout segment}$$

et $g_k : x \mapsto x^k|f(x)|e^{-ax}$ est intégrable sur $]0, +\infty[$ car, g_k est continue par morceaux sur $]0, +\infty[$ et :
  - ⋆ sur $]0, x_0]$, $g_k(x) \leq Mx^{\beta+k}$ et $x \mapsto x^{\beta+k} = \frac{1}{x^{-\beta-k}}$ est intégrable sur $]0, x_0]$ d'après le critère de Riemann en 0 car $-\beta - k < 1 - k < 1$, donc par comparaison g_k est intégrable sur $]0, x_0]$;
  - ⋆ sur $[x_0, +\infty[$, $g_k(x) \leq Mx^{\alpha+k}e^{-ax} \underset{x \to +\infty}{=} o\left(\dfrac{1}{x^2}\right)$ et $x \mapsto \frac{1}{x^2}$ est intégrable sur $[x_0, +\infty[$ d'après le critère de Riemann en $+\infty$ car $2 > 1$, donc par comparaison g_k est intégrable sur $[x_0, +\infty[$.

Ainsi, d'après le théorème de dérivation des intégrales à paramètre, $\mathcal{L}(f)$ est de classe C^∞ sur $\mathbb{R}_+^*$ et on a, pour tout $k \in \mathbb{N}$ et tout $p \in \mathbb{R}_+^*$:

$$\mathcal{L}(f)^{(k)}(p) = (-1)^k \int_0^{+\infty} x^k f(x) e^{-px} \,\mathrm{d}x.$$

(c) D'après les hypothèses sur f, la fonction g est continue par morceaux sur $]0, +\infty[$; $g(x) = \underset{x\to+\infty}{O}(x^{\alpha-1})$ et $g(x) = \underset{x\to 0^+}{O}(x^{\beta-1})$. Comme $\alpha + 1 \in \mathbb{R}$ et $\beta - 1 > -1$, d'après les deux questions précédentes, $\mathcal{L}(g)$ admet pour limite 0 en $+\infty$, $\mathcal{L}(g)$ est de classe C^∞ sur $\mathbb{R}_+^*$ et donc en particulier dérivable sur $\mathbb{R}_+^*$ et on a, pour tout $p \in \mathbb{R}_+^*$:

$$\mathcal{L}(g)'(p) = -\int_0^{+\infty} \underbrace{xg(x)}_{=f(x)} e^{-px} \,\mathrm{d}x = -\mathcal{L}(f).$$

Ainsi, $-\mathcal{L}(g)$ est une primitive de $\mathcal{L}(f)$ et donc on a, $\mathcal{L}(g)$ admettant pour limite 0 en $+\infty$:

$$\begin{aligned}
\mathcal{L}(g)(p) &= -\lim_{q\to+\infty} \mathcal{L}(g)(q) - (-\mathcal{L}(g)(p)) \\
&= [-\mathcal{L}(g)(q)]_p^{+\infty} \\
\mathcal{L}(g)(p) &= \int_p^{+\infty} \mathcal{L}(f)(q) \,\mathrm{d}q.
\end{aligned}$$

2. (a) D'après la question 3.(a) du problème 3, on a $\mathcal{L}(\sin) : p \mapsto \frac{1}{p^2+1}$. De plus sin est continue sur $\mathbb{R}$ donc continue par morceaux sur $]0, +\infty[$ et vérifie $\sin(x) = \underset{x\to+\infty}{O}(x)$ et $\sin(x) = \underset{x\to 0^+}{O}(x)$, d'où, d'après la question précédente, $\mathcal{L}(\mathrm{sinc})$ est dérivable sur $\mathbb{R}_+^*$ et on a, pour tout $p \in \mathbb{R}_+^*$:

$$\begin{aligned}
\mathcal{L}(\mathrm{sinc})(p) &= \int_p^{+\infty} \mathcal{L}(\sin)(q) \,\mathrm{d}q \\
&= \int_p^{+\infty} \frac{1}{q^2+1} \,\mathrm{d}q \\
&= [\arctan(q)]_p^{+\infty} \\
\mathcal{L}(\mathrm{sinc})(p) &= \frac{\pi}{2} - \arctan(p).
\end{aligned}$$

(b) La fonction sinc est continue sur $\mathbb{R}_+$ (en la prolongeant par continuité en 0 par $\mathrm{sinc}(0) = 1$) et l'intégrale $\int_0^{+\infty} \mathrm{sinc}(x) \,\mathrm{d}x$ converge (voire Chapitre Intégration généralisée). Par suite, d'après la question 4.(b) du problème 3, la fonction $\mathcal{L}(\mathrm{sinc})$ est continue sur $\mathbb{R}_+$ et donc en particulier en 0. Ainsi, on a :

$$\begin{aligned}
\int_0^{+\infty} \operatorname{sinc}(x)\,\mathrm{d}x &= \mathcal{L}(\operatorname{sinc})(0) \\
&= \lim_{p\to 0^+} \mathcal{L}(\operatorname{sinc})(p) \\
&= \lim_{p\to 0^+} \left(\frac{\pi}{2} - \arctan(p)\right) \\
\int_0^{+\infty} \operatorname{sinc}(x)\,\mathrm{d}x &= \frac{\pi}{2}
\end{aligned}$$

Annexe
Outils pour la résolution des problèmes

1. Propriétés utiles des intégrales de Wallis

Problème 5. *Intégrales de Wallis*

Pour $n \in \mathbb{N}$, on pose $W_n = \displaystyle\int_0^{\frac{\pi}{2}} \sin^n(t)\,\mathrm{d}t$. On note W la suite $(W_n)_{n\in\mathbb{N}}$.

1. Montrer que W est une suite à valeurs strictement positives, strictement décroissante. En déduire que W est convergente et montrer que sa limite est 0.
   Indication pour ce dernier point : on pourra utiliser le théorème de convergence dominé si on est pressé ; sinon, on pourra fixer $\varepsilon > 0$ et couper intelligemment l'intégrale en deux parties chacune inférieure à $\varepsilon/2$ à partir d'un certain rang.
2. Déterminer une relation de récurrence $(*)$ entre W_{n+2} et W_n pour $n \in \mathbb{N}$.
3. (a) Utiliser $(*)$ pour montrer que $W_{n+1} \underset{n\to+\infty}{\sim} W_n$ et que la suite $((n+1)W_nW_{n+1})_{n\in\mathbb{N}}$ est constante en $\frac{\pi}{2}$.
   (b) En déduire que $W_n \underset{n\to+\infty}{\sim} \sqrt{\dfrac{\pi}{2n}}$ -‘ *on peut donc prouver que la limite de W est nulle d'une troisième façon !*
4. Calculer W_0 et W_1 puis déterminer à l'aide de la relation de récurrence $(*)$ que :

$$\forall\, n \in \mathbb{N}^*,\ W_{2n} = \frac{\binom{2n}{n}}{4^n}.\frac{\pi}{2} \text{ et } \forall\, n \in \mathbb{N},\ W_{2n+1} = \frac{4^n}{\binom{2n}{n}}.\frac{1}{2n+1}.$$

2. Formules de la moyenne

Proposition Annexe 1. Première formule de la moyenne

Soit $a, b \in \mathbb{R}$ avec $a < b$ et $f, g : [a, b] \to \mathbb{R}$. Si f, g sont continues et g positive, alors il existe $c \in [a, b]$ tel que :

$$\int_a^b f(t)g(t)\,\mathrm{d}t = f(c).\int_a^b g(t)\,\mathrm{d}t.$$

Démonstration.

Comme f est continue sur le segment $[a, b]$, d'après le théorème des bornes atteintes, elle est bornée et atteint ses bornes. Ainsi, comme g est positive et par croissance de l'intégrale, on a :

$$\min(f)\int_a^b g(t)\,\mathrm{d}t \le \int_a^b f(t)g(t)\,\mathrm{d}t \le \max(f)\int_a^b g(t)\,\mathrm{d}t$$

Par suite, $\int_a^b f(t)g(t)\,\mathrm{d}t$ est compris entre deux valeurs de la fonction continue $x \mapsto f(x)\int_a^b g(t)\,\mathrm{d}t$; ainsi, d'après le théorème des valeurs intermédiaires, il existe $c \in [a,b]$ tel que :

$$\int_a^b f(t)g(t)\,\mathrm{d}t = f(c)\int_a^b g(t)\,\mathrm{d}t.$$

□

Proposition Annexe 2. Deuxième formule de la moyenne

Soit $a, b \in \mathbb{R}$ avec $a < b$ et $f, g : [a,b] \to \mathbb{R}$. Si f est continue, g est de classe C^1, positive et décroissante, et f est continue, alors il existe $c \in [a,b]$ tel que :

$$\int_a^b f(t)g(t)\,\mathrm{d}t = g(a)\int_a^c f(t)\,\mathrm{d}t$$

Démonstration.

On note F la primitive de f qui s'annule en a. On obtient, par intégration par parties :

$$\int_a^b f(t)g(t)\,\mathrm{d}t = F(b)g(b) - \int_a^b F(t)g'(t)\,\mathrm{d}t.$$

Comme $-g', F$ sont continues et $-g'$ positive, d'après la première formule de la moyenne, il existe $c_0 \in [a,b]$ tel que :

$$-\int_a^b F(t)g'(t)\,\mathrm{d}t = \int_a^b F(t)(-g'(t))\,\mathrm{d}t = F(c_0)\int_a^b (-g'(t))\,\mathrm{d}t = F(c_0)(g(a)-g(b)).$$

Donc,

$$\int_a^b f(t)g(t)\,\mathrm{d}t = g(a)\int_a^{c_0} f(t)\,\mathrm{d}t + g(b)\int_{c_0}^b f(t)\,\mathrm{d}t$$

Si $\int_{c_0}^b f(t)\,\mathrm{d}t \geq 0$, comme g est décroissante, on a :

$$g(a)\int_{c_0}^b f(t)\,\mathrm{d}t \geq g(b)\int_{c_0}^b f(t)\,\mathrm{d}t \geq 0 = g(a)\int_{c_0}^{c_0} f(t)\,\mathrm{d}t$$

et on a, par le même raisonnement, l'inégalité dans l'autre sens si $\int_{c_0}^b f(t)\,\mathrm{d}t \leq 0$. La valeur $g(b)\int_{c_0}^b f(t)\,\mathrm{d}t$ est donc comprise entre deux valeurs de la fonction continue $x \mapsto g(a)\int_{c_0}^x f(t)\,\mathrm{d}t$ donc d'après le théorème des valeurs intermédiaires, il existe $c \in [c_0, b] \subset [a,b]$ tel que $g(b)\int_{c_0}^b f(t)\,\mathrm{d}t = g(a)\int_{c_0}^c f(t)\,\mathrm{d}t$, d'où, par la relation de Chasles :

$$\int_a^b f(t)g(t)\,\mathrm{d}t = g(a)\int_a^{c_0} f(t)\,\mathrm{d}t + \underbrace{g(b)\int_{c_0}^b f(t)\,\mathrm{d}t}_{=g(a)\int_{c_0}^c f(t)\,\mathrm{d}t} = g(a)\int_a^c f(t)\,\mathrm{d}t.$$

□

Chapitre XI
Probabilités discrètes

Table des matières

Dans ce chapitre, on utilisera la notation $\overline{A}$ pour désigner le complémentaire d'une sous-partie A d'un ensemble.

Partie A
Dénombrabilité

Dans cette partie, E et F désignent des ensembles quelconques.

1. Généralités

Définition 1. *Dénombrabilité*

- On dit que E est **au plus dénombrable** s'il est en bijection avec une partie de $\mathbb{N}$.
- On dit que E est **dénombrable** s'il est en bijection avec $\mathbb{N}$.

Exemple 1.

- Un ensemble fini est au plus dénombrable.

  > En effet, si son cardinal est n, alors il est en bijection avec la partie finie de $\mathbb{N}$, $[\![0, n]\!]$.

- L'ensemble $\mathbb{N}$ et $\mathbb{N}^*$ sont dénombrables.

  > Pour la dénombrabilité de $\mathbb{N}$ il suffit de considérer l'identité. Quant à $\mathbb{N}^*$, l'application $f : n \mapsto n+1$ est une bijection de $\mathbb{N}$ dans $\mathbb{N}^*$ (sa réciproque est $g : n \mapsto n-1$).

- L'ensemble $\mathbb{Z}$ est dénombrable.

  > On définit l'application $f : \mathbb{N} \to \mathbb{Z}$ telle que, pour $n \in \mathbb{N}$
  >
  > $$f(n) = \begin{cases} \frac{n}{2} & \text{si } n \text{ pair} \\ -\frac{n+1}{2} & \text{si } n \text{ impair} \end{cases}$$
  >
  > i.e. $f(0) = 0$, $f(1) = -1$, $f(2) = 1$, $f(3) = -2$, $f(4) = 2$, etc...
  > f est bijective de réciproque $f^{-1} : \mathbb{Z} \to \mathbb{N}$:
  >
  > $$f^{-1}(n) = \begin{cases} 2k & \text{si } k \geq 0 \\ -(2k+1) & \text{si } k < 0 \end{cases}$$

Exercice 1.

Montrer que $2\mathbb{N}$ est un ensemble dénombrable.

Correction.

On définit $f : \mathbb{N} \to 2\mathbb{N}$ par :
$$\forall\, n \in \mathbb{N}, \quad f(n) = 2n$$
Alors f est bijective : sa réciproque est $g : 2\mathbb{N} \to \mathbb{N}$ où, pour tout $k \in 2\mathbb{N}$, $g(k) = \frac{k}{2}$.

En fait, on a le résultat général suivant :

Proposition 1.

Soit A une partie infinie de $\mathbb{N}$. Alors A est dénombrable.

Démonstration.

Soit A une partie infinie de $\mathbb{N}$. Alors A est non vide - car l'ensemble vide est un ensemble fini (de cardinal 0), et A est non majoré dans $\mathbb{N}$ car, si on suppose par l'absurde qu'il est majoré par $n \in \mathbb{N}$, alors $A \subset [\![0, n]\!]$ qui est fini, contradiction.

On construit par récurrence une suite strictement croissante (u_n) à valeurs dans A qui parcourt, dans l'ordre, tous les éléments de A :

— **Initialisation.** Comme A est une partie non vide de $\mathbb{N}$, A possède un plus petit élément $\min(A)$. On note alors $u_0 = \min(A)$.

— **Hérédité.** Soit $n \in \mathbb{N}$. On suppose $u_0, \dots, u_n$ construits. Comme A n'est pas majoré par u_n, l'ensemble $\{k \in A \mid k > u_n\}$ est une partie non vide de $\mathbb{N}$ et donc possède un plus petit élément que l'on note u_{n+1} i.e.
$$u_{n+1} = \min(\{k \in A \mid k > u_n\}).$$

Ainsi la fonction de $\mathbb{N}$ dans A associée à la suite (u_n), $\varphi : n \mapsto u_n$, est une bijection.

En effet, elle est :

— *injective.* Pour tout $n \in \mathbb{N}$, par construction, on a $u_{n+1} > u_n$. Par suite la suite (u_n) est strictement croissante et donc la fonction φ l'est aussi. Ainsi, φ est injective.

— *surjective.* Soit $k \in A$. Comme la suite (u_n) est strictement croissante, il existe $m \in \mathbb{N}$ tel que $u_{m+1} > k$. L'ensemble de ces m est une partie non vide de $\mathbb{N}$, donc il en existe un plus petit que l'on note n i.e. $u_{n+1} > k$ et $k \geq u_n$.
Or, par définition, u_{n+1} est le plus petit élément de A strictement plus grand que u_n donc l'inégalité $u_{n+1} > k > u_n$ est impossible. Ainsi, $\varphi(n) = u_n = k$. Par suite, φ est surjective.

□

Proposition 2.

L'ensemble E est au plus dénombrable si, et seulement si, il existe une injection de E dans $\mathbb{N}$.

Démonstration.

Si E est au plus dénombrable, alors il existe une bijection f de E dans $A \subset E$. Par suite, $f : E \to \mathbb{N}$ est injective.

Réciproquement, s'il existe une injection f de E dans $\mathbb{N}$, alors $f : E \to f(E)$ est une bijection de

E dans la partie $f(E)$ de $\mathbb{N}$. Donc E est au plus dénombrable. $\square$

Exemple 2.

L'ensemble $\mathbb{N}^2$ est dénombrable : en effet, l'application $(n,m) \to 2^n 3^m$ est injective par unicité de la décomposition d'un nombre entier en facteurs premiers et de plus, $\mathbb{N}^2$ est infini.

Proposition 3.

Soit $A \subset E$. Si E est dénombrable, alors A est au plus dénombrable. Et de plus, si E est dénombrable et A est infini, alors A est dénombrable.

Démonstration.

On suppose E dénombrable. Alors il existe une bijection $f : E \to \mathbb{N}$. De plus, l'application $\mathrm{id}_{|A} : A \to E$ telle que, pour $a \in A$, $\mathrm{id}_{|A}(a) = a$ est injective. Par suite, l'application $g = \mathrm{id}_{|A} \circ f : A \to \mathbb{N}$ est injective. Il en résulte que A est au plus dénombrable.

De plus, si A est infini, $g : A \to g(A)$ est une bijection et $g(A)$ est donc une partie infinie de $\mathbb{N}$. Par suite, $g(A)$ est dénombrable, donc il existe une bijection $h : g(A) \to \mathbb{N}$. Ainsi, $g \circ h$ est une bijection de A dans $\mathbb{N}$. Donc A est dénombrable. $\square$

Proposition 4.

On suppose E non vide. Alors E est au plus dénombrable si, et seulement si, il existe une surjection de $\mathbb{N}$ sur E.

Démonstration.

- $(\Rightarrow)$. On suppose E au plus dénombrable. Alors il existe $A \subset \mathbb{N}$ et $f : A \to E$ une bijection. Comme E est non vide, on choisit $x_0 \in E$ et on construit l'application $\tilde{f}$: pour tout $n \in \mathbb{N}$,
$$\tilde{f}(n) = \begin{cases} f(n) & \text{si } n \in A \\ x_0 & \text{si } n \notin A. \end{cases}$$
Alors on a $\tilde{f}(\mathbb{N}) = f(A) \cup \{x_0\} = f(A) = E$. Donc $\tilde{f}$ est une surjection de $\mathbb{N}$ sur E.
- $(\Leftarrow)$. On suppose qu'il existe $f : \mathbb{N} \to E$ une surjection. Montrons qu'il existe une application injective g de E dans $\mathbb{N}$. Comme f est surjective, pour tout $x \in E$, $f^{-1}(\{x\})$ est une partie non vide de $\mathbb{N}$. Ainsi, la fonction de E dans $\mathbb{N}$:
$$g : x \mapsto \min(f^{-1}(\{x\})$$
est injective. En effet, pour $x, y \in E$ tels que $g(x) = g(y)$, on a $x = f(g(x)) = f(g(y)) = y$. $\square$

Exercice 2.

Montrer de nouveau que $\mathbb{N}^2$ est dénombrable en exhibant (graphiquement) une surjection de $\mathbb{N}^2$ sur $\mathbb{N}$.

2. Opérations sur les ensembles dénombrables

Proposition 5. *Produit cartésien*

Si E, F sont dénombrables, alors $E \times F$ est dénombrable.

Démonstration.

On suppose E, F dénombrables. Alors il existe des bijections $f : E \to \mathbb{N}$ et $g : F \to \mathbb{N}$. Alors la fonction $(x, y) \mapsto (f(x), g(y))$ est une bijection de $E \times F$ dans $\mathbb{N}^2$ qui est lui-même dénombrable. Ainsi, on peut construire une bijection de $\mathbb{N}$ sur $E \times F$. Par suite, $E \times F$ est dénombrable. $\square$

Exemple 3.

L'ensemble $\mathbb{Q}$ des nombres rationnels est dénombrable.

> En effet, l'application $f : \mathbb{Z} \times \mathbb{N}^* \to \mathbb{Q}$ telle que, pour $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$, $f(p, q) = \frac{p}{q}$ est surjective et l'ensemble $\mathbb{Z} \times \mathbb{N}^*$ est dénombrable comme produit cartésien d'ensembles dénombrables. Par suite, il existe une surjection de $\mathbb{N}$ sur $\mathbb{Q}$, d'où $\mathbb{Q}$ est dénombrable.

Remarque 1.

Par récurrence, on obtient qu'un produit cartésien d'une famille finie (non vide) d'ensembles dénombrables est dénombrable.

Exercice 3.

Soit A, B des parties de E. Montrer que si A, B sont dénombrables, alors $A \cup B$ est dénombrable.

On a la généralisation suivante :

Proposition 6. *Réunion*

Soit $(A_i)_{i \in I}$ une famille de parties de E. Si I est au plus dénombrable et si, pour tout $i \in I$, A_i est au plus dénombrable, alors $\bigcup_{i \in I} A_i$ est au plus dénombrable. Et de plus, s'il existe $i \in I$ tel que A_i est dénombrable, alors $\bigcup_{i \in I} A_i$ est dénombrable.

Démonstration.

On suppose I non vide, au plus dénombrable et pour tout $i \in I$, A_i au plus dénombrable. Alors, pour chaque $i \in I$, il existe une surjection f_i de $\mathbb{N}$ dans A_i (quitte à le supprimer de la liste, on peut supposer A_i non vide). Alors la fonction $f : I \times \mathbb{N} \to \bigcup_{i\in I} A_i$ telle quel, pour $(i,n) \in I \times \mathbb{N}$:

$$f(i,n) = f_i(n).$$

est surjective. En effet, pour tout $a \in \bigcup_{i\in I} A_i$, il existe $i \in \mathbb{I}$ tel que $a \in A_i$ et il existe $n \in \mathbb{N}$ tel que $a = f_i(n)$; par suite, $a)f(i,n)$.
Ainsi, comme $I \times \mathbb{N}$ est dénombrable, il en résulte que $\bigcup_{i\in I} A_i$ est dénombrable. □

3. Exemples d'ensembles non dénombrables

Proposition 7.

Il n'existe aucune surjection de E sur l'ensemble $\mathcal{P}(E)$ des parties de E.

Démonstration.

Soit $f : E \to \mathcal{P}(E)$ une application. On note $A = \{x \in E \mid x \notin f(x)\}$. Alors $A \subset \mathcal{P}(E)$ n'a pas d'antécédent par f. En effet, pour $x \in E$ on a :

— *1er cas :* $x \in A$. Alors $f(x) \neq A$ car comme $x \in A$, par définition de A, $x \notin f(x)$.

— *2nd cas :* $x \notin A$. Alors $f(x) \neq A$ car comme $x \notin A$, par définition de A, $x \in f(x)$.

Dans tous les cas, $f(x) \neq A$.

Il en résulte que f n'est pas surjective. □

Remarque 2.

Il en résulte que $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable : en effet, il n'existe a fortiori aucune bijection de $\mathbb{N}$ dans $\mathcal{P}(\mathbb{N})$.

Exercice 4.

1. Montrer que $\mathcal{P}(E)$ est en bijection avec $\{0,1\}^E$.
2. On considère $X = \{0,1\}^{\mathbb{N}}$. Montrer que X n'est pas dénombrable grâce à l'argument diagonal de Cantor : supposer par l'absurde qu'il existe une énumération (indexée par $\mathbb{N}$) de X, et construire un élément de X qui ne peut pas être dans cette énumération (indice : argument "diagonal" !).
3. En déduire qu'un produit dénombrable d'ensembles au plus dénombrables n'est pas dénombrable en général et également que $\mathcal{P}(\mathbb{N})$ n'est toujours pas dénombrable !

Proposition 8.

L'ensemble $\mathbb{R}$ des nombres réels n'est pas dénombrable.

Partie B
Familles sommables

1. Familles sommables de nombres réels positifs

Dans ce paragraphe, I désigne un ensemble dénombrable.

a. Définitions

Définition 2. *Famille sommable*

Soit $(a_i)_{i\in I}$ de réels positifs est **sommable** si l'ensemble

$$\left\{\sum_{j\in J} a_j \mid J\subset I \text{ et } J \text{ fini}\right\}$$

est majoré ; autrement dit si :

il existe $M\geq 0$ tel que, pour toute partie J fini de I, $\sum_{j\in J} a_j\leq M$.

— Si la famille $(a_i)_{i\in I}$ est sommable, alors on appelle **somme de la famille** $(a_i)_{i\in I}$ et on note $\displaystyle\sum_{i\in I} a_i$ la quantité :

$$\sum_{i\in I} a_i = \sup\left(\left\{\sum_{j\in J} a_j \mid J\subset I \text{ et } J \text{ fini}\right\}\right).$$

— Si la famille $(a_i)_{i\in I}$ n'est pas sommable, on convient que $\sum_{i\in I} a_i = +\infty$.

Exemple 4.

La famille $(\frac{1}{2^n})_{n\in\mathbb{N}}$ est sommable.

En effet, pour toute partie finie J de $\mathbb{N}$, on a $J\subset [\![0,m]\!]$ où $m=\max(J)$, donc :

$$\sum_{j\in J}\frac{1}{2^j}\leq\sum_{i=0}^{m}\frac{1}{2^i} = 2-\frac{1}{2^m}\leq 2.$$

(Démontrer que la somme de la famille est en fait bien égale à 2 grâce à la définition précédente)

Proposition 9.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels positifs. Alors la famille $(u_n)_{n\in\mathbb{N}}$ est sommable si, et seulement si, $\sum u_n$ converge. Dans ce cas, on a :

$$\sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n.$$

b. Propriétés

Proposition 10. *Sous-familles sommable*

Soit $(a_i)_{i\in I}$ une famille de réels positifs. Si $(a_i)_{i\in I}$ est sommable, alors pour tout $J \subset I$ la sous-famille $(a_j)_{j\in J}$ est sommable et

$$\sum_{j\in J} a_j \leq \sum_{i\in I} a_i.$$

Proposition 11.

Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles de réels positifs telles que pour tout $i \in I$, $a_i \leq b_i$. Si $(b_i)_{i\in I}$ est sommable, alors $(a_i)_{i\in I}$ est sommable et

$$\sum_{i\in I} a_i \leq \sum_{i\in I} b_i.$$

Proposition 12.

Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles de réels positifs et $\lambda, \mu \in \mathbb{R}_+$. Si $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ sont sommables, alors la famille $(\lambda a_i + \mu b_i)_{i\in I}$ est sommable et

$$\sum_{i\in I} (\lambda a_i + \mu b_i) = \lambda \sum_{i\in I} a_i + \mu \sum_{i\in I} b_i.$$

Proposition 13. *Changement d'indice*

Soit σ une bijection d'un ensemble J sur I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable si et seulement si, $(a_{\sigma(j)})_{j\in J}$ est sommable. Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}.$$

c. Sommation par paquets

Définition 3. *Partition*

Soit $(I_\omega)_{\omega\in\Omega}$ une famille de partie de I. On dit que $(I_\omega)_{\omega\in\Omega}$ est une **partition** de I si $\bigcup_{\omega\in\Omega} I_\omega = I$ et pour tout $\omega, \omega' \in \Omega$ avec $\omega \neq \omega'$, $I_\omega \cap I_{\omega'} = \emptyset$.

Exemple 5.

Les ensembles $2\mathbb{N}$ et $2\mathbb{N}+1$ forment une partition de $\mathbb{N}$.

Proposition 14. *Sommation par deux paquets*

Soit J, K des parties de I tels que (J,K) est une partition de I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable si, et seulement si $(a_j)_{j\in J}$ et $(a_k)_{i\in K}$ sont sommables. Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{j\in J} a_j + \sum_{k\in K} a_k.$$

Exercice 5.

On admet l'égalité $\sum_{n=1}^{+\infty} \dfrac{1}{n^2} = \dfrac{\pi^2}{6}$. En déduire que $\sum_{n\geq 1} \dfrac{1}{(2n+1)^2}$ converge et déterminer sa somme.

On peut généraliser la proposition précédente de la façon suivante :

Théorème 1. *Sommation par paquets*

Soit $(I_\omega)_{\omega\in\Omega}$ une partition au plus dénombrable de I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

— pour tout $\omega \in \Omega$, la famille $(a_i)_{i\in I_\omega}$ est sommable de somme $\alpha_\omega = \sum_{i\in I_\omega} a_i$,

— la famille $(\alpha_\omega)_{\omega\in\Omega}$ est sommable.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{\omega\in\Omega} \left(\sum_{i\in I_\omega} a_i \right).$$

Corollaire 1.

Soit $(I_n)_{n\in\mathbb{N}}$ une partition au plus dénombrable de I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

— pour tout $n \in \mathbb{N}$, la famille $(a_i)_{i\in I_n}$ est sommable,

— la série $\sum \left(\sum_{i\in I_n} a_i \right)$ est convergente.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} a_i\right).$$

2. Familles sommables de nombres complexes

Dans ce paragraphe, I désigne un ensemble dénombrable.

a. Définitions

Définition 4. *Famille sommable de nombres complexes*

Soit $(a_i)_{i\in I}$ une famille de nombres complexes. On dit que la famille $(a_i)_{i\in I}$ est **sommable** si la famille $(|a_i|)_{i\in I}$ de réels positifs est sommable.

Remarque 3.

Ainsi, d'après ce qu'on a vu précédemment, une suite $(a_n)_{n\in\mathbb{N}}$ est sommable si, set seulement si, la série de terme général $(a_n)_{n\in\mathbb{N}}$ est absolument convergente.

Définition 5. *Somme d'une famille sommable de réels*

Soit $(a_i)_{i\in I}$ une famille sommable de nombres réels. On pose, pour chaque $i\in I$, $a_i^+ = \max(0,a_i)$ et $a_i^- = \min(0,a_i)$. On appelle **somme** de la famille $(a_i)_{i\in I}$ et on note $\sum\limits_{i\in I} a_i$ la quantité :

$$\sum_{i\in I} a_i = \sum_{i\in I} a_i^+ + \sum_{i\in I} a_i^-.$$

Définition 6. *Somme d'une famille sommable de complexes*

Soit $(a_i)_{i\in I}$ une famille sommable de nombres complexes. On appelle **somme** de la famille $(a_i)_{i\in I}$ et on note $\sum\limits_{i\in I} a_i$ la quantité :

$$\sum_{i\in I} a_i = \sum_{i\in I}\mathrm{Re}(a_i) + i\sum_{i\in I}\mathrm{Im}(a_i).$$

Proposition 15.

Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres complexes. La famille $(a_n)_{n\in\mathbb{N}}$ est sommable si, et seulement si, la série $\sum a_n$ est absolument convergente. Dans ce cas, on a :

$$\sum_{n\in\mathbb{N}} a_n = \sum_{i=0}^{+\infty} a_n.$$

b. Propriétés

Proposition 16. *Changement d'indice*

Soit σ une bijection d'un ensemble J sur I et $(a_i)_{i\in I}$ une famille de nombres complexes. Alors $(a_i)_{i\in I}$ est sommable si et seulement si, $(a_{\sigma(j)})_{j\in J}$ est sommable. Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}.$$

Remarque 4.

En combinant les deux précédentes propositions, on obtient que la somme d'une série absolument convergente est invariante par changement bijectif d'indice. Dans le cas d'une série semi-convergente, par exemple $\sum_{n\geq 1} \frac{(-1)^n}{n}$, on peut, en appliquant une permutation d'indice bien choisie, faire converger la série obtenue vers n'importe quel nombre réel, et même la faire diverger !

Proposition 17. *Inégalité triangulaire*

$(a_i)_{i\in I}$ une famille de nombres complexes. Si $(a_i)_{i\in I}$ est sommable, alors

$$|\sum_{i\in I} a_i| \leq \sum_{i\in I} |a_i|.$$

Proposition 18.

Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles à valeurs dans $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$ et $\lambda,\mu\in\mathbb{K}$. Si $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ sont sommables, alors la famille $(\lambda a_i+\mu b_i)_{i\in I}$ est sommable et

$$\sum_{i\in I}(\lambda a_i+\mu b_i) = \lambda\sum_{i\in I} a_i + \mu\sum_{i\in I} b_i.$$

Question.

De quelle structure peut-on munir l'ensemble des familles sommables sur le corps $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$?

c. Sommation par paquets

Théorème 2. *Sommation par paquets*

Soit $(I_\omega)_{\omega\in\Omega}$ une partition au plus dénombrable de I et $(a_i)_{i\in I}$ une famille de complexes. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

— pour tout $\omega \in \Omega$, la famille $(a_i)_{i\in I_\omega}$ est sommable de somme $\alpha_\omega = \sum\limits_{i\in I_\omega} a_i$,

— la famille $(\alpha_\omega)_{\omega\in\Omega}$ est sommable.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{\omega\in\Omega}\left(\sum_{i\in I_\omega} a_i\right).$$

Corollaire 2.

Soit $(I_n)_{n\in\mathbb{N}}$ une partition de I et $(a_i)_{i\in I}$ une famille de complexes. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

— pour tout $n \in \mathbb{N}$, la famille $(a_i)_{i\in I_n}$ est sommable,

— la série $\sum\left(\sum\limits_{i\in I_n} |a_i|\right)$ est convergente.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} a_i\right).$$

3. Applications des familles sommables

a. Sommes doubles

Définition-Proposition 7. *Partitions classiques de $\mathbb{N}^2$*

On appelle :

— **Partition verticale** de $\mathbb{N}^2$, la partition $(V_m)_{m\in\mathbb{N}}$ de $\mathbb{N}^2$ où, pour $m \in \mathbb{N}$,

$$V_m = \{m\} \times \mathbb{N} = \{(m, n) \mid n \in \mathbb{N}\}$$

.

— **Partition horizontale** de $\mathbb{N}^2$, la partition $(H_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$ où, pour $n \in \mathbb{N}$,

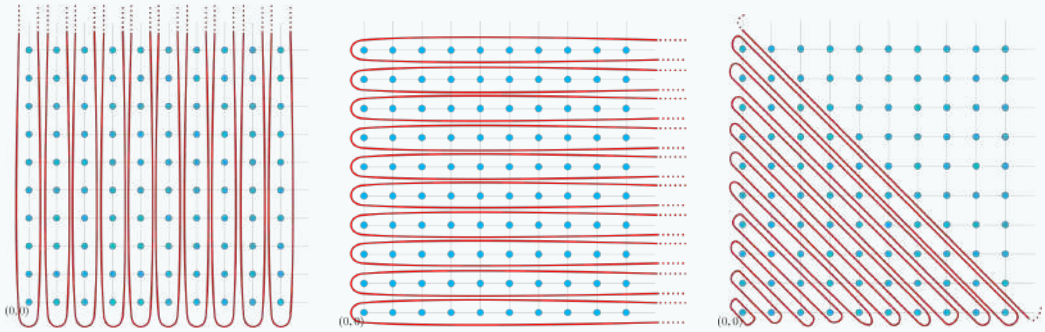
$$H_n = \mathbb{N} \times \{n\} = \{(m, n) \mid m \in \mathbb{N}\}$$

.

— **Partition diagonale** de $\mathbb{N}^2$, la partition $(\Delta_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$ où, pour $k \in \mathbb{N}$,

$$\Delta_k = \{(m, n) \in \mathbb{N}^2 \mid m + n = k\} = \{(p, k - p) \mid p \in [\![0, k]\!]\}$$

.



Partition verticale Partition horizontale Partition diagonale

Théorème 3. *Somme double "verticale"*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels positifs. Alors $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si :

— pour tout $m \in \mathbb{N}$, la série $\sum_{n\geq 0} a_{m,n}$ converge (somme sur V_m) et

— la série $\sum \left(\sum_{n=0}^{+\infty} a_{m,n}\right)$ converge.

Dans ce cas, on a :

$$\sum_{(n,m)\in\mathbb{N}^2} a_{m,n} = \sum_{m=0}^{+\infty}\left(\sum_{n=0}^{+\infty} a_{m,n}\right).$$

Théorème 4. *Somme double "horizontale"*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels positifs. Alors $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si :

— pour tout $n \in \mathbb{N}$, la série $\sum_{m\geq 0} a_{m,n}$ converge (somme sur H_n) et

— la série $\sum \left(\sum_{m=0}^{+\infty} a_{m,n}\right)$ converge.

Dans ce cas, on a :

$$\sum_{(n,m)\in\mathbb{N}^2} a_{m,n} = \sum_{n=0}^{+\infty}\left(\sum_{m=0}^{+\infty} a_{m,n}\right).$$

Théorème 5. *Somme double "diagonale"*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels positifs. Alors $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si :

— la série $\sum \left(\underbrace{\sum_{p=0}^{k} a_{p,k-p}}_{\text{somme sur } \Delta_k} \right)$ converge.

Dans ce cas, on a :

$$\sum_{(n,m)\in\mathbb{N}^2} a_{m,n} = \sum_{k=0}^{+\infty} \left(\sum_{p=0}^{k} a_{p,n-p} \right) = \sum_{k=0}^{+\infty} \left(\sum_{m+n=k} a_{m,n} \right).$$

Corollaire 3. *Somme double de complexes*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de complexes. Si $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable alors :

$$\begin{aligned}
\sum_{(m,n)\in\mathbb{N}^2} a_{m,n} &= \sum_{n=0}^{+\infty} \left(\sum_{m=0}^{+\infty} a_{m,n} \right) \\
&= \sum_{m=0}^{+\infty} \left(\sum_{n=0}^{+\infty} a_{m,n} \right) \\
&= \sum_{k=0}^{+\infty} \left(\sum_{p=0}^{k} a_{p,n-p} \right) \\
&= \sum_{k=0}^{+\infty} \left(\sum_{m+n=k} a_{m,n} \right)
\end{aligned}$$

Exercice 6.

Soit $\alpha \in \mathbb{R}$. Étudier la sommabilité de la famille $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ où, pour $(m,n) \in \mathbb{N}^2$,

$$a_{m,n} = \frac{1}{(m+n+1)^\alpha}.$$

b. Produit de Cauchy

Proposition 19.

Soit $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ des suites de nombres complexes. Si $\sum a_n$ et $\sum b_n$ sont absolument convergentes, alors

$$\sum_{(m,n)\in\mathbb{N}^2} a_m b_n = \left(\sum_{n=0}^{+\infty} a_n \right) \left(\sum_{n=0}^{+\infty} b_n \right).$$

Définition 8. *Produit de Cauchy*

Soit $\sum a_n$ et $\sum b_n$ des séries de nombres complexes. On appelle **produit de Cauchy** des séries $\sum a_n$ et $\sum b_n$ la série $\sum c_n$ de terme général défini, pour $n \in \mathbb{N}$, par :

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Théorème 6.

Soit $\sum a_n$ et $\sum b_n$ des séries de nombres complexes. Si $\sum a_n$ et $\sum b_n$ sont absolument convergentes, alors leur produit de Cauchy $\sum c_n$ est absolument convergent et on a :

$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right) = \sum_{n=0}^{+\infty}\left(\sum_{p+q=n} a_p b_q\right).$$

Exercice 7.

Soit $z \in \mathbb{C}$ tel que $|z| < 1$. Montrer que $\displaystyle\sum_{n=0}^{+\infty}(n+1)z^n = \left(\sum_{n=0}^{+\infty} z^n\right)^2$. En déduire que

$$\sum_{n=0}^{+\infty}(n+1)z^n = \frac{1}{(z-1)^2}.$$

Partie C

Espaces probabilisés

1. Espaces probabilisables

a. Définitions

Définition 9. *Tribu*

Soit Ω un ensemble non vide et $\mathcal{T} \subset \mathcal{P}(\Omega)$. On dit que $\mathcal{T}$ est une **tribu** sur Ω si :

- $\Omega \subset \mathcal{T}$;
- $\mathcal{T}$ est stable par passage au complémentaire i.e. pour tout $A \in \mathcal{T}$, $\overline{A} \in \mathcal{T}$;
- $\mathcal{T}$ est stable par réunion dénombrable i.e. pour toute suite $(A_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathcal{T}$, $\displaystyle\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}$.

Remarque 5.

On trouve parfois la terminologie σ-algèbre pour désigner une tribu et la notation $\mathcal{A}$ au lieu de $\mathcal{T}$.

Définition 10. *Espace probabilisable*

Soit Ω un ensemble non vide et $\mathcal{T}$ une tribu sur Ω. Le couple $(\Omega, \mathcal{T})$ est appelé **espace probabilisable**.

Exemple 6.

Soit Ω un ensemble non vide.
- $\mathcal{T} = \{\emptyset, \Omega\}$ est une tribu appelée **tribu grossière** ;
- $\mathcal{T} = \mathcal{P}(\Omega)$ est une tribu ;
- Si $A \subset \Omega$, $\mathcal{T} = \{\emptyset, A, \overline{A}, \Omega\}$ est une tribu

Exercice 8.

Soit Ω un ensemble non vide et $\mathcal{F} \subset \mathcal{P}(\Omega)$.
1. Montrer que l'intersection de toutes les tribus contenant $\mathcal{F}$ (en existe-t-il ?) est une tribu. On note cette tribu $\sigma(\mathcal{F})$ et on l'appelle **tribu engendrée par** $\mathcal{F}$.
2. Montrer que $\sigma(\mathcal{F})$ est la plus petite tribu contenant $\mathcal{F}$.

Remarque 6. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On utilise la terminologie suivante :
- Ω est l'**univers** ;
- les éléments A de $\mathcal{T}$ sont les **événements** ;

b. Les événements

Proposition 20.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On a les propriétés suivantes :
- $\emptyset \in \mathcal{T}$;
- $\mathcal{T}$ est stable par réunion et intersection finies ;
- $\mathcal{T}$ est stable par intersection dénombrable i.e. pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements de $\mathcal{T}$, alors $\displaystyle\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{T}$;
- pour tout $A, B \in \mathcal{T}$, $A \setminus B \in \mathcal{T}$.

Remarque 7. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $A, B \in \mathcal{T}$ des événements. On utilise la terminologie suivante concernant les événements :

— Ω est l'événement **certain** et $\emptyset$ l'événement **impossible** ;
— $\overline{A}$ est l'événement **contraire** de A ;
— si $A \cap B = \emptyset$, on dit que A et B sont des événements **incompatibles**.

Définition 11. *Système complet d'événements*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $(A_i)_{i \in I}$ une famille au plus dénombrable d'événements. On dit que la famille $(A_i)_{i \in I}$ est un **système complet d'événements** si elle forme une partition de Ω.

2. Espaces probabilisés

a. Définition

Définition 12. *Probabilité*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $P : \mathcal{T} \to [0, 1]$ une application. On dit que P est une **probabilité** sur $(\Omega, \mathcal{T})$ si :

i) $P(\Omega) = 1$;

ii) pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux disjoints, la série $\sum P(A_n)$ converge et :

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} P(A_n). \qquad \textbf{σ-additivité}$$

Remarque 8. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{T})$. Pour $A \in \mathcal{T}$, on dit que $P(A)$ est **la probabilité de l'événement** A.

Définition 13. *Espace probabilisé*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{T})$. Le triplet $(\Omega, \mathcal{T}, P)$ est appelé **espace probabilisé**.

Définition 14. *Événement négligeable/presque sûr*

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $A \in \mathcal{T}$. On dit que :

— l'événement A est **négligeable** si $P(A) = 0$.
— l'événement A est **presque sûr** si $P(A) = 1$.

b. Propriétés

Proposition 21.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $A, B \in \mathcal{T}$. On a les propriétés suivantes :

i) $P(\emptyset) = 0$;

ii) si A et B sont incompatibles, $P(A \cup B) = P(A) + P(B)$;

iii) $P(\overline{A}) = 1 - P(A)$;

iv) si $A \subset B$, $P(B \setminus A) = P(B) - P(A)$;

v) si $A \subset B$, $P(A) \leq P(B)$;

vi) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Exercice 9.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$. Montrer que si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements, alors :

1. $\sum P(A_n)$ converge et $\displaystyle\sum_{n=0}^{+\infty} P(A_n) = 1$.

2. pour tout $B \in \mathcal{T}$, $\sum P(A_n \cap B)$ converge et $\displaystyle\sum_{n=0}^{+\infty} P(A_n \cap B) = P(B)$.

3. Propriétés élémentaires des probabilités

Théorème 7. *Continuité croissante/décroissante*

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$.

- Si $(A_n)_{n\in\mathbb{N}}$ est croissante i.e. pour tout $n \in \mathbb{N}$, $A_n \subset A_{n+1}$, alors :

$$P(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P(A_n).$$

- Si $(A_n)_{n\in\mathbb{N}}$ est décroissante i.e. pour tout $n \in \mathbb{N}$, $A_{n+1} \subset A_n$, alors :

$$P(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P(A_n).$$

Proposition 22.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$.

- pour $n \in \mathbb{N}$, on a :

$$P(\bigcup_{i=0}^{n} A_n) \leq \sum_{i=0}^{n} P(A_i).$$

- On a :

$$P(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n=0}^{+\infty} P(A_i).$$

Corollaire 4.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$. Si, pour tout $n \in \mathbb{N}$, A_n est négligeable, alors $\bigcup\limits_{n\in\mathbb{N}} A_n$ est négligeable.

4. Probabilité sur un univers au plus dénombrable

a. Caractérisation

Théorème 8.

Soit Ω un ensemble non vide au plus dénombrable.

i) Soit P une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ et on pose, pour $\omega \in \Omega$, $p_\omega = P(\{\omega\})$. Alors, pour tout $\omega \in \Omega$, $p_\omega \geq 0$ et :
$$\sum_{\omega\in\Omega} p_\omega = 1.$$

ii) Si $(p_\omega)_{\omega\in\Omega}$ est une suite de réels positifs telle que $\sum\limits_{\omega\in\Omega} p_\omega = 1$, alors il existe une unique probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ telle que, pour tout $\omega \in \Omega$, $P(\{\omega\}) = p_\omega$. Elle est définie, pour $A \in \mathcal{P}(\Omega)$, par :
$$P(A) = \sum_{\omega\in A} p_\omega.$$

b. Exemples de probabilités sur $\mathbb{N}$

Exemple 7.

1. **Loi géométrique.** Soit $p \in]0, 1[$. On définit la probabilité P suivante sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par $p_n = P(\{n\}) = (1-p)^n p$ pour $n \in \mathbb{N}$.
2. **Loi de Poisson.** Soit $\lambda \in \mathbb{R}_+^*$. On définit la probabilité P suivante sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par $p_n = P(\{n\}) = e^{-\lambda}\dfrac{\lambda^n}{n!}$ pour $n \in \mathbb{N}$.

Partie D
Probabilités conditionnelles

Dans cette partie, $(\Omega, \mathcal{T}, P)$ désigne un espace probabilisé.

1. Conditionnement

Définition 15. *Probabilité conditionnelle*

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. Pour $A \in \mathcal{T}$, on appelle **probabilité (conditionnelle) de** A **sachant** B et on note $P_B(A)$ ou encore $P(A|B)$ la quantité :

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

L'application $P_B : \mathcal{T} \to [0,1]$ définie par :

$$P_B : A \mapsto P_B(A)$$

est appelée **probabilité conditionnelle sachant** B.

Théorème 9.

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. La probabilité conditionnelle $P_B : \mathcal{T} \to [0,1]$ sachant B est une probabilité sur l'espace probabilisable $(\Omega, \mathcal{T})$.

2. Formules conditionnelles

a. Formule des probabilités composées

Théorème 10. *Formule des probabilités composées*

i) Soit $A, B \in \mathcal{T}$ tel que $P(B) > 0$. On a :

$$P(A \cap B) = P(A|B)P(B).$$

ii) Soit $n \in \mathbb{N}^*$ et $A_1, \ldots, A_n \in \mathcal{T}$ tels que $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$. On a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \ldots \times P\left(A_n | \bigcap_{i=1}^{n-1} A_i\right).$$

b. Formule des probabilités totales

Théorème 11. *Formule des probabilités totales*

Soit $(B_i)_{i\in I}$ un système complet d'événements. Pour tout événement $A \in \mathcal{T}$, on a :

$$P(A) = \sum_{i\in I} P(A|B_i)P(B_i).$$

avec la convention $P(A|B_i)P(B_i) = 0$ *si* $P(B_i) = 0$.

Démonstration.

Comme $(B_i)_{i\in I}$ est un système complet d'événements, on a :

$$\bigcup_{i\in I} B_i = \Omega \quad \text{et} \quad B_i \cap B_j = \emptyset \text{ si } i \neq j.$$

Par suite, on a :

$$A = \bigcup_{i\in I} (A \cap B_i) \text{ avec } (A \cap B_i) \cap (A \cap B_j) = \emptyset \text{ si } i \neq j.$$

En effet,

$$\bigcup_{i\in I} (A \cap B_i) = \bigcup_{i\in I} A \cap \bigcup_{i\in I} B_i = A \cap \Omega = A;$$

et

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset.$$

Par suite, par les propriétés de la fonction de probabilité P, on a :

$$P(A) = P\left(\bigcup_{i\in I} (A \cap B_i)\right) = \sum_{i\in I} P(A \cap B_i) = \sum_{i\in I} P(A|B_i)P(B_i).$$

□

Exercice 10.

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$ et $P(\overline{B}) > 0$. Montrer que pour tout $A \in \mathcal{T}$, $P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$.

c. Formule de Bayes

Théorème 12. *Formule de Bayes*

Soit $A \in \mathcal{T}$ tel que $P(A) > 0$.

i) Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. On a :

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

ii) Soit $(B_i)_{i\in I}$ un système complet d'événements. On a, pour tout $j \in I$:

$$P(B_j) = \frac{P(A|B_j)P(B_j)}{\sum_{i\in I} P(A|B_i)P(B_i)}.$$

d. Exercice d'application des formules

Exercice 11.

Dans une population, on donne la probabilité définie, pour $n \in \mathbb{N}$ par p_n qu'une famille ait n enfants où, α étant un réel strictement positif :

$$p_n = \alpha \frac{2^n}{n!}.$$

On suppose qu'il est équiprobable qu'un enfant soit une fille ou un garçon.

1. Déterminer le nombre α.
2. Déterminer la probabilité qu'une famille ait au moins un garçon.
3. Déterminer la probabilité qu'une famille ait exactement 2 garçons et 3 filles.
4. On cherche à déterminer la probabilité qu'une famille ait exactement deux enfants sachant qu'elle a au moins deux filles.
   a) Déterminer la probabilité qu'une famille ait exactement deux filles.
   b) Déterminer la probabilité qu'une famille ait au moins deux filles.
   c) Conclure.

Correction.

1. Soit Ω l'ensemble des compositions (d'enfants) des familles de la population. La suite d'événements $(A_n)_{n\in\mathbb{N}}$ définie par :

   $$A_n = \text{"La famille possède exactement } n \text{ enfants"}.$$

   étant un système complet d'événements, on a :

   $$1 = P(\Omega) = P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} A_n = \sum_{n=0}^{+\infty} p_n.$$

   Par suite,

   $$1 = \sum_{n=0}^{+\infty} p_n = \alpha \sum_{n=0}^{+\infty} \frac{2^n}{n!} = \alpha e^2.$$

D'où $\alpha = e^{-2}$.

2. On note B l'événement :

B="La famille possède uniquement des filles (ou aucun enfant)".

Alors $\overline{B}$="La famille possède au moins un garçon" est l'événement dont la probabilité est recherchée.

On a, d'après la formule des probabilités totales :

$$P(B) = \sum_{n=0}^{+\infty} P(B|A_n)P(A_n) = \sum_{n=0}^{+\infty} \frac{1}{2^n} p_n = e^{-2}.e = \frac{1}{e}.$$

3. On note C l'événement recherché, D l'événement "la famille possède 2 garçons et 3 filles (au moins)" et on reprend la suite (A_n) de la question précédente. Alors on a $C = D \cap A_2$ donc :

$$P(C) = P(D \cap A_5) = P(D|A_5)P(A_5) = \left(\binom{5}{2} \frac{1}{2^2} \frac{1}{2^3}\right).e^{-2} \frac{2^5}{5!} = \frac{1}{12e^2}.$$

4. On note F l'événement "une famille a au moins deux filles" et on reprend la suite (A_n) de la question 1.

a) Avec les notations précédentes, ce qu'on doit calculer est :

$$P(F \cap A_2) = P(F|A_2)P(A_2) = \frac{1}{4}.e^{-2} \frac{4}{2} = \frac{1}{2e^2}.$$

b) D'après la formule des probabilités totales, on a :

$$P(F) = 1 - P(\overline{F}) = 1 - \sum_{n=0}^{+\infty} P(\overline{F}|A_n)P(A_n).$$

et on a :

$$P(\overline{F}|A_n) = \binom{n}{0} \frac{1}{2^0} \frac{1}{2^n} + \binom{n}{1} \frac{1}{2^1} \frac{1}{2^{n-1}} = (n+1)\frac{1}{2^n}.$$

Ainsi, on a :

$$\begin{aligned}
P(\overline{F}) &= \sum_{n=0}^{+\infty} P(\overline{F}|A_n)P(A_n) \\
&= \sum_{n=0}^{+\infty} (n+1) \frac{1}{2^n} p_n \\
&= e^{-2} \sum_{n=0}^{+\infty} \frac{n+1}{n!} \\
&= e^{-2} \left(\sum_{n=1}^{+\infty} \frac{n}{n!} + \sum_{n=0}^{+\infty} \frac{1}{n!}\right) \\
&= 2e^{-2} \sum_{n=0}^{+\infty} \frac{1}{n!} \\
&= \frac{2}{e}.
\end{aligned}$$

D'où $P(F) = 1 - \frac{2}{e}$.

c) On a :

$$P(A_2|F) = \frac{P(A_2 \cap F)}{P(F)} = \frac{1}{2e^2} \cdot \frac{1}{1 - \frac{2}{e}} = \frac{1}{2e(e-2)}.$$

3. Événements indépendants

a. Définition et premières propriétés

Définition 16. *Événements indépendants*

Soit $A, B \in \mathcal{T}$. On dit que A et B sont **indépendants** si

$$P(A \cap B) = P(A).P(B)$$

On justifie le choix de la terminologie d'"indépendance" par la proposition suivante :

Proposition 23.

Soit $A, B \in \mathcal{T}$ avec $P(B) > 0$. Les événements A et B sont indépendants si, et seulement si, $P(A|B) = P(A)$.

Proposition 24.

Soit $A, B \in \mathcal{T}$. Si A et B sont indépendants, alors

— $\overline{A}$ et $\overline{B}$ sont indépendants ;

— $\overline{A}$ et B sont indépendants (et A et $\overline{B}$ également).

b. Indépendance d'une famille d'événement

Définition 17. *Indépendance deux à deux/Indépendance*

Soit $(A_i)_{i \in I}$ une famille à valeurs dans $\mathcal{T}$.

- On dit que les événements A_i, pour $i \in I$ sont **deux à deux indépendants** si, pour tous $i, j \in \mathbb{I}$ avec $i \neq j$, on a :

$$P(A_i \cap A_j) = P(A_i)P(A_j).$$

- On dit que les événements A_i, pour $i \in I$ sont **indépendants** (ou *mutuellement indépendants*) si, pour toute partie finie J de I, on a :

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i).$$

Remarque 9.

Il est clair que l'indépendance d'une famille d'événements implique l'indépendance deux à deux de ces événements mais **la réciproque est fausse !**

En effet, si on considère deux lancers de dé et les trois événements suivants sont deux à deux indépendants, mais pas indépendants !

— A_1="La somme des deux lancers est pair" ;

— A_2="Le premier lancer est pair" ;

— A_3="Le second lancer est pair".

Partie E

Variables aléatoires discrètes

1. Variables aléatoires discrètes

a. Définition

Définition 18. *Variable aléatoire discrète*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$. On dit que la fonction X est une **variable aléatoire discrète** sur $(\Omega, \mathcal{T})$ si :

i) l'ensemble $X(\Omega)$ est au plus dénombrable ;

ii) pour tout $x \in X(\Omega)$, $X^{-1}(\{x\})$ est un événement de $\mathcal{T}$, i.e.

$$X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\} \in \mathcal{T}.$$

De plus, si $E \subset \mathbb{R}$, on dit que X est une variable aléatoire discrète **réelle**.

Proposition 25.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$ une variable aléatoire discrète. Pour tout $A \subset E$, $X^{-1}(A) \in \mathcal{T}$.

Pour des questions de lisibilité et de simplicité, on emploie les notations suivantes :

Notation 1.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$ une variable aléatoire discrète.

— Pour $A \subset E$, l'événement $X^{-1}(A)$ est noté $(X \in A)$.

— Pour $x \in E$, l'événement $X^{-1}(\{x\})$ est noté :

$$(X = x).$$

— Si $E \subset \mathbb{R}$, pour $x \in E$, les événements $X^{-1}(]-\infty, x[)$, $X^{-1}(]-\infty, x])$, $X^{-1}(]x, +\infty[)$ et $X^{-1}([x, +\infty[)$ sont notés respectivement :

$$(X < x), \qquad (X \leq x), \qquad (X > x), \qquad (X \geq x).$$

b. Loi de probabilité d'une variable aléatoire

Définition 19. *Loi de probabilité*

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. On appelle **loi de probabilité de** X l'application $P_X : \mathcal{P}(X(\Omega)) \to [0,1]$ définie par :

$$P_X : A \mapsto P(X \in A).$$

Proposition 26.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Le triplet $(X(\Omega), \mathcal{P}(X(\Omega)), P_X)$ est un espace probabilisé.

Proposition 27.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Alors la famille $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements. En particulier, on a :

$$\sum_{x \in X(\Omega)} P(X = x) = 1.$$

Remarque 10.

— D'après ce qui précède, si X une variable aléatoire *discrète* sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, la loi de probabilité P_X est entièrement déterminée par les événements $(X = x)$ pour $x \in X(\Omega)$.

Ainsi, pour déterminer la loi de probabilité P_X de la variable aléatoire X, (il faut et) il suffit de déterminer les valeurs de $P(X = x)$ pour chaque $x \in X(\Omega)$.

— Si X est une variable aléatoire discrète à valeurs dans un ensemble au plus dénombrable E, $X(\Omega)$ n'est pas forcément égal à E. Dans la suite, on étendra la loi P_X de X à $\mathcal{P}(E)$ tout entier, en posant, pour $x \in E \setminus X(\Omega)$, $P_X(\{x\}) := 0$ ou encore $P(X = x) = 0$. Ainsi, P_X définit une loi de probabilité sur $(E, \mathcal{P}(E))$.

Définition 20. *Loi de probabilité discrète*

Soit E un ensemble au plus dénombrable et $(p_x)_{x \in E}$ une famille de réels positifs telle que $\sum_{x \in E} p_x = 1$. On appelle **loi (de probabilité) discrète sur** E la probabilité sur $(E, \mathcal{P}(E))$ définie par la famille $(p_x)_{x \in E}$. On notera souvent $\mathcal{L}$ une telle loi.

Notation 2.

Soit E un ensemble au plus dénombrable, $\mathcal{L}$ une loi discrète sur E définie par une famille $(p_x)_{x \in E}$ et X, Y des variables aléatoires discrètes à valeurs dans E.

— On dit que X **suit la loi discrète** $\mathcal{L}$ et on note $X \sim \mathcal{L}$ si $P_X = \mathcal{L}$ i.e. si pour tout

$x \in E$,

$$P(X = x) = p_x.$$

— On dit que X et Y **suivent la même loi** ou encore que X et Y sont **équidistribuées** et on note $X \sim Y$ si $P_X = P_Y$ i.e. si pour tout $x \in E$,

$$P(X = x) = P(Y = y).$$

c. Fonction d'une variable aléatoire discrète

Proposition-Notation 28.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, F un ensemble et $f : X(\Omega) \to F$ une fonction. On note $f(X)$ la variable aléatoire discrète $f \circ X$.

Proposition 29.

Soit E, F des ensembles, X, Y des variables aléatoires discrètes à valeurs dans un ensemble E et $f : E \to F$ une fonction. Si $X \sim Y$, alors $f(X) \sim f(Y)$.

2. Lois usuelles sur un univers fini

a. Loi uniforme

Définition 21. *Loi uniforme*

Soit E un ensemble fini. On dit qu'une variable aléatoire X **suit une loi uniforme sur** E et on note $X \sim \mathcal{U}(E)$ si $X(\Omega) = E$ et, pour tout $x \in E$,

$$P(X = x) = \frac{1}{\#E}.$$

Si $E = [\![1, n]\!]$ pour un certain $n \in \mathbb{N}^*$, on notera simplement $\mathcal{U}(n)$ la loi uniforme sur $[\![1, n]\!]$.

b. Loi de Bernoulli

Définition 22. *Loi de Bernoulli*

Soit $p \in]0, 1[$. On dit qu'une variable aléatoire X **suit une loi de Bernoulli de paramètre** p et on note $X \sim \mathcal{B}(p)$ si $X(\Omega) = \{0, 1\}$ et

$$P(X = 1) = p, \qquad P(X = 0) = 1 - p.$$

c. Loi binomiale

Définition 23. *Loi binomiale*

Soit $n \in \mathbb{N}^*$ et $p \in]0,1[$. On dit qu'une variable aléatoire X **suit une loi binomiale de paramètres** n **et** p et on note $X \sim \mathcal{B}(n,p)$ si $X(\Omega) = [\![0,n]\!]$ et, pour tout $k \in [\![0,n]\!]$,

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

3. Lois usuelles sur un univers dénombrable

a. Loi géométrique

Définition 24. *Loi géométrique*

Soit $p \in]0,1[$. On dit qu'une variable aléatoire X **suit une loi géométrique de paramètre** p et on note $X \sim \mathcal{G}(p)$ si $X(\Omega) = \mathbb{N}^*$ et, pour tout $k \in \mathbb{N}^*$,

$$P(X=k) = p(1-p)^{k-1}.$$

Remarque 11.

Une variable aléatoire suivant une loi géométrique de paramètre p modélise le temps d'attente du premier succès d'une suite d'épreuves de Bernoulli indépendantes de paramètre p.

Exercice 12.

Soit $p \in]0,1[$ et X une variable aléatoire qui suit une loi géométrique de paramètre p. Déterminer la valeur de $P(X > k)$ pour $k \in \mathbb{N}^*$.

Correction.

On a $X \sim \mathcal{G}(p)$ donc $X(\Omega) = \mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$,

$$P(X=k) = p(1-p)^{k-1}.$$

On remarque que :

$$(X > k) = (X = k+1) \cup (X = k+2) \cup \dots = \bigcup_{n=k+1}^{+\infty} (X = n)$$

Les évènements $(X = n)$ étant incompatibles, on a alors :

$$P(X > k) = \sum_{n=k+1}^{+\infty} P(X=n) = \sum_{n=k+1}^{+\infty} p(1-p)^{n-1} = p(1-p)^k \underbrace{\sum_{n=0}^{+\infty} (1-p)^n}_{=\frac{1}{p}} = (1-p)^k$$

Proposition 30.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ et que $P(X=n)\neq 0$ pour tout $n\in\mathbb{N}^*$. Alors X suit une loi géométrique si, et seulement si, pour tous $k,l\in\mathbb{N}^*$,

$$P(X>k+l|X>k)=P(X>l).$$

Démonstration.

On remarque tout d'abord que, pour tous $k,l\in\mathbb{N}^*$, $(X>k+l)\subset(X>k)$ d'où $(X>k+l)\cap(X>k)=(X>k+l)$, et ainsi :

$$P(X>k+l|X>k)=\frac{P((X>k+l)\cap(X>k))}{P(X>k)}=\frac{P(X>k+l)}{P(X>k)}.$$

($\Rightarrow$) On suppose qu'il existe $p\in]0,1[$ tel que $X\sim\mathcal{G}(p)$. Alors, d'après l'exercice précedent, pour tous $k,l\in\mathbb{N}^*$:

$$P(X>k+l|X>k)=\frac{P(X>k+l)}{P(X>k)}=\frac{(1-p)^{k+l}}{(1-p)^k}=(1-p)^l=P(X>l).$$

($\Leftarrow$) On suppose que pour tous $k,l\in\mathbb{N}^*$,

$$\left(\frac{P(X>k+l)}{P(X>k)}=\right)P(X>k+l|X>k)=P(X>l).$$

On pose $p=1-P(X>1)\in]0,1[$ (car les $P(X=n)$ sont non nuls) et pour $n\in\mathbb{N}^*$, $u_n=P(X>n)>0$. Alors on a :

$$\frac{u_{n+1}}{u_n}=\frac{P(X>n+1)}{P(X>n)}\underbrace{=}_{k=n\text{ et }l=1}P(X>1)=1-p$$

Ainsi, $(u_n)_{n\in\mathbb{N}^*}$ est une suite géométrique de raison $1-p$ et donc, pour tout $n\in\mathbb{N}^*$:

$$P(X>n)=u_n=(1-p)^{n-1}u_1=(1-p)^{n-1}P(X>1)=(1-p)^n$$

On remarque que la formule précédente est toujours valable pour $n=0$ car X étant à valeurs dans $\mathbb{N}^*$, $P(X>0)=1$.
Par suite, comme, pour $k\in\mathbb{N}^*$,

$$(X=k)=(X\leq k)\cap(X>k-1)=\overline{(X>k)}\cap(X>k-1)=(X>k-1)\setminus(X>k)$$

et $(X>k)\subset(X>k-1)$, on a :

$$\begin{aligned}P(X=k)&=P(X>k-1)-P(X>k)\\&=(1-p)^{k-1}-(1-p)^k\\&=(1-p)^{k-1}(1-(1-p))\\P(X=k)&=p(1-p)^{k-1}\end{aligned}$$

Il en résulte que $X\sim\mathcal{G}(p)$.

□

b. Loi de Poisson

Définition 25. *Loi de Poisson*

Soit $\lambda \in \mathbb{R}_+^*$. On dit qu'une variable aléatoire X **suit une loi de Poisson de paramètre** λ et on note $X \sim \mathcal{P}(\lambda)$ si $X(\Omega) = \mathbb{N}$ et, pour tout $k \in \mathbb{N}$,

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Exercice 13.

Soit $\lambda \in \mathbb{R}_+^*$, $(p_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $]0,1[$ et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour chaque $n \in \mathbb{N}$, $X_n \sim \mathcal{B}(n, p_n)$.
Montrer que si $np_n \xrightarrow[n\to+\infty]{} \lambda$, alors, pour tout $k \in \mathbb{N}$,

$$P(X_n = k) \xrightarrow[n\to+\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Remarque 12.

Pour $n \in \mathbb{N}$ et $p \in]0,1[$ tels que $\lambda = np$ est proche de 0, une variable aléatoire qui suit une loi binomiale de paramètres n et p, suit approximativement une loi de Poisson de paramètre $\lambda = np$.

4. Couples de variables aléatoires

Dans ce paragraphe, $(\Omega, \mathcal{T}, P)$ désigne un espace probabilisé. Les variables aléatoires considérées sont définies sur cet espace.

a. Généralités

Définition 26. *Couple de variable aléatoires discrètes*

Soit X, Y des variables aléatoires discrètes sur un espace probabilisable. On appelle **couple de variables aléatoires discrètes sur** $(\Omega, \mathcal{T}, P)$ et on note (X,Y) l'application $\omega \mapsto (X(\omega), Y(\omega))$.
De plus, pour $(A,B) \subset X(\Omega) \times Y(\Omega)$, on note $(X \in A, Y \in B)$ l'évènement $((X,Y) \in A \times B) = (X \in A) \cap (Y \in B)$.

Remarque 13.

En particulier, pour $(x,y) \in X(\Omega) \times Y(\Omega)$, la notation est :

$$(X = x, Y = y) = ((X,Y) = (x,y)) = (X = x) \cap (Y = y).$$

Proposition 31.

Soit X, Y des variables aléatoires discrètes à valeurs dans E, F respectivement. Alors le couple de variables aléatoires (X, Y) est une variable aléatoire discrète à valeurs dans $E \times F$.

Proposition 32.

Soit (X, Y) un couple de variables aléatoires discrètes. Alors la famille d'événements :

$$((X = x) \cap (Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$$

est un système complet d'événements appelé **système complet d'événements associé au couple** (X, Y).

b. Loi conjointe

Définition 27. *Loi conjointe*

Soit (X, Y) un couple de variables aléatoires discrètes. On appelle **loi conjointe de** (X, Y) et on note $P_{(X,Y)}$ la loi de probabilité du couple (X, Y).

Proposition 33.

Soit (X, Y) un couple de variables aléatoires discrètes. La loi $P_{(X,Y)}$ est déterminée par la donnée des $P(X = x, Y = y)$ pour chaque couple $(x, y) \in X(\Omega) \times Y(\Omega)$.

c. Lois marginales

Définition 28. *Loi marginales*

Soit (X, Y) un couple de variables aléatoires discrètes. On appelle **première loi marginale du couple** la loi de probabilité de X et **deuxième loi marginale du couple** la loi de probabilité de Y.

Théorème 13.

Soit (X, Y) un couple de variables aléatoires discrètes. On a :

— pour tout $x \in X(\Omega)$, $P(X = x) = \displaystyle\sum_{y \in Y(\Omega)} P(X = x, Y = y)$.

— pour tout $y \in Y(\Omega)$, $P(Y = y) = \displaystyle\sum_{x \in X(\Omega)} P(X = x, Y = y)$.

d. Lois conditionnelles

Définition 29. *loi conditionnelle*

Soit (X, Y) un couple de variables aléatoires discrètes et $x \in X(\Omega)$ tel que $P(X = x) \neq 0$. On appelle **loi conditionnelle de Y sachant $X = x$** la probabilité notée $P_{(X=x)}$ sur $(\Omega, \mathcal{T})$ déterminée, pour $y \in Y(\Omega)$, par :

$$P_{(X=x)}(Y = y) = P(Y = y|X = x).$$

Exercice 14.

Soit (X, Y) un couple de variables aléatoires discrètes tel que, pour tout $x \in X(\Omega)$, $P(X = x) \neq 0$.

1. Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. Montrer que
$$P((X = x) \cap (Y = y)) = P(X = x)P(Y = y|X = x).$$
2. Soit $y \in Y(\Omega)$. Montrer que
$$P(Y = y) = \sum_{x \in X(\Omega)} P(X = x)P(Y = y|X = x).$$

Remarque 14.

On généralise toutes les notions précédentes aux cas d'un n-uplet de variables aléatoires.

5. Variables aléatoires indépendantes

a. Couple de variables aléatoires indépendantes

Définition 30. *Couple de variables aléatoires indépendantes*

Soit (X, Y) un couple de variables aléatoires discrètes. On dit que X et Y sont **indépendantes** et on note $X \perp\!\!\!\perp Y$ si, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

Proposition 34.

Soit (X, Y) un couple de variables aléatoires discrètes. Les variables X et Y sont indépendantes si, et seulement si, pour tout $(A, B) \subset X(\Omega) \times Y(\Omega)$,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Proposition 35.

Soit (X, Y) un couple de variables aléatoires discrètes et f, g des fonctions définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement. Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.

Proposition 36. *Loi d'une somme de variables aléatoires discrètes*

Soit (X, Y) un couple de variables aléatoires discrètes à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors, pour tout $n \in \mathbb{N}$,

$$P(X + Y = n) = \sum_{k=0}^{n} P(X = k)P(Y = n - k).$$

Exercice 15.

Soit $\lambda, \mu \in \mathbb{R}_+^*$ et (X, Y) un couple de variables aléatoires discrètes tel que $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$. Montrer que si X et Y sont indépendantes, alors $X + Y \sim \mathcal{P}(\lambda + \mu)$.

b. Famille finie de variables aléatoires indépendantes

Définition 31. *Variables aléatoires indépendantes*

Soit $n \in \mathbb{N}^*$ et $(X_k)_{k \in [\![1,n]\!]}$ une famille de variables aléatoires discrètes. On dit que les variables $X_1, \dots, X_n$ sont **indépendantes** (ou mutuellement indépendantes) si, pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$:

$$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{k=1}^{n} P(X_k = x_k).$$

Proposition 37.

Soit $n \in \mathbb{N}^*$ et $(X_k)_{k \in [\![1,n]\!]}$ une famille de variables aléatoires discrètes. Les variables $X_1, \dots, X_n$ sont indépendantes si, et seulement si, pour tout $(A_1, \dots A_n) \subset X_1(\Omega) \times \dots \times X_n(\Omega)$,

$$P(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{k=1}^{n} P(X_k \in A_k).$$

Proposition 38. *Lemme des coalitions*

Soit $n, m \in \mathbb{N}^*$ avec $m < n$, $(X_k)_{k \in [\![1,n]\!]}$ une famille de variables aléatoires discrètes et f, g des fonctions définies sur $X_1(\Omega) \times \dots \times X_m(\Omega)$ et $X_{m+1}(\Omega) \times \dots \times X_n(\Omega)$ respectivement. Si $X_1, \dots, X_n$ sont indépendantes, alors $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

Exercice 16.

Soit $m \in \mathbb{N}^*$, $n_1, ..., n_m \in \mathbb{N}$, $p \in [0,1]$ et $X_1, ..., X_m$ des variables aléatoires telles que, pour $k \in [\![1, m]\!]$, $X_k \sim \mathcal{B}(n_k, p)$.
On suppose $X_1, ..., X_n$ indépendantes. Déterminer la loi de $Y = X_1 + \ldots + X_n$.

c. Famille infinie de variables aléatoires indépendantes

Définition 32.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires discrètes. On dit que les variables X_n sont **indépendantes** si, pour toute partie finie $I = \{i_1, ..., i_k\}$ de $\mathbb{N}$, les variables $X_{i_1}, \ldots, X_{i_k}$ sont indépendantes.

Le théorème suivant (dont la démonstration est admise) va nous permettre de pouvoir "parler sans crainte" de suites de variables aléatoires indépendantes suivant chacune une loi prescrite :

Théorème 14. *Théorème d'existence de Kolmogorov*

Si $(\mathcal{L}_n)_{n\in\mathbb{N}}$ est une suite de lois discrètes, alors il existe un espace probabilisé et une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires discrètes indépendantes sur cet espace telle que, pour tout $n \in \mathbb{N}$, $X_n \sim \mathcal{L}_n$.

Exemple 8.

Soit $p \in [0,1]$ et $(X_n)_{n\in\mathbb{N}^*}$ une suites de variables aléatoires indépendantes suivant chacune une loi de Bernouilli de paramètre p - *le théorème précédent (et il n'est là que pour cela) garantit que cette suite est bien définie !*
La variable $T = \min\{n \in \mathbb{N}^* \cup \{\infty\} \mid X_n = 1\}$ suit une loi géométrique de paramètre p.

Partie F

Espérance, variance

Sauf mention contraire, les variables aléatoires discrètes considérées dans cette partie sont définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ et sont à valeurs réelles.

1. Espérance

a. Définitions et exemples

Définition 33. *Espérance*

Soit X une variable aléatoire discrète .

- Si X est à valeurs dans $\mathbb{R}_+$, on appelle **espérance de** X et on note $E(X)$ la quantité (potentiellement infinie) :
$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$
- Si X est à valeurs dans $\mathbb{R}$, on dit que X **admet une espérance finie** et on note $X \in L^1$, si la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas, on appelle **espérance de** X et on note $E(X)$ la quantité :
$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

Exemple 9.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $X \in L^1$ et $E(X) = \dfrac{1}{p}$, Si $X \sim \mathcal{G}(p)$, alors pour tout $k \in \mathbb{N}^*$, $P(X = k) = p(1-p)^{k-1}$. D'après le critère de D'Alembert, $u_k = kp(1-p)^{k-1}$ est le terme général d'une série convergente car $0 < 1-p < 1$. Par suite, $X \in L^1$ et

$$E(X) = \sum_{k=1}^{+\infty} kp(1-p)^{k-1}.$$

Calculons la valeur de cette somme. La série entière $\sum_{k\geq 0} x^k$ a pour rayon de convergence 1, donc sa somme $f : x \mapsto \frac{1}{1-x}$ est C^1 sur $]-1,1[$ et on a, pour tout $x \in]-1,1[$:

$$\begin{aligned} \frac{1}{(1-x)^2} = f'(x) &= \frac{d}{dx}\sum_{k=0}^{+\infty} x^k \\ &= \sum_{k=0}^{+\infty} \frac{d}{dx} x^k \\ &= \sum_{k=0}^{+\infty} kx^{k-1} \end{aligned}$$

Par suite, on a, comme $0 < 1-p < 1$:

$$\begin{aligned} E(X) &= \sum_{k=1}^{+\infty} kp(1-p)^{k-1} \\ &= p\sum_{k=0}^{+\infty} k(1-p)^{k-1} \\ &= pf'(1-p) = \frac{p}{(1-(1-p))^2} \\ E(X) &= \frac{1}{p} \end{aligned}$$

— Si $X \sim \mathcal{P}(\lambda)$ alors $X \in L^1$ et $E(X) = \lambda$.

Si $X \sim \mathcal{P}(\lambda)$, alors pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$. D'après le critère de D'Alembert, $u_k = k\dfrac{\lambda^k}{k!}$ est le terme général d'une série convergente. Par suite, $X \in L^1$ et

$$E(X) = \sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!}.$$

Calculons la valeur de cette somme. On a :

$$\begin{aligned}
E(X) &= \sum_{k=0}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!} \\
&= p\sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!} \\
&= \lambda e^{-\lambda}\textstyle\sum_{k=1}^{+\infty} \dfrac{\lambda^{k-1}}{(k-1)!} \\
&= \lambda e^{-\lambda}\textstyle\sum_{k=0}^{+\infty} \dfrac{\lambda^{k}}{(k)!} \\
&= \lambda e^{-\lambda}e^{\lambda} \\
E(X) &= \lambda
\end{aligned}$$

Exercice 17.

1. Soit $c \in \mathbb{R}$ et X une variable aléatoire discrète presque sûrement égale à c. Calculer l'espérance de X.
2. Déterminer les espérances des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Exercice 18.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ telle que, pour $n \in \mathbb{N}^*$, $P(X = n) = \dfrac{\alpha}{n^2}$ pour un certain $\alpha > 0$.

1. Déterminer α.
2. Quelle est l'espérance de X ?

b. Propriétés de l'espérance

Théorème 15. *Théorème de transfert*

Soit X une variable aléatoire discrète et $f : X(\Omega) \to \mathbb{R}$ une fonction. La variable aléatoire $f(X)$ est d'espérance finie si, et seulement si, la famille $(f(x)P(X = x))_{x\in X(\Omega)}$ est sommable. Dans ce cas, on a :

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$

Théorème 16. *Linéarité de l'espérance*

Soit X, Y des variables aléatoires discrètes et $\lambda, \mu \in \mathbb{R}$. Si $X, Y \in L^1$, alors la variable aléatoire discrète $\lambda X + \mu Y \in L^1$ et on a :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Proposition 39. *Positivité/Croissance*

Soit X, Y des variables aléatoires discrètes.

— Si X est à valeurs positives, alors $E(X) \geq 0$ et $E(X) = 0$ implique $X = 0$ presque sûrement.

— Si $X, Y \in L^1$ et $X \leq Y$, alors $E(X) \leq E(Y)$.

Proposition 40.

Soit X une variable aléatoire discrète. Alors $X \in L^1$ si, et seulement si, $|X| \in L^1$. Dans ce cas on a $|E(X)| \leq E(|X|)$.

Théorème 17. *Espérance d'un produit de variables indépendantes*

Soit X, Y des variables aléatoires discrètes. Si X et Y sont **indépendantes** et d'espérance finie, alors la variable aléatoire XY est d'espérance finie et on a :

$$E(XY) = E(X)E(Y).$$

Théorème 18. *Inégalité de Markov*

Soit X une variable aléatoire discrète à valeurs positives. Si $X \in L^1$, alors, pour tout $a > 0$, on a :

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Démonstration.

On suppose $X \in L^1$. Soit $a > 0$. On pose $Y = \mathbb{1}_{(X \geq a)}$ i.e. pour $\omega \in \Omega$,

$$Y(\omega) = \begin{cases} 1 & \text{si } X(\omega) \geq a \\ 0 & \text{si } X(\omega) < a \end{cases}$$

Alors Y est une variable aléatoire à valeurs dans $\{0, 1\}$ et donc d'espérance finie ; et on a :

$$E(X) = 0 \times P(Y = 0) + 1 \times P(Y = 1) = P(Y = 1) = P(X \geq a).$$

De plus, on remarque que $X \geq aY$; en effet, pour tout $\omega \in \Omega$:
- si $X(\omega) \geq a$, alors $Y(\omega) = 1$, et donc $X(\omega) \geq aY(\omega)$;
- si $X(\omega) < a$, alors $Y(\omega) = 0$ et comme X est à valeurs positives, $X(\omega) \geq 0 = aY(\omega)$.

Par croissance de l'espérance, on a donc $E(aY) \leq E(X)$ et donc

$$aP(X \geq a) = aE(Y) = E(aY) \leq E(X)$$

D'où le résultat en divisant par $a > 0$. □

2. Variance

a. Moments

Définition 34. *Moments*

Soit X une variable aléatoire discrète et $r \in \mathbb{N}$. On dit que X **admet un moment d'ordre** r et on note :

$$X \in L^r.$$

si la variable aléatoire discrète X^r admet une espérance finie. Dans ce cas, on appelle **moment d'ordre** r la quantité $E(X^r)$.

Proposition 41.

Soit X une variable aléatoire discrète. Si $X \in L^2$, alors $X \in L^1$.

Théorème 19. *Inégalité de Cauchy-Schwarz*

Soit X, Y des variables aléatoires discrètes. Si $X, Y \in L^2$ alors $XY \in L^1$ et on a :

$$E(XY)^2 \leq E(X^2)E(Y^2).$$

b. Variance, écart-type

Définition 35. *Variance*

Soit X une variable aléatoire discrète telle que $X \in L^2$. On appelle **variance de** X et on note $V(X)$ la quantité :

$$V(X) = E\left((X - E(X))^2\right).$$

Théorème 20. *Formule de Kœnig-Huygens*

Soit X une variable aléatoire discrète. Si $X \in L^2$, alors on a :

$$V(X) = E(X^2) - E(X)^2.$$

Exemple 10.

Soit X une variable aléatoire discrète.

- Si $X \sim \mathcal{G}(p)$ alors $X \in L^2$ et $V(X) = \dfrac{1-p}{p^2}$,
- Si $X \sim \mathcal{P}(\lambda)$ alors $X \in L^2$ et $V(X) = \lambda$.

On suppose que $X \sim \mathcal{P}(\lambda)$. Alors

- $X(\Omega) = \mathbb{N}$
- pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$.

$X \in L^2$ si X^2 est d'espérance finie, ce qui est équivalent, d'après le théorème de transfert, à : la famille $(k^2 P(X = k))_{k\in\mathbb{N}}$ est sommable. Comme cette famille est une suite à termes positifs, il suffit alors de montrer que c'est le terme général d'une série convergente.
Or, pour tout $k \geq 1$, $u_k = k^2 P(X = k) = k^2 e^{-\lambda}\frac{\lambda^k}{k!} > 0$ et on a :

$$\frac{u_{k+1}}{u_k} = \frac{\lambda(k+1)}{k^2} \xrightarrow[k\to+\infty]{} 0 < 1$$

donc, d'après la régle de D'Alembert, $\sum u_k$ converge.
Par suite, $(k^2 P(X = k))_{k\in\mathbb{N}}$ est sommable et donc $X \in L^2$.
Ainsi, X admet une variance finie et également une espérance finie. Calculons tout d'abord $E(X(X-1))$ (la variable $X(X-1)$ est bien d'espérance finie car X^2 et X sont d'espérance finie et on a $X(X-1) = X^2 - X$) :

$$\begin{aligned}
E(X(X-1)) &= \sum_{k=0}^{+\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \lambda^2 e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!} \\
&= \lambda^2 e^{-\lambda}\underbrace{\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!}}_{=e^{\lambda}} \\
E(X(X-1)) &= \lambda^2.
\end{aligned}$$

On a, par linéarité : $E(X(X-1)) = E(X^2) - E(X)$ et d'après le théorème de Kœnig-Huygens, $V(X) = E(X^2) - E(X)^2$; par suite :

$$V(X) = E(X(X-1)) + E(X) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Exercice 19.

Déterminer les variances des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Proposition 42.

Soit $\lambda, \mu \in \mathbb{R}$ et X une variable aléatoire discrète. Si $X \in L^2$, alors on a :

$$V(\lambda X + \mu) = \lambda^2 V(X).$$

Définition 36. *Écart-type*

Soit X une variable aléatoire discrète telle que $X \in L^2$. On appelle **écart-type de** X et on note $\sigma(X)$ la quantité :

$$\sigma(X) = \sqrt{V(X)}.$$

Définition 37. *Variable centrée réduite*

Soit X une variable aléatoire discrète telle que $X \in L^2$. On dit que la variable aléatoire X est **centrée réduite** si $E(X) = 0$ et $\sigma(X) = 1$.

Proposition 43.

Soit X une variable aléatoire discrète telle que $X \in L^2$ et $V(X) \neq 0$. Alors la variable aléatoire $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est une variable aléatoire centrée réduite.

Théorème 21. *Inégalité de Bienaymé-Tchebychev*

Soit X une variable aléatoire discrète. Si $X \in L^2$, alors on a, pour tout $\varepsilon > 0$:

$$P\left(|X - E(X)| \geq \varepsilon\right) \leq \frac{V(X)}{\varepsilon^2}.$$

3. Covariance

a. Généralités

Définition 38. *Covariance*

Soit (X, Y) un couple de variables aléatoires discrètes tel que $X, Y \in L^2$. On appelle **covariance du couple** (X, Y) et on note $\mathrm{Cov}(X, Y)$ la quantité :

$$\mathrm{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right).$$

Théorème 22. *Formule de Kœnig-Huygens*

Soit (X, Y) un couple de variables aléatoires discrètes. Si $X, Y \in L^2$, alors on a :

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Proposition 44.

Soit (X, Y) un couple de variables aléatoires discrètes tel que $X, Y \in L^2$. Si X et Y sont indépendantes, alors $\mathrm{Cov}(X, Y) = 0$.

b. Variance d'une somme

Théorème 23.

Soit (X, Y) deux variables aléatoires discrètes. Si $X, Y \in L^2$, alors :

$$V(X + Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y).$$

Soit $n \in \mathbb{N}^*$. On a, plus généralement, si $(X_1, ..., X_n)$ est un n-uplet de variables aléatoires discrètes, on a :

$$V(X_1 + \ldots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leq i < j \leq n} \mathrm{Cov}(X_i, X_j).$$

Corollaire 5.

Soit (X, Y) deux variables aléatoires discrètes telles que $X, Y \in L^2$. Si X et Y sont indépendantes, on a :

$$V(X + Y) = V(X) + V(Y).$$

4. Loi faible des grands nombres

Théorème 24. *Loi faible des grands nombres*

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires discrètes, deux à deux indépendantes, de même loi et admettant un moment d'ordre 2. On pose $m = E(X_1)$ et, pour $n \in \mathbb{N}^*$, $S_n = X_1 + \ldots + X_n$. Alors on a, pour tout $\varepsilon > 0$:

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow[n \to +\infty]{} 0.$$

5. Fonctions génératrices

a. Définition et premières propriétés

Définition 39. *Fonction génératrice*

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. On appelle **fonction génératrice de**

X et on note G_X la fonction :
$$G_X : t \mapsto E(t^X).$$

Théorème 25.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. La fonction génératrice G_X de X est définie sur $[-1,1]$ et on a, pour tout $t \in [-1,1]$:

$$G_X(t) = \sum_{n=0}^{+\infty} t^n P(X=n).$$

Démonstration.

Pour $t \in [-1,1]$, on considère la famille $(t^n P(X=n))_{n\in\mathbb{N}}$. On remarque que pour tout $t \in [-1,1]$ et tout $n \in \mathbb{N}$
$$|t^n P(X=n)| = |t|^n P(X=n) \leq P(X=n)$$
Or $P(X=n)$ est le terme général d'une série convergente (dont la somme vaut 1), donc par comparaison, $\sum t^n P(X=n)$ converge absolument.
Par suite, $(t^n P(X=n))_{n\in\mathbb{N}}$ est une famille sommable, et ainsi, d'après le théorème de transfert, la variable aléatoire t^X est d'espérance finie. Alors, pour $t \in [-1,1]$ $G_X(t) = E(t^X)$ est bien défini et on a :
$$G_X(t) = E(t^X) = \sum_{n=0}^{+\infty} t^n P(X=n)$$

$\square$

Proposition 45.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. La fonction génératrice de X est de classe C^∞ sur $]-1,1[$ et on a, pour tout $k \in \mathbb{N}$,

$$G_X^{(k)}(0) = k!P(X=k).$$

Corollaire 6.

Soit X, Y des variables aléatoires discrètes à valeurs dans $\mathbb{N}$.
- La loi de probabilité P_X de X est entièrement déterminée par sa fonction génératrice G_X.
- les variables X et Y ont la même loi si, et seulement si, $G_X = G_Y$.

Proposition 46.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice. On a :

— La variable X appartient à L^1 si, et seulement si, G est dérivable en 1. Dans ce cas, on a :
$$G'_X(1) = E(X).$$

— La variable X appartient à L^2 si, et seulement si, G est deux fois dérivable en 1. Dans ce cas, on a :
$$G''_X(1) = E(X(X-1)).$$

Exercice 20.

Soit $X \in L^2$ une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice. Exprimer $E(X^2)$ et $V(X)$ en fonction des dérivées de G_X en 1.

Exemple 11.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$, On suppose $X \sim \mathcal{G}(p)$ avec $p \in]0,1[$. Alors :

$$\begin{cases} X(\Omega) = \mathbb{N}^* \\ P(X = k) = p(1-p)^{k-1} \text{ pour tout } k \in \mathbb{N}^* \end{cases}$$

D'après le théorème 25, G_x est définie sur $[-1,1]$ mais on remarque, en posant $a_n = P(X = n) > 0$ pour $n \geq 1$, que $\sum a_n t^n$ est de rayon de convergence $\frac{1}{1-p}$. En effet, comme $|\frac{a_{n+1}}{a_n}| = (1-p) \xrightarrow[n \to +\infty]{} (1-p)$, d'après la règle de D'Alembert pour les séries entières, $R = \frac{1}{1-p}$.
Ainsi, si $X \sim \mathcal{G}(p)$, G_X est définie sur $]-\frac{1}{1-p}, \frac{1}{1-p}[$ (intervalle plus gros que $[-1,1]$ car $1-p<1$). De plus, pour tout $t \in]-\frac{1}{1-p}, \frac{1}{1-p}[$, on a :

$$\begin{aligned} G_X(t) &= \sum_{n=1}^{+\infty} t^n P(X = n) \\ &= \sum_{n=1}^{+\infty} t^n p(1-p)^{n-1} \\ &= pt \sum_{n=1}^{+\infty} (t(1-p))^{n-1} \\ &= pt \sum_{n=0}^{+\infty} ((1-p)t)^n \text{ et } t(1-p) \in]-1,1[\\ &= pt \frac{1}{1-(1-p)t} \\ G_X(t) &= \frac{pt}{1-(1-p)t} \end{aligned}$$

— Si $X \sim \mathcal{P}(\lambda)$ alors $G_X(t) = e^{\lambda(t-1)}$.

Exercice 21.

Déterminer les fonctions génératrices des lois uniforme sur $[\![1,n]\!]$, de Bernoulli et binomiale.

Théorème 26.

Soit X, Y des variables aléatoires discrètes à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors, pour tout $t \in [-1,1]$:

$$G_{X+Y}(t) = G_X(t)G_Y(t).$$

Démonstration.

On suppose X et Y indépendantes. Soit $t \in [-1,1]$. On note $Z = X+Y$; alors Z est une variable aléatoire à valeurs dans $\mathbb{N}$ et on a :

$$G_Z(t) = E(t^Z) = \sum_{n=0}^{+\infty} t^n P(Z=n) = \sum_{n=0}^{+\infty} t^n P(X+Y=n)$$

Or, X et Y étant indépendantes, on a, pour tout $n \in \mathbb{N}$:

$$P(X+Y=n) = P\left(\bigcup_{k=0}^{n} ((X=k)\cap(y=n-k))\right) = \sum_{k=0}^{n} P(X=k)P(Y=n-k).$$

Ainsi :

$$G_Z(t) = \sum_{n=0}^{+\infty} t^n \sum_{k=0}^{n} P(X=k)P(Y=n-k) = \sum_{n=0}^{+\infty}\sum_{k=0}^{n} \underbrace{t^k P(X=k)}_{=a_k}\underbrace{t^{n-k}P(Y=n-k)}_{=b_{n-k}}.$$

où $a_n = t^n P(X=n)$ et $b_n = t^n P(Y=n)$ pour $n \in \mathbb{N}$.
On remarque que $G_X(t) = \sum_{n=0}^{+\infty} a_n$ et $G_Y(t) = \sum_{n=0}^{+\infty} b_n$ et que $\sum a_n, \sum b_n$ convergent absolument (car pour tout $n \in \mathbb{N}$, $|a_n| \leq P(X=n)$). Ainsi, le produit de Cauchy $\sum c_n$ de ces séries converge où $c_n = \sum_{k=0}^{n} a_n b_n$ et on a l'égalité :

$$G_{X+Y}(t) = G_Z(t) = \sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right) = G_X(t)G_Y(t).$$

□

Remarque 15.

Plus génralement, si $X_1, ..., X_n$ sont des variables aléatoires discrètes indépendantes, alors $G_{X_1+...+X_n} = G_{X_1} \times \ldots \times G_{X_n}$.

Exercice 22.

Soit $k, m, n \in \mathbb{N}^*$ tels que $k = nm$ et X, Y des variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes. On pose $Z = X+Y$ et on suppose que $X \sim \mathcal{U}([\![0, n-1]\!])$ et $Z \sim \mathcal{U}([\![0, k-1]\!])$.

1. Déterminer la fonction génératrice de Y.
2. En déduire la loi de Y.

Chapitre XII

Équations différentielles linéaires

Table des matières

Dans ce chapitre, le corps $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$ et I désigne un intervalle non vide de $\mathbb{R}$.

Partie A

Équations différentielles linéaires d'ordre 1

Dans cette partie E désigne un espace vectoriel sur $\mathbb{K}$ de dimension finie.

1. Définitions

Définition 1. *Équation différentielle linéaire d'ordre 1*

Soit $a : I \to \mathcal{L}(E)$ et $b : I \to E$ des applications continues.

- On appelle **équation différentielle linéaire d'ordre** 1 une équation (E) de la forme :
$$x' = a(t)(x) + b(t). \qquad (E)$$
où l'inconnue x est une fonction dérivable de I dans E.
- Une fonction $f : I \to E$ est une **solution** de E si, pour tout $t \in I$,
$$f'(t) = a(t)(f(t)) + b(t).$$
- On appelle **équation homogène** associée à (E), l'équation différentielle (E_h) linéaire d'ordre 1 :
$$x' = a(t)(x). \qquad (E_h)$$

Définition 2. *Traduction matricielle*

Soit $n \in \mathbb{N}^*$, $A : I \to M_n(\mathbb{K})$ et $B : I \to M_{n,1}(\mathbb{K})$.
On appelle **système différentiel linéaire d'ordre** 1 une équation différentielle (S) linéaire d'ordre 1 de la forme :
$$X' = A(t)X + B(t). \qquad (S)$$
où l'inconnue X est une fonction dérivable de I dans $M_{n,1}(\mathbb{K})$.

Question 1.

Justifier la terminologie "système linéaire" employée dans la définition précédente.

Si, pour $t \in I$, on note $A(t) = (a_{ij}(t))_{\substack{1\leq i\leq n\\1\leq j\leq n}}$, $B(t) = (b_i(t))_{1\leq j\leq n}$ et $X(t) = (x_i(t))_{1\leq j\leq n}$, l'équation (S) est équivalente au système :

$$(S)\left\{\begin{array}{ccccccccc} x_1' & = & a_{11}(t)x_1 & + & \ldots & + & a_{1n}(t)x_n & + & b_1(t) \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ x_1' & = & a_{n1}(t)x_1 & + & \ldots & + & a_{nn}(t)x_n & + & b_n(t) \end{array}\right.$$

Remarque 1. *Traduction matricielle*

Ainsi, étant donnée une base $\mathcal{B}$ de E, une équation différentielle $x' = a(t)(x) + b(t)$ linéaire d'ordre 1 est équivalente au système linéaire $X' = A(t)X + B(t)$ où :

- $A(t) = \mathrm{Mat}_{\mathcal{B}}(a(t))$ et $B(t) = \mathrm{Mat}_{\mathcal{B}}(B(t))$;
- l'inconnue $X(t)$ est obtenue de même par $X(t) = \mathrm{Mat}_{\mathcal{B}}(x(t))$.

2. Structure de l'ensemble des solutions

Proposition 1.

Soit $(E) : x' = a(t)(x) + b(t)$ une équation différentielle linéaire d'ordre 1. L'application $x \mapsto \{x'(t) - a(t)(x(t))\}$ est une application linéaire de $C^1(I, E)$ dans $C(I, E)$.

Proposition 2.

Soit (E) une équation différentielle linéaire d'ordre 1.

L'ensemble $\mathcal{S}_h$ des solutions de l'équation homogène (E_h) est un sous-espace vectoriel de $C^1(I, E)$.

Proposition 3.

Soit (E) une équation différentielle linéaire d'ordre 1 et $\mathcal{S}_h$ l'ensemble des solutions de (E_h).

Soit f_p une solution de (E). Alors l'ensemble $\mathcal{S}$ des solutions de (E) s'exprime :

$$\mathcal{S} = f_p + \mathcal{S}_h = \{f_p + f_h \mid f_h \in \mathcal{S}_h\}.$$

Proposition 4. *Superposition*

Soit $(E) : x' = a(t)(x) + b(t)$ une équation différentielle linéaire d'ordre 1.
On suppose que $b = \lambda_1 b_1 + \mu_1 b_2$ avec $\lambda_1, \lambda_2 \in \mathbb{K}$ et $b_1, b_2 \in C(I, E)$.
Si f_1, f_2 sont des solutions respectives de :

$$(E_1) : x' = a(t)(x) + b_1(t) \quad \text{et} \quad (E_2) : x' = a(t)(x) + b_2(t)$$

alors $f = \lambda_1 f_1 + \lambda_2 f_2$ est solution de (E).

3. Problème de Cauchy

Définition 3. *Problème de Cauchy*

Soit $(E) : x' = a(t)(x) + b(t)$ une et t_0. On appelle **problème de Cauchy** la donnée d'une équation différentielle $(E) : x' = a(t)(x) + b(t)$ linéaire d'ordre 1 et d'une **condition initiale** $x(t_0) = x_0$ où $t_0 \in I$ et $x_0 \in E$.
Autrement dit, un problème de Cauchy est un système d'inconnue $x : I \to E$ de la forme :

$$\begin{cases} (E) : & x' = a(t)(x) + b(t) \\ (\text{C.I.}) & x(t_0) = x_0 \end{cases}$$

Remarque 2.

Résoudre un problème de Cauchy revient donc à déterminer toutes les solutions f de $(E) : x' = a(t)(x) + b(t)$ qui vérifient $f(t_0) = x_0$.

Théorème 1. *Théorème de Cauchy linéaire*

Soit $(E) : x' = a(t)(x) + b(t)$ une équation différentielle linéaire d'ordre 1.
Pour tout $(t_0, x_0) \in I \times E$, il existe **une unique solution** f au problème de Cauchy :

$$\begin{cases} (E) : & x' = a(t)(x) + b(t) \\ (\text{C.I.}) & x(t_0) = x_0 \end{cases}$$

Lemme 1.

Soit $(E) : x' = a(t)(x)+b(t)$ une équation différentielle linéaire d'ordre 1, $t_0 \in I$ et $\mathcal{S}_h$ l'ensemble des solutions de (E_h).
L'application

$$\text{eval}_{t_0} : \left| \begin{array}{ccc} \mathcal{S}_h & \to & E \\ f & \mapsto & f(t_0) \end{array} \right.$$

est un isomorphisme d'espaces vectoriels.

Théorème 2.

Soit $(E) : x' = a(t)(x) + b(t)$ une équation différentielle linéaire d'ordre 1 et $\mathcal{S}_h$ l'ensemble des solutions de (E_h). L'espace vectoriel $\mathcal{S}_h$ est de dimension finie et

$$\dim(\mathcal{S}_h) = \dim(E).$$

Partie B

Résolution d'une équation différentielle linéaire d'ordre 1 à coefficients constants

1. Exponentielle d'un endomorphisme, d'une matrice

Dans ce paragraphe, n est un entier naturel non nul et E est un espace vectoriel de dimension finie n sur le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

a. Définitions

Lemme 2.

Soit $a \in \mathcal{L}(E)$ et $A \in M_n(\mathbb{K})$. Les séries $\sum \frac{a^n}{n!}$ et $\sum \frac{A^n}{n!}$ sont convergentes dans, respectivement, $\mathcal{L}(E)$ et $M_n(\mathbb{K})$.

Définition 4.

Soit $a \in \mathcal{L}(E)$ et $A \in M_n(\mathbb{K})$.

— On appelle **exponentielle de l'endomorphisme** a et on note $\exp(a)$ ou encore e^a, l'endomorphisme de E :

$$\exp(a) = \sum_{n=0}^{+\infty} \frac{a^n}{n!}.$$

— On appelle **exponentielle de la matrice** A et on note $\exp(A)$ ou encore e^A, la matrice de $M_n(\mathbb{K})$:

$$\exp(A) = \sum_{n=0}^{+\infty} \frac{A^n}{n!}.$$

Proposition 5.

Soit $a \in \mathcal{L}(E)$ et $\mathcal{B}$ une base de E. On a :

$$\mathrm{Mat}_{\mathcal{B}}\left(\exp(a)\right) = \exp\left(\mathrm{Mat}_{\mathcal{B}}(a)\right)$$

b. Dérivation de $t \mapsto \exp(tA)$

Théorème 3.

Soit $A \in M_n(\mathbb{K})$. L'application $\varphi : \mathbb{R} \to M_n(\mathbb{K})$ telle que :

$$\varphi : t \mapsto \exp(tA),$$

est dérivable sur $\mathbb{R}$ et pour tout $t \in \mathbb{R}$,

$$\varphi'(t) = A.\exp(tA) = A.\varphi(t).$$

Démonstration.

Pour $n \in \mathbb{N}$, on pose $\varphi_n : t \mapsto \frac{(tA)^n}{n!}$. Vérifions les hypothèses du théorème d'interversion dérivation/somme :

- Pour tout $n \in \mathbb{N}$, φ_n est de classe C^1 sur $\mathbb{R}$ et pour tout $t \in \mathbb{R}$:

$$\varphi_n'(t) = \frac{nt^{n-1}A^n}{n!}.$$

— D'après le lemme 2, $\sum \varphi_n$ converge simplement sur $\mathbb{R}$ - vers $t \mapsto \exp(tA)$.

— Soit $a > 0$. On se donne une norme $\|\cdot\|$ sour multiplicative sur $M_n(\mathbb{K})$.
CVN sur $[-a, a]$:
Soit $n \in \mathbb{N}$. Pour tout $t \in [-a, a]$, on a $\|\varphi_0(t)\| = 0$ et, si $n \geq 1$:

$$\|\varphi_n'(t)\| = \left\| \frac{t^{n-1}A^n}{(n-1)!} \right\| \leq \frac{a^{n-1}\|A\|^n}{(n-1)!},$$

donc φ_n' est bornée sur $[-a, a]$ et on a, à partir du rang 1, pour tout $n \geq 1$:

$$\|\varphi_n'\|_\infty \leq \frac{a^{n-1}\|A\|^n}{(n-1)!} = u_n.$$

Or la série $\sum_{n\geq 1} u_n$ converge (d'après la règle de D'Alembert par exemple), donc par comparaison, $\sum \|\varphi_n'\|_\infty$ converge.
Par suite, $\sum \varphi_n'$ converge normalement sur $[-a, a]$ et donc uniformément sur $[-a, a]$.
Par suite, $\sum \varphi_n'$ converge uniformément sur tout segment de $\mathbb{R}$.

Ainsi, d'après le théorème d'interversion dérivation/somme, φ est de classe C^1 sur $\mathbb{R}$ et on a, pour tout $t \in \mathbb{R}$:

$$\varphi'(t) = \sum_{n=0}^{+\infty} \frac{nt^{n-1}A^n}{n!} = A \sum_{n=1}^{+\infty} \frac{t^{n-1}A^{n-1}}{(n-1)!} = A \sum_{n=0}^{+\infty} \frac{(tA)^n}{n!} = A.\exp(tA).$$

□

Corollaire 1.

Soit $A \in M_n(\mathbb{K})$. L'application $\varphi : t \mapsto \exp(tA)$ est de classe C^∞ sur $\mathbb{R}$ et, pour tout $k \in \mathbb{N}$, on a, pour $t \in \mathbb{R}$:

$$\varphi^{(k)}(t) = A^k \exp(tA) = A^k \varphi(t).$$

c. Propriétés de l'exponentielle

Proposition 6.

Soit $A, B \in M_n(\mathbb{K})$. Si A et B commutent, alors A et $\exp(B)$ commutent.

Théorème 4.

Soit $A, B \in M_n(\mathbb{K})$. Si A et B commutent, alors $\exp(A)$ et $\exp(B)$ commutent et :

$$\exp(A+B) = \exp(A)\exp(B)$$

On se sert souvent du théorème précédent pour calculer l'exponentielle d'une matrice A trigonalisable mise sous la forme $A = P(D+N)P^{-1}$ où D est diagonale, N nilpotente et D, N commutent (*remarque : il est toujours possible d'écrire une matrice trigonalisable sous cette forme, il s'agit de la décomposition de Dunford de cette matrice).*

Exercice 1.

Déterminer l'exponentielle de la matrice $A = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}$

Correction.

On $A = P(I_2 + 2N)P^{-1}$ où $N = E_{12}$, $P = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$ et $I_2, 2N$ commutent.

Ainsi, on a $\exp(A) = P\exp(I_2 + 2N)P^{-1} = P\exp(I_2)\exp(2N)P^{-1}$. Or :

- $\exp(I_2) = eI_2$;
- $\exp(2N) = I_2 + 2N$ car N est niloptente d'indice 2 et donc

$$\exp(2N) = \sum_{n=0}^{1} \frac{2^n N^n}{n!} = I_2 + 2N.$$

Ainsi, on a :

$$\exp(A) = Pe(I_2 + 2N)P^{-1} = eA.$$

2. Équation homogène à coefficients constants

Dans ce paragraphe, on considère un système différentielle homogène de la forme $X' = AX$ où $A \in M_n(\mathbb{K})$ (A ne dépend pas du paramètre t).

a. Solution générale du problème de Cauchy

Théorème 5.

Soit $A \in M_n(\mathbb{K})$, $t_0 \in \mathbb{R}$ et $X_0 \in M_{n,1}(\mathbb{K})$. L'unique solution du problème de Cauchy

$$\begin{cases} (E_h) : & X' = AX \\ (\text{C.I.}) & X(t_0) = X_0 \end{cases}$$

est donnée par :

$$f : t \mapsto \exp\left((t-t_0)A\right) X_0.$$

b. Résolution pratique de l'équation homogène

Considérons un système différentiel homogène $(E_h) : X' = AX$ à coefficients constants dans $M_n(\mathbb{R})$.

Utilisation de l'exponentielle de matrice.

- On calcule l'exponentielle de la matrice tA pour $t \in \mathbb{R}$.
  Le calcul est "simple" si la matrice est diagonalisable ou nilpotente par exemple.
- L'ensemble $\mathcal{S}_h$ des solutions de $(E_h) : X' = AX$ est :

$$\mathcal{S}_h = \{f : t \mapsto \exp(tA).C \mid C \in M_{n,1}(\mathbb{R})\}$$

Dans la pratique, le calcul d'une exponentielle de matrice n'est pas aisé. Il en est donc de même pour le calcul explicite des solutions d'un système différentiel homogène d'ordre A. On peut toutefois distinguer des cas où on dispose de méthodes alternatives pour déterminer explicitement les solutions. *Dans chacun des exercices suivants, on effectuera également le calcul de l'exponentielle pour comparer les méthodes de résolution.*

1er cas : A est diagonalisable dans $\mathbb{R}$.

- On détermine les valeurs propres $\lambda_1, ..., \lambda_n \in \mathbb{R}$ (pas forcément deux à deux différentes) et $(V_1, ...V_n)$ une base de vecteurs propres A avec V_i associé à λ_i. On note P la matrice des vecteurs propres de A et $D = P^{-1}AP$.
- Le système $X' = AX$ est équivalent au système $Y' = DY$ où $Y = P^{-1}X$.
- On résout le système diagonale $Y' = DY$ et en utilisant la relation $X = PY$ et on montre ainsi que la famille $(f_1, ..., f_n)$ de fonctions de $\mathbb{R}$ dans $M_{n,1}(\mathbb{R})$ telles que :

$$f_i : t \mapsto e^{\lambda_i t} V_i$$

  forme une base de l'espace $\mathcal{S}_h$ des solutions de (E_h).

Exercice 2.

Résoudre le système différentiel $X' = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} X.$

Correction.

La matrice A est symétrique réelle donc diagonalisable. Les valeurs propres de A sont 1 et 4 avec $m(1) = 2$ et $m(4) = 1$. De plus, (V_1, V_2, V_3) est une base de vecteurs propres avec V_1, V_2 associé à 1 et V_3 associé à 4 où :

$$V_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad V_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad V_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Par suite, $A = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} P^{-1}$ où $P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$.

— Résolution avec la méthode :

Le système $X' = AX$ est équivalent à $Y' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} Y$ avec $Y = P^{-1}X$. On a donc le système diagonal :

$$\begin{cases} y_1' &= y_1 \\ y_2' &= \qquad y_2 \\ y_3' &= \qquad\qquad 4y_3' \end{cases}$$

D'où :

$$\begin{cases} y_1 = & C_1 e^t \\ y_2 = & C_2 e^t \\ y_3 = & C_3 e^{4t} \end{cases}$$

De plus, $X = PY$, donc :

$$X = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} C_1 e^t \\ C_2 e^t \\ C_3 e^{4t} \end{pmatrix} = \begin{pmatrix} C_1 e^t + C_3 e^{4t} \\ C_2 e^t + C_3 e^{4t} \\ -C_2 e^t - C_2 e^t C_3 e^{4t} \end{pmatrix}$$

Autrement dit, comme d'après le théorème de Cauchy linéaire, $\mathcal{S}_h$ est de dimension 3 et que V_1, V_2, V_3 sont linéairement indépendant, la famille (f_1, f_2, f_3) est une base de $\mathcal{S}_h$ où :

$$f_1 : t \mapsto e^t \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}; \quad f_2 : t \mapsto e^t \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}; \quad f_2 : t \mapsto e^{4t} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

— Résolution avec l'exponentielle :

Pour $t \in \mathbb{R}$, on a $tA = P(tD)P^{-1}$, donc :

$$\exp(A) = P\exp(tD)P^{-1} = P \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^{4t} \end{pmatrix} P^{-1} = \frac{1}{3} \begin{pmatrix} 2e^t + e^{4t} & e^{4t} - e^t & e^{4t} - e^t \\ e^{4t} - e^t & 2e^t + e^{4t} & e^{4t} - e^t \\ e^{4t} - e^t & e^{4t} - e^t & 2e^t + e^{4t} \end{pmatrix}$$

Ainsi, l'ensemble $\mathcal{S}_h$ des solutions est :

$$\mathcal{S}_h = \{ f : t \mapsto \frac{1}{3} \begin{pmatrix} 2e^t + e^{4t} & e^{4t} - e^t & e^{4t} - e^t \\ e^{4t} - e^t & 2e^t + e^{4t} & e^{4t} - e^t \\ e^{4t} - e^t & e^{4t} - e^t & 2e^t + e^{4t} \end{pmatrix} .C \mid C \in M_{3,1}(\mathbb{R}) \}.$$

2eme cas : A est diagonalisable dans $\mathbb{C}$ mais pas dans $\mathbb{R}$.

- On procède de la même façon que pour A diagonalisable dans $\mathbb{R}$ mais les vecteurs de la famille $(f_1, ..., f_n)$ ne sont pas tous réels. On doit donc déterminer une famille $(g_1, ..., g_n)$ de fonctions à valeurs réelles qui forment une base de $\mathcal{S}_h$.
- Considérons une valeur propre complexe $\lambda \in \mathbb{C} \setminus \mathbb{R}$ et V un vecteur propre associé. Alors, comme A est à coefficients réels, $\bar{\lambda}$ est également valeur propre et $\bar{V}$ est vecteur propre associé à $\bar{V}$.

  Ainsi, pour $f : t \mapsto e^{\lambda t} V$, $(f, \bar{f})$ est un couple de solution de $X' = AX$ vu comme une équation complexe.
  On pose alors :
  $$g_1 = \frac{1}{2}(f + \bar{f}) = \mathrm{Re}(f) \text{ et } g_2 = \frac{1}{2i}(f - \bar{f}) = \mathrm{Im}(f).$$
  Alors, (g_1, g_2) est une famille libre de solutions à valeurs réelles de $X' = AX$.
- On procède ainsi pour toutes les couples $(\lambda, \bar{\lambda})$ de valeurs propres complexes et on forme, en regroupant avec les vecteurs f_i associés aux valeurs propres réelles, une famille de n vecteurs réels solutions de $X' = AX$ et ainsi base de $\mathcal{S}_h$ (puisqu'il est de dimension n).

Exercice 3.

Résoudre le système différentiel $X' = \begin{pmatrix} 1 & -1 & 2 \\ 2 & -1 & 3 \\ 0 & 0 & 1 \end{pmatrix} X$.

Correction.

La polynôme caractéristique de A est $X^3 - X^2 + X + 1 = (X-1)(X-i)(X+i)$ qui est scindé à racines simples (dans $\mathbb{C}$) donc A est diagonalisable dans $\mathbb{C}$ - mais pas dans $\mathbb{R}$ puisque X^2+1 est un facteur irréductible dans $\mathbb{R}$ du polynôme caractéristique. De plus, (V_1, V_2, V_3) est une base de vecteurs propres avec V_1 associé à 1, V associé à i et $\bar{V}$ associé à $-i = \bar{i}$ où :

$$V_1 = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}, \quad V = \begin{pmatrix} 1 \\ 1-i \\ 0 \end{pmatrix}$$

Par suite, $A = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} P^{-1}$ où $P = \begin{pmatrix} 1 & 1 & 1 \\ 4 & 1-i & 1+i \\ 2 & 0 & 0 \end{pmatrix}$. Ainsi, $X' = AX$ est équivalent à $Y' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix} Y$ avec $Y = P^{-1}X$. On a donc le système diagonal :

$$\begin{cases} y_1' &= y_1 \\ y_2' &= \quad iy_2 \\ y_3' &= \qquad -iy_3' \end{cases}$$

D'où :

$$Y_1(t) = e^t \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad Y_2(t) = e^{it} \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad Y_3(t) = e^{-it} \begin{pmatrix}0\\0\\1\end{pmatrix},$$

Ainsi, en utilisant $X = PY$ on obtient la famille $(f_1, f, \bar{f})$ de solutions de l'équation "complexe" $X' = AX$ où

$$f_1 : t \mapsto e^t \begin{pmatrix}1\\4\\2\end{pmatrix} ; \quad f : t \mapsto e^{it} \begin{pmatrix}1\\1-i\\0\end{pmatrix}.$$

Ainsi, on considère :

$$g_1 : t \mapsto \frac{1}{2}(f + \bar{f}) = \begin{pmatrix}\cos(t)\\\cos(t) - \sin(t)\\0\end{pmatrix} \text{ et } g_2 : t \mapsto \frac{1}{2}(f - \bar{f}) = \begin{pmatrix}\sin(t)\\\cos(t) + \sin(t)\\0\end{pmatrix}$$

Par suite, (f_1, g_1, g_2) est une base de $\mathcal{S}_h$.

3eme cas : A est trigonalisable.

- On trigonalise A sous la forme $A = PTP^{-1}$.
- On résout le système $Y' = TY$ où $X = PY$. Ce système est triangulaire supérieur : on le résout en remontant ligne par ligne les équations.
- On récupère les solutions en utilisant $X = PY$ et on procède de la même façon que dans la méthode précédente si certaines valeurs propres sont complexes.

Exercice 4.

Résoudre le système différentiel $X' = \begin{pmatrix}3 & 0 & -1\\-1 & 2 & 1\\1 & 0 & 1\end{pmatrix} X$.

Correction.

La polynôme caractéristique de A est $X^3 - X^2 + X + 1 = (X-2)^3$ qui est scindé dans $\mathbb{R}$ donc A est trigonalisable dans $\mathbb{R}$. Elle n'est pas diagonalisable car elle possède 2 pour unique valeur propre et $A \neq 2I_2$.

On montre que $E_2(A) = \text{Vect}\left(\begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\1\end{pmatrix}\right)$ En complétant avec un troisième vecteur qui n'appartient pas à $E_2(A)$ - par exemple $\begin{pmatrix}1\\0\\0\end{pmatrix}$, on obtient la trigonalisation :

$$A = P \begin{pmatrix}2 & 0 & 0\\0 & 2 & 1\\0 & 0 & 2\end{pmatrix} P^{-1} \text{ où } P = \begin{pmatrix}0 & 1 & 1\\1 & -1 & 0\\0 & 1 & 0\end{pmatrix}.$$

Ainsi, en posant $Y = P^{-1}X$, $X' = AX$ est équivalent à $Y' = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} Y$. On a donc le système triangulaire :

$$\begin{cases} y_1' &= 2y_1 \\ y_2' &= 2y_2 + y_3 \\ y_3' &= y_3 \end{cases}$$

D'où il existe $C_1, C_3 \in \mathbb{R}$ tels que :

$$y_1 : t \mapsto C_1 e^{2t} \text{ et } y_3 : t \mapsto C_3 e^{2t}$$

Ainsi, on a $y_2' = 2y_2 + C_3 e^{2t}$. On vérifie par la m̂ethode habituelle qu'il existe $C_2 \in \mathbb{R}$ tel que $y_2 : t \mapsto (C_2 + tC_3)e^{2t}$.
Par suite,

$$Y : t \mapsto e^{2t} \begin{pmatrix} C_1 \\ C_2 + tC_3 \\ C_3 \end{pmatrix}$$

Il en résulte que :

$$X = PY : t \mapsto e^{2t} \begin{pmatrix} C_2 + C_3(t+1) \\ C_1 - C_2 - tC_3 \\ C_2 + tC_3 \end{pmatrix} \quad \left(= C_1 e^{2t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + (C_2 + tC_3)e^{2t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + C_3 e^{2t} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right).$$

3. Recherche d'une solution particulière

On considère un système différentiel $(E) : X' = A(t)X + B(t)$. Dans ce paragraphe, on s'intéresse à la recherche effective d'une solution f_p de cette équation.

Méthode de variation des constantes

- On détermine une base $(f_1, ..., f_n)$ de solutions de l'équation homogène $(E_h) : X' = A(t)X$. On note alors, pour $t \in I$:

$$f_p(t) = C_1(t)f_1(t) + \ldots + C_n(t)f_n(t).$$

  où $C_1, ..., C_n$ sont des fonctions dérivables de I dans $\mathbb{K}$.
- Après avoir calculé f_p', on reporte f_p dans l'équation (E) et, en remarquant que, pour tout i, $f_i \in \mathcal{S}_h$, on obtient :

  f_p est solution de (E) si, et seulement si, $C_1'(t)f_1(t) + \ldots + C_n'(t)f_n(t) = B(t)$.
- On résout alors le système précédent afin de déterminer les C_i' puis les C_i par primitivation.

Exercice 5.

Résoudre les systèmes différentiels :

$$X' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} X + \begin{pmatrix} e^{-t} \\ 0 \end{pmatrix} \qquad X' = \frac{1}{2} \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} X + e^t \begin{pmatrix} (t+1)^3 \\ t^2 \end{pmatrix}$$

Correction.

1. On note $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ et $B(t) = \begin{pmatrix} e^{-t} \\ 0 \end{pmatrix}$.
   Tout d'abord, on résout l'équation homogène $X' = AX$. L'ensemble $\mathcal{S}_h$ des solutions de cette équation est :
   $$\mathcal{S}_h = \{t \mapsto \exp(tA).C \mid C \in M_{2,1}(\mathbb{R})\}.$$
   Calculons $\exp(tA)$ pour $t \in \mathbb{R}$. On remarque que A est symétrique réelle donc diagonalisable dans une base orthonormale et que $\frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $\frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sont vecteurs propres de A associés aux valeurs propres 3 et -1 respectivement. Ainsi, on a :
   $$A = PD{}^tP \text{ où } D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \text{ et } P = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$
   Par suite, on a
   $$\exp(tA) = P\exp(tD){}^tP = P\begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix}{}^tP = \frac{1}{2}\begin{pmatrix} e^{3t} + e^{-t} & e^{3t} - e^{-t} \\ e^{3t} - e^{-t} & e^{3t} + e^{-t} \end{pmatrix}$$
   Il en résulte que :
   $$\mathcal{S}_h = \{t \mapsto \quad \mid C_1, C_2 \in \mathbb{R}\}.$$

Partie C

Équations différentielles linéaires scalaires d'ordre supérieur

1. Définitions

Définition 5.

Soit $n \in \mathbb{N}^*$, $(a_i)_{0 \leq i \leq n}$ une famille d'applications continues de I dans $\mathbb{K}$ et $b : I \to \mathbb{K}$ une application continue.
On appelle **équation différentielle linéaire scalaire d'ordre** n, une équation de la forme :

$$(E) : \ y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = b(t).$$

On appelle **équation homogène** associée à (E) l'équation :

$$(E_h) : \ y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = 0.$$

Proposition 7.

Si les a_i et b sont de classe C^k sur I et si $f : I \to \mathbb{K}$ est une solution de (E), alors f est de classe C^{n+k} sur I.

Démonstration.

Par définition de l'équation différentielle (E), une solution de (E) est n fois dérivable sur I.
Pour $k \in \mathbb{N}$, on note $\mathcal{P}_k$ ="si les a_i et b sont de classe C^k sur I, alors toute solution de (E) est de classe C^{n+k} sur I".
Montrons, par récurrence sur $\mathbb{N}$ que pour tout $k \in \mathbb{N}$, $\mathcal{P}_k$ est vraie.

- <u>Initialisation :</u> Montrons $\mathcal{P}_0$. On suppose les a_i et b continues sur I. Soit f une solution de (E) sur I. Alors f est n fois dérivable sur I et ainsi, pour tout $i \in [\![0, n-1]\!]$, $f^{(i)}$ est $n-i$ fois dérivable et donc en particulier continue sur I car $n - i \geq 1$. De plus, on a :

$$f^{(n)} = -a_{n-1}.f^{(n-1)} - \ldots - a_1.f' - a_0(t)f + b$$

donc $f^{(n)}$ est continue sur I comme somme de produits de fonctions continues sur I.
Par suite, f est de classe C^n sur I car n fois dérivable sur I et de dérivée n-ième continue sur I. Il en résulte que $\mathcal{P}_0$ est vraie.

- <u>Hérédité :</u> Soit $k \in \mathbb{N}$. On suppose $\mathcal{P}_k$ vraie. Montrons $\mathcal{P}_{k+1}$.
On suppose les a_i et b sont de classe C^{k+1} sur I et soit f une solution de (E) sur I. Alors en particulier, les a_i et b sont de classe C^k sur I et donc, par hypothèse de récurrence, f est de classe C^{n+k} sur I. Alors pour tout $i \in [\![0, n-1]\!]$, $f^{(i)}$ est de classe C^{n+k-i} sur I et donc de classe C^{k+1} sur I car $n + k - i \geq k + 1$. De plus, on a :

$$f^{(n)} = -a_{n-1}.f^{(n-1)} - \ldots - a_1.f' - a_0(t)f + b$$

donc $f^{(n)}$ est de classe C^{k+1} sur I comme somme de produits de fonctions C^{k+1} sur I. Par suite, f est de C^{n+k+1} sur I. Il en résulte que $\mathcal{P}_{k+1}$ est vraie.

On a donc montré par récurrence que pour tout $k \in \mathbb{N}$, si les a_i et b sont de classe C^k sur I, alors toute solution de (E) est de classe C^{n+k} sur I. $\square$

Proposition 8. *Traduction matricielle*

Soit $(E) : y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = b(t)$ une équation différentielle scalaire d'ordre n.

On considère les applications $A : I \to M_n(\mathbb{K})$ et $B : I \to M_{n,1}(\mathbb{K})$ définies par :

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \\ -a_0(t) & -a_1(t) & \ldots & \ldots & -a_{n-1}(t) \end{pmatrix} \quad \text{et} \quad B(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}$$

Alors, une fonction $f : I \to \mathbb{K}$ est solution de (E), si, et seulement si, la fonction de I dans $\mathbb{K}$ telle que pour $t \in I$:

$\varphi(t) = \begin{pmatrix} f(t) \\ f'(t) \\ \vdots \\ f^{(n)}(t) \end{pmatrix}$ est solution du système différentiel linéaire d'ordre 1 donné par $X' = A(t)X + B(t)$.

Remarque 3.

La matrice $A(t)$ est exactement la transposée de la matrice compagnon du polynôme $X^n + a_{n-1}(t)X^{n-1} + \ldots + a_1(t)X + a_0(t)$.

2. Structure de l'ensemble des solutions

On considère $(E) : y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_1(t)y' + a_0(t)y = b(t)$ une équation différentielle scalaire d'ordre n.

Proposition 9.

Soit $\mathcal{S}$ l'ensemble des solutions de (E) et $\mathcal{S}_h$ l'ensemble des solutions de (E_h). L'ensemble $\mathcal{S}_h$ est un espace vectoriel et pour f_p une solution particulière de (E), on a :

$$\mathcal{S} = f_p + \mathcal{S}_h = \{f_p + f_h \mid f_h \in \mathcal{S}_h\}.$$

Définition 6. *Problème de Cauchy*

Soit $t_0 \in I$ et $(y_0, \dots, y_{n-1}) \in \mathbb{K}^n$. On appelle **problème de Cauchy** est un système d'inconnue $y : I \to \mathbb{K}$ de la forme :

$$\begin{cases} (E) : & y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t) \\ (\text{C.I.}) & y(t_0) = y_0 \\ & \vdots \\ & y^{(n-1)}(t_0) = y_{n-1} \end{cases}$$

Théorème 6.

Soit $t_0 \in I$ et $(y_0, \dots, y_{n-1}) \in \mathbb{K}^n$. Il existe une unique solution au problème de Cauchy :

$$\begin{cases} (E) : & y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t) \\ (\text{C.I.}) & y(t_0) = y_0 \\ & \vdots \\ & y^{(n-1)}(t_0) = y_{n-1} \end{cases}$$

Lemme 3.

Soit $t_0 \in I$ et $\mathcal{S}_h$ l'ensemble des solutions de (E_h).
L'application

$$\left| \begin{array}{ccc} \mathcal{S}_h & \to & \mathbb{K}^n \\ f & \mapsto & (f(t_0), \dots, f^{(n-1)}(t_0)) \end{array} \right.$$

est un isomorphisme d'espaces vectoriels.

Théorème 7.

Soit $(E) : y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_1(t)y' + a_0(t)y = b(t)$ une équation différentielle scalaire d'ordre n.
L'espace vectoriel $\mathcal{S}_h$ des solutions de (E_h) est de dimension finie et

$$\dim(\mathcal{S}_h) = n.$$

Partie D

Méthode et exemples de résolution d'une équation différentielle linéaire scalaire

Contexte et notations :

Soit $n \in \mathbb{N}$. On considère une équation de la forme une équation différentielle linéaire scalaire d'ordre n qui n'est pas sous forme normalisée :

$$(E): \; a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y = b(t).$$

où b et a_i, pour $i \in [\![0, n]\!]$, sont des fonctions définies sur un intervalle I ouvert non vide de $\mathbb{R}$. On note $\mathcal{S}$ l'ensemble des solutions de (E).

On appelle équation homogène associée à (E) et on note (E_h) l'équation différentielle :

$$(E_h): \; a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y = 0.$$

On note $\mathcal{S}_h$ l'ensemble des solutions de (E_h).

On note $Z = \{t \in \mathbb{R} \mid a_n(t) = 0\}$ l'ensemble des zéros de a_n.

On suppose que Z est un ensemble au plus dénombrable de cardinal $m \in \mathbb{N} \cup \{+\infty\}$ tel que les $z \in Z$ sont **isolés** dans I i.e.

$$I \smallsetminus Z = \bigsqcup_{k=1}^{m} I_k.$$

où les I_k sont des intervalles ouverts non vides de $\mathbb{R}$ dont les bornes finies sont des éléments de Z et vérifiant qu'il existe $\delta > 0$ tel que pour tout k, la longueur de I_k est supérieur ou égale à δ.

En particulier, pour $k \in [\![0, m]\!]$ - ou $\mathbb{N}$ si $m = \infty$, a_n ne s'annule pas sur I_k et note :

$$(E)_k: \; y^{(n)} + \frac{a_{n-1}(t)}{a_n(t)}y^{(n-1)} + \ldots + \frac{a_0(t)}{a_n(t)}y = \frac{b(t)}{a_n(t)}.$$

l'équation différentielle normalisée sur l'intervalle I_k et équivalente à (E) sur I_k ; on note $(E_h)_k$ l'équation homogène associée à $(E)_k$.

On note $\mathcal{S}_k$ l'ensemble des solutions de $(E)_k$ et $\mathcal{S}_{kh}$ l'ensemble des solutions de $(E_h)_k$.

1. Une méthode de résolution

On souhaite résoudre l'équation (E) sur I c'est-à-dire on cherche l'ensemble $\mathcal{S}$ des fonctions n-fois dérivables solutions de (E) sur I.

a. Description de la méthode

Méthode de résolution de (E) sur I :

- On détermine l'ensemble Z des zéros de a_n (qui sont isolés) et on condidère les intervalles I_k tels que $I \smallsetminus Z = \bigsqcup_{k=1}^{m} I_k$.

- **Pour chaque intervalle** I_k, a_n ne s'annulant pas sur I_k, on résout l'équation normalisée :

$$(E)_k : \ y^{(n)} + \frac{a_{n-1}(t)}{a_n(t)}y^{(n-1)} + \ldots + \frac{a_0(t)}{a_n(t)}y = \frac{b(t)}{a_n(t)}.$$

  avec la méthode usuelle :
  - ⋆ On résout l'équation homogène $(E_h)_k$ (dont l'ensemble des solutions $\mathcal{S}_{kh}$ est un espace vectoriel de dimension n).
  - ⋆ On détermine une solution particulière f_p de l'équation $(E)_k$.

    Remarque importante : il est recommandé de commencer par essayer de trouver une solution particulière f_p de l'équation (E) sur I directement ; l'avantage est qu'elle sera solution de tous les $(E)_k$!
  - ⋆ On finalise la résolution de $(E)_k$ en écrivant $\mathcal{S}_k = f_p + \mathcal{S}_{hk}$.
- Soit f une fonction n-fois dérivable **sur I**. On remarque alors que f est solution de (E) si, et seulement si, f est solution de $(E)_k$ sur chaque I_k. Les équations (E_k) étant résolues, cela donne une expression explicite de f sur chaque I_k.
  Mais comme f est n-fois dérivable sur I, on doit recoller les "morceaux" en chaque zéro z dans Z :
  - ⋆ f étant continue en z, on calcule les limites $\lim\limits_{t \to z^-} f(t)$ et $\lim\limits_{t \to z^+} f(t)$. Ces limites doivent être finies et égales ; puis on en déduit des conditions sur la forme de f (et notamment sur les constantes venant de la résolution des équations homogènes) ;
  - ⋆ f étant de classe C^1 en z, on effectue le même principe avec $\lim\limits_{t \to z^-} f'(t)$ et $\lim\limits_{t \to z^+} f'(t)$;
  - ⋆ etc ...
  - ⋆ Et attention avec la dernière dérivée : f étant n-fois dérivable mais pas supposée C^n, on doit cette fois égaliser les limites, qui doivent de nouveau être finies, $\lim\limits_{t \to z^-} \dfrac{f^{(n-1)}(t) - f^{(n-1)}(z)}{t-z}$ et $\lim\limits_{t \to z^+} \dfrac{f^{(n-1)}(t) - f^{(n-1)}(z)}{t-z}$.
- On obtient l'ensemble $\mathcal{S}$ des solutions de (E) dont la dimension *en tant qu'espace affine* dépend du nombre de constantes "provenant" des $\mathcal{S}_{hk}$.

b. Structure de l'ensemble des solutions

Proposition 10.

L'ensemble $\mathcal{S}_h$ des solutions de l'équation (E_h) : $a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y = 0$ est un sous-espace vectoriel de l'espace vectoriel $D_n(I, \mathbb{K})$ des fonctions n fois dérivables de I dans $\mathbb{K}$ et on a :

$$\dim(\mathcal{S}_h) \leq mn$$

où on rappelle que m désigne le nombre de zéros de a_n sur I (avec potentiellement $m = +\infty$).

Démonstration.

— L'application $y \mapsto a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \ldots + a_0(t)y$ est une application linéaire de $D_n(I, \mathbb{K})$ dans $\mathcal{F}(I, \mathbb{K})$ comme combinaison linéaire des applications $y \mapsto a_i(t)y^{(i)}$ qui sont linéaires par linéarité de la dérivation et par bilinéarité du produit terme à terme de deux

fonctions.
Par suite, $\mathcal{S}_h$ est un sous-espace vectoriel de $D_n(I, \mathbb{K})$ comme noyau de cette application linéaire.

— On considère l'application φ telle que, pour $f \in \mathcal{S}_h$:

$$\varphi(f) = \left(f_{|I_k}\right)_{k \in [\![1,m]\!]} \in \prod_{i=1}^{m} \mathcal{F}(I_k, \mathbb{K}).$$

Alors, φ est linéaire par linéarité de la restriction et on remarque que si f est solution de (E_h) alors, pour tout $k \in [\![1,m]\!]$, a_n ne s'annulant pas sur I_k, la restriction $f_{|I_k}$ de f sur I_k est solution de l'équation homogène normalisée :

$$(E_h)_k : \ y^{(n)} + \frac{a_{n-1}(t)}{a_n(t)} y^{(n-1)} + \ldots + \frac{a_0(t)}{a_n(t)} y = 0$$

sur I_k.
De plus, pour tout $k \in [\![1,m]\!]$, d'après le théorème 7, l'ensemble $\mathcal{S}_{hk}$ des solutions de $(E_h)_k$ sur I_k est un espace vectoriel de dimension n.

Donc φ est une application linéaire de $\mathcal{S}_h$ dans le produit cartésien $\displaystyle\prod_{k=1}^{m} \mathcal{S}_{hk}$.

Montrons que φ est injective. Soit $f \in \mathrm{Ker}(\varphi)$. Alors f est continue sur I car n fois dérivable sur I avec $n \geq 1$ et pour tout $k \in [\![1,m]\!]$, $f_{|I_k} = 0$.
Montrons alors que $f(z) = 0$ pour tout $z \in Z$. Soit $z \in Z$. Par hypothèse, z est une borne d'un certain I_k qui est ouvert non vide. Par suite, on a, par continuité de f en z :

$$f(z) = \lim_{\substack{t \to z \\ t \in I_k}} \underbrace{f(t)}_{= f_{|I_k}(t)} = \lim_{\substack{t \to z \\ t \in I_k}} 0 = 0.$$

Par suite, f est nulle sur $Z \sqcup \bigsqcup_{i=1}^{m} I_k = I$. Ainsi φ est injective.
Comme φ est une application linéaire injective, on a donc :

$$\dim(\mathcal{S}_h) \leq \dim\left(\prod_{k=1}^{m} \mathcal{S}_{hk}\right) = \sum_{k=1}^{m} \dim(\mathcal{S}_{hk}) = mn.$$

□

Remarque 4.

Comme on va le voir avec les exemples et exercices suivants, la dimension de $\mathcal{S}_h$ peut prendre toutes les valeurs entières possibles entre 0 et mn.

2. Ordre 1 : Exemples de résolutions complètes

a. Exemple avec $\mathcal{S}_h$ de dimension 2

Exemple 1.

L'ensemble des fonctions dérivables solutions de $(E): \ ty' - 2y = 21t^3e^t - 5t$ sur $\mathbb{R}$ est :

$$\mathcal{S} = \left\{ f : t \mapsto 21t^2e^t + 5t + \begin{cases} C_1t^2 & \text{si } t \geq 0 \\ C_2t^2 & \text{si } t < 0 \end{cases} \middle| \ C_1, C_2 \in \mathbb{R} \right\}.$$

La fonction $t \mapsto t$ s'annule sur $\mathbb{R}$ en 0. On considère alors les intervalles de résolution de l'équation normalisée $I_1 =]0, +\infty[$ et $I_2 =]-\infty, 0[$.

— Résolution de l'équation sur I_1 : Sur I_1, l'équation est équivalente à $(E)_1 : y' - \frac{2}{t}y = 21t^2e^t - 5$.

- Résolution de l'équation homogène $(E_h)_1 : y' - \frac{2}{t}y = 0$.

  Une primitive de $t \mapsto \frac{2}{t}$ sur I_1 est $t \mapsto 2\ln(t)$. Comme, pour tout $t \in I_1$, $e^{2\ln(t)} = t^2$, l'ensemble $\mathcal{S}_{h1}$ de $(E_h)_1$ sur I_1 est :

  $$\mathcal{S}_{h1} = \left\{ t \mapsto C_1t^2 \mid C_1 \in \mathbb{R} \right\}.$$

- Recherche d'une solution particulière de $(E)_1$.
  On remarque que, pour tout $t \in \mathbb{R}$:

  $$21t^2e^t - 5 = 21 \times (t^2e^t) - 5 \times 1$$

  Ainsi si g_p et h_p sont des solutions de respectivement $y' - \frac{2}{t}y = t^2e^t$ et $y' - \frac{2}{t} = 1$, alors d'après le principe de superposition, $f_p = 21g_p - 5h_p$ est solution de $(E)_1$.
  Appliquons la méthode de variation de la constante pour chercher g_p et h_p :

  ⋆ Variation de la constante pour $y' - \frac{2}{t}y = t^2e^t$.
  On pose $g_p : t \mapsto C(t)t^2$ où C est une fonction dérivable sur I_1. Alors g_p est solution de l'équation si, et seulement si, $C'(t)t^2 = t^2e^t$ pour tout $t \in I_1$.
  Par suite, pour tout $t \in I_1$, $C'(t) = e^t$ et donc $C : t \mapsto e^t$ convient *(on peut choisir n'importe quelle primitive de C')*.
  Il en résulte que $g_p : t \mapsto t^2e^t$ est solution de $y' - \frac{2}{t}y = t^2e^t$ sur I_1.

  ⋆ Variation de la constante pour $y' - \frac{2}{t}y = 1$.
  On pose $h_p : t \mapsto C(t)t^2$ où C est une fonction dérivable sur I_1. Alors h_p est solution de l'équation si, et seulement si, $C'(t)t^2 = 1$ pour tout $t \in I_1$.
  Par suite, pour tout $t \in I_1$, $C'(t) = \frac{1}{t^2}$ et donc $C : t \mapsto -\frac{1}{t}$ convient.
  Il en résulte que $h_p : t \mapsto -\frac{1}{t}t^2 = -t$ est solution de $y' - \frac{2}{t}y = 1$ sur I_1.

  Ainsi, par principe de superposition, $f_p = 21g_p - 5h_p : t \mapsto 21t^2e^t + 5t$ est solution de $(E)_1$ sur I_1.

  On remarque que la fonction f_p est bien définie et dérivable sur $\mathbb{R}$ donc elle est même solution de (E) sur $\mathbb{R}$!

- Conclusion sur I_1 :
  L'ensemble des solution $\mathcal{S}_1$ de $(E)_1$ sur I_1 est :

  $$\mathcal{S}_1 = \left\{ t \mapsto C_1t^2 + 21t^2e^t + 5t \mid C_1 \in \mathbb{R} \right\}.$$

— Résolution de l'équation sur I_2 : Sur I_2, l'équation est équivalente à $(E)_2 : y' - \frac{2}{t}y = 21t^2e^t - 5$.

- Résolution de l'équation homogène $(E_h)_2 : y' - \frac{2}{t}y = 0$.

Une primitive de $t \mapsto \frac{2}{t}$ sur I_2 est $t \mapsto 2\ln(|t|) = 2\ln(-t)$. Comme, pour tout $t \in I_2$, $e^{2\ln(-t)} = -t^2$, l'ensemble $\mathcal{S}_{h2}$ de $(E_h)_2$ sur I_2 est :

$$\mathcal{S}_{h2} = \left\{ t \mapsto -C_2 t^2 \mid C_2 \in \mathbb{R} \right\}.$$

Quitte à changer C_2 en $-C_2$ (possible car C_2 parcourt $\mathbb{R}$), on peut écrire :

$$\mathcal{S}_{h2} = \left\{ t \mapsto C_2 t^2 \mid C_2 \in \mathbb{R} \right\}.$$

- Recherche d'une solution particulière de $(E)_2$.
  Comme $f_p : t \mapsto 21t^2e^t + 5t$ est solution de (E) sur $\mathbb{R}$, f_p est solution de $(E)_2$ sur I_2.
- Conclusion sur I_2 :
  L'ensemble des solution $\mathcal{S}_2$ de $(E)_2$ sur I_2 est :

$$\mathcal{S}_2 = \left\{ t \mapsto C_2 t^2 + 21t^2e^t + 5t \mid C_2 \in \mathbb{R} \right\}.$$

L'équation (E) étant résolue sur I_1 et I_2, on passe à l'étape de "recollement" des solutions. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Alors f est solution de (E) si, et seulement si, f est dérivable sur $\mathbb{R}$ et est solution de (E) sur I_1 et I_2.
Ainsi, si f est une solution de (E), il existe $C_1, C_2 \in \mathbb{R}$ tels que :

$$f(t) = \begin{cases} C_1t^2 + 21t^2e^t + 5t & \text{pour } t \in I_1 =]0, +\infty[\\ C_2t^2 + 21t^2e^t + 5t & \text{pour } t \in I_2 =]-\infty, 0[\end{cases}$$

On peut remarquer avant de commencer le recollement que si on évalue l'équation différentielle (E) en 0, on obtient $f(0) = 0$.

⋆ Comme f est continue en 0 et $f(0) = 0$, on a $\lim_{x\to 0^-} f(x) = 0 = \lim_{t\to 0^+} f(x)$.
Or on a $\lim_{t\to 0^+} f(x) = \lim_{t\to 0^+}(C_1t^2 + 21t^2e^t + 5t) = 0$ et
$\lim_{t\to 0^-} f(x) = \lim_{t\to 0^-}(C_2t^2 + 21t^2e^t + 5t) = 0$.
À ce stade, aucune condition sur C_1, C_2 supplémentaires ne sont imposées.

⋆ Comme f est dérivable sur 0, on a :

$$\lim_{t\to 0^-} \frac{f(t) - f(0)}{t} = f'(0) = \lim_{t\to 0^+} \frac{f(t) - f(0)}{t}.$$

Or on a $\lim_{t\to 0^+} \frac{f(t)-f(0)}{t} = \lim_{t\to 0^+}(C_1t + 21te^t + 5) = 5$ et
$\lim_{t\to 0^-} \frac{f(t)-f(0)}{t} = \lim_{t\to 0^-}(C_2t + 21te^t + 5) = 5$;
D'où $f'(0) = 5$ et toujours aucune condition supplémentaire sur C_1, C_2.

Ainsi, si f est solution de (E), on a $f : t \mapsto 21t^2e^t + 5t + \begin{cases} C_1t^2 & \text{si } t \geq 0 \\ C_2t^2 & \text{si } t < 0 \end{cases}$; et réciproquement si f est de cette forme, f est bien dérivable et solution de (E).
Il en résulte que l'ensemble $\mathcal{S}$ des solutions de (E) sur $\mathbb{R}$ est :

$$\mathcal{S} = \left\{ f : t \mapsto 21t^2e^t + 5t + \begin{cases} C_1t^2 & \text{si } t \geq 0 \\ C_2t^2 & \text{si } t < 0 \end{cases} \middle| \, C_1, C_2 \in \mathbb{R} \right\}.$$

et on peut remarquer que $\mathcal{S}_h = \left\{ f : t \mapsto \begin{cases} C_1t^2 & \text{si } t \geq 0 \\ C_2t^2 & \text{si } t < 0 \end{cases} \middle| \, C_1, C_2 \in \mathbb{R} \right\}$ est un espace vectoriel de dimension 2.

Remarque : on aurait trouver pu une solution f_p de (E) en premier lieu puis déterminer $\mathcal{S}_h$ et enfin conclure que $\mathcal{S} = f_p + \mathcal{S}_h$ (comme nous le verrons dans l'exemple 3).

Exemple 2.

L'équation (E) : $t^2y' - 2ty = 1$ sur $\mathbb{R}$ n'admet aucune solution et l'ensemble $\mathcal{S}_h$ des solutions de l'équation homogène sur $\mathbb{R}$ est :

$$\mathcal{S}_h = \left\{ f : t \mapsto \begin{cases} C_1t^2 & \text{si } t \geq 0 \\ C_2t^2 & \text{si } t < 0 \end{cases} \;\middle|\; C_1, C_2 \in \mathbb{R} \right\}.$$

La fonction $t \mapsto t^2$ s'annule sur $\mathbb{R}$ en 0. On considère alors les intervalles de résolution de l'équation normalisée $I_1 =]0, +\infty[$ et $I_2 =]-\infty, 0[$.

— Résolution de l'équation sur I_1 : Sur I_1, l'équation est équivalente à $(E)_1 : y' - \frac{2}{t}y = \frac{1}{t^2}$.

- comme pour l'exemple précédent, l'ensemble $\mathcal{S}_{h1}$ des solutions de l'équation homogène sur I_1 est :
$$\mathcal{S}_{h1} = \left\{ t \mapsto C_1t^2 \mid C_1 \in \mathbb{R} \right\}.$$
- Recherche d'une solution particulière de $(E)_1$.
Variation de la constante pour $y' - \frac{2}{t}y = \frac{1}{t^2}$.
On pose $f_p : t \mapsto C(t)t^2$ où C est une fonction dérivable sur I_1. Alors f_p est solution de l'équation si, et seulement si, $C'(t)t^2 = \frac{1}{t^2}$ pour tout $t \in I_1$.
Par suite, pour tout $t \in I_1$, $C'(t) = \frac{1}{t^4}$ et donc $C : t \mapsto -\frac{1}{4t^3}$ convient.
Il en résulte que $f_p : t \mapsto -\frac{1}{4t}$ est solution de $y' - \frac{2}{t}y = \frac{1}{t^2}$ sur I_1.
- Conclusion sur I_1 :
L'ensemble des solution $\mathcal{S}_1$ de $(E)_1$ sur I_1 est :
$$\mathcal{S}_1 = \left\{ t \mapsto C_1t^2 - \frac{1}{4t} \mid C_1 \in \mathbb{R} \right\}.$$

— Résolution de l'équation sur I_2 : Sur I_2, l'équation est équivalente à $(E)_2 : y' - \frac{2}{t}y = \frac{1}{t^2}$.
La résolution est en tout point similaire à celle sur I_1 et on obtient que l'ensemble des solution $\mathcal{S}_2$ de $(E)_2$ sur I_2 est :
$$\mathcal{S}_2 = \left\{ t \mapsto C_2t^2 - \frac{1}{4t} \mid C_2 \in \mathbb{R} \right\}.$$

L'équation (E) étant résolue sur I_1 et I_2, on passe à l'étape de "recollement" des solutions. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Alors f est solution de (E) si, et seulement si, f est dérivable sur $\mathbb{R}$ et est solution de (E) sur I_1 et I_2.
Ainsi, si f est une solution de (E), il existe $C_1, C_2 \in \mathbb{R}$ tels que :

$$f(t) = \begin{cases} C_1t^2 - \frac{1}{4t} & \text{pour } t \in I_1 =]0, +\infty[\\ C_2t^2 - \frac{1}{4t} & \text{pour } t \in I_2 =]-\infty, 0[\end{cases}$$

⋆ Comme f est continue en 0, on a $\lim_{x \to 0^-} f(x) = f(0) = \lim_{t \to 0^+} f(x)$.
Or on a $\lim_{t \to 0^+} f(x) = \lim_{t \to 0^+} (C_1t^2 - \frac{1}{4t}) = -\infty$ et $\lim_{t \to 0^-} f(x) = \lim_{t \to 0^-} (C_2t^2 - \frac{1}{4t}) = +\infty$.
Donc aucune valeur de C_1 et C_2 ne conviennent pour assurer la continuité de f en 0 ; on peut donc s'arrêter ici !

Ainsi, l'équation (E) n'admet aucune solution sur $\mathbb{R}$. On peut tout de même prouver que l'espace $\mathcal{S}_h$ des solutions de l'équation homogène est $\mathcal{S}_h = \left\{ f : t \mapsto \begin{cases} C_1t^2 & \text{si } t \geq 0 \\ C_2t^2 & \text{si } t < 0 \end{cases} \;\middle|\; C_1, C_2 \in \mathbb{R} \right\}$ et c'est un espace vectoriel de dimension 2.

Remarque : on aurait pu éviter tout cette résolution qui conduit à l'absence de solution. En effet, supposons par l'absurde que (E) possède une solution f sur $\mathbb{R}$. Alors, pour tout $t \in \mathbb{R}$, $t^2 f'(t) - 2tf(t) = 1$ et donc, en particulier, pour $t = 0$, on a

$$0 = 0^2 \times f'(t) - 2 \times 0 \times f(t) = 1$$

Contradiction ! Donc (E) ne possède pas de solution sur $\mathbb{R}$.

b. Exemple avec $\mathcal{S}_h$ de dimension 1

Exemple 3.

L'ensemble des fonctions dérivables solutions de (E) : $t^2 y' - y = t(t-1)$ sur $\mathbb{R}$ est :

$$\mathcal{S} = \left\{ f : t \mapsto t + \begin{cases} C_1 e^{-\frac{1}{t}} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases} \;\middle|\; C_1 \in \mathbb{R} \right\}.$$

Changeons de méthode pour cet exemple : déterminons tout de suite une solution de (E) sur $\mathbb{R}$: les coefficients de (E) sont polynomiaux, cherchons une solution polynomiale.
Soit p une fonction polynomiale de degré $n \geq 0$ solution de (E) Alors $t^2 p' - p$ est de degré $n+1$ car $\deg(t^2 p') = 2 + (n-1) = n+1$ et $\deg(-p) = n$. Ainsi, $n + 1 = \deg(t(t-1)) = 2$ d'où $n = 1$.
Soit $a, b \in \mathbb{R}$ et $p : t \mapsto at + b$. Alors p solution de (E) si, et seulement si, pour tout $t \in \mathbb{R}$:

$$at^2 - at - b = t^2 - t$$

or deux fonctions poynomiales sont égales si, et seulement si, leurs coefficients sont égaux d'où : p solution de (E) si, et seulement si, $a = 1$ et $b = 0$ i.e. $p : t \mapsto t$.

Ainsi, la fonction $f_p : t \mapsto t$ est une solution de (E) sur $\mathbb{R}$.

Maintenant, il nous reste à déterminer l'ensemble $\mathcal{S}_h$ des solutions de l'équation homogène.
La fonction $t \mapsto t^2$ s'annule sur $\mathbb{R}$ en 0. On considère alors les intervalles de résolution de l'équation normalisée $I_1 =]0, +\infty[$ et $I_2 =]-\infty, 0[$.

— Résolution de l'équation homogène sur I_1 : Sur I_1, l'équation homogène est équivalente à $(E_h)_1 : y' - \frac{1}{t^2} y = 0$.
Une primitive de $t \mapsto \frac{1}{t^2}$ sur I_1 est $t \mapsto -\frac{1}{t}$. Ainsi, l'ensemble $\mathcal{S}_{h1}$ de $(E_h)_1$ sur I_1 est :

$$\mathcal{S}_{h1} = \left\{ t \mapsto C_1 e^{-\frac{1}{t}} \mid C_1 \in \mathbb{R} \right\}.$$

— Résolution de l'équation homogène sur I_2 : Sur I_2, l'équation homogène est équivalente à $(E_h)_2 : y' - \frac{1}{t^2} y = 0$.
Une primitive de $t \mapsto \frac{1}{t^2}$ sur I_2 est $t \mapsto -\frac{1}{t}$. Ainsi, l'ensemble $\mathcal{S}_{h2}$ de $(E_h)_2$ sur I_2 est :

$$\mathcal{S}_{h2} = \left\{ t \mapsto C_2 e^{-\frac{1}{t}} \mid C_2 \in \mathbb{R} \right\}.$$

L'équation (E_h) étant résolue sur I_1 et I_2, on passe à l'étape de "recollement" des solutions. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Alors f est solution de (E_h) si, et seulement si, f est dérivable sur $\mathbb{R}$ et

est solution de (E_h) sur I_1 et I_2.
Ainsi, si f est une solution de (E_h), il existe $C_1, C_2 \in \mathbb{R}$ tels que :

$$f(t) = \begin{cases} C_1 e^{-\frac{1}{t}} & \text{pour } t \in I_1 =]0, +\infty[\\ C_2 e^{-\frac{1}{t}} & \text{pour } t \in I_2 =]-\infty, 0[. \end{cases}$$

On peut remarquer avant de commencer le recollement que si on évalue l'équation différentielle (E_h) en 0, on obtient $f(0) = 0$.

- ⋆ Comme f est continue en 0 et $f(0) = 0$, on a $\lim_{x \to 0^-} f(x) = 0 = \lim_{t \to 0^+} f(x)$.
  Or on a $\lim_{t \to 0^+} f(x) = \lim_{t \to 0^+} C_1 e^{-\frac{1}{t}} = 0$ et
  $\lim_{t \to 0^-} f(x) = \lim_{t \to 0^-} C_2 e^{-\frac{1}{t}} = \begin{cases} \pm\infty & \text{si } C_2 \neq 0 \\ 0 & \text{si } C_2 = 0 \end{cases}$.
  Par suite, $C_2 = 0$ mais pas de condition sur C_1.
- ⋆ Comme f est dérivable sur 0, on a :

$$\lim_{t \to 0^-} \frac{f(t) - f(0)}{t} = f'(0) = \lim_{t \to 0^+} \frac{f(t) - f(0)}{t}.$$

  Or on a $\lim_{t \to 0^+} \frac{f(t) - f(0)}{t} = \lim_{t \to 0^+} \frac{C_1}{t} e^{-\frac{1}{t}} = 0$ par croissances comparées et
  $\lim_{t \to 0^-} \frac{f(t) - f(0)}{t} = \lim_{t \to 0^-} 0 = 0$;
  D'où $f'(0) = 0$ et toujours aucune condition supplémentaire sur C_1.

Ainsi, si f est solution de (E_h), on a $f : t \mapsto \begin{cases} C_1 e^{-\frac{1}{t}} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases}$; et réciproquement si f est de cette forme, f est bien dérivable et solution de (E_h).
Il en résulte que l'ensemble $\mathcal{S}_h$ des solutions de (E_h) sur $\mathbb{R}$ est :

$$\mathcal{S}_h = \left\{ f : t \mapsto \begin{cases} C_1 e^{-\frac{1}{t}} & \text{si } t > 0 \\ 0 & \text{si } t \leq 0 \end{cases} \;\middle|\; C_1 \in \mathbb{R} \right\}.$$

qui est un espace de dimension 1.

On obtient ainsi l'ensemble $\mathcal{S}$ des solutions de (E) sur $\mathbb{R}$ car $\mathcal{S} = f_p + \mathcal{S}_h$:

$$\mathcal{S} = \left\{ f : t \mapsto t + \begin{cases} C_1 e^{-\frac{1}{t}} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases} \;\middle|\; C_1 \in \mathbb{R} \right\}.$$

Remarque : on aurait pu utiliser la même méthode que dans l'exemple 1.

c. Exemple avec $\mathcal{S}_h$ de dimension 0

Exemple 4.

L'ensemble des fonctions dérivables solutions de (E) : $ty' + y = 1$ sur $\mathbb{R}$ est :

$$\mathcal{S} = \{t \mapsto 1\}.$$

Déterminons une solution de (E) sur $\mathbb{R}$: les coefficients de (E) sont polynomiaux, cherchons une solution polynomiale.
Soit p une fonction polynomiale de degré $n \geq 0$ solution de (E) et notons $a_n \neq 0$ son coeffcient dominant. Alors le monôme de plus haut degré de $tp' + p$ est égal à $(n+1)a_n t^n$. Or $(n+1)a_n \neq 0$ donc $\deg(tp' + p) = n$. Ainsi, $n = \deg(1) = 0$.
Soit $c \in \mathbb{R}$ et $p : t \mapsto c$. Alors p solution de (E) si, et seulement si, pour tout $t \in \mathbb{R}$:

$$t \times 0 + c = 1$$

d'où p solution de (E) si, et seulement si, $c = 1$ i.e. $p : t \mapsto 1$.
Ainsi, la fonction $f_p : t \mapsto 1$ est une solution de (E) sur $\mathbb{R}$.

Maintenant, il nous reste à déterminer l'ensemble $\mathcal{S}_h$ des solutions de l'équation homogène.
La fonction $t \mapsto t$ s'annule sur $\mathbb{R}$ en 0. On considère alors les intervalles de résolution de l'équation normalisée $I_1 =]0, +\infty[$ et $I_2 =]-\infty, 0[$.

— Résolution de l'équation homogène sur I_1 : Sur I_1, l'équation homogène est équivalente à $(E_h)_1 : y' + \frac{1}{t}y = 0$.
  Une primitive de $t \mapsto -\frac{1}{t}$ sur I_1 est $t \mapsto -\ln(t)$. Comme, pour tout $t \in I_1$, $e^{-\ln(t)} = \frac{1}{t}$, l'ensemble $\mathcal{S}_{h1}$ de $(E_h)_1$ sur I_1 est :

$$\mathcal{S}_{h1} = \left\{ t \mapsto C_1 \frac{1}{t} \mid C_1 \in \mathbb{R} \right\}.$$

— Résolution de l'équation homogène sur I_2 : Sur I_2, l'équation homogène est équivalente à $(E_h)_2 : y' + \frac{1}{t}y = 0$.
  Une primitive de $t \mapsto -\frac{1}{t}$ sur I_2 est $t \mapsto -\ln(-t)$. Ainsi, l'ensemble $\mathcal{S}_{h2}$ de $(E_h)_2$ sur I_2 est :

$$\mathcal{S}_{h2} = \left\{ t \mapsto -C_2 \frac{1}{t} \mid C_2 \in \mathbb{R} \right\}.$$

L'équation (E_h) étant résolue sur I_1 et I_2, on passe à l'étape de "recollement" des solutions. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Alors f est solution de (E_h) si, et seulement si, f est dérivable sur $\mathbb{R}$ et est solution de (E_h) sur I_1 et I_2.
Ainsi, si f est une solution de (E_h), il existe $C_1, C_2 \in \mathbb{R}$ tels que :

$$f(t) = \begin{cases} C_1 \frac{1}{t} & \text{pour } t \in I_1 =]0, +\infty[\\ -C_2 \frac{1}{t} & \text{pour } t \in I_2 =]-\infty, 0[. \end{cases}$$

On peut remarquer avant de commencer le recollement que si on évalue l'équation différentielle (E_h) en 0, on obtient $f(0) = 0$.

⋆ Comme f est continue en 0 et $f(0) = 0$, on a $\lim_{x \to 0^-} f(x) = 0 = \lim_{t \to 0^+} f(x)$.
  Or on a $\lim_{t \to 0^+} f(x) = \lim_{t \to 0^+} C_1 \frac{1}{t} = \begin{cases} \pm\infty & \text{si } C_1 \neq 0 \\ 0 & \text{si } C_1 = 0 \end{cases}$ et
  $\lim_{t \to 0^-} f(x) = \lim_{t \to 0^-} C_2 e^{-\frac{1}{t}} = \begin{cases} \pm\infty & \text{si } C_2 \neq 0 \\ 0 & \text{si } C_2 = 0 \end{cases}$.
  Par suite, $C_1 = 0$ et $C_2 = 0$; donc f est la fonction nulle sur $\mathbb{R}$ (qui est bien dérivable donc pas besoin d'aller plus loin).

Ainsi, si f est solution de (E_h), on a $f : t \mapsto 0$; et réciproquement la fonction nulle est bien dérivable et solution de (E_h).
Il en résulte que l'ensemble $\mathcal{S}_h$ des solutions de (E_h) sur $\mathbb{R}$ est :

$$\mathcal{S}_h = \{t \mapsto 0\}$$

qui est un espace de dimension 0.

On obtient ainsi l'ensemble $\mathcal{S}$ des solutions de (E) sur $\mathbb{R}$ car $\mathcal{S} = f_p + \mathcal{S}_h$:

$$\mathcal{S} = \{t \mapsto 1\}.$$

d. Exemple avec $\mathcal{S}_h$ de dimension infinie

Exemple 5.

L'ensemble des fonctions dérivables solutions de (E) : $\cos(t)y' + 2\sin(t)y = \sin(t)$ sur $\mathbb{R}$ est :

$$\mathcal{S} = \left\{ f : t \mapsto \frac{1}{2} + C_k \cos^2(t) \text{ si } t \in \left[\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi\right[\text{ avec } k \in \mathbb{Z} \;\middle|\; \forall\, k \in \mathbb{Z},\ C_k \in \mathbb{R} \right\}.$$

Déterminons tout d'abord l'ensemble $\mathcal{S}_h$ des solutions de l'équation homogène.
La fonction $\cos : \mathbb{R} \to \mathbb{R}$ s'annule sur $\mathbb{R}$ en $\frac{\pi}{2} + k\pi$ pour tout $k \in \mathbb{Z}$. On considère alors les intervalles de résolution de l'équation normalisée $I_k =]\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi[$ pour $k \in \mathbb{Z}$.

— Soit $k \in \mathbb{Z}$. Résolution de l'équation homogène sur I_k : Sur I_k, l'équation homogène est équivalente à $(E_h)_k$: $y' + 2\tan(t)y = 0$.

Une primitive de $t \mapsto -2\tan(t) = 2\frac{\cos'}{\cos}$ sur I_k est $t \mapsto 2\ln(|\cos(t)|)$. Comme, pour tout $t \in I_k$, $e^{2\ln(|\cos(t)|)} = \cos^2(t)$, l'ensemble $\mathcal{S}_{hk}$ de $(E_h)_k$ sur I_k est :

$$\mathcal{S}_{hk} = \left\{t \mapsto C_k \cos^2(t) \mid C_k \in \mathbb{R}\right\}.$$

L'équation (E_h) étant résolue sur chaque I_k, on passe à l'étape de "recollement" des solutions.
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Alors f est solution de (E_h) si, et seulement si, f est dérivable sur $\mathbb{R}$ et est solution de (E_h) sur chaque I_k.
Ainsi, si f est une solution de (E_h), pour tout $k \in \mathbb{Z}$, il existe $C_k \in \mathbb{R}$ tels que :

$$f(t) = C_k \cos^2(t) \;\text{ si } t \in I_k =]\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi[.$$

On peut remarquer avant de commencer le recollement que si on évalue l'équation différentielle (E_h) en $\frac{\pi}{2} + k\pi$ pour $k \in \mathbb{Z}$, on obtient $f(\frac{\pi}{2} + k\pi) = 0$. Soit $k \in \mathbb{Z}$. On note $z_k = \frac{\pi}{2} + k\pi$.

⋆ Comme f est continue en z_k et $f(z_k) = 0$, on a $\lim_{x \to z_k^-} f(x) = 0 = \lim_{t \to z_k^+} f(x)$.

Or on a $\lim_{t \to z_k^+} f(x) = \lim_{t \to z_k^+} C_1 \cos^2(t) = 0$ et
$\lim_{t \to 0^-} f(x) = \lim_{t \to 0^-} C_2 \cos^2(t) = 0$.
À ce stade, aucune condition sur C_1, C_2 supplémentaires ne sont imposées.

⋆ Comme f est dérivable sur 0, on a :

$$\lim_{t \to z_k^-} \frac{f(t) - f(z_k)}{t} = f'(z_k) = \lim_{t \to z_k^+} \frac{f(t) - f(z_k)}{t}.$$

Or on a $\lim_{t \to z_k^+} \frac{f(t) - f(z_k)}{t} = \lim_{t \to z_k^+} C_1 \frac{\cos^2(t)}{t} = 0$ et
$\lim_{t \to z_k^-} \frac{f(t) - f(z_k)}{t} = \lim_{t \to z_k^-} C_2 \frac{\cos^2(t)}{t} = 0$;
D'où $f'(z_k) = 0$ et toujours aucune condition supplémentaire sur C_1, C_2.

Ainsi, si f est solution de (E_h), on a $f : t \mapsto C_k \cos^2(t)$ si $t \in [\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi[$ pour $k \in \mathbb{Z}$; et réciproquement si f est de cette forme, f est bien dérivable et solution de (E).
Il en résulte que l'ensemble $\mathcal{S}_h$ des solutions de (E_h) sur $\mathbb{R}$ est :

$$\mathcal{S} = \left\{ f : t \mapsto C_k \cos^2(t) \text{ si } t \in \left[\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi\right[\text{ avec } k \in \mathbb{Z} \;\middle|\; \forall\, k \in \mathbb{Z},\ C_k \in \mathbb{R} \right\}.$$

Il nous reste à déterminer une solution particulière de (E) sur $\mathbb{R}$ (s'il en existe !). Pour cela, on peut essayer des fonctions sinusoïdales car les coefficients le sont... mais malheureusement, cela ne fonctionne pas.

Essayons alors la méthode de variation de la constante sur un I_k fixé, en espérant en déduire une solution particulière sur $\mathbb{R}$ tout entier. Ici, prenons $I_{-1} =]-\frac{\pi}{2}, \frac{\pi}{2}[$.

- ⋆ Variation de la constante pour $y' + 2\tan(t)y = \tan(t)$.

  On pose $f_p : t \mapsto C(t)\cos^2(t)$ où C est une fonction dérivable sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. Alors f_p est solution de l'équation si, et seulement si, $C'(t)\cos^2(t) = \tan(t)$ pour tout $t \in]-\frac{\pi}{2}, \frac{\pi}{2}[$.
  Par suite, pour tout $t \in]-\frac{\pi}{2}, \frac{\pi}{2}[$,

$$C'(t) = \frac{\tan(t)}{\cos^2(t)} = -\frac{\cos'(t)}{\cos^3(t)}$$

  et donc $C : t \mapsto \frac{1}{2\cos^2(t)}$ convient.
  Il en résulte que $f_p : t \mapsto \frac{1}{2}$ est solution de $y' - \frac{2}{t}y = t^2e^t$ sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

Comme espéré, $f_p : t \mapsto \frac{1}{2}$ est bien dérivable sur $\mathbb{R}$ et solution de (E) sur $\mathbb{R}$.
On obtient ainsi l'ensemble $\mathcal{S}$ des solutions de (E) sur $\mathbb{R}$ car $\mathcal{S} = f_p + \mathcal{S}_h$:

$$\mathcal{S} = \left\{ f : t \mapsto \frac{1}{2} + C_k \cos^2(t) \text{ si } t \in \left[\frac{\pi}{2} + k\pi, \frac{\pi}{2} + (k+1)\pi\right[\text{ avec } k \in \mathbb{Z} \;\middle|\; \forall\, k \in \mathbb{Z},\ C_k \in \mathbb{R} \right\}.$$

e. Exercices

Exercice 6.

Résoudre l'ensemble des fonctions dérivables solutions des équations suivantes sur l'intervalle I :

1. $t^3y' + 2y = t(2t^2 + 1)$
2. $t(t-1)y' - y = 1$ sur $\mathbb{R}$
3. $t^2(t-1)y' - ty = 1 - 2t$ sur $\mathbb{R}$

Partie E

Résolution d'un équation différentielle linéaire scalaire d'ordre 2

Dans cette partie, on adapte et développe la méthode de résolution d'une équation différentielle linéaire scalaire d'ordre n dans le particulier $n = 2$.
On considère alors dans cette partie une équation définie sur un intervalle I de la forme :

$$(E): \ a_2(t)y'' + a_1(t)y'(t) + a_0(t)y = b(t).$$

1. Précision sur la méthode de résolution

Pour résoudre l'équation (E), on utilise la méthode décrite dans la partie précédente. Il reste tout de même deux points qui ne sont pas developpés dans cette méthode : la résolution de l'équation homogène sur chaque intervalle (I_k) et l'obtention d'une solution particulière de l'équation non homogène.
Pour l'ordre 2, des techniques existent pour cela et nous allons les étudier dans la suite ; mais tout d'abord, resumons la méthode dans le cas de l'ordre 2 :

Précisions et résumé de la méthode pour l'ordre 2 :

— On trouve l'ensemble Z des zéros de a_n et on considère les intervalles ouverts I_k qui forment $I \setminus Z$.
— On resout $(E)_k$ sur chaque I_k.
Précisions pour l'équation homogène : L'ensemble $\mathcal{S}_{hk}$ des solutions de l'équation homogène sur I_k est de la forme $\mathcal{S}_{hk} = \mathrm{Vect}(f_1, f_2)$ (Théorème 7).
Ainsi, pour déterminer $\mathcal{S}_{hk}$, il suffit de **trouver deux solutions** f_1 **et** f_2 **non colinéaires** de l'équation homogène $(E_h)_k$.
— On "recolle" les solutions des $(E)_k$ pour former les solutions de (E).

'

2. Recherche d'une première solution d'une équation homogène normalisée

Dans ce paragraphe, on considère une équation différentielle homogène (pas forcément normalisée !) :

$$(E_h): \ a_2(t)y'' + a_1(t)y'(t) + a_0(t)y = 0.$$

et on recherche une première solution f_1 de cette équation. Pour cela, il existe de nombreuses techniques empiriques. En voici quelques-unes présentées sur des exemples :

a. À l'aide de la "forme" des coefficients

Recherche d'une solution du même type que les coefficients
Lorsque les coefficients sont tous de la même "forme" (**y compris celui devant** y'' **!**) : polynômes, fonctions trigonométriques, exponentielles ; il peut être pertinent de rechercher une solution de (E_h) du même type.

Exemple : Considérons l'équation :

$$(E_h):\ (t^2+2t-1)y''+(t^2-3)y'-(2t+2)y=0.$$

1. Soit $t\mapsto P(t)$ une fonction polynomiale solution de (E_h). Montrer que $\deg(P)=2$.
2. Déterminer une solution de (E_h).

1. Soit P un polynôme à coefficients réels de degré $n\geq 0$ et on note $a_n\neq 0$ sont coefficient dominant. Si la fonction polynomiale $P: t\mapsto P(t)$ associée à P est solution de (E_h), alors :

$$Q=(X^2+2X-1)P''+(X^2-3)P'-(2X+2)P=0$$

Or un polynôme est nul si, et seulement si, ses coefficients sont nuls, donc, en particulier, le coefficient c de Q associé au monôme X^{n+1} est nul. De plus, le coeffcients associés à X^n sont : 0 dans $(X^2+2X-1)P''$ car il est de degré $n-2+2=n$; na_n dans $(X^2-3)P'$ et $-2a_n$ dans $-(2X+2)P$. Ainsi, on a :

$$0=c=0+na_n-2a_n=(n-2)a_n$$

Or $a_n\neq 0$ donc $n-2=0$ i.e. $n=2$. Par suite, si $P: t\mapsto P(t)$, alors $\deg(P)=2$.

2. Soit $a,b,c\in\mathbb{R}$ et $P=\mapsto aX^2+bX+c$. On a :

$$(X^2+2X-1)P''+(X^2-3)P'-(2X+2)P=-bX^2-2(a+b+c)X+-(2a+3b+2c)$$

Or, un polynôme est nul si, et seulement si, tous ses coefficients sont nuls donc $(X^2+2X-1)P''+(X^2-3)P'-(2X+2)P=0$ si, et seulement si, $\begin{cases} b &=0\\ a+b+c &=0\\ 2a+3b+c &=0\end{cases}$ i.e. $c=-a$ et $b=0$.
Il en résulte que, pour $P\in\mathbb{R}[X]$, la fonction polynomiale $t\mapsto P(t)$ est solution de (E_h) si, et seulement si, il existe $a\in\mathbb{R}$ tel que $P=a(X^2-1)$.

On en déduit une première solution f_1 "simple" de (E_h) :

$$f_1: t\mapsto t^2-1.$$

Exercice 7.

Déterminer une solution de l'équation homogène suivante :

$$(E_h):\ y''+\frac{1}{t}y'-\frac{1}{t^2}y=0.$$

Correction Indications et résultat :.

On dénormalise l'équation qui est équivalente à $t^2y''+ty'-y=0$ et on cherche une solution polynomiale pour trouver la solution :

$$f_1 : t \mapsto t.$$

b. À l'aide d'une série entière

Recherche d'une solution grâce à un Développement en Série Entière

On considère une série entière $\sum a_n t^n$ de rayon de convergence $R > 0$ et on suppose que sa somme f est solution de (E_h).

On obtient alors, en reportant f dans (E_h), une relation entre les coefficients a_n qui permettent de déterminer explicitement la fonction f.

Exemple : Considérons l'équation :

$$(E_h) : \; ty'' + 2y' + ty = 0.$$

1. Soit $\sum a_n t^n$ une série entière de rayon de convergence $R > 0$ telle que sa somme f est solution de (E_h). Montrer que :

$$a_1 = 0 \text{ et } (n+1)(n+2)a_{n+1} + a_{n-1} = 0 \quad \text{pour tout } n \geq 1.$$

2. En déduire une solution de (E_h).

1. On remarque que $f : t \mapsto \sum_{n=0}^{+\infty} a_n t^n$ est de classe C^∞ sur $]-R, R[$ et on a :

$$\begin{aligned}
tf''(t) + 2f'(t) + tf(t) &= \sum_{n=2}^{+\infty} n(n-1)a_n t^{n-1} + \sum_{n=1}^{+\infty} 2na_n t^{n-1} + \sum_{n=0}^{+\infty} a_n t^{n+1} \\
&= \sum_{n=0}^{+\infty} (n+2)(n+1)a_{n+2} t^{n+1} + 2a_1 \sum_{n=0}^{+\infty} 2(n+2)a_{n+2} t^{n+1} + \sum_{n=0}^{+\infty} a_n t^{n+1} \\
&= 2a_1 + \sum_{n=0}^{+\infty} \left((n+2)(n+3)a_{n+2} + a_n\right) t^{n+1} \\
tf''(t) + 2f'(t) + tf(t) &= 2a_1 t^0 + \sum_{n=1}^{+\infty} ((n+1)(n+2)a_{n+1} + a_{n-1}) a_n t^n.
\end{aligned}$$

Alors f est solution de (E_h) sur $]-R, R[$ si, et seulement si, pour tout $t \in]-R, R[$:

$$2a_1 t^0 + \sum_{n=1}^{+\infty} ((n+1)(n+2)a_{n+1} + a_{n-1}) a_n t^n = 0 \quad (*)$$

Or une somme de série entière est nulle sur un intervalle ouvert si, et seulement si, ses coefficients sont nuls ; d'où f est solution de (E_h) sur $]-R, R[$ si, et seulement si,

$$\begin{cases} a1 = 0 \\ (n+1)(n+2)a_{n+1} + a_{n-1} = 0 \;\; \forall\, n \in \mathbb{N}^*. \end{cases} \quad \Leftrightarrow \quad \begin{cases} a1 = 0 \\ a_{n+1} = -\frac{1}{(n+1)(n+2)} a_{n-1} \;\; \forall\, n \in \mathbb{N}^*. \end{cases}$$

Par récurrence, comme $a_1 = 0$, $a_{2k+1} = 0$ pour tout $k \in \mathbb{N}$.
De même, si on suppose $a_0 = 0$, alors $a_{2k} = 0$ pour tout $k \in \mathbb{N}$ et donc $f = \mathbf{0}$ - qui est bien solution de (E_h) mais on le savait déjà !
Plus intéressant : supposons $a_0 \neq 0$. Alors, par récurrence, $a_{2k} \neq 0$ pour tout $k \in \mathbb{N}$ et donc :

$$\begin{cases} a0 \neq 0 \\ \dfrac{a_{2k+2}}{a_{2k}} = -\dfrac{1}{(2k+2)(2k+3)} \text{ pour tout } n \in \mathbb{N}^*. \end{cases}$$

Par suite, pour $z \in \mathbb{C}^*$:

$$\left|\frac{a_{2k+2}z^{2k+2}}{a_{2k}z^{2k}}\right| = \frac{|z|^{2k}}{(2k+2)(2k+3)} \xrightarrow[k\to+\infty]{} 0.$$

Donc le rayon de convergence de la série entière $\sum a_n t^n$ est $R = +\infty$ et on a, pour tout $n \in \mathbb{N}$:

$$\begin{aligned} a_{2n} &= a_0 \prod_{k=0}^{n-1} \frac{a_{2k+2}}{a_{2k}} \\ &= a_0(-1^n) \prod_{k=0}^{n-1} \frac{1}{(2k+2)(2k+3)} \\ a_{2n} &= a_0 \frac{(-1)^n}{(2n+1)!} \end{aligned}$$

Ainsi, pour tout $t \in \mathbb{R}$, on a :

$$\begin{aligned} f(t) &= \sum_{m=0}^{+\infty} a_m t^m \\ &= \sum_{n=0}^{+\infty} a_{2n} t^{2n} \\ f(t) &= a_0 \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} t^{2n} \end{aligned}$$

Cela rappelle le développement en série entière sur $\mathbb{R}$ de sinus i.e. pour tout $t \in \mathbb{R}$,

$$\sin(t) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1} = t \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} t^{2n};$$

ce qui permet d'écrire :

$$f(t) = \begin{cases} a_0 \dfrac{\sin(t)}{t} & \text{si } t \neq 0 \\ a_0 & \text{si } t = 0 \end{cases}$$

On choisit donc un a_0 "simple", $a_0 = 1$ (ou 33 si on en a envie) et on pose :

$$f_1 : t \mapsto \begin{cases} \dfrac{\sin(t)}{t} & \text{si } t \neq 0 \\ 1 & \text{si } t = 0 \end{cases}$$

Alors f_1 vérifie $(*)$ et donc f_1 est une solution de (E_h) sur $\mathbb{R}$ (car $R = +\infty$). On a donc trouvé une première solution de l'équation homogène.

Exercice 8.

Déterminer une solution de l'équation homogène suivante :

$$(E_h):\ 4ty'' + 2y' - y = 0.$$

3. Recherche d'une seconde solution d'un équation homogène normalisée

Dans ce paragraphe, on considère une équation différentielle homogène (pas forcément normalisée !) :

$$(E_h):\ a_2(t)y'' + a_1(t)y'(t) + a_0(t)y = 0.$$

Une fois une première solution f_1 de cette équation obtenue, on peut en "fabriquer" une deuxième f_2 à partir de f_1 et telle que (f_1, f_2) soit libre. De ce fait, (f_1, f_2) sera une base des solutions de l'équation homogène sur I_k **(attention, pas forcément sur I tout entier par contre !)** ce qui finalisera les résolutions des équations $(E_h)_k$.

a. À l'aide de la méthode de Lagrange

La méthode de Lagrange utilise la même astuce que la méthode de variation de la constante pour l'ordre 1 :

Proposition 11.

Soit f_1 une solution de $(E_h):\ a_2(t)y'' + a_1(t)y'(t) + a_0(t)y = 0$ sur I et z une fonction deux fois dérivable sur I. La fonction $f = f_1 z$ est solution de (E_h) si, et seulement si, z' vérifie l'équation différentielle :

$$a_2(t)f_1(t)x' + (2a_2(t)f_1(t) + a_1(t)f_1(t))x = 0$$

Méthode de Lagrange : On suppose connue une solution f_1 de (E_h).

- On pose $f = f_1 z$ où z est une fonction deux fois dérivable sur I. Alors f est solution de (E_h) si, et seulement si, z' est solution d'une équation différentielle d'ordre 1 :
  - ⋆ que l'on obtient en exprimant $f = f_1 z$ comme solution de (E_h) ;
  - ⋆ que l'on résout pour trouver z' puis z en primitivant.
- On choisit les constantes qui apparaissent dans les résolutions précédentes et on écrit $f_2 = f_1 z$ qui est donc solution de (E_h) et on vérifie que f_1 et f_2 ne sont pas colinéaires (ce qui sera le cas si z n'est pas constant !).

Exercice 9.

Résoudre les équations homogènes :

1. $(E_h):\ (t^2+1)y'' - 2y = 0$; '
2. $(E_h):\ (t+1)y'' - y' - ty = 0$

b. À l'aide du Wronskien

Pour l'utilisation de ce qu'on appellera dans la suite le wronskien, on doit supposer que l'équation est *normalisée* ; ainsi, dans ce paragraphe, par (E_h), on entendra :

$$(E_h): \ y'' + a_1(t)y'(t) + a_0(t)y = 0.$$

Définition 7. *Wronskien*

Soit f, g des solutions de l'équation homogène (E_h).
On appelle **wronskien** du couple (f, g) l'application de I dans $\mathbb{K}$ définie, pour tout $t \in I$ par :

$$W_{f,g}(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix} = f(t)g'(t) - f'(t)g(t).$$

Exercice 10.

Soit f, g des solutions de l'équation homogène (E_h). Montrer que le wronskien $W_{f,g}$ est dérivable sur I et que, pour tout $t \in I$,

$$W'_{f,g}(t) = f(t)g''(t) - f''(t)g(t).$$

Correction.

Soit f, g des solutions de l'équation homogène (E_h). Comme a_0, a_1 sont des fonctions continues sur I, f, g sont de classe C^2 sur I (Remarque 3). Par suite, $W_{f,g} = f'g - fg'$ est de classe C^1 sur I comme différence de produits de fonctions C^1 sur I. De plus, on a :

$$\begin{array}{rcl} W'_{f,g} & = & (fg' - f'g)' \\ & = & (fg')' - (f'g)' \\ & = & f'g' + fg'' - (f''g + f'g') \\ W'_{f,g} & = & fg'' - f''g. \end{array}$$

Proposition 12.

Soit f, g des solutions de l'équation homogène (E_h).
Le wronskien $W_{f,g}$ du couple (f, g) est solution de l'équation différentielle :

$$x' = -a_1(t)x.$$

Démonstration.

Soit f, g des solutions de l'équation homogène (E_h). Ainsi, pour tout $t \in I$, on a :

$$f''(t) = -a_1(t)f'(t) - a_0(t)f(t) \text{ et } g''(t) = -a_1(t)g'(t) - a_0(t)g(t).$$

D'après l'exercice précédent, $W_{f,g}$ est dérivable sur I et $W'_{f,g} = f''g - fg''$. Par suite, pour tout

$t \in I$, on a :

$$\begin{aligned} W'_{f,g}(t) &= f(t)g''(t) - f''(t)g(t) \\ &= f(t)(-a_1(t)g'(t) - a_0(t)g(t)) - (-a_1(t)f'(t) - a_0(t)f(t))g(t) \\ &= -a_1(t)f(t)g'(t) - a_0(t)f(t)g(t) + a_1(t)f'(t)g(t) + a_0(t)f(t)g(t) \\ &= -a_1(t)(f(t)g'(t) - f'(t)g(t)) \\ W'_{f,g}(t) &= -a_1(t)W_{f,g}(t). \end{aligned}$$

Ainsi, $W_{f,g}$ est solution sur I de l'équation différentielle $x' = -a_1(t)x$. □

Corollaire 2.

Soit f, g des solutions de l'équation homogène (E_h).
Il existe $C \in \mathbb{K}$ tel que, pour tout $t \in I$, $W_{f,g}(t) = Ce^{-A_1(t)}$ où A_1 est une primitive de a_1.

Correction.

L'ensemble $\mathcal{S}$ des solutions sur I de l'équation différentielle scalaire d'ordre 1 homogène $x' = -a_1(t)x$ est donné par, étant donné une primitive A_1 de a_1 sur I :

$$\mathcal{S} = \{t \mapsto Ce^{-A_1(t)} \mid C \in \mathbb{R}\}$$

Or, d'après la proposition précédente, pour f, g des solutions sur I de l'équation homogène (E_h) : $y'' + a_1(t)y' + a_0(t)y = 0$, $W_{f,g}$ appartient à $\mathcal{S}$, donc il existe $C \in \mathbb{R}$ tel que $W_{f,g} : t \mapsto Ce^{-A_1(t)}$.

Proposition 13.

Soit f, g des solutions de l'équation homogène (E_h). Les assertions suivantes sont équivalentes :

i) le couple (f, g) est une base de $\mathcal{S}_h$;

ii) il existe $t \in I$ tel que $W_{f,g}(t) \neq 0$;

iii) pour tout $t \in I$ tel que $W_{f,g}(t) \neq 0$.

On utilise le wronskien lorsque l'on dispose déjà d'une solution f_1 de (E_h) : $y'' + a_1(t)y' + a_0(t)y = 0$ afin de déterminer une seconde solution non colinéaire à la première.

Méthode du Wronskien : On suppose connue une solution f_1 de (E_h).

- Soit $f \in \mathcal{S}_h$. Alors W_{f,f_1} est solution d'une équation différentielle d'ordre 1 (à retrouver soi-même à chaque fois). On obtient alors W_{f,f_1} à une constante multiplicative près.
- Sur un intervalle où f_1 ne s'annule pas, on considère la fonction $\dfrac{f}{f_1}$. Alors sa dérivée vérifie :

$$\left(\frac{f}{f_1}\right)' = \frac{-W_{f,f_1}}{f_1^2}.$$

On en déduit alors f puis on détermine une fonction $f_2 \in \mathcal{S}_h$ plus "simple" que f en supprimant ses composantes selon f_1.

- On vérifie que (f_1, f_2) est une famille libre et donc une base de $\mathcal{S}_h$.

Exercice 11.

Résoudre les équations homogènes suivantes :

1. $(E_h) : \ y'' + \dfrac{t^2-3}{t^2+2t-1}y' - \dfrac{2t+2}{t^2+2t-1}y = 0.$
2. $(E_h) : \ y'' + \dfrac{1}{t}y' - \dfrac{1}{t^2}y = 0.$
3. $(E_h) : \ y'' + \dfrac{2}{t}y' + y = 0.$
4. $(E_h) : \ y'' + \dfrac{1}{2t}y' - \dfrac{1}{4t}y = 0.$

4. Recherche d'une solution particulière de l'équation normalisée

Dans ce paragraphe, on s'intéresse à la recherche d'une solution particulière f_p de

$$(E) : y'' + a_1(t)y' + a_0(t)y = b(t)$$

afin de finaliser la résolution une fois les solutions de l'équation homogène obtenues :

a. À l'aide de la méthode de variation des constantes

Proposition 14.

Soit (f_1, f_2) une base de l'équation homogène (E_h).
On pose $f_p : t \mapsto C_1(t)f_1(t) + C_2(t)f_2(t)$ où C_1, C_2 sont des fonctions dérivables. Alors :

La fonction f_p est solution de (E) si, et seulement si,

$$\begin{cases} C_1'(t)f_1(t) + C_2'(t)f_2(t) = 0 \\ C_1'(t)f_1'(t) + C_2'(t)f_2'(t) = b(t) \end{cases}.$$

Correction.

D'aprés la proposition 8, f solution de (E_h) si, et seulement si, $\begin{pmatrix} f \\ f' \end{pmatrix}$ est solution du système différentiel $(S_{E_h}) \ : \ X' = A(t)X$ où $A = \begin{pmatrix} 0 & 1 \\ -a_0(t) & -a_1(t) \end{pmatrix}$.
Par suite, comme l'ensemble $\mathcal{S}_h$ des solutions de (E_h) est $\mathcal{S}_h = \text{Vect}(f_1, f_2)$, l'ensemble des solutions de (S_{E_h}) est $\text{Vect}\,(X_1, X_2)$ où, pour $i = 1, 2$, $X_i = \begin{pmatrix} f_i \\ f_i' \end{pmatrix}$.

On pose $f_p : t \mapsto C_1(t)f_1(t)+C_2(t)f_2(t)$ où C_1, C_2 sont des fonctions dérivables ; $X_p = C_1(t)X_1 + C_2(t)X_2$ et $B(t) = \begin{pmatrix} 0 \\ b(t) \end{pmatrix}$.

Comme, pour $i = 1, 2$, $X_i = \begin{pmatrix} f_i \\ f_i' \end{pmatrix}$ est solution de du système homogène (S_{E_h}), on a :

$$X_i' = A(t)X_i.$$

On a alors :

$$\begin{array}{rcl} X_p' & = & C_1'(t)X_1 + C_2'(t)X_2 + C_1(t)X_1' + C_2(t)X_2' \\ & = & C_1'(t)X_1 + C_2'(t)X_2 + A(t)\left(C_1(t)X_1 + C_2(t)X_2\right) \\ X_p' & = & C_1'(t)X_1 + C_2'(t)X_2 + A(t)X_p. \end{array}$$

Ainsi :

f_p est solution de (E)
si, et seulement si,

X_p est solution de $(S_E) \ : \ X' = A(t)X + B(t)$
si, et seulement si,

$X_p' = A(t)X_p + B(t)$
si, et seulement si,

$C_1'(t)X_1 + C_2'(t)X_2 + A(t)X_p = A(t)X_p + B(t)$
si, et seulement si,

$C_1'(t)X_1 + C_2'(t)X_2 = B(t)$
si, et seulement si,

$$\begin{cases} C_1'(t)f_1(t) + C_2'(t)f_2(t) = 0 \\ C_1'(t)f_1'(t) + C_2'(t)f_2'(t) = b(t) \end{cases} .$$

Méthode de variations des constantes :
Une fois déterminée une base (f_1, f_2) de l'équation homogène (E_h), on pose

$$f_p : t \mapsto C_1(t)f_1(t) + C_2(t)f_2(t)$$

où C_1, C_2 sont des fonctions dérivables. Alors :

La fonction f_p est solution de (E) si, et seulement si,

$$\begin{cases} C_1'(t)f_1(t) + C_2'(t)f_2(t) = 0 \\ C_1'(t)f_1'(t) + C_2'(t)f_2'(t) = b(t) \end{cases}$$

On résout alors ce système pour trouver C_1', C_2' puis C_1, C_2 en primitivant.

Exercice 12.

Résoudre les équations différentielles suivantes :

1. $(E): \; y'' + \frac{1}{t}y' - \frac{1}{t^2}y = \frac{4\ln(t)}{t}$.
2. $(E): \; y'' + \frac{2}{t}y' + y = t$.

Chapitre XIII

Calcul différentiel

Table des matières

Dans tous le chapitre, E, F désignent des espaces vectoriels normés de dimension finie, respectivement p et q. De plus, U, V désignent des ouverts de E et F respectivement.

Partie A

Introduction et rappels

Le but de ce chapitre est de généraliser l'étude des fonctions d'un intervalle $\mathbb{R}$ dans $\mathbb{R}$ aux fonctions d'un ouvert de $\mathbb{R}^p$ dans $\mathbb{R}$. On souhaite donner un sens aux concepts de dérivabilité, tangente, etc... afin de pouvoir déterminer les extrema de telles fonctions ou tracer leurs graphes par exemple.

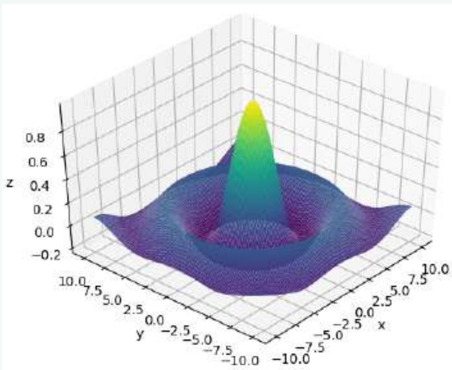
1. Exemple

Exemple 1.

Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ telle que :

$$f : (x, y) \mapsto \frac{\sin(\sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}}$$

Son graphe $\mathcal{G} = \{(x, y, f(x, y)) \mid (x, y) \in \mathbb{R}^2\}$ est une partie de $\mathbb{R}^3$:



2. Continuité

Définition 1. *Rappel : continuité*

Soit $f : U \to F$ une fonction et $a \in U$. On dit que f est continue en a si :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in U, \quad \|x - a\|_E \leq \delta \Rightarrow \|f(x) - f(a)\| \leq \varepsilon.$$

Montrer qu'une fonction est continue en un point.

Soit $f : U \to \mathbb{R}$. Pour montrer que f est continue en a, il suffit très souvent de :

- calculer $|f(x) - f(a)|$;
- Majorer $|f(x) - f(a)|$ par $K\|x - a\|^{\alpha}$ où $K, \alpha > 0$ et la norme $\|\cdot\|$ sur E est choisie judicieusement en fonction de l'expression de f.
  Remarque : on peut choisir la norme que l'on veut car E est de dimension finie et donc toutes ses normes sont équivalentes.

Exercice 1.

Montrer que les fonctions suivantes de $\mathbb{R}^2$ dans $\mathbb{R}$ sont continues en $(0, 0)$:

$$f(x,y) = \begin{cases} \dfrac{x^3y^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases} \qquad g(x,y) = \begin{cases} (x+y)\sin\left(\frac{1}{x+y^2}\right) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases}$$

Montrer qu'une fonction est n'est pas continue en un point.

Soit $f : U \to \mathbb{R}$. Pour montrer que f est n'est pas continue en a (étant donnée une valeur pour $f(a)$), il suffit de déterminer une suite $(u_n)_{n\in\mathbb{N}}$ à valeurs dans E tels que $u_n \xrightarrow[n\to+\infty]{} a$ et

$$f(u_n) \nrightarrow f(a).$$

Exercice 2.

Montrer que les fonctions suivantes ne sont pas continues en 0_E :

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon} \end{cases} \qquad g(x,y,z) = \begin{cases} \dfrac{xy+yz}{x^2+2y^2+2z^2} & \text{si } (x,y,z) \neq (0,0,0) \\ 0 & \text{sinon} \end{cases}$$

Remarque 1.

Pour montrer qu'une fonction de U dans $\mathbb{R}$ n'est pas prolongeable par continuité en a, il suffit donc de trouver deux suites (u_n) et (v_n) qui convergent vers a et telle que $f(u_n)$ et $f(v_n)$ ne convergent pas vers une même limite (ou que l'une des deux ne converge pas du tout !)

Exercice 3.

Déterminer le domaine de définition de la fonction $f : (x, y) \mapsto \dfrac{x^2 - y^2}{x^2 + y^2}$ et montrer qu'elle n'est pas prolongeable par continuité en $(0, 0)$.

Partie B
Dérivées partielles

1. Dérivée suivant un vecteur

Définition 2.

Soit $f : U \to F$ une fonction, $a \in U$ et $u \in E$. On dit que f est **dérivable en a suivant u** si la fonction $t \mapsto \frac{1}{t}(f(a+tu) - f(a))$ admet une limite en 0. Dans ce cas, on note :

$$D_u f(a) = \lim_{t \to 0} \frac{1}{t}(f(a+tu) - f(a)),$$

et $D_u f(a)$ est appelée la **dérivée de f en a selon u**.

Exemple 2.

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ telle que $f : (x, y) \mapsto x^2 + 5xy^3$. Alors

$$D_{(1,1)} f(x_0, y_0) = 2x_0 + 5y_0^3 + 15x_0 y_0^2.$$

Exercice 4.

Soit $f : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ tel que $f(M) = M^2$ et $A, U \in M_n(\mathbb{R})$.
Déterminer la dérivée de f en A suivant U.

Définition 3. *Dérivée partielle en un point*

Soit $\mathcal{B} = (e_1, ..., e_p)$ une base de E, $f : U \to F$ et $a \in U$.
Soit $j \in [\![1, p]\!]$. On dit que f admet une **j-ème dérivée partielle de f en a dans la base $\mathcal{B}$** si f admet une dérivée en a suivant e_j. Dans ce cas, on note :

$$\partial_j f(a) = D_{e_j} f(a) \text{ ou encore } \frac{\partial}{\partial x_j} f(a) = D_{e_j} f(a).$$

Définition 4. *Application dérivée partielle*

Soit $\mathcal{B} = (e_1, ..., e_p)$ une base de E, $f : U \to F$ et $a \in U$.
Si f admet une dérivée partielle en tout point a de U, alors on appelle **j-ème dérivée partielle**

de f dans la base $\mathcal{B}$ l'application $\dfrac{\partial}{\partial x_j} f$ de U dans F telle que :

$$\frac{\partial}{\partial x_j} f : x \mapsto \frac{\partial}{\partial x_j} f(x).$$

Remarque 2.

Pour $E = \mathbb{R}^3$ par exemple, et $f : (x, y, z) \mapsto f(x, y, z)$ on notera

$$\frac{\partial}{\partial x} f, \quad \frac{\partial}{\partial y} f, \quad \frac{\partial}{\partial z} f$$

les dérivées partielles suivant les vecteurs de la base canonique.

Exemple 3.

Pour $f : (x, y, z) \mapsto (x^2 y, 2z^3 + ye^{2x})$ admet des dérivées partielles en tout point de $\mathbb{R}^2$ et on a :

$$\frac{\partial}{\partial x} f(x, y, z) = (2xy, 2ye^{2x}), \ \frac{\partial}{\partial y} f(x, y, z) = (x^2, e^{2x}), \ \frac{\partial}{\partial z} f(x, y, z) = (0, 3z^2)$$

Exercice 5.

Calculer les dérivées partielles de $f : \mathbb{R}^4 \to \mathbb{R}$ suivant les vecteurs de la base canonique où :

$$f(x, y, z, t) = \frac{xy + z}{1 + t^2}$$

Exercice 6.

On considère la fonction $f : (x, y) \mapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) = (0, 0) \\ 0 & \text{sinon} \end{cases}$. Montrer que f admet des dérivées partielles en $(0, 0)$ suivant la base canonique.

Correction.

Tout d'abord f n'est pas continue en $(0, 0)$: en effet, on a :

$$f(1/n, 1/n) = \frac{\frac{1}{n^2}}{\frac{2}{n2}} = \frac{1}{2} \xrightarrow[n \to +\infty]{} \frac{1}{2} \neq 0 = f(0, 0)$$

et $(1/n, 1/n) \xrightarrow[n \to +\infty]{} (0, 0)$.

Ainsi, f n'est pas continue en $(0, 0)$ (on peut montrer également qu'elle n'est pas prolongeable par continuité : on aurait pu prendre n'importe quelle valeur pour $f(0, 0)$, f n'aurait pas été continue en

$(0,0))$.
Malgré cela, on va montrer que f possède des dérivées partielles dans la base canonique (e_1, e_2) de $\mathbb{R}^2$.
On a,

$$\frac{1}{t}(f((0,0)+te_1) - f(0,0)) = \frac{1}{t}(f(1,0) - f(0,0)) = 0 \xrightarrow[t \to 0]{} 0$$

Donc, par définition, f admet une dérivée partielle en $(0,0)$ suivant e_1 et on a :

$$\frac{\partial}{\partial x} f(0,0) = D_{e_1} f(0,0) = 0.$$

Par un calcul similaire, on trouve que f admet une dérivée partielle en $(0,0)$ suivant e_2 et que :

$$\frac{\partial}{\partial y} f(0,0) = D_{e_2} f(0,0) = 0.$$

Ainsi, f admet des dérivées partielles en $(0,0)$ dans la base canonique de $\mathbb{R}^2$... alors qu'elle n'est pas continue en $(0,0)$...

Remarque 3.

Attention ! Contrairement au cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, une fonction peut admettre des dérivées partielles égales en un point mais ne pas être continue en ce point ! (voire la fonction précédente)

2. Matrice jacobienne

Rappel :
Étant donné une fonction $f : U \to F$ et $\mathcal{C} = (\varepsilon_1, \ldots, \varepsilon_q)$ une base de F, on a la décomposition, pour chaque $x \in U$,

$$f(x) = \sum_{i=1}^{q} f_i(x) \varepsilon_i$$

où les fonctions $f_i : U \to \mathbb{R}$ sont les **applications composantes** de f dans la base $\mathcal{C}$.

Exemple : Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ telle que :

$$f(x,y,z) = (x + y^2, 2xyz).$$

Dans la base canonique de $\mathbb{R}^2$, les applications composantes de f sont :

$$f_1(x,y,z) = x + y^2 \text{ et } f_2(x,y,z) = 2xyz.$$

Définition 5. *Matrice jacobienne*

Soit $f : U \to F$ une fonction, $\mathcal{B}, \mathcal{C}$ des bases de E, F respectivement et $a \in U$.
On appelle **matrice jacobienne de f en a dans les bases $\mathcal{B}$ et $\mathcal{C}$** la matrice notée $Jf(a) \in M_{q,p}(\mathbb{R})$ telle que :

$$Jf(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1\leq i\leq q\\1\leq j\leq p}} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dots & \dfrac{\partial f_1}{\partial x_p}(a) \\ \vdots & & \vdots \\ \dfrac{\partial f_q}{\partial x_1}(a) & \dots & \dfrac{\partial f_q}{\partial x_p}(a) \end{pmatrix}.$$

Exemple 4.

Application de passage des coordonnées polaires aux coordonnées cartésiennes
L'application $\varphi : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}^2$ définie, pour $(r, \theta) \in \mathbb{R}_+^* \times \mathbb{R}$, par :

$$\varphi(r, \theta) = (r\cos(\theta), r\sin(\theta)),$$

admet pour matrice jacobienne en (r, θ) dans la base canonique de $\mathbb{R}^2$:

$$J\varphi(r, \theta) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

Exercice 7.

Calculer la matrice jacobienne dans les bases canoniques :

1. $f : (x, y) \mapsto (x + y, x^2y)$ en (x, y) puis en $(0, 0)$;
2. $g : (x, y, z) \mapsto \left(\dfrac{\cos(xy)}{1+z^2}, z^3e^{xy}\right)$ en (x, y, z) puis en $(-1, \pi, 0)$;
3. $h : (x, y, z, t) \mapsto (xy, yz, zt, tx)$ en (x, y, z, t) puis en $(1, 1, 1, 1)$.

Exercice 8.

Donner la matrice jacobienne de l'application S de passage des coordonnées sphériques vers les coordonnées cartésiennes dans la base canonique de $\mathbb{R}^3$ au point (r, θ, φ)

Partie C
Différentiabilité

1. Différentielle d'une application

a. Définitions

Définition 6. *Différentiabilité locale*

Soit $f : U \to F$ et $a \in U$.
On dit que f est **différentiable en** a, s'il existe $\ell \in \mathcal{L}(E, F)$ telle que :

$$f(a+h) = f(a) + \ell(h) + o(\|h\|) \text{ quand } h \to 0_E.$$

Dans ce cas, on dira également que f **admet un développement limité à l'ordre** 1 **en** a.

Proposition 1.

Soit $f : U \to F$ et $a \in U$. Si f est différentiable en a, alors f est continue en a.

Démonstration.

On suppose f différentiable en a. Alors f admet un développement limité à l'ordre 1 en a *i.e.* il existe $\ell \in \mathcal{L}(E, F)$ tel que :

$$f(a+h) = f(a) + \ell(h) + \underset{h \to 0_E}{o}(\|h\|).$$

Comme E est de dimension finie, ℓ est une apllication linéaire continue sur E et donc en 0_E, d'où $\ell(h) \xrightarrow[h \to 0_E]{} \ell(0_E) = 0_F$.
Par suite, pour tout x dans un voisinage de a dans U, on a, par inégalité triangulaire :

$$\|f(x) - f(a)\|_F = \|f(a+h) - f(a)\|_F \leq \|\ell(h)\| + \underset{h \to 0_E}{o}(\|h\|) \xrightarrow[h \to 0_E]{} 0$$

d'où $f(x) \xrightarrow[x \to a]{} f(a)$ et donc f est continue en a. □

Lemme 1.

Soit $f : U \to F$ et $a \in U$. Si f est différentiable en a, alors il existe une **unique** application $\ell \in \mathcal{L}(E, F)$ telle que :

$$f(a+h) = f(a) + \ell(h) + o(\|h\|) \text{ quand } h \to 0_E.$$

Démonstration.

On suppose f différentiable en a. Soit $\ell, \ell' \in \mathcal{L}(E, F)$ tels que :

$$f(a+h) = f(a) + \ell(h) + o(\|h\|) \text{ et } f(a+h) = f(a) + \ell'(h) + o(\|h\|)$$

Montrons que $\ell = \ell'$ *i.e.* pour tout $x \in E$, $\ell(x) = \ell'(x)$.
Soit $x \in E$. Pour $t \in \mathbb{R}^*$, on a $h = tx \to 0_E$ quand $t \to 0$ et $\|h\| = \|tx\| = |t|.\|x\|$. Par suite, on a :

$$f(a+tx) = f(a) + \ell(tx) + \underset{t\to 0}{o}(t) \text{ et } f(a+tx) = f(a) + \ell'(tx) + \underset{t\to 0}{o}(t).$$

Ainsi, par linéarité de ℓ et ℓ' puis en effectuant la différence de ces deux égalités, on obtient :

$$t(\ell(x) - \ell'(x)) = \ell(tx) - \ell'(tx) = \underset{t\to 0}{o}(t)$$

et donc :

$$\ell(x) - \ell'(x) = \underset{t\to 0}{o}(1) \xrightarrow[t\to 0]{} 0$$

d'où $\ell(x) = \ell'(x)$.
Ceci étant vrai pour tous $x \in E$, il en résulte que $\ell = \ell'$; ce qui prouve l'unicité. □

Le lemme précédent justifie la définition suivante :

Définition 7. *différentielle en un point*

Soit $f : U \to F$ et $a \in U$. Si f est différentiable en a, on appelle **différentielle de f en** a et on note $\mathrm{d}f(a)$ l'unique application linéaire telle que :

$$f(a+h) = f(a) + \mathrm{d}f(a)(h) + o(\|h\|) \text{ quand } h \to 0_E.$$

Exercice 9.

Soit I un intervalle ouvert de $\mathbb{R}$ et $f : I \to F$. Caractériser la différentiabilité de f en a en terme de dérivabilité en a et donner dans ce cas une expression de $\mathrm{d}f(a)$.

Correction.

On commence par démontrer le lemme suivant :

> Soit $\varphi : \mathbb{R} \to \mathbb{R}$. On a : $\varphi \in \mathcal{L}(\mathbb{R}, F)$ si, et seulement si, il existe un unique $u \in F$ tel que $\varphi : x \mapsto xu$.
>
> Démonstration.
>
> $\Leftarrow$ S'il existe un unique $v \in F$ tel que $\varphi : x \mapsto xu$ alors pour tout $\lambda, \mu, x, y \in \mathbb{R}$, on a $\varphi(\lambda x + \mu y) = (\lambda x + \mu y)u = \lambda(xu) + \mu(yu) = \lambda\varphi(x) + \mu\varphi(y)$ donc $\varphi \in \mathcal{L}(\mathbb{R}, F)$.
>
> $\Rightarrow$ On suppose $\varphi \in \mathcal{L}(\mathbb{R}, F)$. On pose $u = \varphi(1) \in F$. Alors, pour tout $x \in \mathbb{R}$, on a, par linéarité de φ,
> $$\varphi(x) = \varphi(x.1) = x\varphi(1) = xu;$$
> d'où $\varphi : x \mapsto xu$. Par construction, $u = \varphi(1)$ est unique.

On a :

$$f \text{ est dérivable en } a$$
$$\text{si, et seulement si,}$$
$$h \mapsto \frac{f(a+h)-f(a)}{h} \text{ admet une limite } \ell \in F \text{ quand } h \to 0$$
$$\text{si, et seulement si,}$$
$$\text{il existe } \ell \in F \text{ tel que } \frac{f(a+h)-f(a)}{h} = \ell + \underset{h\to 0}{o}(1)$$
$$\text{si, et seulement si,}$$
$$\text{il existe } \ell \in \mathbb{R} \text{ tel que } f(a+h) = f(a) + h.\ell + \underset{h\to 0}{o}(h)$$
$$\text{si, et seulement si, (en vertu du lemme initial)}$$
$$\text{il existe } L \in \mathcal{L}(\mathbb{R}, F) \text{ tel que } f(a+h) = f(a) + L(h) + \underset{h\to 0}{o}(|h|)$$
$$\text{si, et seulement si,}$$
$$f \text{ est différentiable en } a.$$

Dans ce cas, on a donc :
$$\mathrm{d}f(a) : h \mapsto L(h) = h.\ell = h.f'(a)$$

Définition 8. *Différentiabilité globale*

Soit $f : U \to F$. On dit que f est **différentiable sur** U si f est différentiable en a pour tout $a \in U$. Dans ce cas, appelle **différentielle de** f **sur** U et on note $\mathrm{d}f$ l'application de U dans $\mathcal{L}(E, F)$ telle que :
$$\mathrm{d}f : x \mapsto \mathrm{d}f(x).$$

b. Exemples

Exemple 5.

Soit $f : U \to F$.
- Si f est constante en $c \in F$, alors f est différentiable sur U et $\mathrm{d}f = \mathbf{0}$.
- Si $f \in \mathcal{L}(E, F)$, alors f est différentiable sur E et, pour tout $a \in E$, $\mathrm{d}f(a) = f$.

Proposition 2. *Différentielle d'une application bilinéaire*

Soit $B : E_1 \times E_2 \to F$ une application bilinéaire. Alors B est différentiable sur $E_1 \times E_2$ et on a, pour tout $(x, y) \in E_1 \times E_2$:
$$dB(x, y) : (h, k) \mapsto B(x, k) + B(h, y).$$

Démonstration.

Soit $(x, y) \in E_1 \times E_2$. On a, pour tout $(h, k) \in E_1 \times E_2$:

$$B(x+h, y+k) - B(x, y) = B(x, k) + B(h, y) + B(h, k).$$

Par linéarité des applications $B(x, \cdot)$ et $B(\cdot, y)$, l'application $\ell : E_1 \times E_2 \to F$ telle que :

$$\ell(h, k) = B(x, k) + B(h, y)$$

est linéaire.
De plus, $E_1 \times E_2$ étant de dimension finie, l'application bilinéaire B est continue sur $E_1 \times E_2$. Ainsi, il existe $M \geq 0$ tel que pour tout $(h, k) \in E_1 \times E_2$:

$$\|B(h, k)\|_F \leq M\|h\|_{E_1}\|k\|_{E_2}$$

En considérant sur $E_1 \times E_2$ la norme produit $\|(h, k)\|_\infty = \max(\|h\|_{E_1}, \|k\|_{E_2})$ on obtient :

$$\|B(h, k)\|_F \leq M\|h\|_{E_1}\|k\|_{E_2} \leq M\|(h, k)\|_\infty^2 = o\left(\|(h, k)\|_\infty\right)$$

Ainsi, on a, pour tout $(h, k) \in E_1 \times E_2$:

$$B(x+h, y+k) - B(x, y) = \ell(h, k) + o\left(\|(h, k)\|\right),$$

où ℓ est une application linéaire.
Il en résulte que B est différentiable sur $E_1 \times E_2$ et pour tout $(x, y) \in E_1 \times E_2$:

$$dB(x, y) = \ell : (h, k) \mapsto B(x, k) + B(h, y).$$

□

Plus généralement, on le résultat suivant :

Proposition 3. *Différentielle d'une application multilinéaire*

Soit $M : E_1 \times \ldots \times E_n \to F$ une application multilinéaire. Alors M est différentiable sur $E_1 \times \ldots \times E_n$ et on a, pour tout $(x_1, \ldots, x_n) \in E_1 \times \ldots \times E_n$:

$$dM(x_1, \ldots, x_n) : (h_1, \ldots, h_n) \mapsto \sum_{i=1}^{n} M(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_n).$$

Exemple 6.

Soit E un espace de dimension finie et $\mathcal{B}$ une base de E. L'application $\det_{\mathcal{B}}$ est différentiable sur E.

c. Exercices

Méthode : calculer une différentielle.
Pour montrer qu'une fonction $f : U \to F$ est différentiable en $a \in U$ et calculer $\mathrm{d}f(a)$ avec la définition :

- on calcule $f(a+h)-f(a)$;
- on "récupère" de ce calcul un partie $\ell(h)$ qui dépend linéairement de h ;
- on montre que $f(a+h)-f(a)-\ell(h)=o(\|h\|)$.

Exercice 10.

Soit $n\in\mathbb{N}^*$ et $f: M_n(\mathbb{R})\to M_n(\mathbb{R})$ telle que $f: M\mapsto M^2$. Montrer que f est différentiable sur $M_n(\mathbb{R})$ et calculer sa différentielle.

Exercice 11.

Soit $n\in\mathbb{N}^*$, $\lambda\in\mathbb{R}$ et Montrer que l'application $\exp: M\mapsto \exp(M)$ est différentiable en $A=\lambda I_n$ et calculer sa différentielle.

Remarque : en fait, on peut montrer que exp *est différentiable sur $M_n(\mathbb{R})$ - mais comme en général, A et H ne commutent pas, ce n'est pas aussi simple que dans le cas précédent.*

Démonstration.

On considère une norme $\|\cdot\|$ sous-multiplicative sur $M_n(\mathbb{R})$. Soit $H\in M_n(\mathbb{R})$. Comme $A=\lambda I_n$, A et H commutent, $\exp(A+H)=\exp(A)\exp(H)$. Ainsi, on a :

$$\exp(A+H)-\exp(A)=\exp(A)(\exp(H)-I_n)=\exp(A)\sum_{k=1}^{+\infty}\frac{H^k}{k!}=\exp(A)H+\exp(A)\sum_{k=2}^{+\infty}\frac{H^k}{k!}.$$

L'application $\ell: H\mapsto \exp(A)H$ est linéaire et on a :

$$\begin{aligned}\left\|\sum_{k=2}^{+\infty}\frac{H^k}{k!}\right\| &= \left\|H^2.\sum_{k=2}^{+\infty}\frac{H^{k-2}}{k!}\right\|\\ &\leq \|H\|^2\sum_{k=2}^{+\infty}\frac{\|H\|^{k-2}}{k!}\\ \left\|\sum_{k=2}^{+\infty}\frac{H^k}{k!}\right\| &\leq \|H\|^2\sum_{k=2}^{+\infty}\frac{\|H\|^{k-2}}{(k-2)!}=\|H\|^2e^{\|H\|}\end{aligned}$$

et ainsi $\displaystyle\sum_{k=2}^{+\infty}\frac{H^k}{k!}\underset{h\to 0_E}{=}o(\|H\|)$ car $\dfrac{\|H\|^2e^{\|H\|}}{\|H\|}=\|H\|e^{\|H\|}\xrightarrow[h\to 0_E]{}0$.

Par suite, exp est différentiable en $A=\lambda I_n$ et comme $\exp(A)=\exp(\lambda I_n)=e^{\lambda}I_n$, on a, pour tout $H\in M_n(\mathbb{R})$:

$$d\exp(A)(H)=\exp(A)H=e^{\lambda}H.$$

$\square$

Exercice 12.

Soit $n \in \mathbb{N}^*$ et $\|\cdot\|$ une norme sous-multiplicative sur $E = M_n(\mathbb{R})$.

1. Soit $A \in E$ telle que $\|A\| < 1$.
   (a) Montrer que $\sum A^k$ converge.
   (b) Montrer que $I_n - A$ est inversible et déterminer son inverse.
2. Montrer que la fonction $f : GL_n(\mathbb{R}) \to E$ définie, pour $M \in GL_n(\mathbb{R})$, par $f(M) = M^{-1}$ est différentiable sur $GL_n(\mathbb{R})$ et déterminer sa différentielle en tout point.

Correction.

1. Soit $A \in E$ telle que $\|A\| < 1$.
   (a) On a, pour tout $k \in \mathbb{N}$, $\|A^k\| \leq \|A\|^k$ qui est le terme général d'une série géométrique convergente car $\|A\| < 1$, d'où, par comparaison, $\sum A^k$ converge absolument et donc converge.
   (b) On remarque que :
   $$(I_n - A)\sum_{k=0}^{+\infty} A^k = \sum_{k=0}^{+\infty} A^k - \sum_{k=0}^{+\infty} A^{k+1} = I_n,$$
   donc $I_n - A$ est inversible d'inverse $\sum_{k=0}^{+\infty} A^k$.
2. Soit $M \in GL_n(\mathbb{R})$ et $H \in E$ tel que $\|M^{-1}H\| < 1$. Alors $M + H$ est inversible ; en effet, on a $M + H = M(I_n + M^{-1}H)$, or M est inversible et $I_n + M^{-1}H$ l'est aussi d'après la question précédente avec $A = -M^{-1}H$; d'où $M + H$ est inversible comme produit de matrices inversibles.
   Ainsi, $f(M + H)$ est bien défini, et on a :
   $$\begin{aligned}
   f(M+H) - f(M) &= (M+H)^{-1} - M^{-1} \\
   &= ((I_n + M^{-1}H)^{-1} - I_n)M^{-1} \\
   &= \left(\left(\sum_{k=0}^{+\infty}(-M^{-1}H)^k\right) - I_n\right)M^{-1} \\
   &= \left(\sum_{k=1}^{+\infty}(-M^{-1}H)^k\right)M^{-1} \\
   f(M+H) - f(M) &= -M^{-1}HM^{-1} + \underbrace{\left(\sum_{k=2}^{+\infty}(-M^{-1}H)^k\right)M^{-1}}_{\varepsilon(H)}
   \end{aligned}$$
   Par propriétés du produit matriciel, l'application $\ell : H \mapsto -M^{-1}HM^{-1}$ est linéaire.
   De plus, pour tout $H \in E$ tel que $\|H\| < \|M\|$, on a $\|M^{-1}H\| \leq \|M^{-1}\|.\|H\| < 1$ *(exercice :*

montrer cette affirmation) et :

$$\begin{aligned}
\|\varepsilon(H)\| &\leq \sum_{k=2}^{+\infty} \|(M^{-1}H)^k\| \\
&\leq \left(\|M^{-1}\|.\|H\|\right)^2 \sum_{k=0}^{+\infty} \left(\|M^{-1}\|.\|H\|\right)^k \\
\|\varepsilon(H)\| &\leq \|H\|^2.\frac{\|M^{-1}\|^2}{1-\|M^{-1}\|.\|H\|};
\end{aligned}$$

par suite, $\dfrac{\|\varepsilon(H)M^{-1}\|}{\|H\|} \leq \|H\|.\dfrac{\|M^{-1}\|^3}{1-\|M^{-1}\|.\|H\|} \xrightarrow[\|H\|\to 0]{} 0$.
Ainsi, pour tout $M \in GL_n(\mathbb{R})$, pour tout $H \in E$ tel que $\|H\| < \|M\|$, on a $f(M+H) = f(M) + \ell(M) + o(\|H\|)$ avec ℓ linéaire ; donc f est différentiable sur $G_n(\mathbb{R})$ et :

$$\mathrm{d}f(M) : H \mapsto -M^{-1}HM^{-1}.$$

2. Différentielle et dérivées partielles

Proposition 4. *Différentielle et dérivée suivant un vecteur*

Soit $f : U \to F$ et $a \in U$. Si f est différentiable en a, alors f admet une dérivée en a selon tout vecteur de u et on a :

$$D_u f(a) = \mathrm{d}f(a)(u).$$

Démonstration.

On suppose f est différentiable en a. Alors, pour tout $h \in E$ tel que $a+h \in U$, on a :

$$f(a+h) - f(a) = \mathrm{d}f(a)(h) + \underset{\|h\|\to 0}{o}(\|h\|).$$

Soit $u \in E \setminus \{0_E\}$ (le cas $u = 0_E$ est trivial : toute fonction est dérivable suivant 0_E de dérivée égale à 0_F.
On remarque alors que, pour un réel t strictement positif, $\|tu\| \to 0$ si, et seulement si, $t \to 0$; ainsi, en utilisant la formule précédente appliqué à $h = tu$, on obtient :

$$f(a+tu) - f(a) = \mathrm{d}f(a)(tu) + \underset{t\to 0}{o}(\|tu\|) = t.\,\mathrm{d}f(a)(u) + \underset{t\to 0}{o}(t)$$

Ainsi,

$$\frac{1}{t}(f(a+tu) - f(a)) = \mathrm{d}f(a)(u) + \underset{t\to 0}{o}(1) \xrightarrow[t\to 0]{} \mathrm{d}f(a)(u).$$

Il en résulte que f admet une dérivée en a suivant h et que :

$$D_u f(a) = \lim_{t\to 0} \frac{1}{t}(f(a+tu) - f(a)) = \mathrm{d}f(a)(u).$$

□

Proposition 5. *Lien entre différentielle et dérivées partielles*

Soit $f : U \to F$, $a \in U$ et $\mathcal{B} = (e_1, ..., e_p)$ une base de E.
Si f est différentiable en a, alors, pour tout $h = \sum_{j=1}^{p} h_j e_j \in E$, on a :

$$\mathrm{d}f(a)(h) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a).$$

Démonstration.

On considère une base $\mathcal{B} = (e_1, ..., e_p)$ de E. Comme f est différentiable en a, d'après la proposition 4, f admet des dérivées en a suivant chacun des e_j et on a :

$$\frac{\partial f}{\partial x_j}(a) = D_{e_j} f(a) = \mathrm{d}f(a)(e_j).$$

Ainsi, pour $h = \sum_{j=1}^{p} h_j e_j \in E$, on a, par linéarité de la différentielle :

$$\mathrm{d}f(a)(h) = \mathrm{d}f(a)\left(\sum_{j=1}^{p} h_j e_j\right) = \sum_{j=1}^{p} h_j \,\mathrm{d}f(a)(e_j) = \sum_{j=1}^{p} h_j \frac{\partial f}{\partial x_j}(a).$$

□

Corollaire 1. *Lien entre différentielle et dérivées partielles*

Soit $f : U \to F$, $a \in U$ et $\mathcal{B} = (e_1, ..., e_p), \mathcal{C} = (\varepsilon_1, ..., \varepsilon_q)$ des bases de E et F respectivement.
Si f est différentiable en a, alors :

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\mathrm{d}f(a)) = Jf(a).$$

où $Jf(a)$ est la matrice jacobienne de f en a dans les bases $\mathcal{B}$ et $\mathcal{C}$

3. Opérations sur les applications différentiables

Proposition 6. *Linéarité*

Soit $\lambda, \mu \in \mathbb{R}$, $f, g : U \to F$ et $a \in U$.

— Si f et g sont différentiables en a, alors $\lambda f + \mu g$ est différentiable en a et on a :

$$d(\lambda f + \mu g)(a) = \lambda \,\mathrm{d}f(a) + \mu \,\mathrm{d}g(a).$$

— L'ensemble des applications différentiables en a (resp. sur U) est un espace vectoriel.

Proposition 7. *Produit d'applications différentiables à valeurs réelles*

Soit $f, g : U \to \mathbb{R}$ et $a \in U$.
Si f et g sont différentiables en a, alors fg est différentiable en a et on a :

$$d(fg)(a) = g(a)\,\mathrm{d}f(a) + f(a)\,\mathrm{d}g(a).$$

Correction.

On suppose f et g différentiables en a. On a :

$$\begin{aligned}(fg)(a+h) - (fg)(a) &= f(a+h)g(a+h) - f(a)g(a)\\ &= (f(a) + \mathrm{d}f(a)(h) + o(\|h\|))\,(g(a) + \mathrm{d}g(a)(h) + o(\|h\|)) - f(a)g(a)\\ &= g(a)\,\mathrm{d}f(a)(h) + f(a)\,\mathrm{d}g(a)(h) + (\,\mathrm{d}f(a)(h) + \mathrm{d}g(a)(h))o(\|h\|)\\ &\quad + \mathrm{d}f(a)(h)\,\mathrm{d}g(a)(h) + o(\|h\|^2).\end{aligned}$$

L'application $h \mapsto g(a)\,\mathrm{d}f(a)(h) + f(a)\,\mathrm{d}g(a)(h)$ est linéaire comme combinaison linéaire des applications linéaires $\mathrm{d}f(a)$ et $\mathrm{d}g(a)$. De plus, comme E est de dimension finie, $\mathrm{d}f(a)$ et $\mathrm{d}g(a)$ sont des applications linéaires continues d'où :

$$|\,\mathrm{d}f(a)(h) + \mathrm{d}g(a)(h)| \leq |\,\mathrm{d}f(a)(h)| + |\,\mathrm{d}g(a)(h)| \leq \underbrace{(|||\,\mathrm{d}f(a)||| + |||\,\mathrm{d}g(a)|||)}_{=C}\,\|h\|$$

et

$$|\,\mathrm{d}f(a)(h)\,\mathrm{d}g(a)(h)| = |\,\mathrm{d}f(a)(h)|.|\,\mathrm{d}g(a)(h)| \leq \underbrace{(|||\,\mathrm{d}f(a)|||.|||\,\mathrm{d}g(a)|||)}_{=D}\,\|h\|^2$$

et donc :

$$\frac{\left|(\,\mathrm{d}f(a)(h) + \mathrm{d}g(a)(h))o(\|h\|) + \mathrm{d}f(a)(h)\,\mathrm{d}g(a)(h) + o(\|h\|^2)\right|}{\|h\|} = o(\|h\|) \xrightarrow[h\to 0_E]{} 0$$

Par suite, on a :

$$(fg)(a+h) = (fg)(a) + \ell(h) + o(\|h\|)$$

avec $\ell = g(a)\,\mathrm{d}f(a) + f(a)\,\mathrm{d}g(a) \in \mathcal{L}(E, \mathbb{R})$ d'où fg est différentiable en a et

$$\mathrm{d}(fg)(a) = g(a)\,\mathrm{d}f(a) + f(a)\,\mathrm{d}g(a).$$

Proposition 8. *Règle de la chaîne*

Soit $f : U \to F$, $g : V \to G$ telles que $f(U) \subset V$ et $a \in U$.
Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et on a :

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a).$$

Exercice 13.

Soit $n \in \mathbb{N}^*$ et $E = M_n(\mathbb{R})$. Montrer que l'application $g : E \to \mathbb{R}$ définie, pour $M \in E$, par $g(M) = \mathrm{Tr}(M^2)$ est différentiable sur E et déterminer sa différentielle en tout point.

Corollaire 2. *Jacobienne d'une composée*

Soit $f : U \to F$, $g : V \to G$ telles que $f(U) \subset V$, $a \in U$ et $\mathcal{B}, \mathcal{C}, \mathcal{D}$ des bases de E, F, G respectivement.
Si f est différentiable en a et g est différentiable en $f(a)$, alors :

$$J(g \circ f)(a) = Jg(f(a)) \times Jf(a).$$

Exercice 14.

Soit $f : (x, y) \mapsto (y, x, x + y)$ et $g : (x, y, z) \mapsto (x^2 + y^2, z^2)$. Montrer que $g \circ f$ est différentiable sur $\mathbb{R}^2$ et déterminer sa différentielle en tout point de $\mathbb{R}^2$ de deux manières : en calculant $g \circ f$ et avec le produit des matrices jacobiennes.

Proposition 9. *Différentielle d'une inverse*

Soit $f : U \to \mathbb{R}$ et $a \in U$.
Si f est différentiable en a et $f(a) \neq 0$, alors $\frac{1}{f}$ est différentiable en a et :

$$\mathrm{d}\left(\frac{1}{f}\right)(a) = -\frac{1}{f^2(a)}\,\mathrm{d}f(a).$$

Exercice 15.

Justifier que $f : (x, y) \mapsto \dfrac{x}{x^2 + y^2}$ est C^1 sur $U = R^2 \setminus \{(0, 0)\}$ et déterminer sa différentielle sur U. Est-elle différentiable en $(0, 0)$?

4. Fonctions de classe C^1

a. Définition et premiers exemples

Définition 9. *Fonction de classe C^1*

Soit $f : U \to F$. On dit que f est de **classe C^1 sur** U si f est différentiable sur U et $\mathrm{d}f$ est continue sur U.

Exemple 7.

Soit $f : U \to F$.
- Si f est constante en $c \in F$, alors f est de classe C^1 sur U.
- Si $f \in \mathcal{L}(E, F)$, alors f est de classe C^1 sur E.

Exercice 16.

On suppose que E est un espace euclidien. Soit $f : x \mapsto \|x\|^2$.
1. Montrer que f est différentiable sur E et déterminer sa différentielle.
2. En déduire que f est de classe C^1 sur E.

Correction.

1. On a, pour tous $a, h \in E$,
$$f(a+h) - f(a) = (a+h|a+h) - (a|a) = 2(a|h) + \|h\|^2$$
On pose $\ell : h \mapsto 2(a|h)$. Alors ℓ est une application linéaire de E dans $\mathbb{R}$ par linéarité du produit scalaire par rapport à la deuxième variable et :
$$\frac{\|h\|^2}{\|h\|} = \|h\| \xrightarrow[h \to 0_E]{} 0$$
et d'où $f(a+h) = f(a) + 2(a|h) + o(\|h\|)$ et donc f admet un développement limité à l'ordre 1 en a.
Il en résulte que, pour tout $a \in E$, f est différentiable en a et on a $\mathrm{d}f(a) = \ell : h \mapsto 2(a|h)$.
Ainsi, f est différentiable sur E et $\mathrm{d}f : a \mapsto \mathrm{d}f(a) : h \mapsto 2(a|h)$
2. Montrons que $\mathrm{d}f$ est continue sur E (dans $\mathcal{L}(E, \mathbb{R})$). On remarque que, pour tout $a, b, h \in E$ et $\lambda, \mu \in \mathbb{R}$, par linéarité du produit scalaire par rapport à sa première variable :
$$\mathrm{d}f(\lambda a + \mu b)(h) = 2(\lambda a + \mu b|h) = \lambda \times 2(a|h) + \mu \times 2(b|h) = \lambda\, \mathrm{d}f(a)(h) + \mu\, \mathrm{d}f(b)(h)$$
et donc $\mathrm{d}f(\lambda a + \mu b) = \lambda\, \mathrm{d}f(a) + \mu\, \mathrm{d}f(b)$. Ainsi, $\mathrm{d}f$ est une application linéaire définie sur E qui est de dimension finie, donc $\mathrm{d}f$ est continue sur E.
Par suite, f est de classe C^1 sur E.

b. Opérations sur les fonctions de classe C^1

Proposition 10.

L'ensemble $C^1(U, F)$ est un espace vectoriel.

Question 1.

De quelle structure peut-on munir l'ensemble $C^1(U, \mathbb{R})$?

Proposition 11.

Soit $\mathcal{B} = (e_1, ...e_n)$ une base de E. Pour tout $i \in [\![1, n]\!]$, la i-ème application coordonnée $\varphi_i : x = \sum_{j=1}^{n} x_j e_j \mapsto x_i$ est de classe C^1 de E dans $\mathbb{R}$.

Démonstration.

L'application φ_i est linéaire et donc de classe C^1 sur E (voire exemple 7). □

Théorème 1. *Applications polynomiales*

Les applications polynomiales sur E sont de classe C^1 sur E.

Démonstration.

Une application polynomiale sur E est de classe C^1 sur E comme combinaison linéaire de produits d'applications coordonnées qui sont de classe C^1 sur E d'après la proposition 11. □

Exercice 17.

Justifier que les applications suivantes sont de classe C^1 sur E et déterminer leurs différentielles.

1. $f : (x, y) \mapsto (x^2 + y)^2$ et $E = \mathbb{R}^2$.
2. $g : (x, y, z) \mapsto x^3 + y^3 + z^3 + xyz$ et $E = \mathbb{R}^3$.

Proposition 12.

Soit $f : U \to F$, $g : V \to G$ telles que $f(U) \subset V$. Si f est de classe C^1 sur U et g est de classe C^1 sur V, alors $g \circ f$ est de classe C^1 sur U.

Proposition 13.

Soit $f : U \to \mathbb{R}$. Si f est C^1 et ne s'annule pas sur U, alors $\frac{1}{f}$ est C^1 sur U.

c. Théorème fondamental

Théorème 2. *Théorème fondamental*

Soit $\mathcal{B} = (e_1, ..., e_p)$ une base de E et $f : U \to F$. On a l'équivalence suivante :

i) la fonction f est de classe C^1 sur U.

ii) les dérivées partielles $\dfrac{\partial f}{\partial x_j}$ sont définies et continues sur U.

Exercice 18.

Montrer que la fonction $f : (x, y) \mapsto \dfrac{x^4 + y^4}{x^2 + y^2}$ admet un prolongement C^1 sur $\mathbb{R}^2$.

5. Arcs paramétrés et dérivées le long d'un arc

a. Arc paramétré

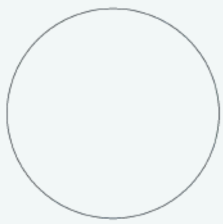
Dans ce paragraphe, I désigne un intervalle non vide de $\mathbb{R}$.

Définition 10. *Arc paramétré*

Une fonction $\gamma : I \to E$ est appelé **arc paramétré** et on appelle **support** de γ l'image $\gamma(I)$ de γ.

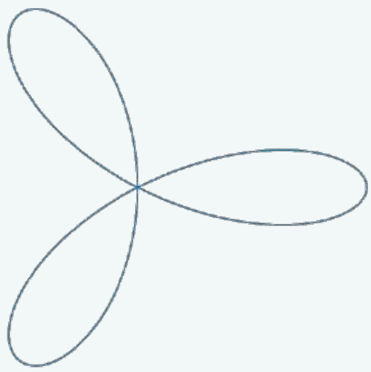
Exemple 8.

1. *Cercle :* L'arc paramétré $\gamma : t \mapsto (\cos(t), \sin(t))$ a pour support le cercle de centre $(0, 0)$ et de rayon 1 :



2. *Cycloïde :* L'arc paramétré $\gamma : t \mapsto (t - \sin(t), 1 - \cos(t))$ a pour support :

3. *Trifolium :* L'arc paramétré $\gamma : t \mapsto (\cos(2t) - \cos(t), \sin(2t) + \sin(t))$ a pour support :



4. *Folium :* L'arc paramétré $\gamma : t \mapsto (\frac{1-t^2}{1+t^2}, t\frac{1-t^2}{1+t^2})$ a pour support :

b. Opérations le long d'un arc

Proposition 14. *Dérivée le long d'un arc*

Soit $f : U \to F$ et $\gamma : I \to U$ un arc paramétré tel que $\gamma(I) \subset U$.

— Soit $t_0 \in I$. Si f est différentiable en $\gamma(t_0)$ et γ est dérivable en t_0, alors $f \circ \gamma$ est dérivable en t_0 et on a :
$$(f \circ \gamma)'(t_0) = \mathrm{d}f(\gamma(t_0))(\gamma'(t_0)).$$

— Si f et γ sont de classe C^1 sur respectivement U et I, alors $f \circ \gamma$ est de classe C^1 sur I.

Démonstration.

— D'après la proposition 8, $f \circ \gamma$ est différentiable en t_0 et donc, d'après l'exercice 9, $f \circ \gamma$

est dérivable en t_0 et, pour tout $h \in \mathbb{R}$:

$$\begin{aligned} h.(f \circ \gamma)'(t_0) &= \mathrm{d}(f \circ \gamma)(t_0)(h) \\ &= (\mathrm{d}f(\gamma(t_0)) \circ \mathrm{d}\gamma(t_0))(h) \\ &= \mathrm{d}f(\gamma(t_0))(\mathrm{d}\gamma(t_0)(h)) \\ &= \mathrm{d}f(\gamma(t_0))(h.\gamma'(t_0)) \\ &= h.\mathrm{d}f(\gamma(t_0))(\gamma'(t_0)) \end{aligned}$$

et donc, pour $h = 1 \in \mathbb{R}$, on obtient $(f \circ \gamma)'(t_0) = \mathrm{d}f(\gamma(t_0))(\gamma'(t_0))$.

— Cela découle directement de la proposition 12.

□

Corollaire 3.

Soit $f : U \to F$, $a, u \in E$ et $\gamma : I \to U$ l'arc paramétré définie, pour $t \in \mathbb{R}$ par $\gamma(t) = a + tu$. Si f est différentiable sur U, alors, pour tout $t \in I$:

$$(f \circ \gamma)'(t) = \mathrm{d}f(\gamma(t))(u).$$

Proposition 15. *Intégration le long d'un arc*

Soit $a, b \in U$ et $f : U \to F$ une fonction de classe C^1 sur U. Pour tout arc paramétré $\gamma : [0,1] \to U$ de classe C^1 sur I tel que $\gamma(0) = a$ et $\gamma(1) = b$, on a :

$$f(b) - f(a) = \int_0^1 \mathrm{d}f(\gamma(t))(\gamma'(t))\,\mathrm{d}t.$$

Démonstration.

Soit $\gamma : [0,1] \to U$ un arc paramétré de classe C^1 sur I tel que $\gamma(0) = a$ et $\gamma(1) = b$. Alors, d'après la proposition 14, $f \circ \gamma$ est dérivable sur $[0,1]$ et on a, pour tout $t \in [0,1]$:

$$(f \circ \gamma)'(t) = \mathrm{d}f(\gamma(t))(\gamma'(t))$$

Ainsi, d'après le théorème fondamental de l'analyse, on a :

$$\begin{aligned} \int_0^1 \mathrm{d}f(\gamma(t))(\gamma'(t))\,\mathrm{d}t &= \int_0^1 (f \circ \gamma)'(t)\,\mathrm{d}t \\ &= f \circ \gamma(1) - f \circ \gamma(0) \\ \int_0^1 \mathrm{d}f(\gamma(t))(\gamma'(t))\,\mathrm{d}t &= f(b) - f(a). \end{aligned}$$

□

Corollaire 4.

Soit $f : U \to F$. Si U **connexe par arcs** alors :
la fonction f est constante sur U si, et seulement si, $\mathrm{d}f = \mathbf{0}$.

Remarque 4.

Si U n'est pas connexe par arcs, l'implication réciproque est fausse en général : par exemple, si $U =]0,1[\sqcup]2,3[$, la fonction qui vaut 0 sur $]0,1[$ et 1 sur $]0,1[$ n'est pas constante sur U mais est différentiable sur U de différentielle nulle sur U.

6. Vecteurs tangents à une partie

Définition 11. *Vecteur tangent*

Soit X une partie de E et x un point de X.

— Soit $v \in E$ un vecteur. On dit que v est **tangent à** X **en** x s'il existe $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\to X$ un arc paramétré dérivable en 0 tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

— On note T_xX l'ensemble des vecteurs tangents à X en x.

Remarque 5.

Si pour v dans E, il existe un arc défini sur $\mathbb{R}$ ou, plus généralement, sur un intervalle dont 0 est un point intérieur, dérivable à 0 avec $\gamma(0) = x$ et $\gamma'(0) = v$, alors v est tangent à X en x : en effet, il suffit de restreindre l'arc à un intervalle de la forme $]-\varepsilon, \varepsilon[$ pour satisfaire à la définition.

Exemple 9.

— Pour tous $X \subset E$ et $x \in X$, 0_E appartient à T_xX.

— Soit $\mathcal{V}$ un sous-espace affine de E dirigé par un sous-espace vectoriel V de E. Alors, pour tout $x \in \mathcal{V}$, $T_x\mathcal{V} = V$.

— On suppose euclidien et on pose $S = S(0_E, 1)$ la sphère unité de E. Alors, pour $x \in S$, $T_xS = \{x\}^\perp$.

Remarque : très souvent, pour déterminer un plan tangent, on conjecture son expression puis on raisonne par double inclusion.

— On considère l'arc γ constant en x de $\mathbb{R}$ dans E : il est dérivable sur $\mathbb{R}$ de dérivée nulle, son image est inclus dans X car $x \in X$ et $\gamma(0) = x$, $\gamma'(0) = 0_E$. D'où $0_E \in T_xX$.

— ⋆ Montrons $V \subset T_x\mathcal{V}$.
Soit $v \in V$. On pose $\gamma : t \mapsto x + tv \in \mathcal{V}$. Alors γ est de classe C^∞ sur $\mathbb{R}$ donc dérivable sur 0 et à valeurs dans $\mathcal{V}$; de plus $\gamma(0) = x$ et $\gamma'(0) = v$ car γ' est constante en v. Par suite, $v \in T_x\mathcal{V}$.

⋆ Montrons $T_x\mathcal{V} \subset V$.

Soit $v \in T_x\mathcal{V}$. Alors il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to \mathcal{V}$ tel que $\gamma(0) = x$ et $\Gamma'(0) = v$.
Pour tout $t \in]-\varepsilon, \varepsilon[$ avec $t \neq 0$, comme $\gamma(t), \gamma(0) \in V$, la combinaison linéaire :

$$\frac{1}{t}(\gamma(t) - \gamma(0)) = \frac{1}{t}\gamma(t) - \frac{1}{t}\gamma(0) \text{ appartient à } V.$$

Or E est de dimension finie, donc le sous-espace vectoriel V est fermé dans E (pour n'importe quelle norme et donc celle qu'on s'est fixée au départ). Par suite, par caractérisation séquentielle des fermés (quitte à prendre une suite t_n tendant vers 0) :

$$\gamma'(0) = \lim_{t \to 0} \underbrace{\frac{1}{t}(\gamma(t) - \gamma(0))}_{\in V} \in V.$$

Ainsi, par double inclusion, on a $T_x\mathcal{V} = V$.

— ⋆ Montrons $\{x\}^\perp \subset T_xS$.
Soit $v \in \{x\}^\perp$. On pose $g : t \mapsto x + tv$. Alors g à valeurs dans $E \setminus \{0_E\}$: en effet, pour tout $t \in \mathbb{R}$, d'après le théorème de Pythgore, on a :

$$\|g(t)\|^2 = \|x + tv\|^2 = \|x\|^2 + |t|\|v^2\| > 0 \text{ car } \|x\| = 1.$$

Ainsi, l'arc $\gamma : t \to \frac{g(t)}{\|g(t)\|}$ est bien défini sur $\mathbb{R}$. Montrons que γ est dérivable sur $\mathbb{R}$ et donc en 0.
On a $\gamma = g \times (\varphi \circ g)$ où $\varphi = \frac{1}{\|\cdot\|} = (\cdot)^{-\frac{1}{2}} \circ \|\cdot\|^2$ est différentiable sur $E \setminus \{0_E\}$ d'où $\varphi \circ g$ est dérivable sur $\mathbb{R}$ d'après la proposition 14 et donc γ est dérivable sur $\mathbb{R}$ comme produit de fonctions dérivables sur $\mathbb{R}$.
De plus, on a, pour tous $a, h \in E$ avec $a \neq 0_E$:

$$\begin{aligned}
\mathrm{d}\varphi(a)(h) &= (\,\mathrm{d}(\cdot)^{-\frac{1}{2}}(\|a\|^2) \circ \mathrm{d}\|\cdot\|^2(a))(h) \\
&= -\tfrac{1}{2}\|a\|^{-3}.(2(a|h)) \\
\mathrm{d}\varphi(a)(h) &= -\frac{(a|h)}{\|a\|^3}.
\end{aligned}$$

Ainsi, on a, d'après la proposition 14 et comme $(x|v) = 0$:

$$(\varphi \circ g)'(t) = \,\mathrm{d}f(g(t))(g'(t)) = -\frac{(x + tv|v)}{\|x + tv\|^3} = -\frac{t\|v\|^2}{\|x + tv\|^3}.$$

Par suite, comme $\|x\| = 1$,

$$\begin{aligned}
\gamma'(0) &= g'(0)\varphi(g(0)) + \gamma(0)(\varphi \circ g)'(0) \\
&= v.\tfrac{1}{\|x\|} + x.0 \\
\gamma'(0) &= v
\end{aligned}$$

Ainsi, l'arc γ est dérivable en 0, $\gamma(0) = x$ et $\gamma'(0) = v$; d'où $v \in T_xS$.

⋆ Montrons $T_xS \subset \{x\}^\perp$.
Soit $v \in T_xS$. Alors il existe $\varepsilon > 0$ et un arc $\gamma :]-\varepsilon, \varepsilon[\to S$ tel que $\gamma(0) = x$ et $\gamma'(0) = v$.
Montrons que $(x|v) = 0$.
Comme $\mathrm{Im}(\gamma) \subset S$, la fonction $f : t \mapsto \|\gamma(t)\|^2$ est constante sur $]-\varepsilon, \varepsilon[$ et donc

dérivable, de dérivée nulle en 0. Ainsi, γ étant dérivable en 0 et $\|\cdot\|^2$ est différentiable sur E, on a, d'après la proposition 14 :

$$0 = f'(0) = \mathrm{d}\|\cdot\|^2(\gamma(0))(\gamma'(0)) = 2(\gamma(0)|\gamma'(0)) = 2(x|v).$$

Donc $v \in \{x\}^\perp$.

Ainsi, par double inclusion, on a $T_xS = \{x\}^\perp$.

7. Fonctions de classe C^k

a. Dérivées partielles successives

Définition 12.

Soit $f : U \to F$, $\mathcal{B} = (e_1, ..., e_p)$ une base de E, $a \in U$, $k \in \mathbb{N}^*$ et $j_1, ..., j_k \in [\![1, p]\!]$.
Sous réserve d'existence, on appelle **dérivée partielle partielle de f d'ordre k selon les indices** $(j_1, ..., j_k)$ le vecteur $D_{j_1}(D_{j_2}(...(D_{j_k}f(a))...))$ et on note :

$$\frac{\partial^k f}{\partial x_{j_1} ... \partial x_{j_k}} = D_{j_1}(D_{j_2}(...(D_{j_k}f(a))...)).$$

Exercice 19.

Calculer les dérivées secondes dans la base canonique de $f : (x, y) \mapsto (x^2y, e^{xy+1})$.

b. Fonction de classes C^k

Définition 13.

Soit $f : U \to F$. Pour $k \in \mathbb{N}^*$, on dit que f est de **classe C^k sur U** si les dérivées partielles de f d'ordre k existent et sont continues sur U.
Si f est de classe C^k sur U pour tout $k \in \mathbb{N}^*$, alors on dit que f est de **classe C^∞ sur U**.

Proposition 16. *Opérations sur les fonctions de classe C^k*

Soit $k \in \mathbb{N}^* \cup \{\infty\}$. Les sommes, composées, produits (avec $F = \mathbb{R}$ pour le produit) de fonctions de classe C^k sur U sont de classe C^k sur U.

Théorème 3. *Applications polynomiales*

Les applications polynomiales sur E sont de classe C^∞ sur E.

c. Théorème de Schwarz

Théorème 4. *Théorème de Schwarz*

Soit $f : U \to F$ de classe C^2 sur U et $\mathcal{B} = (e_1, ..., e_p)$ une base de E. Pour tout $i, j \in [\![1, p]\!]$, on a :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Exercice 20.

Déterminer s'il existe une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 sur $\mathbb{R}^2$ telle que $V = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial t}\right)$; et dans ce cas, déterminer f :

1. $V : (x, t) \mapsto (xt^2, -xt)$
2. $V : (x, t) \mapsto (2xt, x^2 + t)$.

Partie D

Fonctions numériques

Dans cette partie E désigne un espace euclidien et on note $(\cdot|\cdot)$ son produit scalaire.

1. Gradient

a. Définition et premières propriétés

Lemme 2.

Soit $f : U \to \mathbb{R}$ et $a \in U$. Si f est différentiable en a, il existe une unique vecteur $u \in E$ tel que, pour tout $h \in E$:

$$\mathrm{d}f(a)(h) = (u|h).$$

Définition 14. *Gradient d'une application numérique*

Soit $f : U \to \mathbb{R}$ et $a \in U$. Si f est différentiable en a, on appelle **gradient de f en** a et on note $\nabla f(a)$ l'unique vecteur de E tel que, pour tout $h \in E$:

$$\mathrm{d}f(a)(h) = (\nabla f(a)|h).$$

Proposition 17.

Soit $f : U \to \mathbb{R}$. Si f est de classe C^1 sur U, alors $\nabla f : x \mapsto \nabla f(x)$ est continue sur U.

Proposition 18.

Soit $f : U \to \mathbb{R}$, $\mathcal{B} = (e_1, ..., e_p)$ une base de E et $a \in U$. Si f est différentiable en a alors :

$$\nabla f(a) = \sum_{j=1}^{p} \frac{\partial f}{\partial x_j}(a) e_j$$

En particulier, si $E = \mathbb{R}^p$ muni de sa base canonique :

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a) \right).$$

Exercice 21.

Justifier que $f : (x, y) \to y^2(y^2 - x^4)$ est différentiable sur $\mathbb{R}^2$ et déterminer le gradient de f en tout point de $\mathbb{R}^2$.

b. Opérations et gradient

Proposition 19.

Soit $f, g : U \to \mathbb{R}$. Si f et g sont différentiable sur U alors on a :

— $\nabla(f + g) = \nabla f + \nabla g$;

— $\nabla(fg) = f\nabla g + g\nabla f$;

— pour $\varphi : I \to \mathbb{R}$ dérivable sur I et telle que $f(U) \subset I$,
$$\nabla(\varphi \circ f) = (\varphi' \circ f)\nabla f;$$

— si f ne s'annule pas sur U,
$$\nabla\left(\frac{1}{f}\right) = -\frac{1}{f^2}\nabla f.$$

c. Interprétation géométrique

Proposition 20.

Soit $f : U \to \mathbb{R}$ et $a \in U$. On suppose que f est différentiable en a et que $\nabla f(a) \neq 0$. Alors la fonction $h \mapsto D_h f(a)$ restreinte à la sphère unité de E admet un unique maximum en $h_0 = \dfrac{1}{\|\nabla f(a)\|}\nabla f(a)$.

Démonstration.

On a, pour tout $h \in S_E = \{u \in E \mid \|u\| = 1\}$, d'après l'inégalité de Cauchy-Schwarz,
$$D_h f(a) = \mathrm{d}f(a)(h) = (\nabla f(a)|h) \leq \|\nabla f(a)\|.\|h\| = \|\nabla f(a)\|.$$
avec égalité, si, et seulement si, h est colinéaire et de même sens que $\nabla f(a)$

Par suite, $h \mapsto D_h f(a)$ admet un maximum en $h_0 = \dfrac{1}{\|\nabla f(a)\|}\nabla f(a)$. □

2. Vecteurs tangents d'une fonction numérique

Exemple 10.

Soit U un ouvert de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ une fonction de classe C^1 sur U.
Pour $X = \mathcal{G}_f \subset \mathbb{R}^3$ le graphe de f et $M_0 = (x_0, y_0, z_0) \in X$, on a, en considérant $\mathbb{R}^3$ muni de son

produit scalaire canonique :

$$T_{M_0}X = \{n\}^{\perp} \text{ où } n = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1\right).$$

De plus, pour $g : (x, y, z) \mapsto f(x, y) - z$, on a $X = \{(x, y, z) \in U \times \mathbb{R} \mid g(x, y, z) = 0\}$ et $\nabla g(M_0) = n$ d'où :

$$T_{M_0}X = \{\nabla g(M_0)\}^{\perp} = \operatorname{Ker}(\,\mathrm{d}g(M_0)).$$

- ⋆ Montrons $\{n\}^{\perp} \subset T_{M_0}X$.

  Soit $v = (v_x, v_y, v_z) \in \{n\}^{\perp}$. On considère $g : t \mapsto (x_0, y_0) + t(v_x, v_y)$. Alors $(x_0, y_0) \in U$ car $M_0 \in \mathcal{G}_f$ et comme U est ouvert, il existe $\varepsilon > 0$ tel que, pour tout $t \in]-\varepsilon, \varepsilon[$, $g(t) \in U$. On pose, pour $t \in]-\varepsilon, \varepsilon[$, $\gamma : t \mapsto (g(t), f(g(t)))$. Comme $g_{|]-\varepsilon,\varepsilon[}$ est à valeurs dans U, γ est bien défini sur $]-\varepsilon, \varepsilon[$ et à valeurs dans $X = \mathcal{G}_f = \{(a, f(a)) \mid a = (x, y) \in U\}$.
  De plus, γ est dérivable sur $]-\varepsilon, \varepsilon[$ car ses coordonnées le sont : g est dérivable sur $]-\varepsilon, \varepsilon[$ et f est différentiable sur U d'où g et $f \circ g$ sont dérivables sur $]-\varepsilon, \varepsilon[$.

  Dé plus, on a :
  - $\gamma(0) = (g(0), f(g(0))) = (x_0, y_0, f(x_0, y_0)) = M_0$ car, comme $M_0 \in \mathcal{G}_f$, $z_0 = f(x_0, y_0)$.
  - $\gamma'(0) = (g'(0), (f \circ g)'(0)) = (v_x, v_y, (f \circ g)'(0))$. Or, on remarque que, comme $(n|v) = 0$, on a :
  $$\frac{\partial f}{\partial y}(x_0, y_0)v_x + \frac{\partial f}{\partial y}(x_0, y_0)v_y - v_z = 0$$
  d'où :
  $$\begin{array}{rcl} (f \circ g)'(0) & = & \mathrm{d}f(g(0))(g'(0)) \\ & = & \mathrm{d}f(x_0, y_0)(v_x, v_y) \\ & = & (\nabla f(x_0, y_0)|(v_x, v_y)) \\ & = & \dfrac{\partial f}{\partial x}(x_0, y_0)v_x + \dfrac{\partial f}{\partial y}(x_0, y_0)v_y \\ (f \circ g)'(0) & = & v_z \end{array}$$
  et ainsi, $\gamma'(0) = v$.

  Par suite, $v \in T_{M_0}X$. D'où $\{n\}^{\perp} \subset T_{M_0}X$.

- ⋆ Montrons $T_{M_0}X \subset \{n\}^{\perp}$.

  Soit $v = (v_x, v_y, v_z) \in T_{M_0}X$. Alors, il existe $\varepsilon > 0$ et $\gamma :]-\varepsilon, \varepsilon[\to X$ dérivable en 0 et tel que $\gamma(0) = M_0$ et $\gamma'(0) = v$.
  Comme γ est à valeurs dans X, pour tout $t \in]-\varepsilon, \varepsilon[$, il existe $g(t) = (x(t), y(t)) \in \mathbb{R}^2$ tel que $\gamma(t) = (g(t), f(g(t)))$.
  Comme γ définie sur $]-\varepsilon, \varepsilon[$ et dérivable en 0, ses applications coordonnées le sont et donc g et $f \circ g$ sont définies sur $]-\varepsilon, \varepsilon[$ et dérivable en 0. De plus, comme f est différentiable sur U, g est à valeurs dans U et dérivable en 0, on a :
  $$M_0 = (x_0, y_0, z_0) = \gamma(0) = (g(0), f(g(0)),$$
  puis, en utilisant le calcul de $(f \circ g)'(0)$ de l'inclusion précédente :
  $$v = (v_x, v_y, v_z) = \gamma'(0) = (g'(0), (f \circ g)'(0)) = \left(v_x, v_y, \frac{\partial f}{\partial x}(x_0, y_0)v_x + \frac{\partial f}{\partial y}(x_0, y_0)v_y\right).$$

Ainsi,

$$(n|v) = \frac{\partial f}{\partial x}(x_0,y_0)v_x + \frac{\partial f}{\partial y}(x_0,y_0)v_y - \left(\frac{\partial f}{\partial x}(x_0,y_0)v_x + \frac{\partial f}{\partial y}(x_0,y_0)v_y\right) = 0,$$

et donc, $v \in \{n\}^\perp$. Par suite, $T_{M_0}X \subset \{n\}^\perp$.
Ainsi, par double inclusion, on a $T_{M_0}X = \{n\}^\perp$.

Exercice 22.

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction de classe C^1 sur $\mathbb{R}$ et $x_0 \in \mathbb{R}$. Montrer que l'ensemble des vecteurs tangents à $\mathcal{G}_f$ en (x_0, y_0) est égal à la droite vectorielle qui dirige la tangente de f en x_0.

En fait, le résultat final de l'exemple précédent représente la "règle" comme on va le voir avec le théorème suivant. Sa démonstration, plus précisément, la démonstration de l'inclusion "noyau de la différentielle $\subset$ vecteurs tangents" repose sur le fait qu'on peut toujours mettre *localement* une équation $g(x_1, ..., x_n) = 0$ sous la forme $x_n = f(x_1, ..., x_{n-1})$ avec f est classe C^1 lorsque g est de classe C^1 et que sa différentielle en $(x_1, ..., x_n)$ ne s'annule pas. Il s'agit du **théorème des fonctions implicites** dont la démonstration est hors programme ; on admet donc ce théorème ici.

Théorème 5.

Soit $g : U \to \mathbb{R}$ une fonction de classe C^1 sur U, $X = \{y \in U \mid g(y) = 0\}$ l'ensemble des zéros de g et $x \in X$.
Si $\mathrm{d}g(x) \neq \mathbf{0}$, alors :

$$T_xX = \mathrm{Ker}(\mathrm{d}g(x)) = \{\nabla g(x)\}^\perp.$$

Corollaire 5.

Soit $\mathcal{S}$ une surface de $\mathbb{R}^3$ d'équation $\mathcal{S} :\ g(x,y,z) = 0$ où g est de classe C^1 sur $\mathbb{R}^3$ et $(x_0, y_0, z_0) \in \mathcal{S}$. Si $\nabla g(x_0, y_0, z_0) \neq (0,0,0)$, alors $T_{(x_0,y_0,z_0)}\mathcal{S}$ est un plan vectoriel d'équation :

$$T_{(x_0,y_0,z_0)}\mathcal{S} :\ \frac{\partial g}{\partial x}(x_0,y_0,z_0).(x-x_0) + \frac{\partial g}{\partial y}(x_0,y_0,z_0).(y-y_0) + \frac{\partial g}{\partial z}(x_0,y_0,z_0).(z-z_0) = 0$$

Remarque 6.

Ce corollaire est se généralise directement en remplaçant $\mathbb{R}^3$ par $\mathbb{R}^n$ avec $n \geq 2$.

Exercice 23.

Déterminer une équation du plan tangent à $\mathcal{S}$ en un point $M_0 = (x_0, y_0, z_0) \in \mathcal{S}$ pour :

1. $\mathcal{S}$ le paraboloïde d'équation $\mathcal{S} :\ z = x^2 + y^2$.
2. $\mathcal{S}$ la sphére unité euclidienne ;

3. $\mathcal{S}$ le cône de droite génératrice Vect$(1,0,1)$ en $(0,1,1)$.

Correction.

1. On a $\mathcal{S}:\ g(x,y,z)=0$ où $g:(x,y,z)\mapsto x^2+y^2-z$. Alors g est polynomiale et donc de classe C^1 sur $\mathbb{R}^3$. De plus, on a :
$$\nabla g(M_0)=(2x_0,2y_0,-1)\neq(0,0,0),$$
d'où :
$$T_{M_0}\mathcal{S}:\ z=z_0+2x_0(x-x_0)+2y_0(y-y_0).$$
2. On a $\mathcal{S}:\ g(x,y,z)=0$ où $g:(x,y,z)\mapsto x^2+y^2+z^2-1$. Alors g est polynomiale et donc de classe C^1 sur $\mathbb{R}^3$. De plus, comme pour tout $M_0\in S$, $\|M_0\|_2=1$ et donc $M_0\neq(0,0,0)$, on a :
$$\nabla g(M_0)=(2x_0,2y_0,2z_0)\neq(0,0,0),$$
d'où :
$$T_{M_0}\mathcal{S}:\ x_0(x-x_0)+y_0(y-y_0)+z_0(z-z_0)=0.$$
3. On a $\mathcal{S}:\ g(x,y,z)=0$ où $g:(x,y,z)\mapsto x^2+y^2-z^2$. Alors g est polynomiale et donc de classe C^1 sur $\mathbb{R}^3$. De plus, comme pour tout $M_0\in S$ avec $M_0\neq(0,0,0)$:
$$\nabla g(M_0)=(2x_0,2y_0,-2z_0)\neq(0,0,0),$$
d'où :
$$T_{M_0}\mathcal{S}:\ z_0(z-z_0)=x_0(x-x_0)+y_0(y-y_0).$$
Ainsi, en $M_0=(0,1,1)$, on a :
$$T_{(0,1,1)}\mathcal{S}:\ z=y.$$
Ainsi :
$$T_{(0,1,1)}\mathcal{S}=\mathrm{Vect}((0,1,1),(1,0,0)).$$

3. Approximation au second ordre et matrice hessienne

Dans ce paragraphe, on considère une base $\mathcal{B}=(e_1,\ldots,e_n)$ de E.

a. Formule de Taylor-Young au second ordre

Théorème 6. *Formule de Taylor-Young au second ordre*

Soit $f:U\to\mathbb{R}$ une fonction de classe C^2 sur U. Pour tout $a\in U$, on a pour $h=\sum_{j=1}^n h_je_j\in E$:
$$f(a+h)=f(a)+\sum_{j=1}^n h_j\frac{\partial f}{\partial x_j}(a)+\frac{1}{2}\sum_{i\leq i,j\leq n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(a)+\underset{h\to 0_E}{o}(\|h\|^2).$$

Exemple 11. *Cas $E = \mathbb{R}^2$*

Pour $E = \mathbb{R}^2$ muni de sa base canonique et $f : U \to \mathbb{R}$ une fonction de classe C^2 sur U, on a, pour $(x_0, y_0) \in U$, en vertu du théorème de Schwarz :

$$\begin{aligned} f(x_0 + h_1, y_0, h_2) = f(x_0, y_0) \quad &+ \quad h_1.\frac{\partial f}{\partial x}(x_0, y_0) + h_2.\frac{\partial f}{\partial y}(x_0, y_0) \\ &+ \quad \frac{1}{2}\left(h_1^2.\frac{\partial^2 f}{\partial x^2}(x_0, y_0) + h_1 h_2.\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) + h_2^2.\frac{\partial^2 f}{\partial y^2}(x_0, y_0)\right) \\ &+ \quad \underset{(h_1,h_2)\to(0,0)}{o}(h_1^2 + h_2^2) \end{aligned}$$

b. Matrice hessienne

Définition 15. *Matrice hessienne*

Soit $f : U \to \mathbb{R}$ une fonction de classe C^2 sur U et $a \in U$. On appelle **matrice hessienne** ou simplement **hessienne** de f en a et on note $H_f(a)$ la matrice carrée d'ordre n :

$$H_f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{1 \leq i,j \leq n} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

Proposition 21.

Si $f : U \to \mathbb{R}$ est une fonction de classe C^2 sur U, alors, pour tout $a \in U$, $H_f(a) \in S_n(\mathbb{R})$ *i.e.* la hessienne de f en a est symétrique réelle.

Exemple 12. *Cas $E = \mathbb{R}^2$*

Pour $E = \mathbb{R}^2$ muni de sa base canonique et $f : U \to \mathbb{R}$ une fonction de classe C^2 sur U, on a, pour $(x_0, y_0) \in U$:

$$H_f(x_0, y_0) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) & \dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & \dfrac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{pmatrix}$$

Dans la proposition suivante, on résume, grâce au gradient et la hessienne, la formule de Taylor-Young du second ordre en identifiant, à travers une base orthonormale $\mathcal{B}$, l'espace E et $M_{n,1}(\mathbb{R})$. On rappelle que dans ce cas, $(x|y) = {}^t x.y$.

Proposition 22.

Si $f : U \to \mathbb{R}$ est une fonction de classe C^2 sur U, pour tout $a \in U$:

$$f(a+h) = f(a) + (\nabla f(a)|h) + \frac{1}{2}(H_f(a)h|h) + \underset{h \to 0_E}{o}(\|h\|^2)$$

ou matriciellement :

$$f(a+h) = f(a) + {}^t\nabla f(a).h + \frac{1}{2}{}^th.H_f(a).h + \underset{h \to 0_E}{o}(\|h\|^2)$$

Partie E

Optimisation

Dans cette partie E désigne un espace euclidien et on note $(\cdot|\cdot)$ son produit scalaire.

1. Extrema et points critiques

Définition 16. *Extremum local*

Soit $f : U \to \mathbb{R}$ et $a \in U$.
On dit que f admet un **minimum local** (resp. un **maximum local** en a s'il existe un voisinage W de a tel que pour tout $x \in W \cap U$:

$$f(a) \leq f(x) \quad (\text{resp. } f(a) \geq f(x)).$$

Si f admet un minimum ou un maximum local en a, on dit que f admet un **extremum local en** a.

Définition 17. *Point critique*

Soit $f : U \to \mathbb{R}$ et $a \in U$. On dit que a est un **point critique** de f si f est différentiable en a et $\mathrm{d}f(a) = \mathbf{0}$ ou de manière équivalente $\nabla f = 0_E$.

Exemple 13.

La fonction $f : (x, y) \mapsto x^4 + y^4 - x^2 + y^2$ admet des points critiques en $(0, 0)$, $(-\frac{1}{\sqrt{2}}, 0)$ et $(\frac{1}{\sqrt{2}}, 0)$.

2. Extrema libres : étude au premier ordre

Théorème 7.

Soit $f : U \to \mathbb{R}$ un fonction de classe C^1 sur U et $a \in U$. Si f admet un extremum en a, alors a est un point critique de f.

Exemple 14.

La fonction $f : (x, y) \mapsto x^4 + y^4 - x^2 + y^2$ admet des extrema locaux en $(-\frac{1}{\sqrt{2}}, 0)$ et en $(\frac{1}{\sqrt{2}}, 0)$ mais pas en $(0, 0)$.

Remarque 7.

Comme le montre l'exemple précédent, la réciproque du Théorème 7 est fausse ! Une fonction qui admet un point critique n'admet pas forcément d'extremum en ce point.

3. Extrema libres : étude au second ordre

Théorème 8.

Soit $f : U \to \mathbb{R}$ un fonction de classe C^2 sur U et $a \in U$.

— Si f admet un minimum (resp. un maximum) local en a, alors $H_f(a) \in S_n^+(\mathbb{R})$ (resp. $S_n^-(\mathbb{R})$).

— Si f admet un point critique en a et $H_f(a) \in S_n^{++}(\mathbb{R})$ (resp. $S_n^{--}(\mathbb{R})$), alors f admet un minimum (resp. un maximum) local en a.

Corollaire 6. *Cas particulier $E = \mathbb{R}^2$*

Soit $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ un fonction de classe C^2 sur U et $(x_0, y_0) \in U$ un point critique de f. On note $H = H_f(x_0, y_0)$ la hessienne de f en (x_0, y_0).

- Si $\det(H) > 0$ alors f admet un extremum local en (x_0, y_0).
  De plus, si $\mathrm{Tr}(H) > 0$, il s'agit d'un minimum ; sinon, c'est un maximum.
- Si $\det(H) < 0$, f n'admet pas d'extremum en (x_0, y_0).
- Si $\det(H) = 0$, on ne pas conclure directement.

Remarque 8.

Le corollaire précédent est également connu sous le nom de théorème de Monge. Dans la littérature, la hessienne est alors écrite

$$H = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$$

et ainsi, on reformule les conditions du corollaire avec $rt - s^2 = \det(H)$. Dans le cas $rt - s^2 > 0$, on peut même être plus économe : on peut remarquer que sous cette condition, $\mathrm{Tr}(H) = r + t$ et r sont de même signe.

On peut donc reformuler le corollaire avec les notations de Monge :

- Si $rt - s^2 > 0$ alors f admet un extremum local en (x_0, y_0).
  De plus, si $r > 0$, il s'agit d'un minimum ; sinon, c'est un maximum.
- Si $rt - s^2 < 0$, f n'admet pas d'extremum en (x_0, y_0).
- Si $rt - s^2 = 0$, on ne pas conclure directement.

Exemple 15.

Soit $f : (x, y, z) \mapsto x^3 + y^3 - 3xy$ admet deux points critiques en $(0, 0)$ et $(1, 1)$: le premier n'est pas un extremum de f et le second est un minimum local de f.

Exercice 24.

Étudier les extrema de :

1. $f : (x, y) \mapsto y^2(y^2 - x^4)$ sur $\mathbb{R}^2$.
2. $g : (x, y) \mapsto y^2 - x^2 + \dfrac{x^4}{2}$ sur $\mathbb{R}^2$.

4. Extrema liés : optimisation sous contraintes

Théorème 9. *Théorème d'optimisation sous contrainte*

Soit $f, g : U \to \mathbb{R}$ des fonctions de classe C^1 sur U, $X = \{x \in U \mid g(x) = 0\}$ et $a \in X$.
Si la restriction $f_{|X}$ de f à X admet un extremum local en a et $\mathrm{d}g(a) \neq 0$, alors $\mathrm{d}f(a)$ est colinéaire à $\mathrm{d}g(a)$ *i.e.* $\nabla f(a)$ est colinéaire à $\nabla g(a)$.