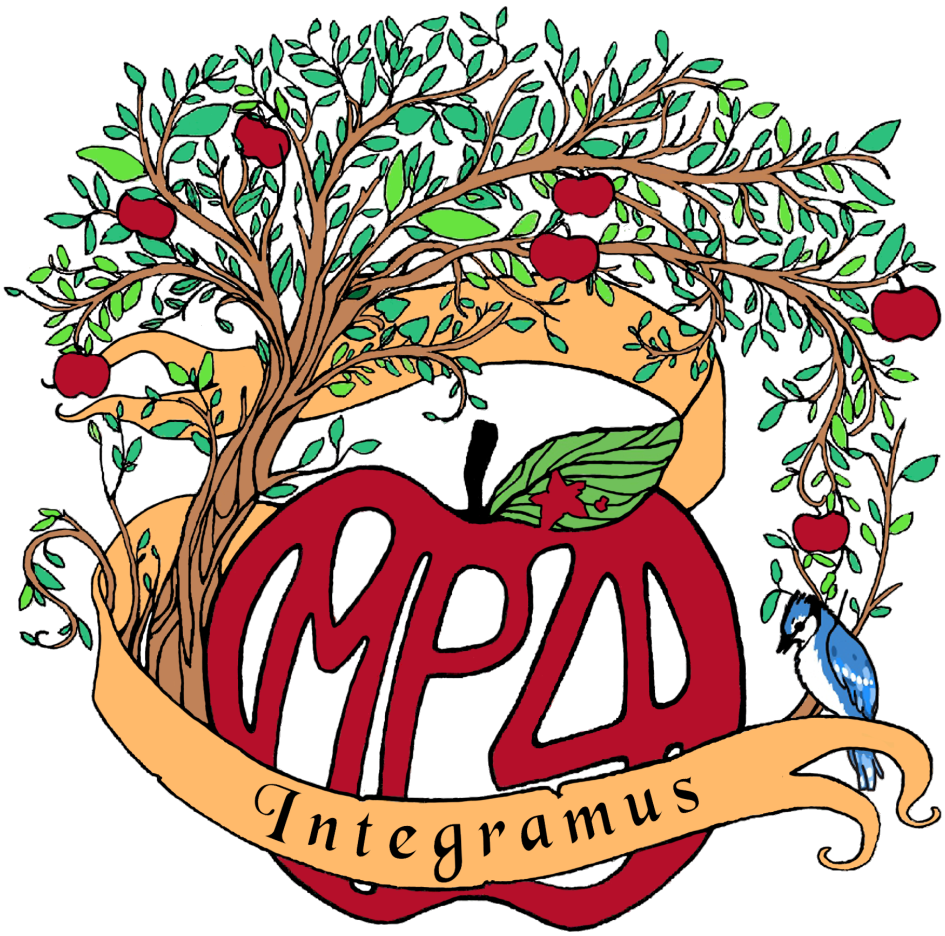
Mathématiques en MP*

Un livre inspiré par le cours de la MP*4 du lycée Louis-le-Grand

Omar Bennouna, Issam Tauil & M.C.

Cours de mathématiques indispensables pour une deuxième année en MP*



Publié sur www.cpge-paradise.com

Préface

Ce livre contient des chapitres indispensables de mathématiques de deuxième année de CPGE filière MP. Les chapitres de ce cours ont étés inspirés des notes de cours de la MP*4 du lycée Louis-le-Grand pendant l'année scolaire 2018/2019. Toutes les propriétés, théorèmes, et différents éléments ont été rédigés de manière assez détaillée et sont accompagnés, lorsque cela est nécessaire, de schémas et dessins qui aideront l'élève à avoir l'intuition de chaque résultat présenté. Tous les exercices du cours ont étés corrigés de manière assez détaillée de manière à ce que l'élève puisse progresser si jamais il n'arrive pas à résoudre un exercice. Tous les cours contiennent un nombre considérable de notions hors programme très utiles pour les concours de haut niveau (X-ENS, Centrale, Mines-Ponts). Nous espérons que le contenu de ce livre sera utile pour les élèves désirant s'approfondir dans les notions vues en deuxième année. Les chapitres actuellement présents dans cette version du livre ne représentent pas l'intégralité des chapitres du cours. Par conséquent, ce livre sera progressivement mis à jour sur le site cpge-paradise.com. À noter qu'une grande partie des résultats hors programmes présentés dans ce livre sont sous forme d'exercice, il est donc souvent utile de mémoriser l'idée derrière (sans trop perdre de temps!) les résultats présentés dans les exercices. Si jamais vous repérez une erreur ou alors quelque chose qui n'est pas claire lors de votre lecture, n'hésitez pas à nous contacter via l'adresse contact@cpge-paradise.com.

Remerciements

Nous remercions nos professeurs de mathématiques de classes préparatoires de première et deuxième année, en particulier Roger Mansuy, Alain Troesh, Yves Duval et Alain Pommellet pour la formation excellente et très agréable qu'ils nous ont offert. Nos deux années au lycée Louis-le-Grand étaient une très belle expérience grâce à eux. Nous sommes très reconnaissants à Alain Pommellet en particulier qui nous a donné son accord pour nous inspirer des notes de son cours en MP*4, Gauthier Thomas, notre professeur du cours des fonctions holomorphes à l'X pour nous avoir fourni une partie de son template LaTeX. Enfin, nous remercions également nos familles qui ont toujours été là pour nous et sans qui nous n'aurions pas pu en arriver là.

Auteurs du livre

Omar Bennouna est actuellement en thèse au département Electrical Engineering and Computer Science du MIT. Après avoir eu son bac au Maroc, il a fait deux ans de classes préparatoires en MPSI2 puis en MP*4 au lycée Louis-le-Grand et intègre l'École polytechnique en 2019 où il se spécialise en mathématiques appliquées.

Issam Tauil est actuellement en un M2 de mathématiques fondamentales au sein d'une université parisienne. Après avoir eu son bac au Maroc, il a fait deux ans de classes préparatoires en MPSI4 puis en MP*2 au lycée Louis-le-Grand puis intègre l'École polytechnique en 2019 où il se spécialise en mathématiques fondamentales. En 2021 puis en 2022, il décroche une médaille d'or et une médaille d'argent à l'IMC (International Mathematics Olympiad).

M.C., polytechnicien et enseignant de mathématiques et d'informatique, est l'initiateur de ce projet. À l'obtention de son baccalauréat scientifique, il se prépare aux concours des Grandes Écoles au Lycée Henri Poincaré de Nancy, à l'issue desquels il intègre l'École polytechnique en 2019, et se spécialise en cybersécurité.

Je dédie ce livre à Amine qui a été
et sera toujours ma plus grande
inspiration.

“Si je suis stupide, je serai la première personne stupide qui va réussir.”

Table des matières

(cc) BY-NC-SA

(cc) BY-NC-SA

(cc) BY-NC-SA

Chapitre 28 : Réduction d'endomorphisme 279

Chapitre 30.1 : Probabilités 333

Chapitre 30.2 : Indépendance, conditionnement 339

Chapitre 30.3 : Variables aléatoires 344

Chapitre 30.4 : Couples de variables aléatoires 354

Chapitre 32 : Exponentielle d'une matrice 357

Chapitre 34 : Espaces préhilbertiens réels 372

(cc) BY-NC-SA

(cc) BY-NC-SA

Chapitre 1

Ordre, inégalités

I Borne supérieure

Rappel I.1.

Toute partie de $\mathbb{R}$ non vide et majorée possède une borne supérieure.

Contre-exemple dans $\mathbb{Q}$: L'ensemble $\mathrm{A} = \{r \in \mathbb{Q} \,|\, r^2 < 2\}$ ne possède pas de borne supérieure rationnelle. En effet, supposons qu'il en ait une et posons $\mathrm{C} = \sup \mathrm{A} \in \mathbb{Q}$. Alors $0 \leq \mathrm{C} \leq \sqrt{2}$ car $\sqrt{2}$ est un majorant de A, et $0 \in \mathrm{A}$. Mieux, $\mathrm{C} < \sqrt{2}$ car $\sqrt{2}$ est irrationnel. Par densité de $\mathbb{Q}$ dans $\mathbb{R}$, il existe $r \in \mathbb{Q}$ tel que $0 \leq \mathrm{C} < r < \sqrt{2}$. Donc $r^2 < 2$, donc C n'est pas le supremum de A.

Proposition I.2.

Soit $c \in \mathbb{R}$. Si A est une partie de $\mathbb{R}$ non vide et majorée, alors on a l'équivalence :

$$c = \sup \mathrm{A} \iff \forall x \in \mathrm{A},\ x \leq c \ \text{ et } \ \forall \varepsilon > 0,\ \exists x \in \mathrm{A},\ c - \varepsilon < x \leq c.$$

Applications

1. Le théorème de la limite monotone pour les suites réelles bornées en est une.
2. Considérons une fonction $f : \mathbb{R}^+ \longmapsto \mathbb{R}$ croissante et majorée. Alors f possède une limite finie en $+\infty$. En effet, posons $\ell = \sup f$, et fixons $\varepsilon > 0$. Il existe alors $x_\varepsilon \in \mathbb{R}^+$ tel que $\ell - \varepsilon < f(x_\varepsilon) \leq \ell$. La croissance de f assure de plus que pour tout $x \geq x_\varepsilon$, $\ell - \varepsilon < f(x) \leq \ell$. Donc, $f(x) \xrightarrow[x \to +\infty]{} \ell$.

Proposition I.3.

Soit A une partie de $\mathbb{R}$ non vide et majorée.

- $\to$ Si $\lambda > 0$, alors $\sup(\lambda \mathrm{A}) = \lambda \sup(\mathrm{A})$.
- $\to$ Si $a \in \mathbb{R}$, $\sup(a + \mathrm{A}) = a + \sup(\mathrm{A})$.

De même, considérons des fonctions f et g d'un ensemble non vide I et à valeurs dans $\mathbb{R}$.

- $\to$ Si $\lambda > 0$, alors $\sup(\lambda f) = \lambda \sup(f)$.
- $\to$ Si $a \in \mathbb{R}$, $\sup(a + f) = a + \sup(f)$.
- $\to$ $\sup(f + g) \leq \sup(f) + \sup(g)$.

Démonstration de l'inégalité. Pour tout $x \in \mathrm{I}$, $f(x) + g(x) \leq \sup(f) + \sup(g)$.
Donc $\sup(f + g) \leq \sup(f) + \sup(g)$

(cc) BY-NC-SA

Corollaire I.4.

Soient a et $b \in \mathbb{R}$ tels que $a \leq b$. Alors l'application

$$\|\cdot\|_\infty : \begin{cases} \mathscr{C}^0([a,b],\mathbb{R}) & \longrightarrow \mathbb{R} \\ f & \longmapsto \sup(|f|) \end{cases}$$

est une norme.

Supremum d'un ensemble paramétré par deux variables

Soient A et B deux ensembles non vides, et $f : \mathrm{A} \times \mathrm{B} \longrightarrow \mathbb{R}$ une fonction majorée. Montrons que

$$\sup f = \sup_{x\in \mathrm{A}}\left(\sup_{y\in \mathrm{B}}(f(x,y))\right) = \sup_{y\in \mathrm{B}}\left(\sup_{x\in \mathrm{A}}(f(x,y))\right).$$

Posons $\mathrm{M} = \sup f$, et pour tout $x \in \mathrm{A}$, posons $\mu_x = \sup\limits_{y\in\mathrm{B}}(f(x,y))$. Pour tout $x \in \mathrm{A}$, et pour tout $y \in \mathrm{B}$, $f(x,y) \leq \mathrm{M}$. Donc, pour tout $x \in \mathrm{A}$, $\mu_x \leq \mathrm{M}$. Donc $\sup\limits_{x\in\mathrm{A}}(\mu_x) \leq \mathrm{M}$.
Par ailleurs, pour tout $(x,y) \in \mathrm{A}\times\mathrm{B}$, $f(x,y) \leq \mu_x \leq \sup\limits_{x\in\mathrm{A}}(\mu_x)$. Donc $\mathrm{M} \leq \sup\limits_{x\in\mathrm{A}}(\mu_x)$. Donc

$$\mathrm{M} = \sup_{x\in\mathrm{A}}(\mu_x)$$

et la deuxième égalité se démontre de la même manière en échangeant les rôles de x et y.

Considérons maintenant un ensemble non vide Ω indexant une famille $(f_\lambda)_{\lambda\in\Omega}$ de fonctions d'une partie non vide I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

Définition I.5.

On dit que

- → $(f_\lambda)_{\lambda\in\Omega}$ est majorée si pour tout $x \in \mathrm{I}$, la famille de réels $(f_\lambda(x))_{\lambda\in\Omega}$ est majorée.
- → $(f_\lambda)_{\lambda\in\Omega}$ est uniformément majorée s'il existe $\mathrm{M} \in \mathbb{R}$ tel que pour tout $\lambda \in \Omega$ et pour tout $x \in \mathrm{I}$, $f_\lambda(x) \leq \mathrm{M}$.
- → $(f_\lambda)_{\lambda\in\Omega}$ est uniformément bornée s'il existe $\mathrm{K} \in \mathbb{R}$ tel que pour tout $\lambda \in \Omega$ et pour tout $x \in \mathrm{I}$, $|f_\lambda(x)| \leq \mathrm{M}$.

Si pour tout $x \in \mathrm{I}$, $(f_\lambda(x))_{\lambda\in\Omega}$ est majorée, on appelle *enveloppe supérieure de la famille* $(f_\lambda)_{\lambda\in\Omega}$ la fonction

$$\sup_{\lambda\in\Omega}(f_\lambda) : \begin{cases} \mathrm{I} & \longrightarrow \mathbb{R} \\ x & \longmapsto \sup\limits_{\lambda\in\Omega}(f_\lambda(x)) \end{cases}$$

Proposition I.6.

Si Ω est fini et si les f_λ sont continues, alors $\sup\limits_{\lambda\in\Omega}(f_\lambda)$ est continue.

Démonstration. Soient f_1 et f_2 deux fonctions continues sur I et à valeurs réelles.

Il est aisé de vérifier – donc loisible de retenir – que $\sup(f_1, f_2) = \dfrac{1}{2}(f_1 + f_2 + |f_1 - f_2|)$, ce qui assure sa continuité.
Le principe de récurrence finie assure alors que $\sup\limits_{\lambda\in\Omega}(f_\lambda)$ est continue.

BY-NC-SA

Contre-exemple avec une famille infinie : Pour tout $n \in \mathbb{N}$ et pour tout $x \in [0,1]$, posons $f_n(x) = 1 - x^n$. La famille de fonctions ainsi définie est uniformément bornée par 1 donc son enveloppe supérieure est bien définie.
Or, $\left(\sup\limits_{n\in\mathbb{N}}(f_n)\right)(1) = 0$ et pour tout $x \in [0,1[$, $\left(\sup\limits_{n\in\mathbb{N}}(f_n)\right)(x) = 1$, donc l'enveloppe supérieure des f_n n'est pas continue.

II Encadrements

Proposition II.1.

Soient $a, b \in \mathbb{R}$. On a l'équivalence

$$a \leq b \iff \forall \varepsilon > 0,\ a \leq b + \varepsilon.$$

Démonstration. ($\Rightarrow$) Cette implication est immédiate.
($\Leftarrow$) Supposons que $a > b$ et posons $\varepsilon = \dfrac{a-b}{2}$. Alors, par hypothèse, $a \leq \dfrac{a+b}{2}$ mais $\dfrac{a+b}{2} < a$ car $b < a$. Donc $a \leq b$.

Encadrements de fractions
Si $0 \leq a' \leq a \leq a''$ et $0 < b' \leq b \leq b''$ alors $\dfrac{a'}{b''} \leq \dfrac{a}{b} \leq \dfrac{a''}{b'}$

Valeurs absolues

→ Si $x \in \mathbb{R}$, $|x| \leq \mathrm{M} \iff -\mathrm{M} \leq x \leq \mathrm{M}$.

→ **Inégalité triangulaire :** Soit n un entier supérieur ou égal à 2. Soit $(z_1, \ldots, z_n) \in (\mathbb{C}^*)^n$. Alors

$$||z_1| - |z_2|| \leq |z_1 + z_2| \leq |z_1| + |z_2| \qquad \text{et} \qquad \left|\sum_{k=1}^{n} z_k\right| \leq \sum_{k=1}^{n} |z_k|.$$

Démonstration de l'inégalité triangulaire. Avec deux complexes,

$$|z_1 + z_2|^2 = |z_1|^2 + 2\,\mathrm{Re}(z_1\overline{z_2}) + |z_2|^2 \leq |z_1|^2 + 2\,|z_1 z_2| + |z_2|^2 = (|z_1| + |z_2|)^2.$$

d'où la majoration. Remarquons qu'en l'appliquant à nouveau, il vient que

$$|z_1| = |z_1 + z_2 - z_2| \leq |z_1 + z_2| + |z_2| \qquad \text{et} \qquad |z_2| = |z_2 + z_1 - z_1| \leq |z_1 + z_2| + |z_1|,$$

d'où la minoration. Une récurrence permet de généraliser la majoration à n complexes.

Extension : Si la série $\sum\limits_{n\geq 0} |z_n|$ converge, alors $\sum\limits_{n\geq 0} z_n$ converge également et $\left|\sum\limits_{n=0}^{+\infty} z_n\right| \leq \sum\limits_{n\geq 0} |z_n|$.

Application

Soit $(u_n) \in (\mathbb{C}^*)^{\mathbb{N}}$ telle que $u_n \xrightarrow[n\to+\infty]{} \ell \neq 0$. Alors $\dfrac{1}{u_n} \xrightarrow[n\to+\infty]{} \dfrac{1}{\ell}$.

En effet, pour tout $n \in \mathbb{N}$, $\left|\dfrac{1}{u_n} - \dfrac{1}{\ell}\right| = \dfrac{|u_n - \ell|}{|\ell u_n|}$. Par ailleurs, il existe $n_\ell \in \mathbb{N}$ tel que pour tout entier

(cc) BY-NC-SA

$n \geq n_\ell$, $|\ell| - |u_n| \leq |u_n - \ell| \leq \dfrac{|\ell|}{2}$. Donc

$$0 \leq \left|\frac{1}{u_n} - \frac{1}{\ell}\right| \leq \frac{2\,|u_n - \ell|}{|\ell|^2} \xrightarrow[n\to+\infty]{} 0,$$

ce qu'il fallait démontrer.

Cas d'égalité de l'inégalité triangulaire

Les points d'affixes $z_1, \dots, z_n$ sont sur une même demi-droite d'origine O. En effet, raisonnons par récurrence en introduisant, pour tout entier $n \geq 2$, l'assertion

$$\mathcal{P}_n : « \forall (z_1, \dots, z_n) \in (\mathbb{C}^*)^n,\ \left(\left|\sum_{k=1}^{n} z_k\right| = \sum_{k=1}^{n} |z_k| \Longrightarrow \forall k \in [\![1; n]\!],\ \exists \lambda_k \in \mathbb{R}^+,\ z_k = \lambda_k z_1\right) »$$

étant entendu que les implications réciproques sont immédiates.

→ Soient $z_1, z_2 \in \mathbb{C}^*$ tels que $|z_1 + z_2| = |z_1| + |z_2|$. En élevant l'égalité au carré, il vient que $\mathrm{Re}(z_1\overline{z_2}) = |z_1\overline{z_2}|$. Donc la partie imaginaire de $z_1\overline{z_2}$ est nulle. Mieux, il existe $k \in \mathbb{R}^+$ tel que $z_1\overline{z_2} = k$, donc $\overline{z_2} = \dfrac{k}{|z_1|^2}\overline{z_1}$ donc $z_2 = \dfrac{k}{|z_1|^2} z_1$. Donc $\mathcal{P}_2$ est vraie.

→ Soit n un entier supérieur ou égal à 2 tel que $\mathcal{P}_n$ soit vraie.
Soit $(z_1, \dots, z_n, z_{n+1}) \in (\mathbb{C}^*)^{n+1}$ tel que $\left|\displaystyle\sum_{k=1}^{n+1} z_k\right| = \displaystyle\sum_{k=1}^{n+1} |z_k|$. Alors,

$$\left|\sum_{k=1}^{n+1} z_k\right| \leq |z_{n+1}| + \left|\sum_{k=1}^{n} z_k\right| \leq \sum_{k=1}^{n+1} |z_k|\,.$$

L'égalité du majorant et du minorant assure que cet encadrement est une égalité, si bien que $\cancel{|z_{n+1}|} + \left|\displaystyle\sum_{k=1}^{n} z_k\right| = \cancel{|z_{n+1}|} + \displaystyle\sum_{k=1}^{n} |z_k|$. Or, $\mathcal{P}_n$ est vraie, donc pour tout $k \in [\![1; n]\!]$, il existe $\lambda_k \in \mathbb{R}^+$ tel que $z_k = \lambda_k z_1$.
Par le même raisonnement qu'à l'initialisation, l'égalité $\left|z_{n+1} + \displaystyle\sum_{k=1}^{n} z_k\right| = |z_{n+1}| + \left|\displaystyle\sum_{k=1}^{n} z_k\right|$ assure l'existence de $\lambda_{n+1} \in \mathbb{R}^+$ tel que $z_{n+1} = \lambda_{n+1} z_1$.
Donc $\mathcal{P}_{n+1}$ est vraie.

→ Finalement, le principe de récurrence assure que le cas d'égalité est vérifié si, et seulement si, les vecteurs d'affixes $z_1, \dots, z_n$ sont colinéaires et de même sens, donc que les points d'affixes $z_1, \dots, z_n$ appartiennent à une même demi-droite d'origine O.

Exercice II.2.

Soit p un entier supérieur ou égal à 3. Soit $(u_n) \in \mathbb{C}^{\mathbb{N}}$ telle que pour tout $n \in \mathbb{N}$, $u_{n+p} = \dfrac{1}{p}\displaystyle\sum_{k=n}^{n+p-1} u_k$.

1. Montrer que toute racine du polynôme caractéristique de $(u_n)_{n\in\mathbb{N}}$ est simple.
2. Montrer que $(u_n)_{n\in\mathbb{N}}$ converge.

(cc) BY-NC-SA

Cas d'égalité de l'inégalité triangulaire avec des sommes de séries

Soit $(z_n) \in (\mathbb{C}^*)^{\mathbb{N}}$ telle qu $\sum\limits_{n\geq 0} |z_n|$ converge. Donnons une condition nécessaire et suffisante pour que

$$\left|\sum_{n=0}^{+\infty} z_n\right| = \sum_{n=0}^{+\infty} |z_n| .$$

Montrons qu'il s'agit de la même condition que dans le cas réel *i.e.* pour tout $n \in \mathbb{N}$, il existe $\lambda_n \in \mathbb{R}^+$ tel que $z_n = \lambda_n z_0$.

C'est une condition suffisante car pour tout $\mathrm{N} \in \mathbb{N}$, $\left|\sum\limits_{n=0}^{\mathrm{N}} z_n\right| = |z_0| \sum\limits_{n=0}^{\mathrm{N}} \lambda_n = \sum\limits_{n=0}^{\mathrm{N}} |\lambda_n z_0| = \sum\limits_{n=0}^{\mathrm{N}} |z_n|$.

Montrons qu'elle est nécessaire. Supposons qu'il existe $\mathrm{N} \in \mathbb{N}^*$ tel que $\left|\sum\limits_{n=0}^{\mathrm{N}} z_n\right| < \sum\limits_{n=0}^{\mathrm{N}} |z_n|$.

Alors

$$\left|\sum_{n=0}^{+\infty} z_n\right| \leq \left|\sum_{n=0}^{\mathrm{N}} z_n\right| + \left|\sum_{n=\mathrm{N}+1}^{+\infty} z_n\right| < \sum_{n=0}^{+\infty} |z_n| ,$$

ce qui est faux. Donc pour tout $\mathrm{N} \in \mathbb{N}$, $\left|\sum\limits_{n=0}^{\mathrm{N}} z_n\right| = \sum\limits_{n=0}^{\mathrm{N}} |z_n|$. Le cas d'égalité de l'inégalité triangulaire assure alors que tous les termes de $(z_n)_{n\in\mathbb{N}}$ sont sur une même demi-droite d'origine O.

Exercice II.3.

Soit n un entier supérieur ou égal à 2. Soit $(a_0, \ldots, a_{n-1}) \in \mathbb{C}^n$, et posons $\mathrm{P} = \mathrm{X}^n + \sum\limits_{k=0}^{n-1} a_k \mathrm{X}^k$ et $\rho = \max\{|z| \mid \mathrm{P}(z) = 0\}$.

1. Montrer que $\rho \leq 1 + \max\limits_{k\in[\![0;n-1]\!]} (|a_k|)$.
2. Montrer que $\rho \leq \max\left(1, \sum\limits_{k=0}^{n-1} |a_k|\right)$.

Application II.4.

Considérons une famille $(\mathrm{P}_\lambda)_{\lambda\in\Omega}$ de polynômes unitaires et de même degré $n \geq 2$ définie par

$$\forall \lambda \in \Omega, \quad \mathrm{P}_\lambda = \mathrm{X}^n + \sum_{k=0}^{n-1} a_{k,\lambda} \mathrm{X}^k.$$

Si la famille de complexes $(a_{k,\lambda})_{k\in[\![0;n-1]\!],\lambda\in\Omega}$ ainsi définie est bornée, alors les racines des P_λ forment une partie bornée du plan.

(cc) BY-NC-SA

III Inégalités classiques

Théorème (Inégalité de Cauchy-Schwarz) III.1.

Soit $n \in \mathbb{N}^*$. Soient $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n) \in \mathbb{R}^n$. Alors

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}.$$

Démonstration. Examinons la différence $\left(\sum_{i=1}^{n} x_i y_i\right)^2 - \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{j=1}^{n} y_j^2\right)$. En développant, les termes $x_i^2 y_i^2$ se simplifient.
De la somme au carré il ne reste que les doubles produits $2x_iy_ix_jy_j$ avec $i < j$, et du produit de sommes il ne reste que les produits mixtes symétriquement regroupés $x_i^2y_j^2 + x_j^2y_i^2$ avec $i < j$ Ainsi,

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 - \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{j=1}^{n} y_i^2\right) = \sum_{1 \leq i < j \leq n} \left(2(x_iy_j)(x_jy_i) - (x_i^2y_j^2 + x_j^2y_i^2)\right) = -\sum_{1 \leq i < j \leq n} (x_iy_j - x_jy_i)^2 \leq 0.$$

Identité de Lagrange : Pour $n = 2$, on a l'identité remarquable valable dans tout anneau commutatif

$$(x_1y_1 + x_2y_2)^2 + (x_1y_2 - x_2y_1)^2 = (x_1^2 + x_2^2)(y_1^2 + y_2^2).$$

Inégalité de Cauchy-Schwarz pour les complexes : À l'aide de l'inégalité triangulaire, elle devient immédiate. Soient $(z_1, \dots, z_n)$ et $(\omega_1, \dots, \omega_n) \in \mathbb{C}^n$. Alors

$$\left|\sum_{k=1}^{n} z_k \omega_k\right| \leq \sum_{k=1}^{n} |z_k|\,|\omega_k| \leq \sqrt{\sum_{k=1}^{n} |z_i|^2}\sqrt{\sum_{k=1}^{n} |\omega_k|^2}.$$

Cas d'égalité de l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$

Corollaire III.2.

Les propositions suivantes sont équivalentes :

→ L'inégalité de Cauchy-Schwarz est une égalité.

→ Tous les mineurs d'ordre 2 de la matrice $Z := \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{pmatrix}$ sont nuls.

→ $\mathrm{rg}(Z) \leq 1$.

→ $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ sont colinéaires.

L'exercice suivant nécessite des notions présentées dans un chapitre ultérieur

Exercice III.3.

Soit n un entier supérieur ou égal à 2. On considère l'espace des colonnes $\mathcal{M}_{n,1}(\mathbb{R})$ muni du produit scalaire canonique $\langle \cdot | \cdot \rangle$ et on note $\| \cdot \|$ sa norme euclidienne.
Soit $A \in \mathcal{S}_n^{++}$. Montrer que pour tout $X \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$\|X\|^2 \leq \sqrt{\langle AX|X\rangle}\sqrt{\langle A^{-1}X|X\rangle}.$$

(cc) BY-NC-SA

Exercice III.4.

Soit n un entier supérieur ou égal à 2. Notons E l'ensemble $\left\{ \mathrm{X} = (x_1, \dots, x_n) \in (\mathbb{R}^+)^n \;\middle|\; \sum_{i=1}^{n} x_i = 1 \right\}$.

Calculer $m = \min_{\mathrm{X} \in \mathrm{E}} \left(\sum_{1 \leq i \neq j \leq n} x_i x_j \right)$ et $\mathrm{M} = \max_{\mathrm{X} \in \mathrm{E}} \left(\sum_{1 \leq i \neq j \leq n} x_i x_j \right)$.

Inégalité de Cauchy-Schwarz pour les intégrales
Soient f et g deux fonctions continues par morceaux sur un segment réel $[a, b]$. Alors,

$$\left| \int_a^b f(t) g(t) \mathrm{d}t \right| \leq \sqrt{\int_a^b |f(t)|^2 \,\mathrm{d}t} \sqrt{\int_a^b |g(t)|^2 \,\mathrm{d}t}.$$

Démonstration. Soit $n \in \mathbb{N}^*$. Pour tout $k \in [\![1; n]\!]$, posons $t_k = a + k \times \dfrac{b-a}{n}$.
Pour tout $k \in [\![1; n]\!]$, posons $x_k = f(t_k)$ et $y_k = g(t_k)$. D'après l'inégalité de Cauchy-Schwarz,

$$\left| \frac{b-a}{n} \sum_{k=1}^{n} x_k y_k \right| \leq \sqrt{\frac{b-a}{n} \sum_{k=1}^{n} |x_k|^2} \sqrt{\frac{b-a}{n} \sum_{k=1}^{n} |y_k|^2}.$$

En passant à la limite $n \longrightarrow +\infty$, on obtient ce qu'on voulait démontrer.

Inégalité de Minkowski : $\sqrt{\sum_{i=1}^{n} (x_i + y_i)^2} \leq \sqrt{\sum_{i=1}^{n} x_i^2} + \sqrt{\sum_{i=1}^{n} y_i^2}$. En effet,

$$\sum_{i=1}^{n} (x_i + y_i)^2 = \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2 + 2 \sum_{i=1}^{n} x_i y_i \underset{\substack{\uparrow \\ \text{Cauchy-Schwarz}}}{\leq} \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2 + 2 \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2} = \left(\sqrt{\sum_{i=1}^{n} x_i^2} + \sqrt{\sum_{i=1}^{n} y_i^2} \right)^2,$$

et le cas d'égalité est également réalisé lorsque $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ sont colinéaires.
On a un énoncé analogue avec les intégrales de fonctions continues par morceaux.

Inégalité de réarrangement : Supposons que $x_1 \leq \cdots \leq x_n$ et $y_1 \leq \cdots \leq y_n$. Alors, pour toute $\sigma \in \mathfrak{S}_n$,

$$\sum_{i=1}^{n} x_{\sigma(i)} y_j \leq \sum_{i=1}^{n} x_i y_j.$$

Démonstration. L'ensemble $\mathfrak{S}_n$ est fini donc il existe une partie P de $\mathfrak{S}_n$ tel que pour toute $\sigma \in \mathrm{P}$, $\sum_{i=1}^{n} x_{\sigma(i)} y_j$ est maximale.
Si l'identité est dans P, il n'y a rien à démontrer. Supposons le contraire, et introduisons l'application

$$\begin{aligned} \varphi : \mathrm{P} &\longrightarrow [\![1; n]\!] \\ \sigma &\longmapsto \min\{ j \in [\![1; n]\!] \mid \sigma(j) \neq j \} \end{aligned}$$

puis, posons $i = \max(\varphi)$, et notons σ une permutation de P où ce maximum est atteint. Enfin, posons $k = \sigma^{-1}(i)$. Par construction $\sigma(i) > i$ et $k > i$. Alors, $(x_{\sigma(i)} - x_i)(y_k - y_i) \geq 0$, donc, en développant

$$x_i y_k + x_{\sigma(i)} y_i \leq x_{\sigma(i)} y_k + x_i y_i.$$

(cc) BY-NC-SA

Considérons la permutation $\tilde{\sigma} = \tau_{\sigma(k),\sigma(i)} \circ \sigma$. L'inégalité se réécrit à l'aide de $\tilde{\sigma}$:

$$x_{\sigma(k)}y_k + x_{\sigma(i)}y_i \leq x_{\tilde{\sigma}(k)}y_k + x_{\tilde{\sigma}(i)}y_i$$

et, étant donné que σ et $\tilde{\sigma}$ coïncident en tout point de $[\![1;n]\!]$ distinct de i et de k, alors

$$\sum_{i=1}^{n} x_{\sigma(i)}y_j \leq \sum_{i=1}^{n} x_{\tilde{\sigma}(i)}y_j,$$

donc $\tilde{\sigma} \in \mathrm{P}$. Mais, par construction, $\varphi(\tilde{\sigma}) > i = \max(\varphi)$ ce qui est faux. Donc l'identité est dans P, ce qu'il fallait démontrer.

IV Étude de fonctions

1. Fonctions homographiques

Soient $c \neq 0$, et $a, b, c \in \mathbb{R}$ tels que $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$. La fonction f définie par

$$\forall x \in \mathbb{R} \setminus \left\{\frac{c}{d}\right\}, \qquad f(x) = \frac{ax+b}{cx+d}$$

est une fonction homographique. Elle est strictement monotone car sa dérivée est de signe constant :

$$\forall x \in \mathbb{R} \setminus \left\{\frac{c}{d}\right\}, \qquad f'(x) = \frac{ad-bc}{(cx+d)^2}.$$

2. Accroissements finis

Théorème (Théorème des accroissements finis) IV.1.

Soient a et $b \in \mathbb{R}$ tels que $a < b$. Soit $f \in \mathscr{C}([a,b], \mathbb{C})$, telle que f soit dérivable sur $]a,b[$. Si $f([a,b]) \subset \mathbb{R}$, alors il existe $c \in]a,b[$ tel que $f(b) - f(a) = f'(c)(b-a)$.

Contre-exemple dans $\mathbb{C}$: Un tel c n'existe pas si $a = 0$, $b = 2\pi$, et si on choisit f telle que pour tout $x \in [0, 2\pi]$, $f(x) = e^{ix}$.

Théorème (Inégalité des accroisements finis) IV.2.

Soient a et $b \in \mathbb{R}$ tels que $a < b$. Soit $f \in \mathscr{C}([a,b], \mathbb{C})$, telle que f soit dérivable sur $]a,b[$. Si $|f'|$ est bornée par $\mathrm{M} \geq 0$, alors

$$|f(b) - f(a)| \leq \mathrm{M}(b-a).$$

Si de plus, f' est continue sur $[a,b]$, on a l'égalité $f(b) - f(a) = \displaystyle\int_a^b f'(t)\mathrm{d}t$ donc

$$|f(b) - f(a)| \leq \int_a^b |f'(t)| \,\mathrm{d}t \leq \|f'\|_{\infty,[a,b]} (b-a).$$

(cc) BY-NC-SA

Remarque. Il est aisé de démontrer que $\left|\int_a^b f(t)\mathrm{d}t\right| \leq \int_a^b |f(t)|\,dt$ à partir de l'inégalité analogue réelle.
En effet, il existe $\theta \in \mathbb{R}$ tel que $\int_a^b f(t)\mathrm{d}t = \left|\int_a^b f(t)\mathrm{d}t\right| e^{i\theta}$.
Donc $\int_a^b e^{-i\theta} f(t)\mathrm{d}t = \left|\int_a^b f(t)\mathrm{d}t\right| = \left|\int_a^b e^{-i\theta} f(t)\mathrm{d}t\right|$. Puis, en introduisant les fonctions réelles u et v telles que $u + iv = e^{-i\theta} f$, il vient

$$\int_a^b u(t)\mathrm{d}t + i\int_a^b v(t)\mathrm{d}t = \left|\int_a^b f(t)\mathrm{d}t\right|,$$

donc $\int_a^b v(t)\mathrm{d}t = 0$. Ainsi

$$\left|\int_a^b f(t)\mathrm{d}t\right| \leq \left|\int_a^b u(t)\mathrm{d}t\right| \leq \int_a^b |u(t)|\,\mathrm{d}t \leq \int_a^b \left|e^{-i\theta} f(t)\right|\mathrm{d}t = \int_a^b |f(t)|\,\mathrm{d}t.$$

3. Puissances

Exemple

Soit $(a_n) \in (\mathbb{R}^+)^{\mathbb{N}}$. Soit $r \in]0,1[$ tel que $\sum_{n\geq 0} a_n^r$ converge. Montrons que $\sum_{n\geq 0} a_n$ converge.
Il suffit de remarquer qu'il existe $n_1 \in \mathbb{N}$ tel que pour tout entier $n \geq n_1$, $0 \leq a_n \leq a_n^r < 1$, car $a_n^r \xrightarrow[n\to+\infty]{} 0$. D'où la convergence de $\sum_{n\geq 0} a_n$.

Clairement, cet exemple sert à rappeler que $q \leq q^r$ si $q \in [0,1]$.

Lemme IV.3.

Soit $r \in]0,1[$. Alors, pour tous $x, y \in \mathbb{R}^+$,

$$(x+y)^r \leq x^r + y^r.$$

Démonstration. Sans perte de généralité, soient $x > 0$ et $y > 0$. Considérons la fonction φ définie par

$$\forall t > 0, \varphi(t) = t^r + 1 - (1+t)^r.$$

φ est dérivable sur $\mathbb{R}^{+*}$ et pour tout $t > 0$, $\varphi'(t) = r(t^{r-1} - (1+t)^{r-1}) \geq 0$, donc φ est croissante.
Or $\varphi(0) = 0$ donc $\varphi\left(\frac{y}{x}\right) \geq 0$, ce qu'il fallait démontrer.
Exercice similaire : On montre pareillement que si $r \geq 1$, pour tous $x, y \geq 0$, $(x+y)^r \geq x^r + y^r$.

4. Trigonométrie

Applications

1. Une récurrence et une formule d'addition permettent de montrer que pour tout $n \in \mathbb{N}^*$ et pour tout $x \in \mathbb{R}$, $|\sin(nx)| \leq n\,|\sin(x)|$.
2. L'exploitation de la concavité de la fonction sinus sur $\left[0, \frac{\pi}{2}\right]$ permet de montrer que pour tout $x \in \left[0, \frac{\pi}{2}\right]$, $\sin(x) \geq \frac{2}{\pi}x$.

(cc) BY-NC-SA

3. Un calcul aisé permet de montrer que pour tout $x \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, et pour tout $n \in \mathbb{N}^*$,

$$\left|\sum_{k=0}^{n} e^{ikx}\right| \leq \frac{1}{\left|\sin\left(\frac{x}{2}\right)\right|}.$$

En effet, si $x \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, et $n \in \mathbb{N}^*$

$$\begin{aligned}
\mathrm{S}_n(x) := \sum_{k=0}^{n} e^{ikx} &= \frac{e^{i(n+1)x} - 1}{e^{ix} - 1} \\
&= \frac{e^{i(n+1)x/2}\big(e^{i(n+1)x/2} - e^{-i(n+1)x/2}\big)}{e^{ix/2}\big(e^{ix/2} - e^{-ix/2}\big)} \\
&= e^{inx/2} \times \frac{\sin\left(\frac{(n+1)x}{2}\right)}{\sin\left(\frac{x}{2}\right)}
\end{aligned}$$

d'où le résultat. On dit alors que $(\mathrm{S}_n)_{n\in\mathbb{N}^*}$ est uniformément bornée en n à x fixé.

(cc) BY-NC-SA

Correction de l'exercice II.2. :

1. Le polynôme caractéristique de $(u_n)_{n\in\mathbb{N}}$ est

$$\mathrm{P} = \mathrm{X}^p - \frac{1}{p}\sum_{k=0}^{p-1}\mathrm{X}^k = \mathrm{X}^p - \frac{1}{p}\frac{\mathrm{X}^p-1}{\mathrm{X}-1} = \frac{\overbrace{p\mathrm{X}^{p+1}-(p+1)\mathrm{X}^p+1}^{\text{noté Q}}}{p(\mathrm{X}-1)}$$

et $\mathrm{Q}' = p(p+1)\mathrm{X}^{p-1}(\mathrm{X}-1)$. 1 est une racine simple de P car c'est une racine double de Q. Par ailleurs, toute racine de P distincte de 1 est une racine de Q avec la même multiplicité dans les deux polynômes. Or 0 n'est pas racine de P, donc toute racine de P distincte de 1 est simple.
Donc toute racine de P est simple.

2. Soit z une racine de P. Supposons que $|z| > 1$. Alors, pour tout entier naturel $k < p$, $|z|^p > |z|^k$, donc $p\,|z|^p > \sum_{k=0}^{p-1}|z|^k \underset{\substack{\uparrow\\ \text{inégalité triangulaire}}}{\geq} \left|\sum_{k=0}^{p-1} z^k\right| = p\,|z|^p$ ce qui est faux.

Donc $|z| \leq 1$. Si $z \in \mathbb{U}$, $p\,|z|^p = \sum_{k=0}^{p-1}|z|^k$. Le cas d'égalité de l'inégalité triangulaire assure alors que pour tout $k \in [\![1;p]\!]$, z^k appartient à la demi-droite réelle positive. Donc $z = 1$.
Les $p-1$ autres racines de P sont de modules strictement inférieurs à 1. Or, $(u_n)_{n\in\mathbb{N}}$ est une combinaison linéaire de suites géométriques dont les raisons sont les racines de P, et ces suites sont convergentes. Donc $(u_n)_{n\in\mathbb{N}}$ converge.

Correction de l'exercice II.3. :

1. Soit z une racine de P tel que $\rho = |z|$. Alors $z^n = -\sum_{k=0}^{n-1} a_k z^k$, donc

$$\rho^n = \left|\sum_{k=0}^{n-1} a_k z^k\right| \underset{\substack{\uparrow\\ \text{inégalité triangulaire}}}{\leq} \sum_{k=0}^{n-1}|a_k|\,\rho^k \leq \underbrace{\max_{k\in[\![0;n-1]\!]}(|a_k|)}_{\text{noté M}} \sum_{k=0}^{n-1}\rho^k.$$

Si $\rho \leq 1$ alors, l'inégalité à démontrer est vraie. Sinon,

$$\cancel{\rho^n} \leq \mathrm{M} \times \frac{\rho^n - 1}{\rho - 1} \leq \frac{\mathrm{M}\cancel{\rho^n}}{\rho - 1},$$

donc $\rho \leq 1 + \mathrm{M}$.

2. D'après le calcul de la question précédente,

$$\rho^n \leq \sum_{k=0}^{n-1}|a_k|\,\rho^k.$$

→ Si $\rho \leq 1$, il n'y a rien à démontrer.
→ Si $\rho > 1$, alors $\rho \leq \sum_{k=0}^{n-1}\frac{|a_k|}{\rho^{n-1-k}} \leq \sum_{k=0}^{n-1}|a_k|$.

Correction de l'exercice III.3. :

Soit $\mathrm{X} \in \mathcal{M}_{n,1}(\mathbb{R})$. Pour toute matrice diagonale $\mathrm{D} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}^{+*})$, et pour tout $\mathrm{U} =$

CC BY-NC-SA

$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$ on a

$$\|\mathrm{U}\|^2 = \sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} \left| \sqrt{d_i} u_i \times \frac{u_i}{\sqrt{d_i}} \right| \underset{\substack{\uparrow \\ \text{Cauchy-Schwarz}}}{\leq} \overbrace{\sqrt{\sum_{i=1}^{n} d_i u_i^2}}^{\sqrt{\langle \mathrm{DU}|\mathrm{U}\rangle}} \underbrace{\sqrt{\sum_{i=1}^{n} \frac{u_i^2}{d_i}}}_{\sqrt{\langle \mathrm{D}^{-1}\mathrm{U}|\mathrm{U}\rangle}} .$$

Or A est une matrice symétrique réelle définie positive, donc il existe $\mathrm{P} \in \mathcal{O}_n(\mathbb{R})$ tel que $\underbrace{\mathrm{P}^\top \mathrm{AP}}_{\text{notée } \Delta}$ soit une matrice diagonale à coefficients diagonaux strictement positifs.
Donc

$$\|\mathrm{X}\|^2 = \left\|\mathrm{P}^\top \mathrm{X}\right\|^2 \leq \sqrt{\langle \Delta \mathrm{P}^\top \mathrm{X}|\mathrm{P}^\top \mathrm{X}\rangle}\sqrt{\langle \Delta^{-1}\mathrm{P}^\top \mathrm{X}|\mathrm{P}^\top \mathrm{X}\rangle} = \sqrt{\langle \mathrm{P}\Delta \mathrm{P}^\top \mathrm{X}|\mathrm{X}\rangle}\sqrt{\langle \mathrm{P}\Delta^{-1}\mathrm{P}^\top \mathrm{X}|\mathrm{X}\rangle},$$

ce qu'il fallait démontrer.

Correction de l'exercice III.4. :

Pour tout $\mathrm{X} \in \mathrm{E}$, $\sum_{1\leq i\neq j\leq n} x_i x_j \geq 0$. Or $(\underbrace{0,\dots,0}_{n-1 \text{ zéros}},1) \in \mathrm{E}$ et réalise ce minorant, donc $m = 0$. Soit $\mathrm{X} \in \mathrm{E}$.
$\sum_{1\leq i\neq j\leq n} x_i x_j = \left(\sum_{i=1}^{n} x_i\right)^2 - \sum_{i=1}^{n} x_i^2 = 1 - \sum_{i=1}^{n} x_i^2$.

Or, par l'inégalité de Cauchy-Schwarz, $n\sum_{i=1}^{n} x_i^2 \geq \left(\sum_{i=1}^{n} x_i\right)^2 = 1$.
Donc, pour tout $\mathrm{X} \in \mathrm{E}$,

$$\sum_{1\leq i\neq j\leq n} x_i x_j \leq 1 - \frac{1}{n} = \frac{n-1}{n}.$$

Ce majorant est atteint en $\left(\frac{1}{n},\dots,\frac{1}{n}\right) \in \mathrm{E}$. En effet, $\sum_{1\leq i\neq j\leq n} \frac{1}{n^2} = \frac{2}{n^2}\binom{n}{2} = \frac{n-1}{n}$, donc

$$\mathrm{M} = \frac{n-1}{n}.$$

Le nombre de couples (i,j) de $[\![1;n]\!]^2$ tel que $i \neq j$ est le double du nombre de couples (i,j) de $[\![1;n]\!]^2$ tel que $i < j$, qui vaut le nombre d'issues possibles lors du choix simultané de 2 éléments distincts d'un ensemble à n éléments, et ce nombre de combinaisons vaut $\binom{n}{2}$.

(cc) BY-NC-SA

Chapitre 2

Fonctions convexes

I Définitions, pentes, régularité

Dans cette partie, on note I un intervalle de $\mathbb{R}$, non trivial et on considère f une fonction de I dans $\mathbb{R}$. Pour tout $(a, b) \in \mathbb{R}^2$, on note

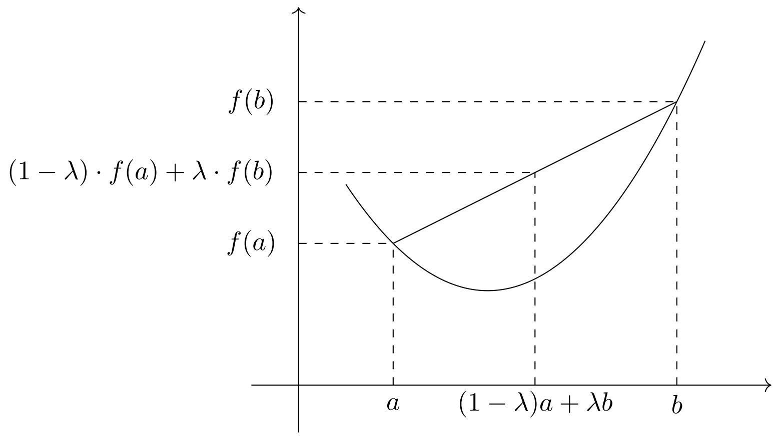
$$[a, b] = \{(1-t)a + tb,\ t \in [0, 1]\}.$$

Définition I.1.

On dit que la fonction f est convexe lorsque

$$\forall (a, b) \in \mathrm{I}^2\ ,\ \forall \lambda \in [0, 1]\ ,\ f\left((1-\lambda)a + \lambda b\right) \leq (1-\lambda) f(a) + \lambda f(b).$$

Remarque. Graphiquement, la définition ci-haut se voit facilement sur un dessin : une fonction convexe sur un intervalle a, b est en dessous de la droite joignant les points $(a, f(a))$ et $(b, f(b))$ sur l'intervalle $[a, b]$. Ceci est aussi valable pour tout intervalle inclus dans I.



La droite qui joint $(a, f(a))$ et $(b, f(b))$ est la courbe de la fonction $g : x \mapsto \dfrac{f(b)-f(a)}{b-a} \cdot (x-a) + f(a)$. On a ainsi $g((1-\lambda)a + \lambda b) = (1-\lambda) \cdot f(a) + \lambda f(b)$.

Avant de faire l'exercice ci-dessous, le lecteur souhaitant revoir la définition et propriétés des ensembles convexes est invité à voir le chapitre 11.9 sur la convexité.

Exercice I.2.

Montrer que f est convexe si et seulement si l'ensemble $\mathrm{Epi}(f) = \{(x, y) \in \mathrm{I} \times \mathbb{R}, f(x) \leq y\}$ est convexe.

Théorème (Théorème des trois pentes) I.3.

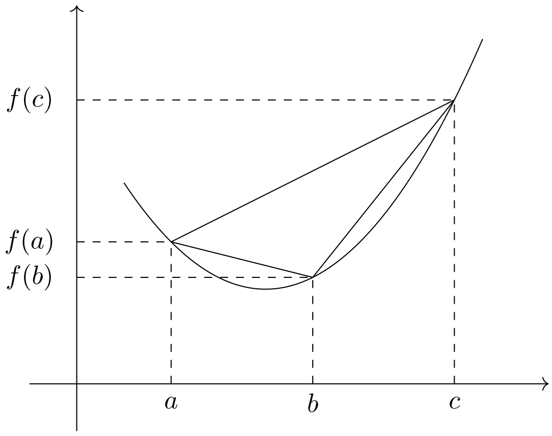
On note pour tous $a, b \in \mathrm{I}$ distincts, $p_{a,b}(f) = \frac{f(b)-f(a)}{b-a}$, la pente de la droite joignant les points $(a, f(a))$ et $(b, f(b))$. f est convexe sur I si et seulement si, pour tous $a, b, c \in \mathrm{I}$, si $a < b < c$, alors on a

$$p_{a,b}(f) \leq p_{a,c}(f) \leq p_{b,c}(f).$$

(cc) BY-NC-SA

Démonstration.

→ (⇒) Observons tout d'abord le dessin suivant, qui nous donne une idée de la preuve.



Lorsque f est convexe, la pente de la droite passant par deux points $(a, f(a)), (b, f(b))$ est croissante lorsque a et/ou b sont translatés vers la droite, chose qu'on observe sur le dessin ci-dessus. Montrons à présent ce résultat. Pour le montrer, il suffit de remarquer que étant donné que $b \in]a, c[$, on peut dire qu'il existe $\lambda \in]0, 1[$ tel que $b = \lambda a + (1 - \lambda)c$, et alors, on écrit

$$\begin{aligned} p_{a,b}(f) &= \frac{f(b) - f(a)}{b - a} = \frac{f(\lambda a + (1-\lambda)c) - f(a)}{(1-\lambda)(c-a)} \\ &\leq \frac{\lambda f(a) + (1-\lambda)f(c) - f(a)}{(1-\lambda)(c-a)} = \frac{f(c) - f(a)}{c - a} = p_{a,c}(f), \end{aligned}$$

$$\begin{aligned} p_{b,c}(f) &= \frac{f(b) - f(c)}{b - c} = \frac{f(\lambda a + (1-\lambda)c) - f(c)}{\lambda \underbrace{(a-c)}_{<0}} \\ &\geq \frac{\lambda f(a) + (1-\lambda)f(c) - f(c)}{\lambda(a-c)} = \frac{f(a) - f(c)}{a - c} = p_{a,c}(f). \end{aligned}$$

→ (⇐) Soient $a, b \in \mathrm{I}$ tels que $a < b$, $\lambda \in]0, 1[$, et $m = \lambda a + (1-\lambda)b$. L'inégalité supposée vraie donne $p_{a,m}(f) \leq p_{a,b}(f)$. On a alors

$$\begin{aligned} p_{a,m}(f) \leq p_{a,b}(f) &\Longrightarrow \frac{f(m) - f(a)}{m - a} \leq \frac{f(b) - f(a)}{b - a} \\ &\Longrightarrow \frac{f(m) - f(a)}{(1-\lambda)(b-a)} \leq \frac{f(b) - f(a)}{b - a} \\ &\Longrightarrow f(m) \leq f(a) + (1-\lambda)(f(b) - f(a)) = \lambda f(a) + (1-\lambda)f(b) \\ &\Longrightarrow f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b), \end{aligned}$$

ce qui est bien le résultat voulu. Les cas $\lambda \in \{0, 1\}$ et $a = b$ sont évidents, nous ne les traiterons donc pas.

□

Remarque. Le résultat démontré ci-dessus permet de montrer que la pente de la droite passant par deux points $(x, f(x))$ et $(a, f(a))$, lorsque f est convexe, est croissante en x. En effet, rigoureusement,

(cc) BY-NC-SA

cela signifie que la fonction

$$g : \begin{cases} \mathrm{I} \setminus \{a\} & \longrightarrow \mathbb{R} \\ x & \longmapsto \dfrac{f(x) - f(a)}{x - a} \end{cases}$$

est croissante lorsque f est convexe. En effet, pour tout $x, y \in \mathrm{I} \setminus \{a\}$ tels que $x < y$, on a

$$g(x) = \frac{f(x) - f(a)}{x - a} = p_{a,x}(f) \leq p_{a,y}(f) = g(y).$$

On va maintenant énoncer quelques propriétés de régularité des fonctions convexes.

Proposition I.4.

Si f est convexe dans I, alors les propositions suivantes sont vraies.

1. En tout point dans l'intérieur de I (différent des bords de I), f admet une dérivée à gauche et à droite. De plus, pour tout a dans l'intérieur de I, en notant f'_g et f'_d respectivement les dérivées à droite et à gauche de f en a, on a $f'_g(a) \leq f'_d(a)$.
2. f est continue en tout point dans l'intérieur de I.

Démonstration.

1. Soit a dans l'intérieur de I. Il s'agit de montrer que le taux d'accroissement $x \longmapsto p_{a,x}(f)$ admet une limite à droite et à gauche de a. Posons encore une fois

$$g : \begin{cases} \mathrm{I} \setminus \{a\} & \longrightarrow \mathbb{R} \\ x & \longmapsto \dfrac{f(x) - f(a)}{x - a}. \end{cases}$$

   g est croissante sur $\mathrm{I} \setminus \{a\}$. g admet donc des limites à droite et à gauche de a. L'implication entre monotonie et existence des limites à droite et à gauche est assez connue, mais nous la montrons quand même. Soit $b \in \mathrm{I}$ tel que $b > a$. La restriction de g à $\mathrm{I} \cap]-\infty, a[$ est croissante majorée par $g(b)$. On en déduit que g admet alors une limite à gauche de a. L'existence de la limite à droite peut être obtenue de la même manière. De plus, la croissance de g nous donne (par comparaison des limites à droite et à gauche) directement l'inégalité $f'_g(a) \leq f'_d(a)$.
2. D'après le point précédent, on sait que $g(x) \xrightarrow[x \to a^+]{} f'_d(a)$. Il existe alors $\eta > 0$ tel que pour tout $x \in]a, a+\eta[$

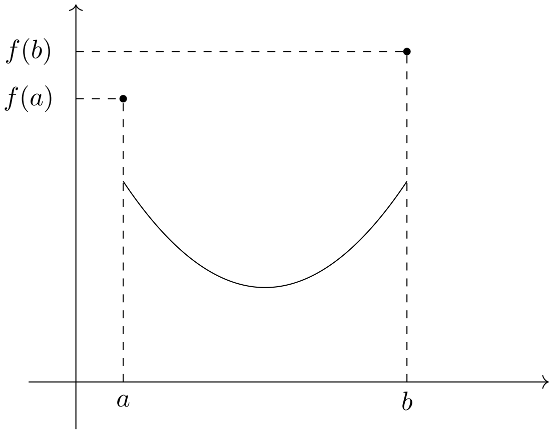
$$\left| \frac{f(x) - f(a)}{x - a} - f'_d(a) \right| \leq 1 \text{ et alors } |f(x) - f(a)| \leq |x - a| + |f'_d(a)(x - a)| \xrightarrow[x \to a^+]{} 0.$$

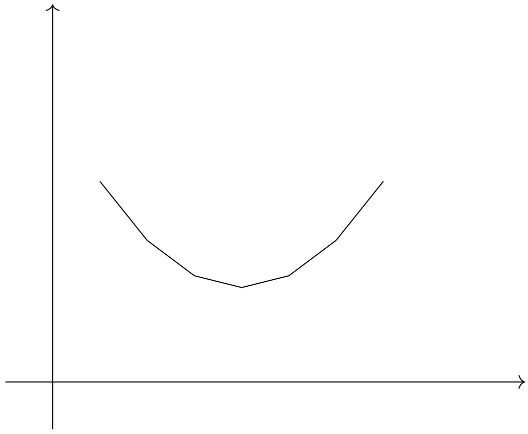
   f est donc continue à droite de a. La continuité à gauche peut être démontrée de la même manière.

$\square$

Remarque. Attention, il est possible que f soit convexe sur I, mais discontinue sur les bords de l'intervalle où elle est définie. En effet, on peut remarquer par exemple que la fonction ci-dessous à gauche, bien que discontinue aux bords, est convexe. Le lecteur pourra vérifier que la définition de la convexité s'applique bien dans ce cas. De même, il est possible qu'une fonction soit convexe mais pas forcément dérivable sur tout point de I. On peut voir de phénomène sur la figure de droite ci-dessous, par la présence de plusieurs points anguleux.

(cc) BY-NC-SA





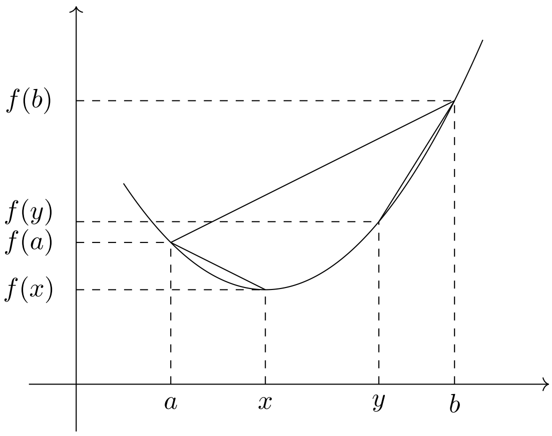
Corollaire I.5.

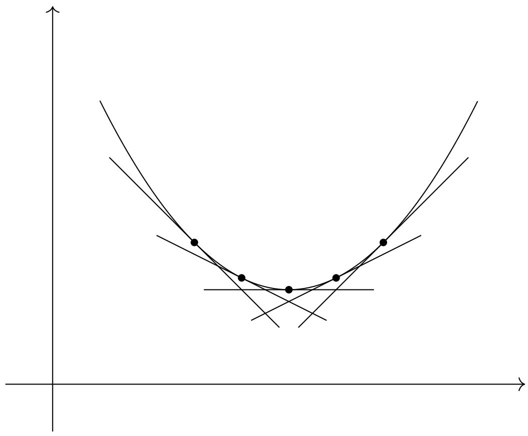
Si f est convexe et dérivable sur I, alors f' est croissante sur I.

Démonstration. Soit $a, b \in$ I tels que $a < b$. f est dérivable, donc pour tout point dans l'intérieur de I, les dérivées à droite et à gauche de f sont égales. Notre intuition pour montrer ce résultat est de comparer les dérivées de f respectivement à droite de a et à gauche de b. Pour ce faire, on va utiliser le théorème des trois pentes (théorème I.3) pour comparer la pente des tangentes en a et en b. En effet, on a pour tout $x, y \in$ I tels que $a < x < y < b$,

$$p_{a,x}(f) \leq p_{a,y}(f) \leq p_{b,y}(f).$$

On peut observer cette inégalité sur le dessin de gauche ci-dessous.





Ensuite, il suffit de faire tendre x vers a à droite et y vers b à gauche dans l'inégalité $p_{a,x}(f) \leq p_{b,y}(f)$ pour obtenir $\underbrace{f'_d(a)}_{=f'(a)} \leq \underbrace{f'_g(b)}_{=f'(b)}$, ce qui donne bien que f' est croissante sur I. $\square$

Cette propriété de croissance de la dérivée peut être facilement vue graphiquement, comme on peut la voir sur la figure de droite ci-dessus.

Remarquons que sous les mêmes hypothèses, la dérivée de f est définie dans l'intérieur de I, et qu'uniquement les dérivées de gauche et de droite respectivement sont définies sur les bords de droite et de gauche de I. Le lecteur est donc invité à montrer que la dérivée à gauche du bord de droite de I est supérieure à la dérivée de droite du bord gauche de I, et que si b et c sont respectivement les bords de gauche et de droite de I, alors pour tout $a \in$ I où f est dérivable, $f'_d(b) \leq f'(a) \leq f'_g(c)$. On peut également montrer de la même manière, même en l'absence d'hypothèse de dérivabilité de f, que f'_d et f'_g sont croissantes sur I.

(cc) BY-NC-SA

Proposition I.6.

Supposons que f est dérivable sur I. Si f' est croissante sur I, alors f est convexe sur I.

Démonstration. Supposons que f' est croissante sur I et supposons par l'absurde que f n'est pas convexe. D'après le théorème des 3 pentes (théorème I.3), il existe $a, b, c \in$ I tels que $a < b < c$ et $p_{a,b}(f) > p_{b,c}(f)$. En appliquant le théorème des accroissements finis sur $]a, b[$ et $]b, c[$, on peut affirmer qu'il existe $u \in]a, b[$ et $v \in]b, c[$ tels que

$$f'(u) = \frac{f(b) - f(a)}{b - a} = p_{a,b}(f) \text{ et } f'(v) = \frac{f(b) - f(c)}{b - c} = p_{b,c}(f).$$

Le fait que f' est croissante sur I nous donne que $f'(u) \leq f'(v)$, ce qui est absurde car cela implique que $p_{a,b}(f) \leq p_{b,c}(f)$. $\square$

Corollaire I.7.

Si f est deux fois dérivable sur I, alors f est convexe si et seulement si f'' est positive sur I.

Démonstration. Il suffit d'utiliser le fait que lorsque f est deux fois dérivable, qu'il y a équivalence entre f'' positive et f' croissante. $\square$

Proposition (position par rapport à la tangente) I.8.

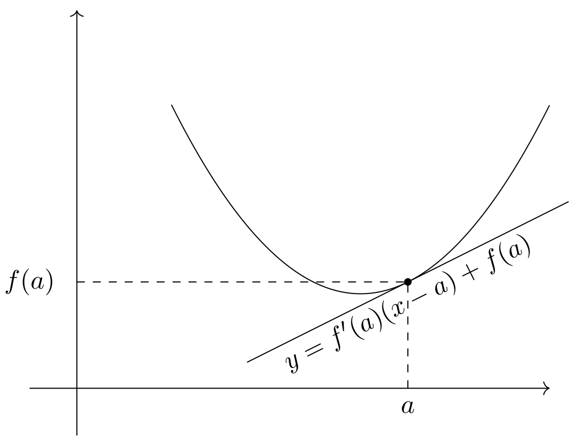
Si f est convexe sur I, alors pour tout $a \in$ I, si f est dérivable en a, alors f est au dessus de sa tangente en a, c'est-à-dire que pour tout $x \in$ I,

$$f(x) \geq f'(a)(x - a) + f(a).$$

Démonstration. On utilise encore le théorème des trois pentes de la manière suivante. On a pour tout $x, y \in$ I tels que $x > y > a$, $p_{a,x}(f) \geq p_{a,y}(f)$. En passant à la limite lorsque y tend vers a à droite, on obtient que

$$p_{a,x}(f) \geq f'(a) \;\; i.e. \;\; f(x) \geq f'(a)(x - a) + f(a).$$

Le cas $x < a$ se fait de manière identique, et le cas $x = a$ est évident. On a donc bien le résultat voulu. $\square$



(cc) BY-NC-SA

Remarques.

→ Attention, ici, on a uniquement supposé que f est dérivable en a, mais pas autre-part.

→ Même lorsque f n'est pas dérivable en a, on peut montrer que la même inégalité est toujours vraie en substituant f' par f'_d ou f'_g.

II Inégalités de convexité

1. Fonctions strictement convexes

On conserve les notations de la partie précédente : on note I un intervalle de $\mathbb{R}$, non trivial et on considère f une fonction de I dans $\mathbb{R}$.

Définition II.1.

On dit que f est strictement convexe lorsque pour tous $a, b \in \mathrm{I}$ distincts,

$$\forall \lambda \in]0,1[,\ f((1-\lambda)a+\lambda b) < (1-\lambda)f(a)+\lambda f(b).$$

Remarques.

→ Lorsque f est strictement convexe, f ne peut être plate nulle part, c'est-à-dire qu'aucune portion de sa courbe ne doit être un segment.

→ Lorsque f est strictement convexe, l'inégalité des 3 pentes s'applique et devient stricte.

Proposition II.2.

Les propositions suivantes sont vraies.

1. Si f est dérivable sur I et f' est strictement croissante, alors f est strictement convexe.
2. Si f est deux fois dérivable sur I et f'' est strictement positive sur I, alors f est strictement convexe.

Démonstration. La preuve est quasi-identique aux preuves précédentes pour la convexité et est donc laissée comme exercice au lecteur. □

Remarque. On sait que si f est dérivable, alors f est croissante sur I si et seulement si f' est positive sur I. Par contre, si f est strictement croissante, alors ceci n'est pas équivalent au fait que f' est strictement positive sur I. En effet, la stricte croissance de f est équivalente à une condition qui est plus faible que la stricte positivité de la dérivée de f, qui est la suivante : f' est positive sur I, et $\mathrm{Z}(f)=\{x\in\mathrm{I},\ f'(x)=0\}$ est d'intérieur vide. Montrons cette équivalence.

→ Supposons que f' est positive sur I, et $\mathrm{Z}(f)=\{x\in\mathrm{I},\ f'(x)=0\}$ est d'intérieur vide. Si f n'est pas strictement croissante, alors il existe $\alpha,\beta\in\mathrm{I}$ tels que $\alpha<\beta$ et $f(\alpha)=f(\beta)$. On en déduit donc que f est constante sur $[\alpha,\beta]$, et que donc sa dérivée est nulle sur $]\alpha,\beta[$, *i.e.* $]\alpha,\beta[\subset \mathrm{Z}(f)$, ce qui est absurde.

→ Supposons que f est strictement croissante sur I. On a alors f' est positive sur I. Supposons que $\mathrm{Z}(f)$ n'est pas d'intérieur vide, i.e. il existe $\alpha,\beta\in\mathrm{I}$ tels que $\alpha<\beta$ et $]\alpha,\beta[\subset\mathrm{Z}(f)$, i.e. f' est nulle sur $]\alpha,\beta[$ et est constante sur cet intervalle, ce qui est absurde car f est strictement croissante.

(cc) BY-NC-SA

2. Inégalité fondamentale

Proposition II.3.

Si f est convexe, alors pour tout $n \in \mathbb{N}^*$, $x_1, \dots, x_n \in \mathrm{I}$ et $\lambda_1, \dots, \lambda_n \in [0,1]$, tels que $\lambda_1 + \dots + \lambda_n = 1$,

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

Si de plus, $x_1, \dots x_n$ sont non tous égaux, $\lambda_1, \dots, \lambda_n$ sont tous strictement positifs, et f est strictement convexe, alors l'inégalité devient stricte. En effet, lorsque f est strictement convexe et pour tout $k \in [\![1;n]\!]$, $\lambda_k > 0$, l'inégalité est une égalité si et seulement si $x_1 = \dots = x_n$.

Démonstration. Nous allons montrer l'inégalité dans le cas où f est convexe. Montrons le résultat par récurrence. Lorsque $n = 1$, le résultat est évident. Soit $n \in \mathbb{N}^*$. Supposons que le résultat est vérifié pour n. Soit $x_1, \dots, x_{n+1} \in \mathrm{I}$, et $\lambda_1, \dots, \lambda_{n+1} \in [0,1]$ tels que $\lambda_1 + \dots + \lambda_{n+1} = 1$. On suppose sans perte de généralité que $\lambda_{n+1} < 1$ (le cas $\lambda_{n+1} = 1$ est évident). On a

$$\begin{aligned} f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) &= f\left((1-\lambda_{n+1})\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k + \lambda_{n+1} x_{n+1}\right) \\ &\underset{(1)}{\leq} (1-\lambda_{n+1}) f\left(\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k\right) + \lambda_{n+1} f(x_{n+1}) \\ &\underset{(2)}{\leq} (1-\lambda_{n+1})\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} f(x_k) + \lambda_{n+1} f(x_{n+1}) = \sum_{k=1}^{n+1} \lambda_k f(x_k). \end{aligned}$$

L'inégalité (1) est vraie par convexité de f, et (2) est vraie par hypothèse de récurrence étant donné que f est convexe, $1 - \lambda_{n+1} = \lambda_1 + \dots + \lambda_n$, et

$$\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} = \sum_{k=1}^{n} \frac{\lambda_k}{\underbrace{\lambda_1 + \dots + \lambda_n}_{\in]0,1[}} = 1.$$

On a donc bien montré l'inégalité voulue.

Montrons à présent par récurrence que le cas d'égalité, lorsque f est strictement convexe et $\lambda_1, \dots, \lambda_n \in]0,1[$, implique $x_1 = \dots = x_n$ (l'implication réciproque est évidente). Lorsque $n = 1$, la propriété est vérifiée. Soit $n \in \mathbb{N}^*$ tel que la propriété est vérifiée pour n. Montrons qu'elle est vérifiée pour $n+1$. On veut d'abord essayer d'utiliser l'inégalité de convexité stricte pour le cas $n = 2$, comme on l'a fait dans la preuve pour l'inégalité. On a

$$\sum_{k=1}^{n+1} \lambda_k x_k = \underbrace{(1-\lambda_{n+1})}_{]0,1[} \sum_{k=1}^{n} \underbrace{\frac{\lambda_k}{1-\lambda_{n+1}}}_{\mu_k} x_k + \lambda_{n+1} x_{n+1}.$$

On a donc, si $x_{n+1} \neq \sum_{k=1}^{n} \mu_k x_k$, alors étant donné que $\lambda_{n+1} \in]0,1[$ et $1 - \lambda_{n+1} \in]0,1[$, l'inégalité de stricte

BY-NC-SA

convexité donne

$$\begin{aligned} f\left(\sum_{k=1}^{n+1}\lambda_k x_k\right) &= f\left((1-\lambda_{n+1})\sum_{k=1}^{n}\mu_k x_k + \lambda_{n+1}x_{n+1}\right) \\ &< (1-\lambda_{n+1})f\left(\sum_{k=1}^{n}\frac{\lambda_k}{1-\lambda_{n+1}}x_k\right) + \lambda_{n+1}f(x_{n+1}) \\ &\leq (1-\lambda_{n+1})\sum_{k=1}^{n}\frac{\lambda_k}{1-\lambda_{n+1}}f(x_k) + \lambda_{n+1}f(x_{n+1}) = \sum_{k=1}^{n+1}\lambda_k f(x_k), \end{aligned}$$

ce qui est absurde. On en déduit donc qu'on a bien $x_{n+1} = \sum_{k=1}^{n}\mu_k x_k$, et alors en substituant x_{n+1} par $\sum_{k=1}^{n}\mu_k x_k$, l'égalité devient

$$f\left(\sum_{k=1}^{n}\mu_k x_k\right) = \sum_{k=1}^{n}\mu_k f(x_k),$$

ce qui implique par hypothèse de récurrence, étant donné que pour tout $k \in [\![1; n]\!]$, $\mu_k > 0$, que $x_1 = \cdots = x_n$ et que finalement $x_1 = \cdots = x_n = x_{n+1}$, ce qui est bien le résultat voulu. Remarquons qu'on aurait aussi pu se passer de la récurrence en supposant sans perte de généralité que $x_{n+1} = \max_{k\in[\![1;n+1]\!]} x_k$. $\square$

Exemple. En fait, avec cette inégalité, on peut facilement montrer l'inégalité arithmético-géométrique. En effet, l'énoncé de cet inégalité est le suivant. Pour tout $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n \in \mathbb{R}^+$, on a

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \ldots x_n},$$

avec égalité si et seulement si tous les x_i sont égaux.

Lorsque l'un des x_i est nul, l'inégalité est facilement vérifiable. Supposons donc que $x_1, \ldots, x_n \in \mathbb{R}_+^*$. L'inégalité est équivalente à

$$\frac{e^{\ln x_1} + \cdots + e^{\ln x_n}}{n} \geq e^{\frac{\ln x_1 + \cdots + \ln x_n}{n}}.$$

Il est donc facile de voir qu'il suffit d'utiliser la stricte convexité de l'exponentielle pour obtenir le résultat voulu.

Exercice II.4.

Soit $a, b, c \in \mathbb{R}_+^*$. Calculer

$$\inf_{x,y\in\mathbb{R}_+^*}\left(ax + by + \frac{c}{xy}\right).$$

(cc) BY-NC-SA

Exercice II.5.

Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue.

1. Montrer que g est convexe si et seulement si

$$\forall x, y \in \mathbb{R},\ g\left(\frac{x+y}{2}\right) \leq \frac{g(x)+g(y)}{2}.$$

2. On ne suppose plus que g est continue, mais plutôt que g est bornée sur tout segment de $\mathbb{R}$ et qu'elle vérifie l'inégalité ci-dessus. Montrer que g est convexe.

III Compléments

1. Inégalités de Hölder et de Minkowski

Proposition (Inégalité de Hölder) III.1.

Soient $p, q \in \mathbb{R}_+^*$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Pour tout $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$, on a

$$\sum_{k=1}^{n} |x_k y_k| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}}.$$

Démonstration. On suppose sans perte de généralité que pour tout $k \in [\![1; n]\!]$, $x_k \neq 0$ et $y_k \neq 0$ (le lecteur pourra s'assurer qu'on peut effectivement faire cette hypothèse). On pose alors pour tout $k \in [\![1; n]\!]$, $e^{s_k} = |x_k|$ et $e^{t_k} = |y_k|$. On a alors pour tout $k \in [\![1; n]\!]$, en utilisant la convexité de l'exponentielle,

$$|x_k y_k| = e^{s_k + t_k} = e^{\frac{1}{p} p s_k + \frac{1}{q} q t_k} \leq \frac{1}{p} e^{p s_k} + \frac{1}{q} e^{q t_k} = \frac{1}{p} |x_k|^p + \frac{1}{q} |y_k|^q$$

Posons $\mathrm{A} = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \neq 0$ et $\mathrm{B} = \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}} \neq 0$. En substituant dans l'inégalité ci-dessus pour tout $k \in [\![1; n]\!]$ x_k par $\dfrac{x_k}{\mathrm{A}}$ et y_k par $\dfrac{y_k}{\mathrm{B}}$ et en sommant, on obtient

$$\frac{1}{\mathrm{AB}} \sum_{k=1}^{n} |x_k y_k| \leq \frac{1}{p} \underbrace{\frac{\sum_{k=1}^{n} |x_k|^p}{\mathrm{A}^p}}_{=1} + \frac{1}{q} \underbrace{\frac{\sum_{k=1}^{n} |x_k|^q}{\mathrm{B}^q}}_{=1} = 1,$$

et finalement en multipliant par AB des deux côtés,

$$\sum_{k=1}^{n} |x_k y_k| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}}.$$

$\square$

Remarque. Cette inégalité est également vérifiée lorsqu'on remplace les sommes par des intégrales.

(cc) BY-NC-SA

Autrement dit, pour tout f, g fonctions continues sur un intervalle $[a, b]$, on a

$$\int_a^b |f(x)g(x)|\,dx \leq \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q\,dx\right)^{\frac{1}{q}}.$$

la preuve de cette inégalité se fait de manière identique, en substituant somme par intégrale.

Lemme (Quasi-linéarité) III.2.

Soit $n \in \mathbb{N}^*$ et L une fonction de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$, continue et bornée sur les ensembles bornés, telle que pour tout $y \in \mathbb{R}^n$, l'application $x \longmapsto \mathrm{L}(x, y)$ est linéaire. Si $\mathrm{B} \subset \mathbb{R}^n$ est un borné de $\mathbb{R}^n$, alors $f : x \longmapsto \sup_{y \in \mathrm{B}} \mathrm{L}(x, y)$ est sous-additive, c'est-à-dire que

$$\forall x_1, x_2 \in \mathbb{R}^n,\ f(x_1 + x_2) \leq f(x_1) + f(x_2).$$

Démonstration. Il suffit d'écrire les définition. En effet, on a pour tout $y \in \mathrm{B}$ et $x_1, x_2 \in \mathbb{R}^n$,

$$\mathrm{L}(x_1 + x_2, y) = \mathrm{L}(x_1, y) + \mathrm{L}(x_2, y) \leq \sup_{y \in \mathrm{B}} \mathrm{L}(x_1, y) + \sup_{y \in \mathrm{B}} \mathrm{L}(x_2, y) = f(x_1) + f(x_2).$$

En passant alors à la borne supérieure en $y \in \mathrm{B}$ dans la partie gauche de l'inégalité, on obtient

$$f(x_1 + x_2) \leq f(x_1) + f(x_2).$$

□

Proposition (Inégalité de Minkowski) III.3.

Soit $p > 1$. On a alors pour tout $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$,

$$\left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p\right)^{\frac{1}{p}}$$

Démonstration. En fait, cette inégalité est équivalente au fait de dire que la fonction

$$f : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x_1 \ldots, x_n) & \longmapsto \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} \end{cases}$$

est sous-additive. Pour prouver ce résultat, nous allons utiliser le lemme III.2, en montrant qu'il existe $\mathrm{L} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ linéaire en son premier argument, ainsi que bornée sur les bornés, et un ensemble borné $\mathrm{B} \subset \mathbb{R}^n$ tels que pour tout $x_1, \ldots, x_n \in \mathrm{R}$, $f(x_1, \ldots, x_n) = \sup_{z \in \mathrm{B}} \mathrm{L}(x_1, \ldots, x_n, z)$. Posons pour tout $x_1, \ldots, x_n, z_1, \ldots, z_n \in \mathbb{R}$,

$$\mathrm{L}(x_1, \ldots, x_n, z_1, \ldots, z_n) = \sum_{k=1}^n x_k z_k,$$

(cc) BY-NC-SA

puis considérons $q > 1$ tel que $\dfrac{1}{p} + \dfrac{1}{q} = 1$, et le sous-ensemble borné de $\mathbb{R}^n$

$$\mathrm{B} = \left\{ (z_1, \ldots, z_n) \in \mathbb{R}^n,\ \sum_{k=1}^{n} |z_k|^q = 1 \right\}.$$

L est clairement linéaire en son premier argument et bornée sur les bornés. On veut alors montrer que pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n \setminus \{(0, \ldots, 0)\}$,

$$\sup_{(z_1,\ldots,z_n)\in \mathrm{B}} \mathrm{L}(x_1, \ldots, x_n, z_1, \ldots, z_n) = f(x_1, \ldots, x_n).$$

En utilisant l'inégalité de Hölder (proposition III.1), on a

$$\sum_{k=1}^{n} x_k z_k \leq \sum_{k=1}^{n} |x_k z_k| \leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}} \underbrace{\left(\sum_{k=1}^{n} |z_k|^q \right)^{\frac{1}{q}}}_{=1} = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}}.$$

Il suffit maintenant de montrer que cette borne supérieure est atteinte. Trouvons donc $(z_1, \ldots, z_n) \in \mathrm{B}$ réalisant cette borne supérieure. Dans le but de faire apparaitre $|x_k|^p$, $k \in [\![1; n]\!]$, on est tentés de poser pour tout $k \in [\![1; n]\!]$, $z_k = \operatorname{signe}(x_k) |x_k|^{p-1}$. Cependant, $z_1, \ldots, z_n$ tels que défini risque de ne pas appartenir à B, donc il faut ajouter une constante de normalisation. Posons donc pour tout $k \in [\![1; n]\!]$, $z_k = \lambda \operatorname{signe}(x_k) |x_k|^{p-1}$ où λ est une constante positive qui nous permet de garantir que $(z_1, \ldots, z_n) \in \mathrm{B}$. On a alors

$$1 = \sum_{k=1}^{n} |z_k|^q = \sum_{k=1}^{n} \lambda^q |x_k|^{q(p-1)} = \lambda^q \underbrace{\sum_{k=1}^{n} |x_k|^p}_{\neq 0} \text{ et donc } \lambda = \frac{1}{\left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{q}}}.$$

On en déduit que

$$\sum_{k=1}^{n} x_k z_k = \sum_{k=1}^{n} \lambda \operatorname{signe}(x_k) |x_k|^{p-1} x_k = \lambda \sum_{k=1}^{n} |x_k|^p = \frac{\sum_{k=1}^{n} |x_k|^p}{\left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{q}}} = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}}.$$

On a donc bien que

$$\sup_{(z_1,\ldots,z_n)\in \mathrm{B}} \mathrm{L}(x_1, \ldots, x_n, z_1, \ldots, z_n) = \sup_{(z_1,\ldots,z_n)\in \mathrm{B}} \sum_{k=1}^{n} x_k y_k = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}} = f(x_1, \ldots, x_n),$$

et alors le lemme III.2 permet de conclure. □

Remarques.

→ Attention, l'inégalité de Minkowski n'est pas valable lorsque $p < 1$. En effet, par exemple, pour tout $p < 1$, lorsque $n = 2$, dans le cas où $(x_1, x_2) = (1, 0)$ et $(y_1, y_2) = (0, 1)$, le côté gauche de l'inégalité est égal à $2^{\frac{1}{p}}$ et le côté droit à 2, et on a clairement $2^{\frac{1}{p}} > 2$.

→ Bien entendu, cette inégalité admet également une version continue : pour toutes fonctions f, g continues sur un intervalle $[a, b]$, et $p > 1$,

$$\left(\int_a^b |f(x) + g(x)|^p \, dx \right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, dx \right)^{\frac{1}{p}}.$$

(cc) BY-NC-SA

2. Inégalité de Jensen

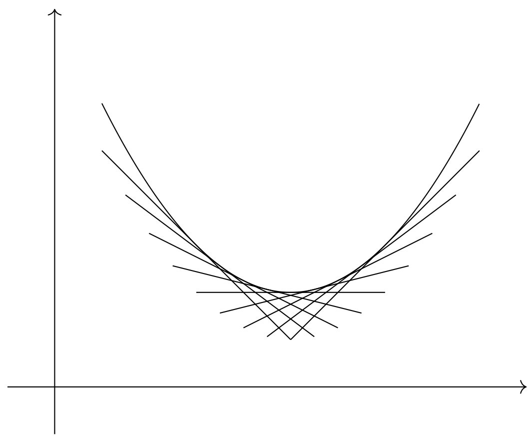
Proposition (Inégalité de Jensen) III.4.

Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et f une fonction continue de $[a, b]$ dans $\mathbb{R}$ où $[a, b]$ est un intervalle non vide de $\mathbb{R}$ et non réduit à un singleton. Les deux propositions suivantes sont vraies.

1. Si g est convexe, alors il existe une famille de fonctions affines $(\varphi_\lambda)_{\lambda\in\Lambda}$ où Λ est un ensemble, telle que pour tout $x \in \mathbb{R}$, $g(x) = \sup\limits_{\lambda\in\Lambda} \varphi_\lambda(x)$.
2. Si g est convexe et $a < b$, alors l'inégalité
$$g\left(\frac{1}{b-a}\int_a^b f(x)dx\right) \leq \frac{1}{b-a}\int_a^b g\circ f(x)dx$$
est vérifiée.

Démonstration.

1. Pour simplifier, on suppose que g est dérivable en tout point de $\mathbb{R}$. En rejetant un coup d'oeil à la proposition I.8, on est tentés de proposer la famille $(\varphi_\lambda)_{\lambda\in\mathbb{R}}$ des tangentes à la courbe en tout point de $\mathbb{R}$, c'est-à-dire
$$\forall \lambda \in \mathbb{R},\ \forall x \in \mathbb{R},\ \varphi_\lambda(x) = g'(\lambda)(x-\lambda) + g(\lambda).$$
On sait d'après la proposition I.8 que pour tout $x \in \mathbb{R}$ et tout $a \in \mathbb{R}$, $g(x) \geq \varphi_a(x)$. De plus, pour tout $x \in \mathbb{R}$, $g(x) = \varphi_x(x)$, et alors, on a bien que $g(x) = \sup\limits_{\lambda\in\mathbb{R}} \varphi_\lambda(x)$. Maintenant, on peut voir qu'on peut faire exactement la même chose en se passant de l'hypothèse de dérivabilité de g, et en substituant g' par g'_d ou g'_g. En fait, on peut voir cette propriété graphiquement : une fonction convexe (lorsqu'elle est dérivable) est l'enveloppe supérieure de la famille de ses tangentes.



2. Pour tout $\lambda \in \mathbb{R}$, φ_λ étant une fonction affine, on a
$$\varphi_\lambda\left(\frac{1}{b-a}\int_a^b f(x)dx\right) = \frac{1}{b-a}\int_a^b \varphi_\lambda\circ f(x)dx \leq \frac{1}{b-a}\int_a^b g\circ f(x)dx.$$
En passant à la borne supérieure en $\lambda \in \mathbb{R}$ à gauche de l'inégalité, on obtient bien
$$g\left(\frac{1}{b-a}\int_a^b f(x)dx\right) \leq \frac{1}{b-a}\int_a^b g\circ f(x)dx.$$

(cc) BY-NC-SA

□

Remarques.

→ En fait, cette inégalité peut être vue comme la version continue de l'inégalité de la proposition II.3 lorsque les poids (la valeur des λ_i) attribués à chaque terme sont égaux. Lorsqu'on attribue des poids quelconques à chaque élément de la somme (qui est ici une somme continue, c'est-à-dire une intégrale), on obtient une inégalité plus générale : pour toute fonction continue $m : [a,b] \longrightarrow \mathbb{R}^+$ telle que

$$\int_a^b m(x)dx = 1,$$

On peut affirmer que

$$g\left(\int_a^b m(x)f(x)dx\right) \leq \int_a^b m(x)g\circ f(x)dx.$$

Cette inégalité se démontre exactement de la même manière que la précédente, à quelques détails mineurs près. Le lecteur est fortement encouragé à la démontrer lui-même. Ici, pour tout $x \in [a,b]$, on attribue à $f(x)$ le poids $m(x)$, contrairement au cas précédent où on attribuait à $f(x)$ le poids $\dfrac{1}{b-a}$.

→ Lorsque g est strictement convexe, alors il y a égalité si et seulement si f est constante sur $[a,b]$.

→ La réciproque du point (1) de la proposition III.4 est également vraie. En effet, soit $(\varphi_\lambda)_{\lambda\in\Lambda}$ une famille de fonctions affines telles que pour tout $x \in \mathbb{R}$, $g(x) = \sup\limits_{\lambda\in\Lambda} \varphi_\lambda(x)$. On a alors, $\mathrm{Epi}(g) = \bigcap\limits_{\lambda\in\Lambda} \mathrm{Epi}(\varphi_\lambda)$. De plus, les fonctions de la famille $(\varphi_\lambda)_{\lambda\in\Lambda}$ étant toutes convexes, on peut affirmer en utilisant le résultat de l'exercice I.2 que $\mathrm{Epi}(\varphi_\lambda)$ est convexe pour tout $\lambda \in \Lambda$. $\mathrm{Epi}(g)$ est donc l'intersection d'ensembles convexes, et est alors convexe. En utilisant encore une fois le résultat de l'exercice I.2, on peut affirmer que g est convexe. Remarquons qu'on a uniquement utilisé dans cette preuve le fait que les fonctions de la famille $(\varphi_\lambda)_{\lambda\in\Lambda}$ sont convexes. En particulier, cette preuve nous permet de dire que pour toute famille de fonctions convexes $(\varphi_\lambda)_{\lambda\in\Lambda}$, la fonction $x \longmapsto \sup\limits_{\lambda\in\Lambda} \varphi_\lambda(x)$, lorsqu'elle est bien définie en tout point de $\mathbb{R}$ (*i.e.* la famille $(\varphi_\lambda(x))_{\lambda\in\Lambda}$ est majorée pour tout $x \in \mathbb{R}$), est convexe.

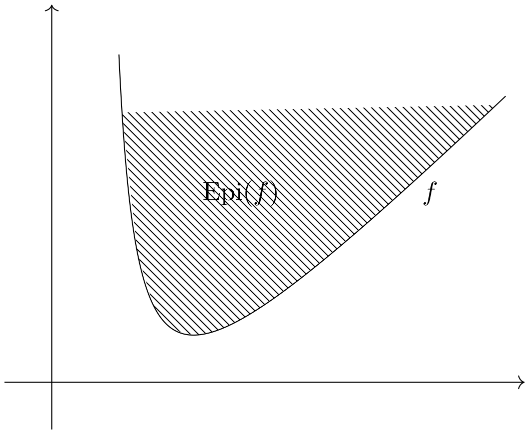
→ Les inégalités de Hölder, Minkowski et Jensen sont également vérifiées pour les variables aléatoires. En particulier, pour tout espace probabilisé Ω, variables aléatoires $\mathrm{X} : \Omega \longrightarrow \mathrm{E}$ et $\mathrm{Y} : \Omega \longrightarrow \mathrm{E}$ où E est un sous-ensemble fini de $\mathbb{R}$, fonction convexe $g : \mathbb{R} \longrightarrow \mathbb{R}$, et $p, q \in \mathbb{R}_+^*$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$, on a les inégalités suivantes.

$$\begin{aligned}\mathbb{E}(|\mathrm{XY}|) &\leq \mathbb{E}\left(|\mathrm{X}|^p\right)^{\frac{1}{p}} \mathbb{E}\left(|\mathrm{Y}|^q\right)^{\frac{1}{q}}\\ \mathbb{E}(|\mathrm{X}+\mathrm{Y}|^p)^{\frac{1}{p}} &\leq \mathbb{E}\left(|\mathrm{X}|^p\right)^{\frac{1}{p}} + \mathbb{E}\left(|\mathrm{Y}|^p\right)^{\frac{1}{p}}\\ g(\mathbb{E}(\mathrm{X})) &\leq \mathbb{E}(g(\mathrm{X})).\end{aligned}$$

(cc) BY-NC-SA

Correction de l'exercice I.2 :

On rappelle qu'un ensemble A est convexe si et seulement si pour tout $a, b \in \mathrm{A}$ et $\lambda \in [0,1]$, $\lambda a+(1-\lambda)b \in \mathrm{A}$. L'énoncé de cet exercice est assez intuitif, et peut se reformuler de la manière suivante : f est une fonction convexe si et seulement si la partie au-dessus de sa courbe est convexe. Pour en avoir une intuition plus forte, observons le dessin suivant.



Montrons à présent ce résultat.

→ ($\Rightarrow$) Supposons que f est convexe. Soit $(x_1, y_1), (x_2, y_2) \in \mathrm{Epi}(f)$. On a alors pour tout $\lambda \in [0,1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda y_1 + (1-\lambda)y_2.$$

On en déduit donc qu'on a bien $\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) \in \mathrm{Epi}(f)$.

→ ($\Leftarrow$) Supposons que $\mathrm{Epi}(f)$ est convexe. Soit $x_1, x_2 \in \mathrm{I}$. On sait que $(x_1, f(x_1)) \in \mathrm{Epi}(f)$ et $(x_2, f(x_2)) \in \mathrm{Epi}(f)$. On a donc pour tout $\lambda \in [0,1]$,

$$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in \mathrm{Epi}(f) \ \textit{i.e.}\ f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2),$$

ce qui est bien le résultat voulu.

Correction de l'exercice II.4 :

Il suffit d'utiiser'inégalité arithmético-géométrique. En effet, a on a pour tout $x, y \in \mathbb{R}_+^*$,

$$ax + by + \frac{c}{xy} \geq 3\sqrt[3]{abc},$$

avec égalité si et seulement si $ax = by = \dfrac{c}{xy}$, c'est à dire que

$$x = \sqrt[3]{\frac{bc}{a^2}} \text{ et } y = \sqrt[3]{\frac{ac}{b^2}}.$$

On en déduit donc que

$$\inf_{x,y \in \mathbb{R}_+^*} \left(ax + by + \frac{c}{xy} \right) = 3\sqrt[3]{abc},$$

et que cette borne inférieure est atteinte lorsque $x = \sqrt[3]{\dfrac{bc}{a^2}}$ et $y = \sqrt[3]{\dfrac{ac}{b^2}}$.

(cc) BY-NC-SA

Correction de l'exercice II.5 :

1. L'implication directe est évidente : il s'agit de la définition de la convexité avec $\lambda = \frac{1}{2}$. Montrons l'implication réciproque. Supposons que

$$\forall x, y \in \mathbb{R},\ g\left(\frac{x+y}{2}\right) \leq \frac{g(x)+g(y)}{2}.$$

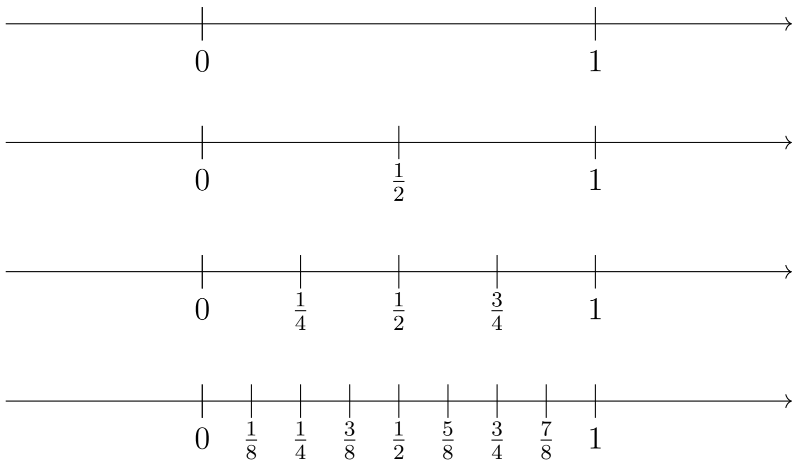
On remarque qu'on peut itérer cette inégalité de la manière suivante : pour tous $x, y \in \mathbb{R}$,

$$g\left(\frac{\frac{x+y}{2}+y}{2}\right) \leq \frac{g\left(\frac{x+y}{2}\right)+g(y)}{2} \leq \frac{g(x)+3g(y)}{4}\ \ i.e.\ g\left(\frac{x+3y}{4}\right) \leq \frac{g(x)+3g(y)}{4}.$$

Notre intuition nous pousse à chercher les $\lambda \in [0,1]$ tels que

$$\forall x, y \in \mathbb{R},\ g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y).$$

On pose Λ l'ensemble des $\lambda \in [0,1]$ qui vérifient l'inégalité ci-dessus. On peut facilement montrer que pour tout $\lambda_1, \lambda_2 \in \Lambda$, $\frac{\lambda_1+\lambda_2}{2} \in \Lambda$. On a de plus, $1, 0, \frac{1}{2} \in \Lambda$. On veut maintenant trouver tous les éléments de Λ. Regardons à quoi rassemble Λ sur un segment. On sait que pour chaque paire de points de Λ sur le segment, le milieu du segment délimité par ces deux points est aussi dans Λ. On peut donc retrouver les éléments de Λ en ajoutant tous milieux de segments qu'on peut trouver à chaque étape, comme sur le dessin ci-dessous.



En regardant le dessin ci-dessus, on peut intuitivement en déduire qu'on peut montrer que pour tout $n \in \mathbb{N}$ et $k \in [\![0; 2^n]\!]$

$$\forall x, y \in \mathbb{R},\ g\left(\frac{k}{2^n}x + \left(1-\frac{k}{2^n}\right)y\right) \leq \frac{k}{2^n}g(x) + \left(1-\frac{k}{2^n}\right)g(y),$$

c'est-à-dire que $\frac{k}{2^n} \in \Lambda$. Pour le montrer, il suffit de faire une récurrence, dont l'idée est exactement décrite pas les dessins ci-dessus. Pour $n = 0$, la propriété est vraie. Soit $n \in \mathbb{N}$ tel que la propriété est vérifiée pour n. Soit $k \in [\![0; 2^{n+1}]\!]$. Si k est pair, alors en posant $k = 2\ell, \ell \in \mathbb{N}$, on a $\frac{k}{2^{n+1}} = \frac{\ell}{2^n} \in \Lambda$. Sinon, on pose $k = 2\ell+1, \ell \in \mathbb{N}$. On a alors (on cherche parmi les milieux des segments obtenus par la subdivisons précédente)

$$\frac{\overbrace{\frac{\ell}{2^n}}^{\in\Lambda} + \overbrace{\frac{\ell+1}{2^n}}^{\in\Lambda}}{2} \in \Lambda \text{ et donc } \frac{k}{2^{n+1}} \in \Lambda,$$

BY-NC-SA

ce qui est bien la propriété désirée. On en déduit donc que Λ contient tous les nombres rationnels entre 0 et 1 dont le dénominateur est une puissance de 2. On va donc approcher tout nombre $\lambda \in [0,1]$ par une suite à valeurs dans Λ. Soit $\lambda \in [0,1]$. On pose pour tout $n \in \mathbb{N}$, $\lambda_n = \dfrac{\lfloor 2^n\lambda \rfloor}{2^n}$. La suite (λ_n) est à valeurs dans Λ, et de plus, on a

$$\lambda_n = \frac{\lfloor 2^n\lambda \rfloor}{2^n} = \lambda - \frac{\overbrace{2^n\lambda - \lfloor 2^n\lambda \rfloor}^{\in[0,1]}}{2^n} \xrightarrow[n\to+\infty]{} \lambda.$$

On a alors pour tous $x, y \in \mathbb{R}$ et $n \in \mathbb{N}$,

$$g\left(\lambda_n x + (1-\lambda_n)y\right) \leq \lambda_n g(x) + (1-\lambda_n)g(y).$$

En passant à la limite à gauche et a droite lorsque n tends vers l'infini, et par continuité de g, on obtient finalement

$$g\left(\lambda x + (1-\lambda)y\right) \leq \lambda g(x) + (1-\lambda)g(y).$$

2. Pour montrer que g est convexe, il suffit de montrer qu'elle est continue et ensuite d'utiliser la question précédente pour obtenir la convexité. Soit $x \in \mathbb{R}$. On veut montrer que g est continue en x. Quitte à remplacer g par $t \longmapsto g(t+x) - g(x)$, on suppose sans perte de généralité que $x = 0$ et $g(x) = 0$ (le lecteur pourra vérifier que cette nouvelle fonction vérifie toujours les propriétés vérifiées par g). On a alors pour tout $y \in \mathbb{R}$

$$g\left(\frac{y+0}{2}\right) \leq \frac{g(y)+g(0)}{2} = \frac{g(y)}{2},$$

et donc par une récurrence immédiate, pour tout $n \in \mathbb{N}$, $g\left(\dfrac{y}{2^n}\right) \leq \dfrac{g(y)}{2^n}$. On a de plus pour tout $y \in \mathbb{R}$,

$$g\left(\frac{y-y}{2}\right) \leq \frac{1}{2}\left(g(y) + g(-y)\right) \ \ i.e.\ g(-y) \geq -g(y).$$

Pour montrer la continuité de g en 0, on va montrer que si $|y|$ est petit, alors $|g(y)|$ est petit. Soit M un majorant de $|g|$ sur $[-1,1]$. Soit $n \in \mathbb{N}^*$ et $y \in \left[-\frac{1}{2^n}, \frac{1}{2^n}\right]$.

→ Si $g(y) \geq 0$, alors on peut affirmer que

$$|g(y)| \leq \frac{|g(2^n y)|}{2^n} \leq \frac{\mathrm{M}}{2^n},$$

→ Sinon, si $g(y) < 0$, alors

$$0 \leq -g(y) \leq g(-y) \leq \frac{g(-2^n y)}{2^n} \text{ et donc } |g(y)| \leq \frac{|g(-2^n y)|}{2^n} \leq \frac{\mathrm{M}}{2^n}.$$

Ceci nous permet alors de dire que

$$\forall n \in \mathbb{N}^*,\ \forall y \in \mathbb{R},\ y \in \left[-\frac{1}{2^n}, \frac{1}{2^n}\right] \Longrightarrow |g(y)| \leq \frac{\mathrm{M}}{2^n},$$

ce qui nous donne directement la continuité, car implique la définition de la continuité de g en 0

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ y \in]-\eta, \eta[\Longrightarrow |g(y)| < \varepsilon.$$

(cc) BY-NC-SA

Chapitre 3

Cardinaux

I Généralités

Définition I.1.

Soient E et F deux ensembles. On dit que E et F ont même cardinal ou sont équivalents lorsqu'il existe une application $f : \mathrm{E} \longrightarrow \mathrm{F}$ qui soit bijective.

Remarquons que la relation introduite ci-dessus définit une relation d'équivalence.

Exercice I.2.

Soit E un ensemble. Montrer que E et $\mathcal{P}(\mathrm{E})$ ne sont pas équivalents.

1. Cardinaux finis et infinis

Définition I.3.

Un ensemble E est dit fini s'il existe $n \in \mathbb{N}$ tel que E est équivalent à $[\![0;n]\!]$. Le cardinal de E est alors $n+1$. Dans ce cas, on note $|\mathrm{E}| = n+1$.

Convention : On convient que l'ensemble vide est fini et que $|\varnothing| = 0$.

Proposition I.4.

Soient E et F deux ensembles finis et f une application de E dans F. Si E et F sont équivalents et $f : \mathrm{E} \longrightarrow \mathrm{F}$ est surjective ou injective, alors elle est bijective.

Définition I.5.

On dit qu'un ensemble est infini lorsqu'il n'est pas fini.

Exercice I.6.

Soit E un ensemble.

1. Montrer que si E est infini, alors il existe une application $f : \mathbb{N} \longrightarrow \mathrm{E}$ qui soit injective.
2. Montrer que E est infini si, et seulement si, il existe une application $f : \mathrm{E} \longrightarrow \mathrm{E}$ qui soit injective, mais non surjective.

(cc) BY-NC-SA

II Ensembles dénombrables

Définition II.1.

Soit E un ensemble. On dit que E est dénombrable lorsqu'il est fini ou s'il existe une bijection f de $\mathbb{N}$ dans E.

Remarque 1. Dans la littérature, il existe deux définitions de la dénombrabilité. L'une d'entre elles est celle ci-dessus, l'autre se réduit uniquement à l'existence d'une bijection avec $\mathbb{N}$.
Remarque 2. Si E et F sont deux ensembles équivalents, alors E est dénombrable si, et seulement si, F l'est.
Exemples.

→ $\mathbb{N}^*$ est dénombrable. En effet, il suffit de considérer la bijection $f : x \longmapsto x+1$ de $\mathbb{N}$ dans $\mathbb{N}^*$.

→ $\mathbb{Z}$ est dénombrable. En effet, il suffit de considérer la bijection $f : x \longmapsto \begin{cases} 2\,|x|-1 & \text{si } x<0 \\ 2x & \text{sinon} \end{cases}$ de $\mathbb{Z}$ dans $\mathbb{N}$.

Remarque. Plutôt que de retenir cette bijection à l'aide de sa formule explicite, il est loisible de retenir comment elle est construite. Trouver une bijection entre un ensemble et $\mathbb{N}$, c'est trouver comment énumérer ses éléments. Pour énumérer les éléments de $\mathbb{Z}$, il suffit de poser que 0 est le 0-ième élément, 1 est le premier élément, -1 est le deuxième élément, 2 est le troisième élément et ainsi de suite : il s'agit d'énumérer tour à tour les entiers positifs et les entiers négatifs.

Proposition II.2.

Soit E et F deux ensembles.

1. Toute partie de $\mathbb{N}$ est dénombrable.
2. S'il existe une application $f : \mathrm{E} \longrightarrow \mathrm{F}$ qui soit injective et si F est dénombrable, alors E est dénombrable.
3. S'il existe une application $f : \mathrm{F} \longrightarrow \mathrm{E}$ qui soit surjective et si F est dénombrable, alors E est dénombrable.
4. $\mathbb{N} \times \mathbb{N}$ est dénombrable.
5. Le produit cartésien de deux ensembles dénombrables est dénombrable.
6. $\mathbb{Q}$ est dénombrable.
7. Soit I un ensemble dénombrable. Soit $(\mathrm{E}_i)_{i\in\mathrm{I}}$ une famille d'ensembles dénombrables. Alors $\bigcup_{i\in\mathrm{I}} \mathrm{E}_i$ est dénombrable.

Démonstrations.

1. Soit A une partie de $\mathbb{N}$. Si A est fini, il n'y a rien à prouver. On suppose donc que A est infini. Nous allons définir récursivement une bijection de $\mathbb{N}$ dans A.

   → **Initialisation :** on pose $f(0) = \min \mathrm{A}$.

   → **Hérédité :** Soit $n \in \mathbb{N}$, supposons que $\mathrm{F}(0), f(1), \dots, f(n)$ soient correctement définis. On pose alors $f(n+1) = \min \mathrm{A} \setminus \{f(0), \dots, f(n)\}$.

   Montrons que f ainsi définie est bien une bijection de $\mathbb{N}$ dans A.

   → Injectivité
   si $m > n$, alors m est le minimum d'un ensemble qui ne contient pas $f(n)$, donc $f(n) \neq f(m)$.

(cc) BY-NC-SA

→ Surjectivité
Pour montrer la surjectivité de f, on suppose par l'absurde qu'il existe un élément $a \in \mathrm{A}$ qui n'est pas atteint par f, *i.e.* $a \notin f(\mathbb{N})$. On a alors pour tout $n \in \mathbb{N}$, $a \in \mathrm{A} \setminus \{f(0), ..., f(n)\}$, donc

$$a > \min \mathrm{A} \setminus \{f(0), ..., f(n)\} = f(n+1),$$

ce qui est absurde, car la quantité $\min \mathrm{A} \setminus \{f(0), ..., f(n)\}$ peut être arbitrairement grande pour n assez grand.

Le lecteur pourra vérifier que f est croissante pour s'exercer.

2. Si F est fini, alors E aussi. Si F est infini, alors il est en bijection avec $\mathbb{N}$. Soit ϕ une bijection de F dans $\mathbb{N}$. $\phi \circ f : \mathrm{E} \to \mathbb{N}$ est injective, donc en posant $\mathrm{A} = \phi \circ f(\mathrm{E})$, l'application $\phi \circ f : \mathrm{E} \to \mathrm{A}$ est bijective. Or A est inclus dans F et F est dénombrable, donc A l'est et finalement E aussi puisqu'ils sont en bijection.

3. f étant surjective, on peut affirmer que pour tout $x \in \mathrm{E}$, $f^{-1}(\{x\})$ est non vide et donc écrire F sous forme de l'union d'ensembles disjoints suivante $\mathrm{F} = \bigcup\limits_{x \in \mathrm{E}} f^{-1}(\{x\})$. On définit une application $\mathrm{S} : \mathrm{E} \longrightarrow \mathrm{F}$ de la manière suivante : pour tout $x \in \mathrm{E}$, $\mathrm{S}(x) \in f^{-1}(\{x\})$ (on choisit un élément quelconque de $f^{-1}(\{x\})$, ce qui est possible, car cet ensemble est non vide). Les ensembles $(f^{-1}(\{x\}))_{x \in \mathrm{E}}$ étant disjoints, on a pour tous $x, y \in \mathrm{E}$, si $x \neq y$ alors $\mathrm{S}(x) \neq \mathrm{S}(y)$. Donc $\mathrm{S} : \mathrm{E} \to \mathrm{F}$ est injective. Il suffit alors d'appliquer le point (2).

4. On considère l'application de $\mathbb{N} \times \mathbb{N}$ dans $\mathbb{N}$ $f : (m, n) \longmapsto 2^m(2n+1)$. Montrons qu'elle est injective. Soit $m, n, m', n' \in \mathbb{N}$. On suppose sans perte de généralité que $m \geq m'$. Si $f(m, n) = f(m', n')$, alors on a $2^{m-m'}(2n + 1) = (2n' + 1)$. $2^{m-m'}$ divise un nombre impair, donc nécessairement $m = m'$ et par conséquent $n = n'$. Donc f est injective. On peut alors appliquer le point (2).

5. **Lemme.** E est dénombrable si et seulement si existe une surjection de $\mathbb{N}$ dans E.
Appliquons ce lemme. Soit E et F dénombrables. Il existe deux applications $s_{\mathrm{E}} : \mathbb{N} \to \mathrm{E}$ et $s_{\mathrm{F}} : \mathbb{N} \to \mathrm{F}$ surjectives. On considère alors l'application

$$s : \begin{cases} \mathbb{N} \times \mathbb{N} & \longrightarrow \mathrm{E} \times \mathrm{F} \\ (m, n) & \longmapsto (s_{\mathrm{E}}(m), s_{\mathrm{E}}(n)). \end{cases}$$

Cette application est surjective et $\mathbb{N} \times \mathbb{N}$ est dénombrable, on peut donc appliquer le point (3).

6. Il suffit de considérer l'application

$$f : \begin{cases} \mathbb{Z} \times \mathbb{N}^* & \longrightarrow \mathbb{Q} \\ (p, q) & \longmapsto \dfrac{p}{q}. \end{cases}$$

Cette application est surjective et $\mathbb{Z} \times \mathbb{N}^*$ est dénombrable d'après le point (5), donc $\mathbb{Q}$ est dénombrable d'après le point (3).

7. On pose $\mathrm{A} = \bigcup\limits_{i \in \mathrm{I}} \mathrm{E}_i$. Pour tout $i \in \mathrm{I}$ E_i est dénombrable, donc on peut considérer une surjection $s_i : \mathbb{N} \to \mathrm{E}_i$. On considère donc l'application

$$s : \begin{cases} \mathrm{I} \times \mathbb{N} & \longrightarrow \mathrm{A} \\ (i, n) & \longmapsto s_i(n). \end{cases}$$

Cette application est une surjection et $\mathrm{I} \times \mathbb{N}$ est dénombrable, on peut alors appliquer le point (3).

(cc) BY-NC-SA

III Exercices

Exercice III.1.

Soit $n \in \mathbb{N}^*$ et $X \subset [\![1; 2n]\!]$. On suppose que $|X| \geq n+1$. Montrer qu'il existe $a, b \in X^2$ tel que $a \neq b$ et $a|b$.

Exercice III.2.

Montrer que $\mathbb{R}$ n'est pas dénombrable.

Exercice III.3.

Soit Λ un ensemble. Soit $(I_\lambda)_{\lambda\in\Lambda}$ une famille d'intervalles ouverts non vides de $\mathbb{R}$. Montrer que si les intervalles $(I_\lambda)_{\lambda\in\Lambda}$ sont deux à deux disjoints, alors Λ est dénombrable.

Exercice III.4.

On dit qu'une partie Ω de $\mathbb{R}$ est ouverte au sens topologique lorsque pour tout $x \in \Omega$, il existe $\varepsilon > 0$ tel que $]x-\varepsilon, x+\varepsilon[\subset \Omega$. Montrer qu'un intervalle est ouvert au sens topologique si et seulement s'il est ouvert au sens de l'ordre.

Exercice III.5.

Soit $f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$. Montrer que $A = \{x \in \mathbb{R},\ f(x) \neq 0\}$ est ouvert.

Exercice III.6.

Soit Ω un ouvert non vide de $\mathbb{R}$. Montrer que Ω est réunion dénombrable d'intervalles ouverts deux à deux disjoints.

Exercice III.7.

Soit $f : \mathbb{R} \to \mathbb{R}$ monotone et Δ_f l'ensemble des points de discontinuité de f. Montrer que Δ_f est dénombrable.

Exercice III.8.

Soit $f : \mathbb{R} \to \mathbb{R}$ convexe. Montrer que $A = \{a \in \mathbb{R},\ \text{f n'est pas dérivable en a}\}$ est dénombrable.

Exercice III.9.

On dit que $\alpha \in \mathbb{C}$ est algébrique s'il existe $P \in \mathbb{Q}[X] \setminus \{0\}$ tel que $P(\alpha) = 0$. Montrer que l'ensemble des nombres algébriques est dénombrable.
Remarque. la définition ci dessus du fait que α soit algébrique est équivalente à $\exists P \in \mathbb{Z}[X]\setminus\{0\}$ $P(\alpha) = 0$.

(cc) BY-NC-SA

Correction de l'exercice I.2. :

Supposons qu'il exsite une application surjective $S : E \longrightarrow \mathcal{P}(E)$. On pose $A = \{x \in E,\ x \notin S(x)\}$.
Si A admet un antécédent a par S, alors si $a \in S(a)$, $a \notin A = S(a)$ ce qui est absurde. De même, si $a \notin S(a)$, alors $a \in A = S(a)$ ce qui est encore absurde. Il n'exsite pas de surjection de E dans $\mathcal{P}(E)$, ils ne sont donc pas équivalents.

Correction de l'exercice I.6. :

1. Nous allons construire $f : \mathbb{N} \to E$ récursivement.
   - → **Initialisation :** E est infini donc non vide. On peut donc choisir $a_0 \in E$ et on pose $f(0) = a_0$.
   - → **Hérédité :** Soit $n \in \mathbb{N}$. Supposons que $f(0), \ldots, f(n)$ sont correctement définis. $A_n = E \setminus \{f(0), \ldots, f(n)\}$ est non vide car E est infini, on peut donc choisir $a_{n+1} \in A_n$ et on pose $f(n+1) = a_{n+1}$.

   L'application f définie ci dessus est injective. En effet, si $m > n$, alors $f(m) \in E \setminus \{f(0), \ldots, f(m-1)\}$ ce qui donne que $f(m) \neq f(n)$.
2.
   - → (⇐) Soit $f : E \to E$ injective, mais non surjective. E est fini équivalent à lui même, donc la proposition I.4 nous permet d'affirmer que f est bijective donc surjective, ce qui est absurde.
   - → (⇒) Supposons que E est infini. D'après la question précédente, on dispose d'une injection $i : \mathbb{N} \to E$. On définit

     $$f : \begin{cases} E & \longrightarrow E \\ x & \longmapsto \begin{cases} x & \text{si } x \in E \setminus i(\mathbb{N}) \\ i(n+1) & \text{si } x = i(n) \text{ avec } n \in \mathbb{N}. \end{cases} \end{cases}$$

     Montrons que f est injective et non surjective.
     Soit $(x, y) \in E^2$ tel que $f(x) = f(y)$.
     - Si $x \in i(\mathbb{N})$, alors y aussi, car sinon on disposerait d'un entier $n \in \mathbb{N}$ tel que $f(x) = i(n)$, et donc $f(y) = y = f(x) = i(n) \in i(\mathbb{N})$ ce qui est absurde. On dispose alors de $a, b \in \mathbb{N}$ tels que $f(x) = i(a+1)$, $f(y) = i(b+1)$, $x = i(a)$ et $x = i(b)$. On a alors $i(a+1) = i(b+1)$. Par injectivité de i, on a $a = b$ et donc $i(a) = i(b)$ et finalement $x = y$.
     - Si $x \notin i(\mathbb{N})$, alors y aussi, sinon on disposerait d'un $n \in \mathbb{N}$ tel que $f(y) = i(n+1)$ et donc $f(x) = x = f(y) = i(n+1) \in i(\mathbb{N})$ ce qui est absurde. L'égalité $f(x) = f(y)$ devient donc $x = y$. D'où l'injectivité de f.

     En ce qui concerne la non-surjectivité, il suffit de remarquer que $i(0)$ n'admet pas d'antécédent. En effet, s'il existe $x \in E$ tel que $f(x) = i(0)$, alors nécessairement, on dispose d'un entier $n \in \mathbb{N}$ tel que $x = i(n)$, et donc on a $i(n+1) = i(0)$. Ceci nous donne que $n = -1$ ce qui est absurde.

Correction de l'exercice III.1. :

Pour tout $x \in X$, on écrit $x = 2^{k(x)}(2\ell(x) + 1)$. Les nombres de la forme $2\ell(x) + 1$ sont tous dans $[\![1; 2n]\!]$ et il y en a au plus n. Par le lemme des tiroirs (car $|X| \geq n + 1 > n$), il existe donc $a \neq b$ éléments de X tels que $2\ell(a) + 1 = 2\ell(b) + 1$. En supposant sans perte de généralité que $k(a) \geq k(b)$, on obtient que $a = 2^{k(a)-k(b)}b$, donc $b|a$.

Correction de l'exercice III.2. :

On va montrer que $[0, 1[$ n'est pas dénombrable. On suppose par l'absurde qu'on dispose d'une énumération dénombrable de $[0, 1[$, *i.e.* il existe une suite injective $(x_n)_{n\in\mathbb{N}}$ telle que $[0, 1[= \{x_n\}_{n\in\mathbb{N}}$.
On pose pour tout $n \in \mathbb{N}$ l'écriture décimale propre de x_n comme $x_n = 0, x_{1,n}x_{2,n}x_{3,n}\ldots$. On construit alors une suite $(y_q)_{q\in\mathbb{N}^*} \in \{1, \ldots, 8\}^{\mathbb{N}^*}$ telle que pour tout $q \in \mathbb{N}^*$, $y_q \neq x_{q,q}$. On pose $y = 0, y_1y_2y_3\ldots$

(cc) BY-NC-SA

Par hypothèse, il existe $n_0 \in \mathbb{N}^*$ tel que $y = x_{n_0}$. L'écriture de y_{n_0} étant également propre, on obtient que $y_{n_0} = x_{n_0,n_0}$ ce qui est absurde.

Correction de l'exercice III.3. :

Pour tout $\lambda \in \Lambda$, il existe $r_\lambda \in \mathrm{I}_\lambda \cap \mathbb{Q}$. Pour tous $\lambda, \lambda' \in \Lambda$, si $\lambda \neq \lambda'$, par hypothèse $\mathrm{I}_\lambda \cap \mathrm{I}_{\lambda'} = \emptyset$ donc $r_\lambda \neq r_{\lambda'}$. L'application $\lambda \longmapsto r_\lambda$ est une injection de Λ dans $\mathbb{Q}$, donc d'après la propriété 2 de la proposition II.2, Λ est dénombrable.

Correction de l'exercice III.4. :

→ ($\Leftarrow$) On pose $\Omega =]a, b[$ avec $a < b$. Soit $x \in \Omega$. On pose $\varepsilon = \frac{1}{2}\min(x-a, b-x)$. On a alors $]x-\varepsilon, x+\varepsilon[\subset \Omega$. Ω est donc ouvert au sens topologique.

→ ($\Rightarrow$) Soit I un intervalle non vide dont les bornes sont $a, b \in \mathbb{R}$ tel que $a < b$. Supposons que I est ouvert au sens topologique. Pour tout $\varepsilon > 0$ $]a-\varepsilon, a+\varepsilon[\not\subset \mathrm{I}$. De même, $]b-\varepsilon, b+\varepsilon[\not\subset \mathrm{I}$. On a alors par hypothèse $a, b \notin \mathrm{I}$. I est donc ouvert au sens de l'ordre. Le cas des bornes infinies est laissé comme exercice au lecteur.

Correction de l'exercice III.5. :

Soit $x_0 \in \mathbb{R}$ tel que $f(x_0) \neq 0$. On suppose sans perte de généralité que $f(x_0) > 0$. Par continuité de f, il existe $\eta > 0$ tel que pour tout $x \in]x_0-\eta, x_0+\eta[$

$$|f(x) - f(x_0)| < \frac{1}{2} f(x_0).$$

Ceci implique que pour tout $x \in]x_0-\eta, x_0+\eta[$,

$$f(x) > \frac{f(x_0)}{2} > 0.$$

On en déduit que pour tout $x_0 \in \mathrm{A}$, il existe $\eta > 0$ tel que $]x_0-\eta, x_0+\eta[\subset \mathrm{A}$. A est alors ouvert.

Correction de l'exercice III.6. :

Uniquement pour cet exercice, pour tout $x, y \in \Omega$, on utilise l'abus de notation suivant

$$[x, y] = [y, x] = [\min(x, y), \max(y, x)].$$

On introduit la relation d'équivalence $\sim$ sur Ω définie par

$$x \sim y \Longleftrightarrow [x, y] \subset \Omega.$$

Le lecteur pourra vérifier par lui-même qu'il s'agit bien d'une relation d'équivalence.
Soit I une classe d'équivalence selon $\sim$. Si $x, y \in \mathrm{I}$, alors pour tout $z \in [x, y]$, $[x, z] \subset [x, y] \subset \Omega$, donc alors $[x, y] \subset \mathrm{I}$. I est donc convexe, et par conséquent, c'est un intervalle.
Montrons que I est ouvert. Soit $x \in \mathrm{I}$. Ω étant ouvert, il existe $\varepsilon > 0$ tel que $]x-\varepsilon, x+\varepsilon[\subset \Omega$. On peut alors affirmer que pour tout $y \in]x-\varepsilon, x+\varepsilon[$, $[x, y] \subset \Omega$ et par conséquent $y \in \mathrm{I}$. I est alors ouvert.
Conclusion : Ω est réunion disjointe de classes d'équivalences selon $\sim$ qui sont toutes des intervalles ouverts. D'après l'exercice III.3, cette réunion est dénombrable.

Correction de l'exercice III.7. :

Pour tout $x \in \mathbb{R}$, on note $f(x^+)$ et $f(x^-)$ respectivement la limite à droite et à gauche de f en x.
Supposons sans perte de généralité que f est croissante. Soit $a \in \mathbb{R}$. On sait que $f(a^-) \leq f(a) \leq f(a^+)$.

(cc) BY-NC-SA

Ainsi, pour tout $a \in \Delta_f$, $f(a^-) < f(a^+)$.
Posons pour tout $a \in \Delta_f$, $\mathrm{I}_a =]f(a^-), f(a^+)[$. Soit $b \in \Delta_f$ différent de a. Si $a < b$, alors $f(a^+) \leq f(b^-)$. De même, si $a > b$, alors $f(a^-) \geq f(b^+)$. Les intervalles $(\mathrm{I}_a)_{a \in \Delta_f}$ sont ouverts, non vides et disjoints. D'après l'exercice III.3, on déduit que Δ_f est dénombrable.

Correction de l'exercice III.8. :

Soit $a \in \mathrm{A}$. On pose $\mathrm{I}_a =]f'_g(a), f'_d(a)[$. On pose pour tous $x \neq y$, $p_x(y) = \frac{f(x)-f(y)}{x-y}$. Si $a, b \in \mathrm{A}$ avec $a < b$, alors pour tout $x \in]a, b[$ (voir le chapitre sur les fonctions convexes)

$$f'_d(a) \leq p_a(x) = p_x(a) \leq p_x(b) = p_b(x) \leq f'_g(b),$$

et alors $\mathrm{I}_a \cap \mathrm{I}_b = \emptyset$ et pour tout $x \in \mathrm{A}$, I_a est non vide. Il suffit donc d'appliquer l'exercice III.3.

Correction de l'exercice III.9. :

On note $\overline{\mathbb{Q}}$ l'ensemble des nombres algébriques. On pose pour tout $n \in \mathrm{N}$,

$$\mathrm{H}_n = \left\{ \mathrm{P}(\mathrm{X}) = \sum_{k=0}^{d} a_k \mathrm{X}^k \in \mathbb{Z}[\mathrm{X}],\ d \leq n \text{ et } \forall k \in [\![1; d]\!],\ |a_k| \leq n \right\}.$$

Un simple raisonnement combinatoire permet d'affirmer que pour tout $n \in \mathbb{N}$, $|\mathrm{H}_n| \leq (2n+1)^{n+1}$. H_n est donc fini pour tout n. Or pour tout $\mathrm{P} \in \mathbb{Z}[\mathrm{X}]$, pour n assez grand $\mathrm{P} \in \mathrm{H}_n$. On peut donc affirmer que $\mathbb{Z}[\mathrm{X}] = \bigcup_{n \in \mathbb{N}} \mathrm{H}_n$. $\mathbb{Z}[\mathrm{X}]$ est union d'ensembles finis, donc d'après la propriété 7 de la proposition II.2, $\mathbb{Z}[\mathrm{X}]$ est dénombrable.
Enfin, on écrit $\overline{\mathbb{Q}} = \bigcup_{\mathrm{P} \in \mathbb{Z}[\mathrm{X}] \setminus \{0\}} \mathrm{Z}(\mathrm{P})$ (on note pour tout P, Z(P) l'ensemble des zéros de P). $\overline{\mathbb{Q}}$ est donc réunion dénombrable d'ensembles finis, donc d'après la proposition 7 de la propriété II.2, $\overline{\mathbb{Q}}$ est dénombrable.

(cc) BY-NC-SA

Chapitre 4

Valeurs d'adhérence

I Extractions

Définition I.1.

On appelle extraction ou extractrice toute application (injective) strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$.

Exemple. Les fonctions suivantes sont des extractrices.

→ $n \longmapsto 2n$

→ $n \longmapsto 2n+1$

→ $n \longmapsto 2^n$

→ $n \longmapsto 4^n$

→ $n \longmapsto (n+1)!$ (attention, $n \longmapsto n!$ n'est pas une extractrice car $0! = 1!$)

Proposition I.2.

Soit φ, ψ des extractions. Les propositions suivantes sont vraies.

1. Pour tout $n \in \mathbb{N}$, $\varphi(n) \geq n$.
2. Soit $\mathrm{A} \subset \mathbb{N}$ infini. Il existe une (unique) extraction γ telle que $\gamma(\mathbb{N}) = \mathrm{A}$.
3. $\psi \circ \varphi$ est une extractrice. Plus généralement, si $(\varphi_i)_{i \in [\![1;k]\!]}$ est une famille d'extractrices, alors $\varphi_1 \circ \cdots \circ \varphi_k$ l'est aussi.
4. Considérons une famille d'extractrices $(\varphi_i)_{i \in \mathbb{N}^*}$. $\gamma(n) = (\varphi_1 \circ \cdots \circ \varphi_n)(n)$ (avec $\gamma(0) = 0$) est une extraction, appelée extraction diagonale.

1. Il suffit de le faire par récurrence.

   → **Initialisation :** Il est clair que $\varphi(0) \geq 0$.

   → **Hérédité :** Soit $n \in \mathbb{N}$ et supposons que $\varphi(n) \geq n$. On a alors

   $$\varphi(n+1) > \varphi(n) \geq n \text{ et donc } \varphi(n+1) \geq n+1,$$

   d'où le résultat voulu.

2. **Existence :** A est une partie infinie de $\mathbb{N}$, on peut donc considérer $(a_i)_{i \in \mathbb{N}}$ une énumération croissante de A. Il est alors facile de montrer que $\gamma : n \longmapsto a_n$ est une extractrice.

   Unicité : Si γ et φ sont deux extractrices vérifiant $\gamma(\mathbb{N}) = \varphi(\mathbb{N}) = \mathrm{A}$, alors en supposant par l'absurde que $\gamma \neq \varphi$, on peut considérer le plus petit entier k tel que $\varphi(k) \neq \gamma(k)$. Sans perte de généralité, on peut supposer que $\varphi(k) > \gamma(k)$. On en déduit que pour tout $\ell < k$, $\varphi(\ell) = \gamma(\ell) < \gamma(k)$ et pour tout $\ell \geq k$, $\varphi(\ell) \geq \varphi(k) > \gamma(k)$. On en déduit alors que $\gamma(k) \notin \varphi(\mathbb{N})$ et alors que $\varphi(\mathbb{N}) \neq \mathrm{A}$ ce qui est absurde.

3. Il suffit d'utiliser le fait que la composition d'applications strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ est strictement croissante.

(cc) BY-NC-SA

4. En effet, γ est à valeurs dans $\mathbb{N}$ et on a pour tout $n \in \mathbb{N}$,

$$\begin{aligned}\gamma(n+1) &= \varphi_1 \circ \cdots \circ \varphi_n(\varphi_{n+1}(n+1)) \\ &\geq \varphi_1 \circ \cdots \circ \varphi_n(n+1) \\ &\geq \varphi_1 \circ \cdots \circ \varphi_n(n) + 1 \geq \gamma(n) + 1 > \gamma(n).\end{aligned}$$

Définition I.3.

Soit $(u_n) \in \mathrm{E}^{\mathbb{N}}$ où E est un ensemble non vide. On dit que $(v_n) \in \mathrm{E}^{\mathbb{N}}$ est une suite extraite de (u_n) s'il existe une extractrice φ telle que pour tout $n \in \mathbb{N}$, $v_n = u_{\varphi(n)}$. On notera $\mathrm{Ext}(u_n)$ l'ensemble des suites extraites de (u_n).

Exemple. Pour toute suite (u_n), les suites (u_{2n}) et (u_{2n+1}) sont des suites extraites de (u_n).

Proposition I.4.

Si (w_n) est une suite extraite de (v_n) et (v_n) une suite extraite de (u_n) alors (w_n) est extraite de (u_n).

Remarque. Cette propriété est généralisable à plusieurs extractions par récurrence.

Proposition I.5.

Soit (u_n) une suite réelle. Les positions suivantes sont vraies.

1. (u_n) est non majorée si et seulement s'il existe une extractrice φ telle que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} +\infty$
2. (u_n) ne tend pas vers 0 si et seulement s'il existe une extractrice φ et $\varepsilon > 0$ tels que pour tout $n \in \mathbb{N}$, $\left|u_{\varphi(n)}\right| > \varepsilon$.
3. Si (u_n) est à valeurs dans $\mathbb{N}$, alors (u_n) ne tend pas vers l'infini si et seulement s'il existe une extractrice φ telle que $(u_{\varphi(n)})$ est constante.

Démonstration.

1. L'implication de droite à gauche est facile. Montrons l'implication de gauche à droite. Supposons que u_n est non majorée et construisons une extractrice φ vérifiant pour tout $n \in \mathbb{N}$, $u_{\varphi(n)} \geq n$.
   - → (u_n) est non bornée, il existe donc k tel que $u_k \geq 0$. On pose alors $\varphi(0) = k$.
   - → Supposons que pour tout $\ell \in [\![0; n]\!]$, φ est définie de manière à ce qu'on ait $u_{\varphi(\ell)} \geq \ell$. Encore une fois, (u_n) est non bornée donc il existe $k \geq \varphi(n)$ tel que $u_k \geq n+1$, on pose donc $\varphi(n+1) = k$.

   On a donc bien pour tout $n \in \mathbb{N}$, $u_{\varphi(n)} \geq n$ et donc $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} +\infty$.
2. Encore une fois, l'implication de droite à gauche est facile. Montrons donc l'implication de gauche à droite. Le fait que (u_n) tend vers 0 est équivalent à

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ |u_n| \leq \varepsilon.$$

   Écrivons donc la négation de cette proposition

$$\exists \varepsilon > 0,\ \forall \mathrm{N} \in \mathbb{N},\ \exists n \geq \mathrm{N},\ |u_n| > \varepsilon.$$

   Considérons donc $\varepsilon > 0$ vérifiant la proposition ci-dessus. Il est alors clair que l'ensemble

$$\mathrm{A} = \left\{n \in \mathbb{N},\ \left|u_{\varphi(n)}\right| > \varepsilon\right\}$$

CC BY-NC-SA

est infini. On sait donc d'après le point 2 de la proposition I.2 qu'il existe une extractrice φ telle que $\varphi(\mathbb{N}) = \mathrm{A}$, et alors pour tout $n \in \mathbb{N}$, $\left|u_{\varphi(n)}\right| \geq n$.

3. Encore une fois, l'implication de droite à gauche est facile. Montrons donc l'implication de gauche à droite. Supposons que (u_n) ne tend pas vers $+\infty$. Le fait que u_n tend vers l'infini est équivalent à la proposition
$$\forall \mathrm{M} > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ u_n \geq \mathrm{M}.$$
La négation de cette proposition est donc
$$\exists \mathrm{M} > 0,\ \forall \mathrm{N} \in \mathbb{N},\ \exists n \geq \mathrm{N},\ u_n < \mathrm{M}.$$
Considérons alors M vérifiant la propriété. La proposition ci-dessus nous permet d'affirmer que l'ensemble
$$\mathrm{A} = \{n \in \mathbb{N},\ u_n \in [\![0;\mathrm{M}]\!]\}$$
est infini. En posant pour tout $i \in [\![0;\mathrm{M}]\!]$, $\mathrm{A}_i = \{n \in \mathbb{N},\ u_n = i\}$, on peut écrire $\mathrm{A} = \bigcup_{i=0}^{\mathrm{M}} \mathrm{A}_i$. A est infini donc il existe $i \in [\![0;\mathrm{M}]\!]$ tel que A_i est infini. En utilisant le point 2 de la propriété I.2, on peut affirmer l'existence d'une extractrice φ telle que $\varphi(\mathbb{N}) = \mathrm{A}_i$ et alors $\forall n \in \mathbb{N},\ u_{\varphi(n)} = i$ ce qui est bien le résultat voulu.

II Valeurs d'adhérence

Définition et Proposition II.1.

Soit $u = (u_n) \in \mathbb{C}^{\mathbb{N}}$ et $a \in \mathbb{C}$. Les propositions suivantes sont équivalentes.

1. $\forall \varepsilon > 0,\ \forall \mathrm{N} \in \mathbb{N},\ \exists n \geq \mathrm{N},\ |u_n - a| < \varepsilon$.
2. Pour tout $\varepsilon > 0$, l'ensemble $\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |u_n - a| < \varepsilon\}$ est infini.
3. Il existe une extractrice φ telle que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$.

Si ces propositions sont vérifiées, on dit que a est une valeur d'adhérence de (u_n) et on note $\mathrm{Adh}(u_n)$ l'ensemble des valeurs d'adhérence de (u_n).

Démonstration.

→ (1) ⇒ (2) Supposons qu'il existe $\varepsilon > 0$ tel que $\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |u_n - a| < \varepsilon\}$ est fini. En posant $\mathrm{N} = \max \mathrm{A}$, on voit que pour tout $n > \mathrm{N}$, $n \notin \mathrm{A}$. On en déduit alors que
$$\exists \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ |u_n - a| \geq \varepsilon$$
ce qui est absurde.

→ (2) ⇒ (3) Construisons par récurrence une extractrice φ telle que $\forall n \in \mathbb{N},\ \left|u_{\varphi(n)} - a\right| \leq \dfrac{1}{n+1}$.

- L'ensemble $\mathrm{A}(1)$ est non vide, il existe donc $k \in \mathbb{N}$ tel que $|u_k - a| \leq 1$. On pose alors $\varphi(0) = k$.
- Soit $n \in \mathbb{N}$. Supposons que pour tout $\ell \in [\![0;n]\!]$, $\varphi(\ell)$ est défini tel que $\left|u_{\varphi(\ell)} - a\right| \leq \dfrac{1}{\ell+1}$. L'ensemble $\mathrm{A}\left(\dfrac{1}{n+1}\right)$ est une partie de $\mathbb{N}$ infinie donc non bornée, il existe donc $k \in \mathrm{A}\left(\dfrac{1}{n+1}\right)$ tel que $k > \varphi(n)$. On pose alors $\varphi(n+1) = k$.

On a donc bien que pour tout $n \in \mathbb{N}$, $\left|u_{\varphi(n)} - a\right| \leq \dfrac{1}{n+1}$ et donc $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$.

(cc) BY-NC-SA

→ (3) ⇒ (1) Le point (3) est équivalent à

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \left|u_{\varphi(n)} - a\right| < \varepsilon.$$

Soit $\varepsilon > 0$ et $\mathrm{N} \in \mathbb{N}$. D'après ce qui précéde, il existe $\mathrm{M} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{M}$, $\left|u_{\varphi(n)} - a\right| < \varepsilon$. En posant $n = \varphi(\max(\mathrm{M}, \mathrm{N}))$, on voit que $n \geq \varphi(\mathrm{N}) \geq \mathrm{N}$ et $|u_n - a| < \varepsilon$, ce qui est bien le résultat voulu.

Exemples.

1. Les valeurs d'adhérence de la suite $(u_n) = ((-1)^n)$ sont 1 et -1. En effet, $1 \in \mathrm{Adh}(u_n)$ et $-1 \in \mathrm{Adh}(u_n)$ car $u_{2n} \xrightarrow[n\to+\infty]{} 1$ et $u_{2n+1} \xrightarrow[n\to+\infty]{} -1$. De plus, si $a \notin \{-1, 1\}$, alors l'ensemble $\{n \in \mathbb{N},\ |u_n - a| < \varepsilon\}$ avec $\varepsilon = \dfrac{1}{2}\min(|a-1|, |a+1|)$ est fini (vide), donc a ne peut pas être une valeur d'adhérence de (u_n).
2. Plus généralement, si (u_n) est T-périodique avec $\mathrm{T} \in \mathbb{N}^*$ alors $\mathrm{Adh}(u_n) = \underbrace{\{u_0, \ldots, u_{\mathrm{T}-1}\}}_{\mathrm{B}}$. En effet, pour tout $k \in [\![0; \mathrm{T}-1]\!]$, $u_{n\mathrm{T}+k} \xrightarrow[n\to+\infty]{} u_k$ et donc $u_k \in \mathrm{Adh}(u_n)$. On en déduit donc que $\mathrm{B} \subset \mathrm{Adh}(u_n)$. Réciproquement, si $a \notin \mathrm{B}$, alors pour tout $b \in \mathrm{B}$, $|a - b| \geq \min\limits_{x\in\mathrm{B}} |a - x| > 0$ et donc l'ensemble $\{n \in \mathbb{N},\ |u_n - a| < \varepsilon\}$ avec $\varepsilon = \min\limits_{x\in\mathrm{B}} |a - x|$ est fini (vide) donc a ne peut pas être une valeur d'adhérence de (u_n). On en déduit alors qu'on a bien $\mathrm{Adh}(u_n) = \mathrm{B}$.
3. Si (u_n) est à valeurs strictement positives, alors $0 \notin \mathrm{Adh}(u)$ si et seulement si $\exists \varepsilon > 0,\ \forall n \in \mathbb{N},\ u_n \geq \varepsilon$. En effet, si $0 \notin \mathrm{Adh}(u_n)$, il existe ε tel que $\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ u_n < \varepsilon\}$ est fini. On en déduit alors que $\mathrm{A}(\varepsilon)$ est majoré et donc en posant $\mathrm{N} = \max \mathrm{A}(\varepsilon)$, on peut affirmer que pour tout $n \in \mathbb{N}$, $u_n \geq \min(\varepsilon, |u_0|, \ldots, |u_\mathrm{N}|) > 0$. La réciproque est vraie par définition.

Exercice II.2.

Trouver une suite réelle (u_n) telle que $\mathrm{Adh}(u_n) = \mathbb{R}$.

Proposition II.3.

Soit $(u_n) \in \mathbb{C}^\mathbb{N}$ et $\ell \in \mathbb{C}$. Si (u_n) converge vers ℓ, alors $\mathrm{Adh}(u_n) = \{\ell\}$.

Démonstration. Soit φ une extractrice et $\varepsilon > 0$. On dispose de $\mathrm{N} \geq 0$ tel que pour tout $n \geq \mathrm{N}$ $|u_n - \ell| < \varepsilon$ et donc, étant donné que pour tout $k \in \mathbb{N}$, $\varphi(k) \geq k$, on peut affirmer que pour tout $n \geq \mathrm{N}$, $|u_{\varphi(n)} - \ell| < \varepsilon$ et donc $u_{\varphi(n)} \longrightarrow \ell$. On en déduit donc que $\mathrm{Adh}(u_n) = \{\ell\}$.

III Bolzano-Weierstraß

Théorème (Théorème de Bolzano-Weierstraß) III.1.

Soit (u_n) une suite réelle. Si (u_n) est bornée, alors elle admet une valeur d'adhérence.

Démonstration. Pour montrer ce théorème, on aura besoin du lemme suivant.

Lemme (Lemme des pics) III.2.

Toute suite réelle (u_n) admet une suite extraite monotone.

Démonstration du lemme. Posons $\mathrm{A} = \{n \in \mathbb{N},\ \forall k \geq n,\ u_n \geq u_k\}$. Deux cas se présentent.

(cc) BY-NC-SA

→ Si A est infini, alors d'après le point 2 de la proposition I.2, on peut considérer une extractrice φ telle que $\varphi(\mathbb{N}) = \mathrm{A}$ et de remarquer que par définition de A, $(u_{\varphi(n)})$ est décroissante.

→ Si A est fini (ou vide), alors étant donné que A est une partie finie de $\mathbb{N}$, elle est bornée. Posons donc $\mathrm{N} = 1 + \max \mathrm{A}$. Construisons une extractrice φ par récurrence telle que $(u_{\varphi(n)})$ est strictement croissante.

- On pose $\varphi(0) = \mathrm{N}$.
- Soit $n \in \mathbb{N}$ et supposons que pour tout $\ell \in [\![0; n]\!]$, $\varphi(\ell)$ est bien définie et que $u_{\varphi(0)} \leq u_{\varphi(1)} \leq \cdots \leq u_{\varphi(n)}$. $\varphi(n) \geq \varphi(0) = \mathrm{N}$ donc par construction $\varphi(n) \notin \mathrm{A}$ *i.e.* il existe $k > \varphi(n)$ tel que $u_{\varphi(n)} < u_k$. On pose alors $\varphi(n+1) = k$. On a alors bien que $u_{\varphi(0)} \leq u_{\varphi(1)} \leq \cdots \leq u_{\varphi(n)} \leq u_{\varphi(n+1)}$.

On en déduit donc que (u_n) admet forcément une suite extraite croissante ou décroissante.

Revenons à la démonstration du théorème. (u_n) admet une suite extraite monotone. (u_n) est réelle bornée donc cette suite extraite est également bornée, donc convergente. (u_n) admet donc bien une valeur d'adhérence.

Remarque. Si $(u_n) \in \mathbb{R}^{\mathbb{N}}$ n'est pas bornée, alors elle admet une sous-suite qui tend vers $+\infty$ ou $-\infty$.

On peut étendre le théorème de Bolzano-Weierstraß à $\mathbb{C}$.

Proposition III.3.

Si (u_n) est une suite complexe bornée, alors elle admet une valeur d'adhérence.

Démonstration. Écrivons pour tout $n \in \mathbb{N}$, $u_n = x_n + iy_n$ avec (x_n) et (y_n) deux suites réelles. On a pour tout $n \in \mathbb{N}$, $|x_n| \leq |u_n|$ et $|y_n| \leq |u_n|$ donc x_n et y_n sont bornées. D'après le théorème de Bolzano-Weierstraß, il existe une suite extraite et $x \in \mathbb{R}$ tels que $x_{\varphi(n)} \xrightarrow[n\to+\infty]{} x$. La suite $(y_{\varphi(n)})$ est bornée donc encore une fois d'après le théorème de Bolzano-Weierstraß, il existe une extractrice ψ et $y \in \mathbb{R}$ tels que $y_{\varphi(\psi(n))} \xrightarrow[n\to+\infty]{} y$. On en déduit alors que finalement $u_{\varphi\circ\psi(n)} = x_{\varphi\circ\psi(n)} + iy_{\varphi\circ\psi(n)} \xrightarrow[n\to+\infty]{} x + iy$. (u_n) admet donc bien une valeur d'adhérence.

Remarque. On verra plus tard dans le chapitre de topologie que ce théorème peut être étendu aux suites à valeurs dans un espace vectoriel normé (réel ou complexe) de dimension finie.

Proposition III.4.

Soit (z_n) une suite complexe bornée. Les propositions suivantes sont équivalentes.

1. (z_n) converge.
2. (z_n) admet exactement une valeur d'adhérence.
3. (z_n) admet au plus une valeur d'adhérence.

Démonstration.

→ (1) ⇒ (2) Cette implication est équivalente à l'énoncé de la proposition II.3.

→ (2) ⇒ (3) Cette implication est évidente.

→ (3) ⇒ (1) (z_n) est bornée donc d'après le théorème de Bolzano-Weierstraß, (z_n) admet une valeur d'adhérence. (z_n) admet donc exactement une valeur d'adhérence qu'on notera z. Montrons que $z_n \xrightarrow[n\to+\infty]{} z$. Supposons le contraire. Il existe donc par définition $\varepsilon > 0$ tel que $\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |z_n - z| < \varepsilon\}$ est fini. On en déduit alors que l'ensemble $\mathrm{A}' = \mathbb{N} \setminus \mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |z_n - z| \geq \varepsilon\}$ est infini. D'après le point 2 de la proposition I.2, il existe une extractrice φ telle que $\varphi(\mathbb{N}) = \mathrm{A}'$. On a alors pour tout $n \in \mathbb{N}$, $\left|z_{\varphi(n)} - z\right| \geq \varepsilon$. La suite $(z_{\varphi(n)})$ est bornée donc d'après le théorème de Bolzano-Weierstraß, il existe une extractrice ψ et $a \in \mathbb{C}$ tels que $z_{\varphi(\psi(n))} \xrightarrow[n\to+\infty]{} a$. De plus, on

(cc) BY-NC-SA

a clairement que $|a-z| \geq \varepsilon > 0$ donc $a \neq z$ et $a \in \mathrm{Adh}(z_n)$ ce qui est absurde. On en déduit donc finalement que (z_n) est bien convergente (vers z).

IV Compléments

Dans cette partie, X est égal à $\mathbb{R}$ ou $\mathbb{C}$.

Définition IV.1.

1. On appelle boule ouverte (resp. fermée) de X de centre $a \in \mathrm{X}$ et de rayon $\varepsilon > 0$ l'ensemble

$$\mathrm{B}(a,\varepsilon) = \{x \in \mathrm{X},\ |x-a| < \varepsilon\} \text{ (resp. } \mathrm{B}_f(a,\varepsilon) = \{x \in \mathrm{X},\ |x-a| \leq \varepsilon\})$$

2. Une partie O de X est dite ouverte dans X lorsque pour tout $a \in \mathrm{O}$, il existe $\varepsilon > 0$ tel que $\mathrm{B}(a,\varepsilon) \subset \mathrm{O}$.
3. Une partie $\mathrm{F} \subset \mathrm{X}$ est dite fermée dans X lorsque $\mathrm{X} \setminus \mathrm{F}$ est ouvert dans X.

Remarque.

→ Il est aisé de vérifier qu'une intersection quelconque de fermés est un fermé.

→ Pour tout $a \in \mathrm{X}$ et $\varepsilon > 0$, $\mathrm{B}(a,\varepsilon)$ est ouvert dans X et $\mathrm{B}_f(a,\varepsilon)$ est fermé dans X.

→ Pour tout $a \in \mathbb{R}$ et $\varepsilon > 0$, on a $\mathrm{B}(a,\varepsilon) =]a-\varepsilon, a+\varepsilon[$.

Exemples.

→ Les intervalles fermés et les ensembles finis sont des parties fermées de $\mathbb{R}$.

→ Les intervalles ouverts sont des parties ouvertes de $\mathbb{R}$.

Proposition IV.2.

Soit F une partie de X. F est fermée si et seulement si pour toute suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans F, si (x_n) est convergente, alors sa limite est dans F.

Démonstration.

($\Rightarrow$) Supposons que F soit fermé. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans F convergente, supposons que la limite ℓ de (x_n) est dans $\mathrm{X} \setminus \mathrm{F}$. $\mathrm{X} \setminus \mathrm{F}$ étant ouvert, il existe $\varepsilon > 0$ tel que $\mathrm{B}(\ell,\varepsilon) \subset \mathrm{X} \setminus \mathrm{F}$. Or par construction, il existe $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|x_n - \ell| \leq \frac{\varepsilon}{2}$ et donc $x_n \in \mathrm{B}(\ell,\varepsilon) \subset \mathrm{X} \setminus \mathrm{F}$ ce qui est absurde.

($\Leftarrow$) Réciproquement, supposons que toute suite à valeurs dans F convergente converge dans F. supposons de plus que F n'est pas fermé, *i.e.* $\mathrm{X} \setminus \mathrm{F}$ n'est pas ouvert. Il existe donc $a \in \mathrm{X} \setminus \mathrm{F}$ tel que pour tout $\varepsilon > 0$, $\mathrm{B}(a,\varepsilon) \not\subset \mathrm{X} \setminus \mathrm{F}$ *i.e.* $\mathrm{B}(a,\varepsilon) \cap \mathrm{F} \neq \varnothing$. Cette propriété étant vraie pour $\varepsilon = \dfrac{1}{n}$ et tout $n \in \mathbb{N}^*$, on peut définir une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans F qui vérifie pour tout $n \in \mathbb{N}^*$, $|a - x_n| \leq \frac{1}{n}$. (x_n) est à valeurs dans F et converge vers a qui n'est pas dans F, ce qui est en contradiction avec les hypothèses.

Proposition IV.3.

1. Pour toute partie F de X, F est fermée si et seulement s'il existe $(u_n) \in \mathrm{X}^{\mathbb{N}}$ telle que $\mathrm{Adh}(u_n) = \mathrm{F}$.
2. Lorsque $\mathrm{X} = \mathbb{R}$, pour toute suite $(u_n) \in \mathrm{X}^{\mathbb{N}}$ bornée, $\mathrm{Adh}(u_n)$ contient sa borne supérieure et sa borne inférieure.

(cc) BY-NC-SA

Démonstration

1. Procédons par double implication.
   - $\to$ ($\Leftarrow$) Soit (u_n) une suite à valeurs dans X. Posons $\mathrm{F} = \mathrm{Adh}(u_n)$ et montrons que F est fermé. Pour cela, on va montrer que $\mathrm{X} \setminus \mathrm{F}$ est ouvert. Soit $\ell \in \mathrm{X} \setminus \mathrm{F}$. Par définition, on sait qu'il existe $\varepsilon > 0$ tel que l'ensemble $\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |u_n - \ell| < \varepsilon\}$ est fini. On veut alors montrer que $\mathrm{B}(\ell, \varepsilon) \subset \mathrm{F} \setminus \mathrm{A}$. Soit $\ell' \in \mathrm{B}(\ell, \varepsilon) = \{x \in \mathrm{X},\ |x - \ell| < \varepsilon\}$. En posant $\delta = \dfrac{1}{2} \min\left(|\ell' - (\ell + \varepsilon)|, |\ell' - (\ell - \varepsilon)|\right)$, on voit clairement que $\mathrm{B}(\ell', \delta) \subset \mathrm{B}(\ell, \varepsilon)$ et alors
   $$\mathrm{A}'(\delta) = \{n \in \mathbb{N},\ |u_n - \ell'| < \delta\} \subset \{n \in \mathbb{N},\ |u_n - \ell| < \varepsilon\} = \mathrm{A}(\varepsilon).$$
   On en déduit alors que nécessairement $\mathrm{A}'(\delta)$ est fini et que donc par définition (proposition II.1) $\ell' \in \mathrm{X} \setminus \mathrm{Adh}(u_n)$, et donc $\mathrm{B}(\ell, \varepsilon) \subset \mathrm{X} \setminus \mathrm{F}$. Ceci nous permet alors d'affirmer que $\mathrm{X} \setminus \mathrm{F}$ est ouvert *i.e.* que F est fermé.
   - $\to$ ($\Rightarrow$) Cette implication sera faite en exercice dans la suite du chapitre.
2. La suite (u_n) est bornée, il existe donc $\mathrm{T} \in \mathbb{R}^+$ tel que pour tout $n \in \mathbb{N}$, $|u_n| \leq \mathrm{T}$. Pour tout $a \in \mathrm{Adh}(u_n)$, il existe une extractrice φ telle que $u_{\varphi(n)} \xrightarrow[n \to +\infty]{} a$. En passant donc à la limite dans l'inégalité précédente, on obtient que $|a| \leq \mathrm{T}$ et on en déduit que $\mathrm{Adh}(u_n)$ est borné. De plus, d'après le point précédent, $\mathrm{Adh}(u_n)$ est non vide fermé. Utilisons enfin le lemme suivant.

Lemme IV.4.

Si F est une partie fermée non vide de $\mathbb{R}$, alors F contient $\sup \mathrm{F}$ et $\inf \mathrm{F}$.

Démonstration du lemme. Il suffit de voir que $\sup \mathrm{F}$ et $\inf \mathrm{F}$ peuvent être vus comme limite d'une suite à valeurs dans F et que F étant fermé, cette limite est forcément dans F d'après la proposition IV.2.

On en déduit donc d'après ce lemme que $\mathrm{Adh}(u_n)$ contient ses bornes.

Exercice IV.5.

Soit (u_n) une suite bornée à valeurs dans $\mathbb{R}$. D'après la proposition précédente, $\mathrm{Adh}(u_n)$ contient ses bornes. Montrer alors que si m et M sont respectivement les bornes supérieure et inférieure de $\mathrm{Adh}(u_n)$, on a

$$m = \lim_{n \longrightarrow \infty} \inf\{u_k,\ k \geq n\} \text{ et } \mathrm{M} = \lim_{n \longrightarrow \infty} \sup\{u_k,\ k \geq n\}.$$

Remarque. En général, on note pour toute suite (u_n)

$$\limsup u_n = \lim_{n \longrightarrow \infty} \sup\{u_k,\ k \geq n\} \text{ et } \liminf u_n = \lim_{n \longrightarrow \infty} \inf\{u_k,\ k \geq n\}.$$

Exercice IV.6.

Soit F un fermé borné de X.

1. Montrer que $\forall \varepsilon > 0,\ \exists n_\varepsilon \geq 1,\ \exists z_1, \ldots, z_{n_\varepsilon} \in \mathrm{F},\ \mathrm{F} \subset \displaystyle\bigcup_{i=1}^{n_\varepsilon} \mathrm{B}(z_i, \varepsilon)$.
2. En déduire qu'il existe une suite $(u_n) \in \mathrm{X}^{\mathbb{N}}$ telle que $\mathrm{F} = \mathrm{Adh}(u_n)$.
3. Montrer que la propriété de la question 2 reste vraie lorsque F n'est plus forcément borné.

(cc) BY-NC-SA

Exercice IV.7.

Soit (u_n) une suite réelle telle que $u_{n+1} - u_n \xrightarrow[n\to+\infty]{} 0$. Montrer que $\mathrm{Adh}(u_n)$ est un intervalle fermé ou vide.

V Suites de Cauchy

Définition V.1.

Une suite $(z_n) \in \mathbb{C}^{\mathbb{N}}$ est dite de Cauchy si

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \geq 0,\ \forall m, n \geq n_\varepsilon,\ |z_m - z_n| < \varepsilon.$$

Remarques.

→ Dans cette définition, m et n sont indépendants.

→ La négation de cette définition peut être formulée par des extractrices. En effet, la négation de cette définition est équivalente à ce qui suit : il existe φ, ϕ deux extractrices et $\varepsilon > 0$ tels que pour tout $n \geq 0$ $|z_{\varphi(n)} - z_{\phi(n)}| > \varepsilon$. Le lecteur est invité à montrer qu'il s'agit bien de la négation du fait qu'une suite soit de Cauchy.

Proposition V.2.

Soit (z_n) une suite à valeurs dans $\mathbb{C}$.

1. Si (z_n) est de Cauchy, alors $\forall p \geq 1,\ z_{n+p} - z_n \xrightarrow[n\to+\infty]{} 0$.
2. Si la limite de la suite $(v_n) = \left(\sum_{k=0}^{n} |z_{k+1} - z_k|\right)_{n\in\mathbb{N}}$ est finie, alors (z_n) est de Cauchy.
3. Si (z_n) converge alors elle est de Cauchy.
4. Si (z_n) est de Cauchy, alors elle est bornée.
5. Si (z_n) est de Cauchy et possède une valeur d'adhérence, alors elle converge.

Démonstration.

1. Soit $p \geq 1$. Écrivons la définition du fait qu'on ait $z_{n+p} - z_n \xrightarrow[n\to+\infty]{} 0$.

$$\forall \varepsilon > 0, \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ |z_{n+p} - z_n| < \varepsilon.$$

Soit $\varepsilon > 0$. (z_n) est de Cauchy, il existe donc $\mathrm{N} \in \mathbb{N}$ tel que pour $m, n \geq \mathrm{N}$, $|z_n - z_m| < \varepsilon$. En particulier, pour tout $n \geq \mathrm{N}$, $n + p \geq \mathrm{N}$ et donc $|z_{n+p} - z_n| < \varepsilon$, d'où le résultat.

2. Pour tout $\mathrm{N} \in \mathbb{N}$, notons $\sum_{k=\mathrm{N}}^{\infty} |z_{k+1} - z_k|$ la limite de la suite $\left(\sum_{k=\mathrm{N}}^{n} |z_{k+1} - z_k|\right)_{n\geq\mathrm{N}}$. On remarque que

$$\sum_{k=\mathrm{N}}^{\infty} |z_{k+1} - z_k| = \sum_{k=0}^{\infty} |z_{k+1} - z_k| - \sum_{k=0}^{\mathrm{N}-1} |z_{k+1} - z_k| \xrightarrow[\mathrm{N}\to+\infty]{} 0.$$

Considérons donc $\mathrm{N} \geq 0$ tel que $\sum_{\mathrm{N}}^{\infty} |z_{k+1} - z_k| < \varepsilon$. On a alors pour tout $n, m \geq \mathrm{N}$, en supposant

(cc) BY-NC-SA

sans perte de généralité que $m \geq n$,

$$|z_n - z_m| = \left|\sum_{k=n}^{m-1} z_{k+1} - z_k\right| \leq \sum_{k=n}^{m-1} |z_{k+1} - z_k| \leq \sum_{k=n}^{+\infty} |z_{k+1} - z_k| \leq \sum_{k=\mathrm{N}}^{+\infty} |z_{k+1} - z_k| < \varepsilon,$$

donc (z_n) est bien une suite de Cauchy.

3. Soit ℓ la limite de (z_n). Soit $\varepsilon > 0$. (z_n) converge vers ℓ, il existe donc $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|z_n - \ell| < \frac{\varepsilon}{2}$. On a alors pour tout $m, n \geq \mathrm{N}$,

$$|z_n - z_m| \leq |z_n - \ell| + |z_m - \ell| + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

La suite (z_n) est donc bien de Cauchy.

4. Appliquons la définition d'une suite de Cauchy pour $\varepsilon = 1$. (z_n) est de Cauchy, il existe donc $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|u_n - u_\mathrm{N}| < 1$. En particulier, pour tout $n \in \mathbb{N}$,

$$|u_n| \leq \max(|u_0|, \ldots, |u_{\mathrm{N}-1}|, |u_\mathrm{N}| + 1).$$

5. Soit z une valeur d'adhérence de (z_n) et φ une extractrice telle que $z_{\varphi(n)} \xrightarrow[n\to+\infty]{} z$. Soit $\varepsilon > 0$ et soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|z_{\varphi(n)} - z| < \frac{\varepsilon}{2}$. (z_n) est de Cauchy, il existe donc N' tel que pour tout $n, m \geq \mathrm{N}'$, $|z_n - z_m| < \frac{\varepsilon}{2}$. En posant alors $\mathrm{M} = \max(\mathrm{N}, \mathrm{N}')$, on voit que pour tout $n \geq \mathrm{M}$,

$$|z_n - z| \leq \left|z_n - z_{\varphi(n)}\right| + \left|z_{\varphi(n)} - z\right| \underset{(*)}{\leq} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

L'inégalité $(*)$ est vraie car $n \geq \mathrm{M}$ et φ est une extractrice donc d'après le point 1 de la proposition I.2, $\varphi(n) \geq n$ et alors $\varphi(n) \geq \mathrm{M}$.

Remarque. La réciproque du point (1) est fausse. En effet, la suite $(z_n) = (\ln n)_{n\in\mathbb{N}}$ vérifie,

$$\forall p \geq 1,\ z_{n+p} - z_n = \ln\left(\frac{n+p}{n}\right) \xrightarrow[n\to+\infty]{} 0$$

Si (z_n) est de Cauchy, alors, on a nécessairement $z_{2n} - z_n \xrightarrow[n\to+\infty]{} 0$ ce qui est faux car pour tout $n \in \mathbb{N}$,

$$z_{2n} - z_n = \ln(2n) - \ln n = \ln 2 > 0$$

Théorème V.3.

Soit (z_n) une suite à valeurs dans $\mathbb{C}$. (z_n) est de Cauchy si et seulement si (z_n) converge.

Démonstration. La suite (z_n) est de Cauchy, elle est donc bornée d'après le point 4 de la proposition V.2 et alors, d'après le théorème Bolzano-Weierstraß, (z_n) possède une valeur d'adhérence. (z_n) est alors une suite de Cauchy qui admet une valeur d'adhérence, elle converge donc d'après le point 5 de la proposition V.2.

(cc) BY-NC-SA

Correction de l'exercice II.2. :

D'après le chapitre 2 (Cardinaux), il existe f une bijection entre $\mathbb{N}$ et $\mathbb{Q}$. Considérons donc la suite $(u_n) = (f(n))_{n\in\mathbb{N}}$. Pour tout $x \in \mathbb{R}$, il est clair que pour tout $\varepsilon > 0$,$\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |u_n - x| < \varepsilon\}$ est infini, car il y a un nombre infini de nombres rationnels dans $\mathrm{B}(x, \varepsilon)$. On en déduit donc que x est bien une valeur d'adhérence de (u_n) et que finalement $\mathrm{Adh}(u_n) = \mathbb{R}$.

Remarquons qu'on peut procéder de la même manière pour $\mathbb{C}$ en utilisant le fait que

$$\mathbb{Q} + i\mathbb{Q} = \{r_1 + ir_2,\ r_1, r_2 \in \mathbb{Q}\}$$

est aussi dénombrable.

Correction de l'exercice IV.5. :

La démonstration pour m étant identique, on va uniquement faire la démonstration pour M. Le lecteur est invité à reproduire cette démonstration pour la limite inférieure m à titre d'exercice. Posons $(v_n) = (\sup\{u_k,\ k \geq n\})_{n\in\mathbb{N}}$. La suite (v_n) est décroissante car la suite d'ensembles non vides majorés $(\mathrm{A}_n) = (\{u_k,\ k \geq n\})_{n\in\mathbb{N}}$ est décroissante pour l'inclusion. De plus u_n est minorée, donc en posant A un minorant de (u_n), on peut facilement voir que A minore également (v_n). (v_n) est donc une suite décroissante minorée, on peut donc affirmer qu'elle converge vers sa borne inférieure. On en déduit alors que

$$\mathrm{M} = \lim_{n\longrightarrow+\infty} v_n = \inf_{n\in\mathbb{N}} v_n = \inf_{n\in\mathbb{N}} \sup\{u_k,\ k \geq n\}.$$

On en déduit alors par définition de la borne inférieure que pour tout $\varepsilon > 0$, il existe $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$ on ait $\mathrm{M} \leq v_n \leq \mathrm{M} + \varepsilon$ et par définition de la borne supérieure, il existe $k \geq n$ tel que $v_n - \varepsilon \leq u_k \leq v_n$. On en déduit donc que pour tout $\varepsilon > 0$ il existe $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$ il existe $k \geq n$ tel que

$$\mathrm{M} - \varepsilon \leq v_n - \varepsilon \leq u_k \leq v_n \leq \mathrm{M} + \varepsilon.$$

n peut être pris arbitrairement grand, ce qui nous permet d'affirmer que

$$\forall\varepsilon > 0,\ \forall\mathrm{M} > 0,\ \exists k \geq \mathrm{M},\ |u_k - \mathrm{M}| \leq \varepsilon,$$

ce qui nous permet de dire d'après la définition II.1 que M est bien une valeur d'adhérence de (u_n).

Il reste à montrer que M est bien la plus grande valeur d'adhérence de (u_n). Soit $a \in \mathrm{Adh}(u_n)$ et φ une extractrice telle que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$. On a alors pour tout $n \in \mathbb{N}$

$$u_{\varphi(n)} \leq \sup\{u_k, k \geq \varphi(n)\} = v_{\varphi(n)}$$

en passant à la limite des deux côtés de l'inégalité, on obtient que $a \leq \mathrm{M}$, ce qui est bien le résultat voulu.

Correction de l'exercice IV.6. :

1. Soit $\varepsilon > 0$. Supposons par l'absurde que pour tout $n \in \mathbb{N}$ et $z_1, \dots, z_n \in \mathrm{X}$, $\mathrm{F} \not\subset \bigcup_{i=1}^{n} \mathrm{B}(z_i, \varepsilon)$. Construisons alors la suite $(z_n)_{n\in\mathbb{N}^*}$ de la manière suivante.

   → $z_1 \in \mathrm{F}$

   → Pour tout $n \in \mathbb{N}$, $z_{n+1} \in \mathrm{F} \setminus \bigcup_{i=1}^{n} \mathrm{B}(z_i, \varepsilon)$. Ceci est possible car $\mathrm{F} \setminus \bigcup_{i=1}^{n} \mathrm{B}(z_i, \varepsilon)$ est toujours non vide par hypothèse.

   La suite (z_n) vérifie clairement pour tout $n, m \in \mathbb{N}$, tels que $n \neq m$, $|z_n - z_m| \geq \varepsilon$. Or (z_n) est à valeurs dans F qui est borné, donc (z_n) est bornée et alors d'après le théorème de Bolzano-Weierstraß,

(cc) BY-NC-SA

il existe $x \in \mathrm{X}$ et une extractrice φ telle que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} x$. On a alors

$$\varepsilon \leq \left|u_{\varphi(n+1)} - u_{\varphi(n)}\right| \xrightarrow[n\to+\infty]{} 0,$$

ce qui est absurde. On a donc bien le résultat voulu.

2. Pour tout $\varepsilon > 0$, posons n_ε et $z_{1,\varepsilon}, \ldots, z_{n_\varepsilon,\varepsilon}$ fournis par la question précédente, tels que $\mathrm{F} \subset \bigcup_{i=1}^{n_\varepsilon} \mathrm{B}(z_{i,\varepsilon}, \varepsilon)$ et posons pour tout $k \in \mathbb{N}$, $\mathrm{N}_k = n_{2^{-0}} + \cdots + n_{2^{-k}}$. Construisons alors la suite $(u_n)_{n\in\mathbb{N}^*}$ de la manière suivante.

   → Pour tout $j \in [\![1; n_{2^{-0}}]\!]$, $u_j = z_{j,2^{-0}}$.

   → Pour tout $k \in \mathbb{N}$ et $j \in [\![1; n_{2^{-(k+1)}}]\!]$, $u_{\mathrm{N}_k+j} = z_{j,2^{-(k+1)}}$.

   Il s'agit de la concaténation des suites finies $(z_{1,2^{-0}}, \ldots, z_{n_{2^{-0}},2^{-0}}), \ldots, (z_{1,2^{-n}}, \ldots, z_{n_{2^{-n}},2^{-n}}), \ldots$
   Montrons maintenant que $\mathrm{Adh}(u_n) = \mathrm{F}$. Soit $x \in \mathrm{F}$. Construisons une extractrice φ vérifiant pour tout $n \in \mathbb{N}$, $\left|u_{\varphi(n)} - x\right| < 2^{-n}$ de la manière suivante.

   → On sait que $\mathrm{F} \subset \bigcup_{i=1}^{n_{2^{-0}}} \mathrm{B}(z_i, 2^{-0})$, il existe donc $i \in [\![1; n_{2^{-0}}]\!]$ tel que $x \in \mathrm{B}(z_i, 2^{-0}) = \mathrm{B}(u_i, 2^{-0})$ *i.e.* $|u_i - x| \leq 2^{-0}$. On pose donc $\varphi(0) = i$.

   → Soit $n \in \mathbb{N}$. Supposons que pour tout $k \in [\![1; n]\!]$, $\varphi(k)$ est défini de manière à ce qu'on ait $\left|u_{\varphi(k)} - x\right| \leq 2^{-k}$ et $\varphi(k) > \varphi(k-1)$. Soit $j \geq n$ tel que $\mathrm{N}_j \geq \varphi(n)$. Par hypothèse, il existe $i \in [\![1; n_{2^{-(j+1)}}]\!]$ tel que

   $$x \in \mathrm{B}(z_{i,2^{-(j+1)}}, 2^{-(j+1)}) = \mathrm{B}(u_{i+\mathrm{N}_j}, 2^{-(j+1)}) \text{ i.e. } \left|u_{i+\mathrm{N}_j} - x\right| \leq 2^{-(j+1)} \leq 2^{-(n+1)}.$$

   On pose alors $\varphi(n+1) = \mathrm{N}_j + i$. Cette construction nous donne bien que $\varphi(n+1) \geq \mathrm{N}_j + 1 > \varphi(n)$ et $\left|u_{\varphi(n+1)} - x\right| < 2^{-(n+1)}$.

   On en déduit donc par construction que pour tout $n \in \mathbb{N}$ $\left|u_{\varphi(n)} - x\right| < 2^{-n}$ et alors que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} x$. On a alors bien que $\mathrm{Adh}(u_n) \supset \mathrm{F}$.

   Montrons l'inclusion réciproque. Soit a une valeur d'adhérence de (u_n) et φ une extractrice telle que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$. La suite $(u_{\varphi(n)})$ est à valeurs dans F donc d'après la proposition IV.2 sa limite est également dans F *i.e.* $a \in \mathrm{F}$. On en déduit alors qu'on a bien $\mathrm{Adh}(u_n) \subset \mathrm{F}$ et finalement $\mathrm{Adh}(u_n) = \mathrm{F}$.

3. Nous ne détaillerons pas cette question mais donnerons uniquement l'idée générale. Pour montrer que la propriété est vraie lorsque F n'est pas forcément borné, on commence par poser pour tout fermé borné H, les éléments $z_{1,\varepsilon,\mathrm{H}}, \ldots, z_{n_\varepsilon,\varepsilon,\mathrm{H}}$ les éléments fournis par la question précédente tels que $\mathrm{H} \subset \bigcup_{i=1}^{n_\varepsilon} \mathrm{B}(z_{i,\varepsilon,\mathrm{H}})$. Ensuite, on considère $a \in \mathrm{F}$ et une suite de fermés $(\mathrm{F}_n)_{n\in\mathbb{N}} = (\mathrm{F} \cap \mathrm{B}_f(a, 2^n))_{n\in\mathbb{N}}$. Un fois cela fait, on peut facilement montrer que la suite résultant de la concaténation des suites finies $(z_{1,2^{-0},\mathrm{F}_0}, \ldots, z_{n_{2^{-0}},2^{-0},\mathrm{F}_0}), \ldots, (z_{1,2^{-n},\mathrm{F}_n}, \ldots, z_{n_{2^{-n}},2^{-n},\mathrm{F}_n}), \ldots$ vérifie bien la propriété voulue.

Correction de l'exercice IV.7. :

Posons $\mathrm{F} = \mathrm{Adh}(u_n)$. Supposons que $\mathrm{F} \neq \varnothing$. On veut montrer que F est un intervalle, c'est à dire que pour tout $a, b \in \mathrm{F}$ tels que $a < b$, $]a, b[\subset \mathrm{F}$. Considérons donc $a, b \in \mathrm{F}$ tels que $a < b$ et $z \in]a, b[$. Supposons par l'absurde que $z \notin \mathrm{Adh}(u_n)$, *i.e.* qu'il existe $\varepsilon > 0$ tel que $\mathrm{A}(\varepsilon) = \{n \in \mathbb{N},\ |u_n - z| < \varepsilon\}$ est fini.

$\mathrm{A}(\varepsilon)$ est une partie majorée de $\mathbb{N}$, on peut donc considérer M son maximum. Soit $\eta > 0$ tel que $\eta < \min(z - a, b - z, \varepsilon)$ et soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|u_{n+1} - u_n| < \eta$. a et b sont des valeurs

(cc) BY-NC-SA

d'adhérence de (u_n), il existe donc $n_1, n_2 \in \mathbb{N}$ tels que $n_1 \geq \max(\mathrm{N}, \mathrm{M}+1)$ et $n_2 \geq \max(\mathrm{N}, \mathrm{M}+1)$ tels que $|u_{n_2} - a| < \eta$ et $|u_{n_1} - b| < \eta$. Supposons sans perte de généralité que $n_1 \leq n_2$ et posons $\mathrm{B} = \{k \in [\![n_1; n_2]\!],\ u_k \geq z\}$. B est non vide car

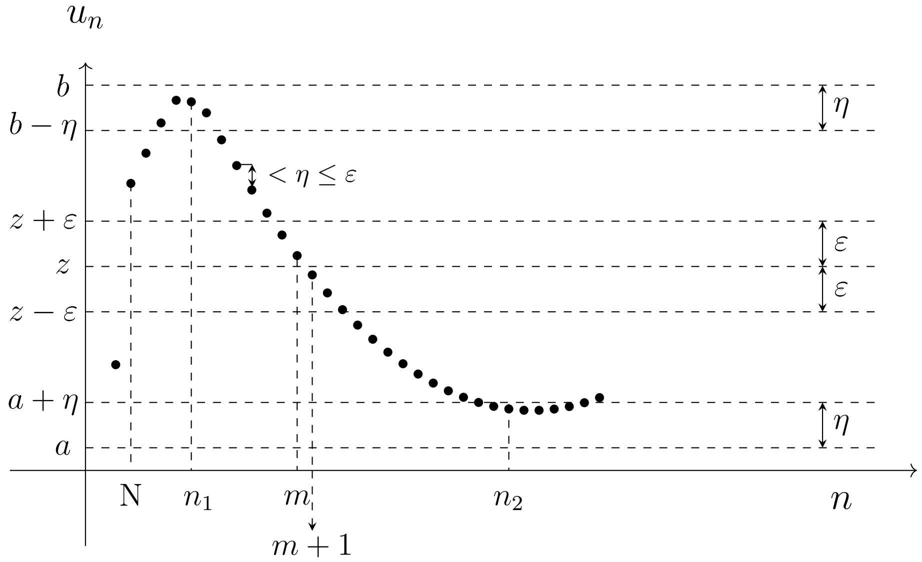
$$u_{n_1} \geq b - \eta \geq b - \min(z-a, b-z, \varepsilon) \geq b - b + z = z,$$

et alors $n_1 \in \mathrm{B}$. De plus B est une partie majorée de $\mathbb{N}$, elle admet donc un maximum qu'on note m. m est strictement inférieur à n_2 car

$$u_{n_2} < a + \varepsilon < a + \min(z-a, b-z, \varepsilon) \leq a + z - a = z.$$

On a alors $u_m \geq z \geq u_{m+1}$ et $m \geq \mathrm{N}$ donc $|u_{m+1} - u_m| < \eta \leq \varepsilon$, ce qui nous permet d'écrire $u_m \geq z \geq u_{m+1} > u_m - \varepsilon$ et alors que $|u_m - z| < \varepsilon$ *i.e.* $m \in \mathrm{A}(\varepsilon)$ ce qui est absurde car $m > \mathrm{M}$ et M est le maximum de $\mathrm{A}(\varepsilon)$. On en déduit alors que $z \in]a, b[$ et que finalement $\mathrm{Adh}(u_n)$ est un intervalle.

Pour comprendre un peu mieux l'intuition de cette démonstration, regardons le dessin ci-dessous.



On voit que lorsque $n \geq \mathrm{N}$, la suite (u_n) fait des pas très petits, ce qui fait qu'elle ne peut pas passer d'une position au-dessus de $z+\varepsilon$ à une position au-dessous de $z-\varepsilon$ sans passer par la bande $]z-\varepsilon, z+\varepsilon[$.

Remarque. En utilisant le résultat de l'exercice IV.5, on peut déduire de cet exercice que lorsque (u_n) est bornée et $u_{n+1} - u_n \xrightarrow[n\to+\infty]{} 0$, on a $\mathrm{Adh}(u_n) = [\liminf u_n,\ \limsup u_n]$.

On peut montrer également (nous ne le ferons pas, mais le lecteur est encouragé à le faire) que lorsque (u_n) n'est pas forcément bornée et que les limites $\limsup u_n$ et $\liminf u_n$ ne sont pas forcément finies, cette propriété reste vraie.

Chapitre 5

Intégrales généralisées

Ce chapitre a deux objectifs : le premier est de munir le lecteur d'outils importants pour les concours et le second de lui donner un recul sur les intégrales en général. En particulier, nous allons revoir rapidement la définition d'une intégrale comme vue en première année, et voir comment généraliser cette définition à des cas plus complexes.

Le lecteur est encouragé suite à la lecture de chapitre à s'imaginer dans la situation suivante : quelqu'un qui n'est pas un mathématicien vous demande "qu'est-ce qu'une intégrale ?". Savoir répondre d'une manière très simple, détachée des notations et formalismes mathématiques utilisés, prouve que vous avez une bonne compréhension.

Préambule : convergence séquentielle

On commence par quelques rappels et éléments qui seront utiles dans la suite.

Rappel .4.

Soit I un intervalle de $\mathbb{R}$, et a un élément de I (ou un bord de I), et soit f une application de I dans $\mathbb{R}$. f possède une limite finie en a si et seulement si pour toute suite $(u_n)_{n\in\mathbb{N}}$ à valeurs dans I, si (u_n) converge vers a, alors $f(u_n)$ est convergente.

Démonstration.

→ ($\Rightarrow$) Supposons que f admet une limite finie en a, qu'on note ℓ. Soit (u_n) une suite à valeurs dans I qui converge vers a. La définition de la limite nous donne

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \mathrm{I},\ |x-a| < \eta \Longrightarrow |f(x)-\ell| < \varepsilon. \tag{1}$$

On va essayer d'utiliser cette proposition pour montrer que la définition de limites pour les suites est vérifiée pour $(f(u_n))$. Soit $\varepsilon > 0$, et η tel que pour tout $x \in \mathrm{I}$, $|x-a| < \eta \Longrightarrow |f(x)-\ell| < \varepsilon$. La définition de limite d'une suite appliquée à (u_n) nous donne

$$\forall \varepsilon' > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ |u_n - a| < \varepsilon'.$$

Considérons donc $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|u_n - a| < \eta$. On a alors pour tout $n \geq \mathrm{N}$, en utilisant la proposition (1), pour tout $n \geq \mathrm{N}$, $|f(u_n)-\ell| < \varepsilon$, ce qui nous dit bien que $(f(u_n))$ est convergente et que sa limite est ℓ.

→ ($\Leftarrow$) Supposons que pour toute suite (u_n) à valeurs dans I convergente vers a, $(f(u_n))$ est convergente. Montrons d'abord, peu importe le choix de la suite (u_n) convergente vers a, $(f(u_n))$ converge vers la même limite. Soit (u_n) et (v_n) deux suites à valeurs dans I convergentes vers a. Posons ℓ et ℓ' les limites respectives de $(f(u_n))$ et $(f(v_n))$, et considérons la suite (w_n) définie par

$$\forall n \in \mathbb{N},\ w_n = \begin{cases} u_{\frac{n}{2}} & \text{si } n \text{ est pair} \\ v_{\frac{n-1}{2}} & \text{sinon.} \end{cases}$$

On a clairement que (w_n) converge vers a et donc $(f(w_n))$ est convergente. Notons ℓ'' la limite de cette suite. $(f(w_{2n}))$ et $(f(w_{2n+1}))$ convergent également vers ℓ'', mais sont égales par définition respectivement à $(f(u_n))$ et $(f(v_n))$, et ces dernières convergent respectivement vers ℓ et ℓ' et donc

(cc) BY-NC-SA

on en déduit que $\ell = \ell' = \ell''$, ce qui montre bien l'unicité de la limite. Montrons maintenant que $f(x)$ converge vers ℓ lorsque x tend vers a. Supposons le contraire. On a

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x \in \mathrm{I},\ |x - a| < \eta \text{ et } |f(x) - \ell| \geq \varepsilon.$$

La proposition ci-dessus nous permet de considérer $\varepsilon > 0$ et une suite (x_n) tel que pour tout $n \in \mathbb{N}$, $|x_n - a| < \dfrac{1}{n+1}$ et $|f(x_n) - \ell| \geq \varepsilon$. On en déduit donc que (x_n) converge vers a, mais que $(f(x_n))$ ne converge pas vers ℓ, ce qui est absurde.

□

Proposition .5.

Soit I un intervalle de $\mathbb{R}$, et a un élément de I (ou un bord de I), et soit f une application de I dans $\mathbb{R}$. f possède une limite finie en a si et seulement si la proposition

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x, y \in \mathrm{I},\ |x - a| < \eta \text{ et } |y - a| < \eta \Longrightarrow |f(x) - f(y)| < \varepsilon \qquad (2)$$

est vérifiée.

Démonstration.

→ (⇒) Supposons que f admet une limite finie en a qu'on notera ℓ. Soit $\varepsilon > 0$. On peut alors considérer $\eta > 0$ tel que pour tout $x \in \mathrm{I}$ tel que $|x - a| < \eta$, on ait $|f(x) - \ell| < \dfrac{\varepsilon}{2}$. On a alors pour tout $x, y \in \mathrm{I}$ tels que $|x - a| < \eta$ et $|y - a| < \eta$,

$$|f(x) - f(y)| \leq |f(x) - a| + |f(y) - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

ce qui est bien le résultat voulu.

→ (⇐) Supposons que (2) est vérifiée. Cette propriété ressemble fortement au critère de Cauchy, nous allons donc l'utiliser pour montrer l'implication voulue. Soit (x_n) une suite à valeurs dans I qui converge vers a. Soit $\varepsilon > 0$. Considérons $\eta > 0$ tel que (2) est vérifiée. Soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|x_n - a| < \eta$. On a alors pour tout $m, n \geq \mathrm{N}$, $|f(x_n) - f(x_m)| < \varepsilon$. On a montré que

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall m, n \geq \mathrm{N},\ |f(x_n) - f(x_m)| < \varepsilon,$$

c'est-à-dire que la suite $(f(x_n))$ est de Cauchy, ce qui nous permet de dire que $(f(x_n))$ est convergente (voir théorème V.3 du chapitre des valeurs d'adhérence) et de conclure grâce au rappel 1.

□

Remarque. La proposition ci-dessus est aussi valable lorsqu'on remplace a par $+\infty$ ou $-\infty$. Par exemple, lorsqu'on remplace a par $+\infty$, la proposition devient : f possède une limite finie en $+\infty$ si et seulement si

$$\forall \varepsilon > 0,\ \exists \mathrm{M} > 0,\ \forall x, y \in \mathrm{I},\ x > \mathrm{M} \text{ et } y > \mathrm{M} \Longrightarrow |f(x) - f(y)| < \varepsilon.$$

Le lecteur est encourage à écrire la proposition pour $a = -\infty$.

I Définitions et premières propriétés

Dans ce chapitre, on souhaite donner une définition plus générale de l'intégrale par rapport à celle vue en première année. Pour faire cela, on commence tout d'abord par introduire quelques définitions et propriétés.

(cc) BY-NC-SA

Définition I.1.

Soit $a, b \in \mathbb{R}$ tels que $a < b$ et f une fonction de $[a, b]$ dans $\mathbb{C}$. On dit que f est continue par morceaux sur $[a, b]$ lorsqu'il existe $n \in \mathbb{N}^*$ et $a_1, \ldots, a_n \in [a, b]$ tels que $a = a_1 < a_2 < \cdots < a_{n-1} < a_n = b$ et pour tout $i \in [\![1; n-1]\!]$, f est continue sur $]a_i, a_{i+1}[$ prolongeable en une fonction continue sur $[a_i, a_{i+1}]$.

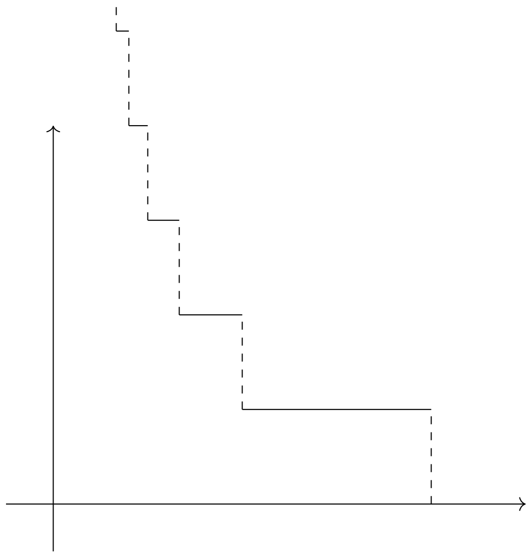
On a défini la continuité par morceaux uniquement pour les intervalles fermés bornés. On étend cette définition aux intervalles ouverts (ou ouverts que d'un seul côté), de la manière suivante.

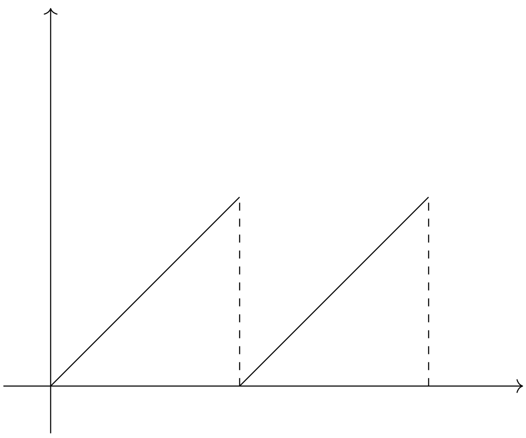
Définition I.2.

Soit $a, b \in \mathbb{R}$ tels que $a < b$, $\mathrm{I} \in \{]a, b[, [a, b[,]a, b]\}$ f une fonction de I dans $\mathbb{C}$. On dit que f est continue par morceaux lorsque pour tout intervalle fermé borné J inclus dans I, f est continue par morceaux sur J.

Remarque. Bien entendu, la définition ci-dessus est aussi valable lorsque a ou b est infini.

Exemple. La fonction $x \longmapsto \left\lfloor \dfrac{1}{x} \right\rfloor$ (à gauche) est continue par morceaux sur $]0, +\infty[$, et $x \longmapsto x - \lfloor x \rfloor$ (à droite) continue par morceaux sur $[0, 2]$.





Attention. Une fonction continue par morceaux sur un intervalle borné n'implique pas forcément qu'elle est bornée, contrairement à une fonction continue (ou continue par morceaux) sur un segment. En effet, $x \longmapsto \left\lfloor \dfrac{1}{x} \right\rfloor$ est continue par morceaux sur $]0, 1]$ mais n'est pas bornée.

On va maintenant définir une notion d'intégrale plus générale que celle vue en première année, mais tout d'abord, donnons une réponse détaillée à la question "qu'est-ce qu'une intégrale ?" d'après le cours de première année. En première année, l'intégrale est en général définie de trois manières équivalentes. Lorsque a, b sont deux réels tels que $a < b$, et $f : [a, b] \longrightarrow \mathbb{R}$ une fonction continue, on peut définir l'intégrale de f entre a et b de trois manières équivalentes.

→ La première définition, la plus utilisée en pratique, est $\displaystyle\int_a^b f(x)dx = \mathrm{F}(b) - \mathrm{F}(a)$, où F est une primitive de f.

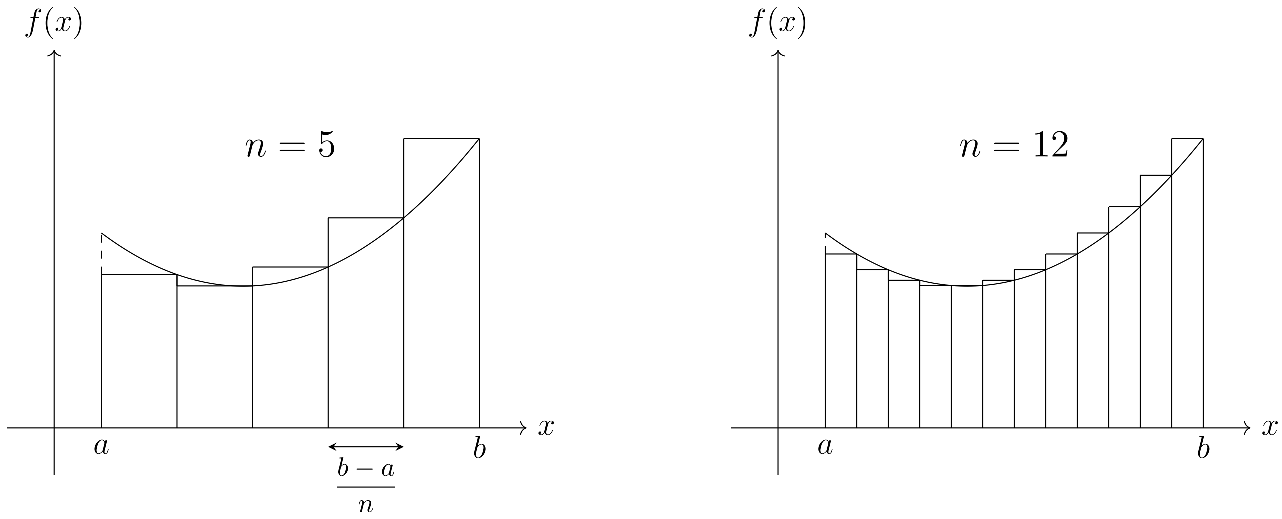
→ La deuxième, plus intrinsèque et intuitive, consiste à définir $\displaystyle\int_a^b f(x)dx$ comme la surface (algébrique) en dessous de la courbe de f entre a et b.

(cc) BY-NC-SA

→ La troisième, encore une fois plus intrinsèque que la première et plus rigoureuse que la deuxième, consiste à définir l'intégrale comme

$$\int_a^b f(x)dx = \lim_{n\to+\infty} \sum_{k=1}^{n} \frac{b-a}{n} f\left(a + \frac{k}{n}(b-a)\right). \tag{3}$$

On peut voir graphiquement pourquoi la deuxième définition de l'intégrale f sur $[a,b]$ est équivalente à la troisième.



Sur chacune des deux figures ci-dessus, la surface du $k-$ième rectangle est égale à $\frac{b-a}{n} f\left(a + \frac{k}{n}(b-a)\right)$, et donc lorsque n est fini, la quantité (3) est bien égale à la surface des rectangles sur la figure, et est intuitivement égale à la surface algébrique sous la courbe de f lorsque n tend vers l'infini. Lorsque f est à valeurs dans $\mathbb{C}$, on définit tout simplement l'intégrale de f comme le nombre complexe de partie réelle l'intégrale de la partie réelle de f et de partie imaginaire l'intégrale de la partie imaginaire de f. Le lecteur curieux est encouragé à montrer que la première définition est équivalente à la troisième.

Ces trois définitions d'intégrale s'étendent facilement aux fonctions continues par morceaux sur un segment, par la somme des intégrales du prolongement de f sur chaque morceau d'une subdivision finie du segment fermé où elle est définie. Cependant, ces définitions sont uniquement valables sur un intervalle fermé, et elles sont difficiles à appliquer pour des fonctions continues par morceaux sur un intervalle ouvert (ou semi-ouvert), surtout lorsque f est irrégulière aux bords, c'est-à-dire lorsqu'elle tend vers l'infini aux bords par exemple. On va donc définir cette intégrale d'une manière assez naturelle : il s'agira de la limite de l'intégrale de f, comme définie auparavant, sur un segment inclus dans l'intervalle où elle est définie, lorsque la taille ce segment tend vers celle de l'intervalle tout entier. Plus rigoureusement, on définit cette intégrale généralisée de la manière suivante. On verra que ceci généralise bien les définitions précédentes.

(cc) BY-NC-SA

Définition I.3.

Soit $a, b \in \mathbb{R}$ tels que $a < b$ et $f : [a, b[\longrightarrow \mathbb{C}$ une fonction continue par morceaux sur $[a, b[$. On considère la fonction

$$\mathrm{F} : \begin{cases} [a, b[& \longrightarrow \mathbb{C} \\ x & \longmapsto \displaystyle\int_a^x f(t)dt. \end{cases}$$

On dit que f admet une intégrale généralisée sur $[a, b[$ lorsque F admet une limite finie en b. Dans ce cas, on dit aussi que l'intégrale $\displaystyle\int_a^b f(t)dt$ converge, et par définition $\displaystyle\int_a^b f(t)dt = \lim_{x \to b^-} \mathrm{F}(x)$.

Remarques.

→ On définit cette intégrale généralisée de la même manière cette intégrale lorsque $b = +\infty$ ou sur l'intervalle $]a, b]$ (il suffit de permuter les rôles de a et b dans ce cas).

→ Bien entendu, les relations suivantes, qui sont vérifiées pour l'intégrale usuelle, le sont aussi pour cette nouvelle définition de l'intégrale : lorsque f admet une intégrale généralisée sur $[a, b[$ et $c \in [a, b[$,

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt \text{ et } \int_a^b f(t)dt = -\int_b^a f(t)dt.$$

Proposition I.4.

Soit $a, b \in \mathbb{R}$ tels que $a < b$ et $f : [a, b[\longrightarrow \mathbb{C}$ une fonction continue par morceaux sur $[a, b[$. Pour tout $a' \in [a, b[$, on a

$$\int_a^b f(t)dt \text{ converge} \Longleftrightarrow \int_{a'}^b f(t)dt \text{ converge.}$$

Démonstration. Il suffit de voir que les deux applications

$$\mathrm{F}_1 : \begin{cases} [a, b[& \longrightarrow \mathbb{C} \\ x & \longmapsto \displaystyle\int_a^x f(t)dt \end{cases} \text{ et } \mathrm{F}_2 : \begin{cases} [a, b[& \longrightarrow \mathbb{C} \\ x & \longmapsto \displaystyle\int_{a'}^x f(t)dt \end{cases}$$

diffèrent d'une constante. En effet, on a pour tout $x \in [a, b[$,

$$\mathrm{F}_1(x) - \mathrm{F}_2(x) = \int_a^x f(t)dt - \int_{a'}^x f(t)dt = \underbrace{\int_a^{a'} f(t)dt}_{\text{constante}}.$$

La suite de la preuve est immédiate. □

Remarques.

→ En fait, la proposition ci-dessus montre que l'existence de l'intégrale de f sur $[a, b[$ dépend uniquement du comportement de f au voisinage de b. On dira donc quelques fois par abus de langage que f admet une intégrale généralisée en b au lieu de "f admet une intégrale généralisée sur $[a, b[$".

→ Lorsque f est définie et continue par morceaux sur $]a, b[$ au lieu de $[a, b[$, on peut facilement étendre la définition I.3 en disant que f admet une intégrale généralisée sur $]a, b[$ lorsqu'il existe $c \in]a, b[$ tel que f admet une intégrale généralisée au sens de la définition I.3 sur $]a, c]$ et $[c, b[$ et alors dans ce cas, l'intégrale de f sur $]a, b[$ est la somme de ses intégrales sur $]a, c]$ et $[c, b[$.

(cc) BY-NC-SA

→ En première année, on voit que pour une fonction continue sur un intervalle $[a,b]$, pour tout $c \in [a,b]$, l'application définie sur $[a,b]$, $x \longmapsto \int_c^x f(t)dt$ est la primitive de f qui s'annule en c. Cette propriété reste la même lorsque c est remplacé par $+\infty$ ou $-\infty$: par exemple, lorsque f est continue par morceaux et admet une intégrale généralisée sur $[a,+\infty[$, alors $x \longmapsto \int_{+\infty}^x f(t)dt$ est la primitive de f qui tend vers 0 en $+\infty$.

Proposition I.5.

La somme et la différence de deux fonctions qui admettent une intégrale généralisée sur un intervalle admettent une intégrale généralisée sur ce même intervalle.

Démonstration. La preuve de cette propriété ne présente pas de difficulté et est laissée comme exercice au lecteur. □

La définition suivante est propre à ce cours et n'existe très probablement pas ailleurs. Nous l'utilisons pour expliquer rigoureusement quelques éléments du cours.

Définition I.6.

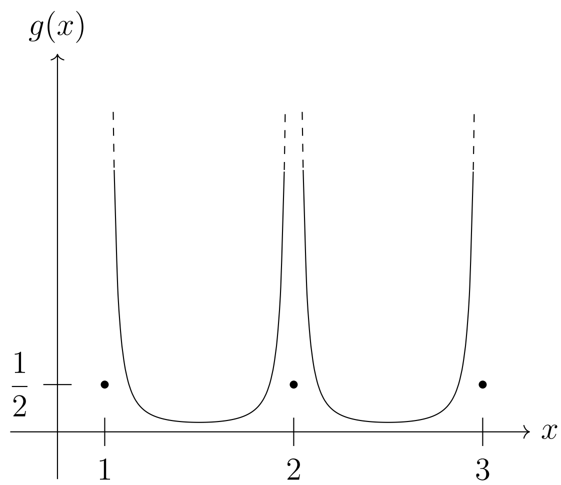
Soit $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in \mathbb{R} \cup \{+\infty\}$ tels que $a < b$, I un intervalle de bords a et b (ouvert, fermé, ou semi-ouvert) et soit f une fonction de I dans $\mathbb{C}$. On dit que f est presque continue par morceaux sur $]a,b[$, lorsqu'il existe $n \in \mathbb{N}^*$ et $a = s_1 < s_2 < \cdots < s_n = b$ telle que pour tout $i \in [\![1; n-1]\!]$, f est continue par morceaux sur $]s_i, s_{i+1}[$.

Remarque. Il est naturel de se poser la question suivante : pourquoi est-ce que par exemple, une fonction presque continue par morceaux n'est pas nécessairement continue par morceaux ? Pour voir que ce n'est pas la cas, il suffit de remarquer qu'une fonction continue par morceaux sur un intervalle fermé borné est bornée, alors qu'une fonction presque continue par morceaux ne l'est pas forcément.

Par exemple, la fonction dans la figure de droite, définie par

$$g : \begin{cases} [1,3] & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} \dfrac{1}{(x-1)^2(x-2)^2} & \text{si } x \in]1,2[\\ \dfrac{1}{(x-2)^2(x-3)^2} & \text{si } x \in]2,3[\\ \dfrac{1}{2} & \text{sinon} \end{cases} \end{cases}$$

est presque continue par morceaux sur $[1,3]$, mais n'est pas continue par morceaux.



En effet, il est facile de voir que g est continue par morceaux sur $]1,2[$ et $]2,3[$ et donc presque continue par morceaux sur $[1,3]$, mais n'est pas continue par morceaux, car n'est pas bornée sur $[1,3]$.

La définition suivante explique pourquoi nous avons besoin de la définition ci-dessus.

(cc) BY-NC-SA

Définition I.7.

Soit I un intervalle de $\mathbb{R}$ de bords $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in \mathbb{R} \cup \{+\infty\}$ tels que $a < b$ et f une fonction de I dans $\mathbb{C}$ presque continue par morceaux. Soit $n \in \mathbb{N}^*$ et $s_1, \dots, s_n \in \mathbb{R} \cup \{+\infty, -\infty\}$ tels que $a = s_1 < s_2 < \cdots < s_n = b$ et pour tout $i \in [\![1; n-1]\!]$, f est continue par morceaux sur $]s_i, s_{i+1}[$. On dit que f admet une intégrale généralisée sur I, lorsque f admet une intégrale généralisée sur $]s_i, s_{i+1}[$ au sens de la définition I.3 pour tout $i \in [\![1; n-1]\!]$. Dans ce cas, on note

$$\int_a^b f(t)dt = \sum_{i=1}^{n-1} \int_{s_i}^{s_{i+1}} f(t)dt.$$

Remarques.

→ La définition ci-dessus est en général utilisée implicitement pour les intégrales qui ne sont pas continues par morceaux, mais presque continues par morceaux (en général, cette propriété de presque continuité par morceaux n'est pas citée). Par exemple, lorsqu'on demande d'étudier l'intégrale de la fonction $x \longmapsto \frac{1}{\sqrt{|x-1||x-2||x-3|}}$ sur $[0, 4]$, on demande en fait tout simplement d'étudier les intégrales de cette fonction sur les intervalles où f est continue par morceaux, qui sont ici $]0, 1[,]1, 2[,]3, 4[$. Nous verrons dans une partie ultérieure de ce chapitre comment faire cette étude en pratique.

→ Lorsqu'on veut étudier l'intégrale d'une fonction continue par morceaux sur un intervalle de la forme $[a, b[$ ou $]a, b]$, il suffit d'appliquer directement la définition I.3, c'est-à-dire étudier la convergence de la fonction $x \longmapsto \int_a^x f(t)dt$ en b à gauche ou de $x \longmapsto \int_x^b f(t)dt$ en a à droite. Lorsqu'on veut faire la même chose sur un intervalle ouvert de la forme $]a, b[$ où f est continue par morceaux, il suffit de le faire **séparément** pour $]a, c]$ et $[c, b[$ où c est un élément arbitraire de $]a, b[$. On pourra alors éviter des erreurs comme celle-ci : la fonction tangente a une intégrale sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ égale a 0 car elle est impaire, ce qui est complètement faux car cette intégrale n'est de toute manière pas définie au sens de la définition I.3. En effet, l'application $x \longmapsto \int_0^x \tan t \, dt$ diverge vers l'infini lorsque x tend vers $\frac{\pi}{2}$.

L'exercice suivant a pour but d'aider le lecteur à se familiariser avec les définitions qu'on vient de voir.

Exercice I.8.

Est-ce que les fonctions suivantes admettent une intégrale généralisée ?

1. $x \longmapsto \frac{1}{x}$ sur $]0, 1]$.
2. $x \longmapsto \frac{1}{x^\alpha}$, $\alpha \in]1, +\infty[$ sur $[1, +\infty[$.
3. $x \longmapsto \tan x$ sur $\left[0, \frac{\pi}{2}\right[$.

II Cas des fonctions positives

1. Propriétés générales

On va maintenant appliquer cette généralisation de l'intégrale pour les fonctions positives.

Encore une fois, la proposition ci-dessous, lorsqu'on permute a et b, et qu'on remplace l'intervalle $[a, b[$

(cc) BY-NC-SA

par $]a, b]$, reste valable.

Proposition II.1.

Soit f une fonction continue par morceaux sur $[a, b[$, où $a < b$, et $b \in \mathbb{R} \cup \{+\infty\}$. Si f est positive sur $[a, b[$, alors les propositions suivantes sont équivalentes.

1. f admet une intégrale généralisée sur $[a, b[$.
2. L'ensemble $\mathrm{A} = \left\{ \int_c^d f(t)dt, c, d \in \mathbb{R},\ a \leq c < d < b \right\}$ est borné.
3. L'application $x \longmapsto \int_a^x f(t)dt$ est bornée sur $[a, b[$.

Démonstration.

→ (1) ⇔ (3) Si f est positive, alors l'application $\mathrm{F} : x \longmapsto \int_a^x f(t)dt$ est croissante sur $[a, b[$. On en déduit donc que F admet une limite à gauche de b si et seulement si F est majorée sur $[a, b[$.

→ (2) ⇒ (3) Si l'ensemble A est borné, alors l'ensemble $\left\{ \int_a^x f(t)dt, x \in [a, b[\right\}$ est borné, ce qui est exactement la proposition (3).

→ (3) ⇒ (2) Soit $\mathrm{M} > 0$ un majorant de l'application $\mathrm{F} : x \longmapsto \int_a^x f(t)dt$ sur $[a, b[$. On a alors pour tout $c, d \in \mathbb{R}$ tels que $a \leq c < d < b$,

$$0 \underset{\substack{\uparrow \\ f \text{ positive}}}{\leq} \int_c^d f(t)dt \underset{\substack{\uparrow \\ f \text{ positive}}}{\leq} \int_a^d f(t)dt = \mathrm{F}(d) \leq \mathrm{M},$$

ce qui donne directement que A est majoré par M.

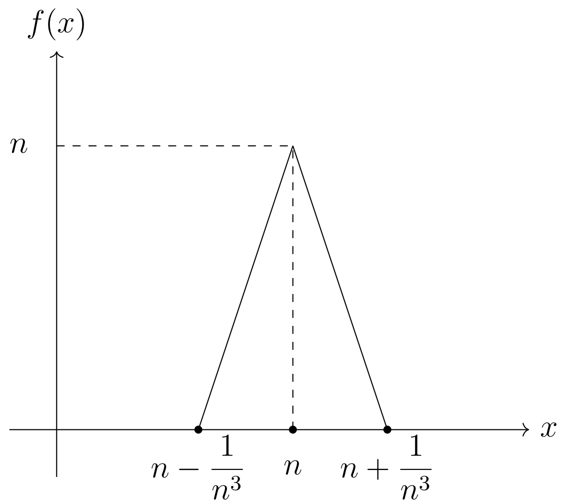
□

Remarque. Attention, même lorsque $b = +\infty$, il est possible que f admette une intégrale généralisée sur $[a, +\infty[$ mais ne soit pas bornée (bien que son intégrale le soit). On peut le voir via l'exemple suivant. Considérons la fonction

$$f : \begin{cases} \mathbb{R}^+ & \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} n^4(x-n)+n & \text{si } x \in \left[n - \frac{1}{n^3}, n\right] \text{ pour tout } n \in \mathbb{N}^*, \\ n - n^4(x-n) & \text{si } x \in \left]n, n + \frac{1}{n^3}\right] \text{ pour tout } n \in \mathbb{N}^*, \\ 0 & \text{sinon.} \end{cases} \end{cases}$$

Comme on peut le voir sur la figure à droite, il est facile de calculer l'intégrale de cette fonction sur un segment, et il est également facile de voir que cette intégrale est bornée. En effet, la courbe de f est une succession de triangles comme celui sur la figure à droite, et donc pour tout $n \in \mathbb{N}^*$, l'intégrale de f de 0 à x pour tout $x \in \left[n + \frac{1}{n^3}, n + 1 - \frac{1}{(n+1)^3}\right]$ est égale à la somme des surfaces de ces triangles de rang inférieur ou égal à n, c'est-à-dire

$$\int_0^x f(t)dt = \sum_{k=1}^{n} \frac{1}{2}k \times \frac{2}{k^3} = \sum_{k=1}^{n} \frac{1}{k^2} \leq \sum_{k=1}^{+\infty} \frac{1}{k^2}.$$



(cc) BY-NC-SA

On conclut alors que l'application $x \longmapsto \int_0^x f(t)dt$ est bornée sur $[0,+\infty[$ par $\sum_{k=1}^{+\infty} \frac{1}{k^2}$ (qui est une série convergente), et que donc f admet bien une intégrale généralisée sur $[0,+\infty[$ bien qu'elle ne soit pas bornée.

On énonce maintenant quelques propriétés très utiles pour l'étude des intégrales en pratique.

Proposition (Comparaison d'intégrales) II.2.

Soit $a, b \in \mathbb{R}$ tels que $a < b$ et f et g deux fonctions positives continues par morceaux sur $[a, b[$. Les propositions suivantes sont vraies.

1. Si pour tout $x \in [a, b[,\ f(x) \leq g(x)$, alors g admet une intégrale généralisée sur $[a, b[$ implique que f admet une intégrale généralisée sur $[a, b[$.
2. Si $f(x) \underset{x \to b}{=} \mathrm{O}(g(x))$, alors g admet une intégrale généralisée sur $[a, b[$ implique que f admet une intégrale généralisée sur $[a, b[$.
3. Si $f \underset{x \to b}{\sim} g$, alors g admet une intégrale généralisée sur $[a, b[$ si et seulement si f admet une intégrale généralisée sur $[a, b[$.

Démonstration. On considère les deux fonctions suivantes

$$\mathrm{F} : \begin{cases} [a, b[& \longrightarrow \mathbb{R} \\ x & \longmapsto \int_a^x f(t)dt \end{cases} \quad \text{et } \mathrm{G} : \begin{cases} [a, b[& \longrightarrow \mathbb{R} \\ x & \longmapsto \int_a^x g(t)dt. \end{cases}$$

1. Supposons que pour tout $x \in [a, b[,\ f(x) \leq g(x)$ et que g admet une intégrale généralisée sur $[a, b[$. On a pour tout $x \in [a, b[,\ \mathrm{F}(x) \leq \mathrm{G}(x)$. La proposition II.1 nous permet d'affirmer qu'étant donné que g est positive et admet une intégrale généralisée sur $[a, b[$, alors G est majorée sur $[a, b[$, et donc F aussi, ce qui nous permet de dire que f, étant positive, admet bien une intégrale généralisée sur $[a, b[$.
2. Supposons que $f(x) \underset{x \to b}{=} \mathrm{O}(g(x))$ et que g admet une intégrale généralisée sur $[a, b[$. Soit $\mathrm{M} \geq 0$ un majorant de G. Il existe alors $\varepsilon > 0$ et $\mathrm{C} \geq 0$ tels que pour tout $x \in [b-\varepsilon, b[,\ f(x) \leq \mathrm{C}g(x)$. On a alors pour tout $x \in [a, b[$,
   - $\rightarrow$ Si $x \leq b - \varepsilon$, alors $\mathrm{F}(x) = \int_a^x f(t)dt \leq \int_a^{b-\varepsilon} f(t)dt$.
   - $\rightarrow$ Si $x \in]b-\varepsilon, b[$, alors
   $$\mathrm{F}(x) = \int_a^x f(t)dt = \int_a^{b-\varepsilon} f(t)dt + \int_{b-\varepsilon}^x f(t)dt \leq \int_a^{b-\varepsilon} f(t)dt + \mathrm{C}\int_{b-\varepsilon}^x g(t)dt \leq \int_a^{b-\varepsilon} f(t)dt + \mathrm{CM}.$$

   On en déduit alors que pour tout $x \in [a, b[$,
   $$0 \leq \mathrm{F}(x) \leq \max\left(\int_a^{b-\varepsilon} f(t)dt + \mathrm{CM}, \int_a^{b-\varepsilon} f(t)dt\right) = \int_a^{b-\varepsilon} f(t)dt + \mathrm{CM}.$$

   Ceci signifie que F est majorée sur $[a, b[$, et que donc f admet bien une intégrale généralisée sur $[a, b[$.
3. Si $f \underset{x \to b}{\sim} g$, alors $f(x) \underset{x \to b}{=} \mathrm{O}(g(x))$ et $g(x) \underset{x \to b}{=} \mathrm{O}(f(x))$. On peut alors utiliser le point précédent pour avoir l'équivalence voulue.

□

Remarque. Bien entendu, cette propriété reste valable lorsqu'on remplace b par $+\infty$ ou lorsqu'on étudie l'intégrale sur un intervalle de la forme $]-\infty, b]$.

(cc) BY-NC-SA

Notons aussi que la positivité est très importante, on reviendra sur ce point ci-dessous en exhibant un contre-exemple.

2. Critère de Cauchy

La propriété suivante est quelques fois utile en exercices.

Théorème (Critère de Cauchy) II.3.

Soit $a, b \in \mathbb{R}$ tels que $a < b$, et f une fonction de $[a, b[$ dans $\mathbb{C}$ continue par morceaux. f admet une intégrale généralisée sur $[a, b[$ si et seulement si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x, y \in [a, b[,\ |x - b| < \eta \text{ et } |y - b| < \eta \Longrightarrow \left|\int_x^y f(t)dt\right| \leq \varepsilon.$$

Démonstration. Pour montrer cette équivalence, il suffit d'appliquer la proposition .5 à la fonction $\mathrm{F} : x \longmapsto \displaystyle\int_a^x f(t)dt$. □

Remarque.

→ Bien entendu, cette proposition est toujours valable lorsqu'on permute les rôles de a et b et qu'on remplace $[a, b[$ par $]a, b]$, ou si on remplace b par $+\infty$ ou a par $-\infty$. Par exemple, lorsqu'on remplace b par $+\infty$, la proposition devient : Si f est continue par morceaux sur $[a, +\infty[$, alors f admet une intégrale généralisée sur $[a, +\infty[$ si et seulement si

$$\forall \varepsilon > 0,\ \exists \mathrm{M} > 0,\ \forall x, y \in [a, b[,\ x > \mathrm{M} \text{ et } y > \mathrm{M} \Longrightarrow \left|\int_x^y f(t)dt\right| \leq \varepsilon.$$

→ Ce théorème peut être interprété de la manière suivante : une fonction f continue par morceaux admet une intégrale généralisée sur un intervalle $[a, b[$ si et seulement l'ensemble des valeurs possibles de l'intégrale de f sur un intervalle inclus dans $[a, b[$ deviennent de plus en plus proches de 0 plus les bornes de cet intervalle sont proches de b.

Exercice II.4.

Soit f une fonction continue par morceaux sur $\mathbb{R}$ à valeurs dans $\mathbb{C}$. On considère l'application de $\mathbb{R}^2$ dans $\mathbb{C}$, $\mathrm{G} : (x, y) \longmapsto \displaystyle\int_x^y f(t)dt$. Montrer que les deux propositions suivantes sont équivalentes.

1. f admet une intégrale généralisée sur $\mathbb{R}$.
2. $\mathrm{G}(x, y)$ admet une limite finie lorsque x tend vers $-\infty$ et y vers $+\infty$, *i.e.*

$$\exists \ell \in \mathbb{C},\ \forall \varepsilon > 0,\ \exists \mathrm{A} < 0,\ \exists \mathrm{B} > 0,\ \forall x, y \in \mathbb{R},\ x < \mathrm{A} \text{ et } y > \mathrm{B} \Longrightarrow |\mathrm{G}(x, y) - \ell| < \varepsilon.$$

III Fonctions intégrables

On définit à présent une propriété plus forte que la convergence de l'intégrale d'une fonction qui permet de faire des manipulations plus flexibles lors de l'étude d'une intégrale.

BY-NC-SA

Définition III.1.

Soit I un intervalle de $\mathbb{R}$ et f une fonction continue par morceaux de I dans $\mathbb{C}$. On dit que f est intégrable sur I lorsque $|f|$ admet une intégrale généralisée sur I ou d'une manière équivalente, l'ensemble

$$\mathrm{A} = \left\{ \int_c^d |f(t)|\, dt,\ c, d \in \mathrm{I} \right\}$$

est borné.

On énonce maintenant le fait que le fait d'être intégrable est plus fort que le fait d'admettre une intégrale généralisée.

Proposition III.2.

Si f est une fonction d'un intervalle I dans $\mathbb{R}$ continue par morceaux. Si f est intégrable sur I, alors elle admet une intégrale généralisée sur I.

Démonstration. Supposons que f est intégrable. On pose pour tout $x \in \mathbb{R}$, $f^+(x) = \max(0, f(x))$ et $f^-(x) = \max(0, -f(x))$ les parties positives et négatives de f respectivement. f^+ et f^- sont toutes les deux positives, et on a $|f| = f^+ + f^-$. On a alors pour tout $x \in \mathrm{I}$, $0 \leq f^+(x) \leq |f(x)|$ et $0 \leq f^-(x) \leq |f(x)|$. On peut alors appliquer la proposition II.2 à f^+ et f^-, ce qui nous permet de dire que f^+ et f^- admettent toutes les deux une intégrale généralisée sur I. On en déduit donc que $f^+ - f^-$ admet également une intégrale généralisée sur I, en on conclut en remarquant que $f = f^+ - f^-$. $\square$

Remarques.

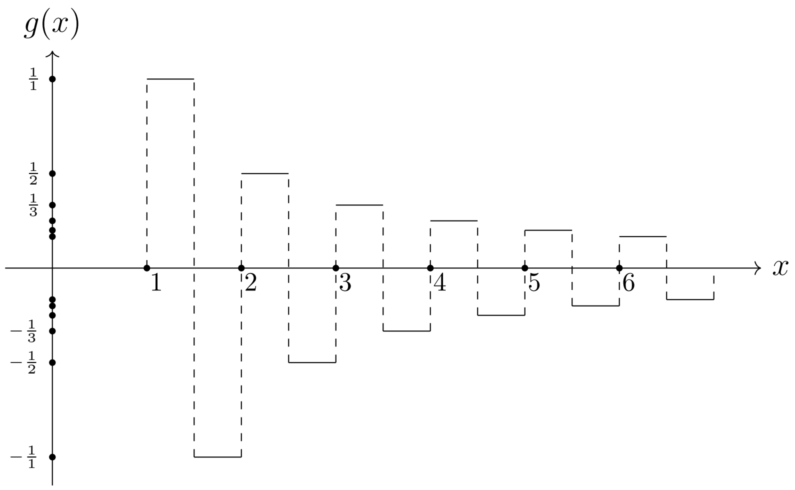
→ Ce résultat est aussi valable lorsque f est à valeurs dans $\mathbb{C}$. Pour le voir, il suffit de remarquer que la valeur absolue des parties réelles et imaginaires de f sont majorées par $|f|$, et qu'elles sont donc intégrables, ce qui implique d'après ce qui précède qu'elles admettent une intégrale généralisée sur I, et donc de même pour f.

→ Comme pour les intégrales généralisées, le fait qu'une fonction f continue par morceaux sur un intervalle $[a, b[$ soit intégrable dépend uniquement du comportement de f en b. On dira donc quelques fois par abus de langage que f est intégrale en b au lieu de "f est intégrable sur $[a, b[$".

Attention, une fonction qui admet une intégrale généralisée sur un intervalle n'est pas forcément intégrable. En effet, on peut facilement construire un contrexemple sur $[1, +\infty[$ (ou même sur un intervalle ouvert borné). Considérons la fonction

$$g : \begin{cases} [1, +\infty[& \longrightarrow \mathbb{R} \\ x & \longmapsto \begin{cases} \dfrac{1}{n} & \text{si } x \in \left[n, n + \dfrac{1}{2}\right[,\ \text{pour tout } n \in \mathbb{N}^*, \\ -\dfrac{1}{n} & \text{si } x \in \left[n + \dfrac{1}{2}, n + 1\right[,\ \text{pour tout } n \in \mathbb{N}^*. \end{cases} \end{cases}$$

BY-NC-SA



On a pour tout $x \geq 1$,

$$\begin{aligned}
\int_1^x g(t)dt &= \sum_{k=1}^{\lfloor x \rfloor -1} \left(\int_k^{k+\frac{1}{2}} g(t)dt + \int_{k+\frac{1}{2}}^{k+1} g(t)dt \right) + \int_{\lfloor x \rfloor}^x g(t)dt \\
&= \sum_{k=1}^{\lfloor x \rfloor -1} \frac{1}{2}\left(\frac{1}{k} - \frac{1}{k}\right) + \int_{\lfloor x \rfloor}^x g(t)dt = \int_{\lfloor x \rfloor}^x g(t)dt \\
&= \begin{cases} \dfrac{x - \lfloor x \rfloor}{\lfloor x \rfloor} & \text{si } x - \lfloor x \rfloor \leq \dfrac{1}{2} \\ \dfrac{1-(x - \lfloor x \rfloor)}{\lfloor x \rfloor} & \text{si } x - \lfloor x \rfloor > \dfrac{1}{2}. \end{cases}
\end{aligned}$$

Dans les deux cas ci-dessus, on peut facilement vérifier que

$$\underbrace{-\frac{1}{\lfloor x \rfloor}}_{\xrightarrow[x\to+\infty]{}0} \leq \int_1^x g(t)dt \leq \underbrace{\frac{1}{\lfloor x \rfloor}}_{\xrightarrow[x\to+\infty]{}0},$$

et donc g admet bien une intégrale généralisée sur $[1, +\infty[$. De plus, on a encore une fois pour tout $x \in [1, +\infty[$

$$\int_1^x |g(t)|\, dt = \sum_{k=1}^{\lfloor x \rfloor -1} \frac{1}{k} + \int_{\lfloor x \rfloor}^x |g(t)|\, dt \geq \sum_{k=1}^{\lfloor x \rfloor -1} \frac{1}{k} \xrightarrow[x\to+\infty]{} +\infty,$$

et alors g n'est pas intégrable sur $[1, +\infty[$.

Notation. Lorsque I est un intervalle, on note $\mathrm{L}^1(\mathrm{I}, \mathbb{C})$, ou $\mathrm{L}^1(\mathrm{I})$ lorsqu'il n'y a pas d'ambiguïté sur l'ensemble d'arrivée, l'ensemble des fonctions intégrables de I dans $\mathbb{C}$.

Proposition III.3.

Pour tout intervalle I de $\mathbb{R}$, $\mathrm{L}^1(\mathrm{I})$ est un $\mathbb{C}-$espace vectoriel. De plus, pour tout $f \in \mathrm{L}^1(\mathrm{I})$ et g continue par morceaux sur I, si g est bornée sur I, alors $f \times g \in \mathrm{L}^1(\mathrm{I})$.

Démonstration. Pour montrer que $\mathrm{L}^1(\mathrm{I})$ is un $\mathbb{C}-$espace vectoriel, il suffit de montrer qu'il s'agit d'un sous-espace vectoriel de l'ensemble des fonctions de I dans $\mathbb{C}$. Pour cela, il suffit de montrer que pour tout $f, g \in \mathrm{L}^1(\mathrm{I})$ et $\lambda \in \mathbb{C}$, $f + \lambda g \in \mathrm{L}^1(\mathrm{I})$. Ceci est facile à voir. En effet, l'intégrale de $|f + \lambda g|$ sur tout segment de I est bornée par $\int_{\mathrm{I}} |f(t)|\, dt + |\lambda| \int_{\mathrm{I}} |g(t)|\, dt$, et donc $|f + \lambda g|$ est bien intégrable.
Pour le deuxième point, il suffit de voir que $|f \times g|$ est majorée par $\|g\|_\infty \times f$ sur I, où $\|g\|_\infty$ est la borne

(cc) BY-NC-SA

supérieure de $|g|$ sur I, et que $\|g\|_\infty \times |f|$ est intégrable sur I. La proposition II.2 permet de conclure. □

Remarque. Si f est une fonction continue par morceaux et bornée sur I, et que I est borné, alors $f \in \mathrm{L}^1(\mathrm{I})$. Le lecteur ayant un doute est encouragé à le montrer lui même.

Exercice III.4.

Étudier la convergence de l'intégrale $\displaystyle\int_0^1 \frac{1}{t} - \left\lfloor \frac{1}{t} \right\rfloor dt$, et calculer sa valeur si elle converge.

IV Intégrales de référence

Nous allons présenter ici quelques intégrales connues qui permettent en général de prouver la convergence d'intégrales grâce à la proposition II.2.

Proposition (Intégrales de Riemann) IV.1.

Soit $\alpha \in \mathbb{R}$.

1. L'intégrale $\displaystyle\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge si et seulement si $\alpha \in]1, +\infty[$.
2. L'intégrale $\displaystyle\int_0^1 \frac{1}{t^\alpha} dt$ converge si et seulement si $\alpha \in]-\infty, 1[$.

Démonstration. Montrons les deux résultats.

1. Soit $x \geq 1$ et $\alpha \in \mathbb{R} \setminus \{1\}$. On a

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

   Ceci nous permet de dire que lorsque $\alpha \neq 1$, l'intégrale $\displaystyle\int_1^{+\infty} \frac{1}{t^\alpha} dt$ converge si et seulement si $\alpha \in]1, +\infty[$. Il reste à traiter le cas $\alpha = 1$, qui est assez facile. En effet, on a

$$\int_1^x \frac{1}{t} dt = \ln x \xrightarrow[x \to +\infty]{} +\infty.$$

   On a donc bien le résultat voulu.

2. La preuve est quasi identique que celle du point précédent. Soit $\alpha \neq 1$. On a pour tout $x \in]0, 1]$,

$$\int_x^1 \frac{1}{t^\alpha} dt = \frac{1}{1-\alpha} - \frac{1}{1-\alpha} x^{1-\alpha} \xrightarrow[x \to 0^+]{} \begin{cases} +\infty & \text{si } \alpha > 1 \\ \dfrac{1}{1-\alpha} & \text{si } \alpha < 1. \end{cases}$$

   On traite enfin le cas $\alpha = 1$. Pour tout $x \in]0, 1]$, on a

$$\int_x^1 \frac{1}{t} = -\ln x \xrightarrow[x \to 0^+]{} +\infty,$$

   et alors on a bien le résultat voulu.

□

En fait, le résultat ci-dessus nous donne un moyen pratique (qui fonctionne souvent, mais pas toujours) qui permet souvent d'étudier la convergence d'intégrales.

(cc) BY-NC-SA

Proposition IV.2.

Soit f une fonction continue par morceaux positive sur $\mathbb{R}$. Les propriétés suivantes sont vérifiées.

1. S'il existe $\gamma < 1$ tel que $t \longmapsto t^\gamma f(t)$ est majorée sur $]0,1]$, alors f est intégrable sur $]0,1]$. De même, s'il existe $\gamma \geq 1$ tel que $t \longmapsto t^\gamma f(t)$ est minorée par un nombre strictement positif sur $]0,1]$, alors f n'est pas intégrable sur $]0,1]$.
2. S'il existe $\gamma > 1$ tel que $t \longmapsto t^\gamma f(t)$ est majorée sur $[1,+\infty[$, alors f est intégrable sur $[1,+\infty[$. De même, s'il existe $\gamma \leq 1$ tel que $t \longmapsto t^\gamma f(t)$ est minorée par un nombre strictement positif sur $[1,+\infty[$, alors f n'est pas intégrable sur $[1,+\infty[$.

Bien entendu, cette proposition reste valable même lorsqu'on remplace la borne fermée des intervalles (ici égale à 1) par n'importe quel réel strictement positif.

Démonstration.

1. Supposons qu'il existe $\gamma < 1$ tel que $t \longmapsto t^\gamma f(t)$ est majorée sur $]0,1]$, soit M un majorant de cette fonction. On a alors pour tout $t \in]0,1]$, $t^\gamma f(t) \leq \mathrm{M}$, *i.e.* $f(t) \leq \dfrac{\mathrm{M}}{t^\gamma}$, et alors $f(t) = \mathrm{O}\left(\dfrac{1}{t^\gamma}\right)$. $t \longmapsto \dfrac{1}{t^\gamma}$ est intégrable sur $]0,1]$ d'après la proposition IV.1, et donc d'après la proposition II.2, on peut dire que f est intégrable sur $]0,1]$.

   Supposons maintenant qu'il existe $\gamma \geq 1$ tel que $t \longmapsto t^\gamma f(t)$ est minorée sur $[1,+\infty[$ par un réel strictement positif qu'on note m. On a alors pour tout $t \in [1,+\infty[$, $t^\gamma f(t) \geq m$, *i.e.* $\dfrac{1}{t^\gamma} \leq \dfrac{1}{m} f(t)$, et alors $\dfrac{1}{t^\gamma} = \mathrm{O}(f(t))$. On en déduit que f ne peut pas être intégrable, car si elle l'était, $t \longmapsto \dfrac{1}{t^\gamma}$ le serait aussi d'après la proposition II.2, ce qui est impossible d'après la proposition IV.1.
2. La preuve du second point est quasi identique à celle du premier point, mais nous la faisons quand même dans le but de mettre le lecteur à l'aise avec ce type de raisonnement. Le lecteur à l'aise est libre de sauter cette preuve. Supposons qu'il existe $\gamma > 1$ tel que $t \longmapsto t^\gamma$ est majorée sur $[1,+\infty[$, et notons M un majorant de cette fonction. On a alors pour tout $t \in [1,+\infty[$, $f(t) \leq \dfrac{\mathrm{M}}{t^\gamma}$, et donc $f(t) = \mathrm{O}\left(\dfrac{1}{t^\gamma}\right)$, ce qui nous permet de dire que f est intégrable sur $[1,+\infty[$ d'après les propositions II.2 et IV.1.

   Supposons maintenant qu'il existe $\gamma \leq 1$ tel que $t \longmapsto t^\gamma f(t)$ est minorée par un nombre réel strictement positif qu'on notera m. On a alors pour tout $t \in [1,+\infty[$, $t^\gamma f(t) \geq m$, et alors $\dfrac{1}{t^\gamma} = \mathrm{O}(f(t))$, ce qui nous donne directement que f n'est pas intégrable sur $[1,+\infty[$, car si ce n'était pas le cas, $t \longmapsto \dfrac{1}{t^\gamma}$ serait intégrable d'après la proposition II.2, ce qui est absurde d'après la proposition IV.1.

$\square$

Exercice IV.3.

Étudier la convergence de l'intégrale $\displaystyle\int_0^{\frac{1}{2}} |\ln t|^\beta \, dt$ pour tout $\beta \in \mathbb{R}$.

Exercice IV.4.

Pour $\alpha, \beta \in \mathbb{R}$, étudier la convergence de l'intégrale $\displaystyle\int_2^{+\infty} \frac{1}{t^\alpha (\ln t)^\beta} dt$.

(cc) BY-NC-SA

Exercice IV.5.

Soit $\lambda \in \mathbb{C}$ un nombre complexe de partie réelle strictement positive, P un polynôme réel non constant, et f une fonction de $\mathbb{R}^+$ dans $\mathbb{R}$ telle que $f(t) \underset{t \to +\infty}{=} \mathrm{O}(\mathrm{P}(t))$. Soit $\alpha > 0$. Étudier la convergence de l'intégrale $\int_0^{+\infty} e^{-\lambda t^\alpha} f(t) dt$.

Exercice IV.6.

Étudier pour $\alpha \in \mathbb{R}$ et $a \in \mathbb{R}$ la convergence de l'intégrale $\int_a^{a+1} \frac{1}{|t-a|^\alpha} dt$.

Exercice IV.7.

Étudier la convergence des intégrales suivantes.

1. $\int_1^{+\infty} \frac{1}{(\ln t)^{\ln t}} dt$
2. $\int_0^{+\infty} \sin \frac{1}{t^2} dt$
3. $\int_0^{+\infty} \arctan(1+t^3) - \arctan(t^3) dt$
4. $\int_0^{+\infty} \frac{1}{\sqrt{x\,|x-1|\,|x-2|}} dt$

Exercice IV.8.

Soit P un polynôme réel de degré 3. Etudier la convergence de l'intégrale $\int_{-\infty}^{+\infty} \frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} dt$.

Exercice IV.9.

Soit $a \in \mathbb{R}$. Étudier la convergence de l'intégrale $\int_0^1 \frac{t^a - 1}{\ln t} dt$.

V Intégration par parties

On peut utiliser la formule d'intégration par parties pour déduire des propriétés de convergences d'intégrales. On commence par rappeler la formule.

Rappel (Intégration par parties) V.1.

Soit $a, b \in \mathbb{R}$ tels que $a < b$ et f, g deux fonctions de $[a, b]$ dans $\mathbb{C}$ continument dérivables. L'égalité suivante est vérifiée.

$$\int_a^b f'(t)g(t) = [f(t)g(t)]_a^b - \int_a^b f(t)g'(t)dt$$

BY-NC-SA

Démonstration. Il suffit de mettre les deux intégrales du même côté de l'égalité et d'utiliser la formule de la dérivée d'un produit de deux fonctions $(fg)' = f'g + g'f$. ☐

On énonce maintenant la propriété que nous permet cette formule de déduire.

Proposition V.2.

Soit $a \in \mathbb{R}$ et $b \in \mathbb{R} \cup \{+\infty\}$, tels que $a < b$. Soit f, g deux fonctions de $[a, b[$ dans $\mathbb{C}$ continument dérivables. Si $t \longmapsto f(t)g(t)$ admet une limite finie en b, alors l'intégrale $\int_a^b f'(t)g(t)$ converge si et seulement si l'intégrale $\int_a^b f(t)g'(t)dt$ converge.

Démonstration. Par symétrie des rôles de f et g, il suffit de montrer l'implication dans un seul sens. Montrons l'implication de droite à gauche. Supposons que $t \longmapsto f(t)g(t)$ admet une limite finie en b. La formule d'intégration par partie nous donne pour tout $x \geq a$,

$$\int_a^x f'(t)g(t)dt = [f(t)g(t)]_a^x - \int_a^x f(t)g'(t)dt.$$

Le premier terme du côté droit de l'égalité admet une limite finie, on peut alors clairement voir que si le second admet une limite finie en b, alors le côté gauche aussi, ce qui permet de conclure. ☐

Bien entendu, encore une fois, cette proposition reste valable lorsqu'on remplace $[a, b[$ par $]a, b]$ et qu'on permute les rôles de a et b (et aussi lorsque a est remplacé par $-\infty$).

Exercice V.3.

Étudier la convergence de l'intégrale $\int_0^{+\infty} \frac{\sin t}{t} dt$.

Exercice V.4.

Montrer la convergence et calculer la valeur de $\int_0^{+\infty} \frac{(\sin t)^2}{t^2} dt$. *Indice : On pourra admettre que* $\int_0^{+\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}$.

Exercice V.5.

Soit $\alpha \in \mathbb{R}$. Étudier la convergence de l'intégrale $\int_1^{+\infty} \frac{e^{it}}{t^\alpha} dt$.

Exercice V.6.

Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$ continue et périodique de période strictement positive. Montrer que $\int_1^{+\infty} \frac{f(t)}{t} dt$ converge si et seulement si f admet une primitive périodique.

VI Découpe

La proposition suivante est assez utile en exercices.

(cc) BY-NC-SA

Proposition VI.1.

Soit f une fonction de $\mathbb{R}^+$ dans $\mathbb{C}$ continue par morceaux, et soit (a_n) une suite croissante à valeurs positives qui diverge vers $+\infty$. L'intégrale $\int_0^{+\infty} |f(t)|\,dt$ converge si et seulement si la série $\sum_{n=0}^{+\infty}\int_{a_n}^{a_{n+1}} |f(t)|\,dt$ converge.

Démonstration.

→ $(\Rightarrow)$ Si $\int_0^{+\infty} |f(t)|\,dt$, alors par définition, pour tout $\mathrm{N} \in \mathbb{N}^*$,

$$\sum_{n=0}^{\mathrm{N}}\int_{a_n}^{a_{n+1}} |f(t)|\,dt = \int_{a_0}^{a_{\mathrm{N}+1}} |f(t)|\,dt \leq \int_0^{+\infty} |f(t)|\,dt.$$

La série $\sum_{n=0}^{+\infty}\int_{a_n}^{a_{n+1}} |f(t)|\,dt$ à termes positifs est majorée, elle est donc convergente.

→ $(\Leftarrow)$ Supposons que la série $\sum_{n=0}^{+\infty}\int_{a_n}^{a_{n+1}} |f(t)|\,dt$ est convergente. Soit $x \in \mathbb{R}^+$. Soit $\mathrm{N} \in \mathbb{N}^*$ tel que $a_{\mathrm{N}+1} \geq x$. On a

$$\int_0^{x} |f(t)|\,dt \leq \int_0^{a_{\mathrm{N}+1}} |f(t)|\,dt \leq \int_0^{a_0} |f(t)|\,dt + \sum_{n=0}^{+\infty}\int_{a_n}^{a_{n+1}} |f(t)|\,dt.$$

$x \longmapsto \int_0^{x} |f(t)|\,dt$ est croissante majorée, elle est alors convergence.

□

Remarque. Attention, le fait qu'on considère l'intégrale du module de f plutôt que l'intégrale de f est nécessaire pour que cette propriété soit vraie.
Exemple. Comme promis, on revient pour exhiber un contre-exemple à II.2 quand la condition de positivité n'est pas vérifiée. Prenons $x > 1$ et posons sur $[1, +\infty[$ les applications

$$f : t \longmapsto \frac{\sin(t)}{\sqrt{t}} \text{ et } g : t \longmapsto f(t)(1 + f(t)).$$

Remarquons que $t \longmapsto \frac{-\cos(t)}{\sqrt{t}}$ admet une limite finie en $+\infty$ et que $\int_1^{\infty} \frac{\cos(t)}{2t\sqrt{t}}$ converge bien et est même intégrable, étant donnée que l'intégrande est (continue et) dominée, en valeure absolue, par $t \longmapsto \frac{1}{2t\sqrt{t}}$, qui est bien intégrable.
On déduit donc d'après V.2 que $\int_1^{+\infty} f(t)dt$ converge bien. Cependant, même si $f \underset{+\infty}{\sim} g$, $\int_1^{+\infty} g(t)dt$ ne converge pas. En effet, supposant par absurde l'inverse, on a alors que $\int_1^{+\infty} f^2(t)dt$ converge.
Posons alors pour tout $n \in \mathbb{N}^*$, $a_n = \pi n$, on a d'après la proposition VI.1 et par positivité de f^2 que

(cc) BY-NC-SA

$\displaystyle\sum_{n=1}^{+\infty}\int_{a_n}^{a_{n+1}} f^2(t)dt$ converge. Or, toujours par positivité de f^2, remarquons que pour tout $\mathrm{N} \geq 1$ on a

$$\begin{aligned}
\sum_{n=1}^{+\infty}\int_{a_n}^{a_{n+1}} f^2(t)dt &\geq \sum_{n=1}^{+\infty}\int_{\pi n+\frac{\pi}{4}}^{\pi n+\frac{3\pi}{4}} f^2(t)dt \geq \sum_{n=1}^{\mathrm{N}}\int_{\pi n+\frac{\pi}{4}}^{\pi n+\frac{3\pi}{4}} f^2(t)dt \geq \sum_{n=1}^{\mathrm{N}}\int_{\pi n+\frac{\pi}{4}}^{\pi n+\frac{3\pi}{4}} \frac{1}{2t}dt \\
&\geq \frac{1}{4}\sum_{n=1}^{\mathrm{N}}\int_{\pi n+\frac{\pi}{4}}^{\pi n+\frac{3\pi}{4}} \left(\frac{1}{t}+\frac{1}{t}\right)dt \geq \frac{1}{4}\sum_{n=1}^{\mathrm{N}}\int_{\pi n+\frac{\pi}{4}}^{\pi n+\frac{3\pi}{4}} \left(\frac{1}{t}+\frac{1}{t+\frac{\pi}{2}}\right)dt \\
&\geq \frac{1}{4}\sum_{n=1}^{\mathrm{N}}\int_{\pi n+\frac{\pi}{4}}^{\pi(n+1)+\frac{\pi}{4}} \frac{1}{t}dt \geq \frac{1}{4}\int_{\pi+\frac{\pi}{4}}^{\pi(\mathrm{N}+1)+\frac{\pi}{4}} \frac{1}{t}dt \\
&\geq \frac{1}{4}\ln\left(\pi(\mathrm{N}+1)+\frac{\pi}{4}\right) - \frac{1}{4}\ln\left(\frac{5\pi}{4}\right) \xrightarrow[\mathrm{N}\to+\infty]{} +\infty
\end{aligned}$$

ce qui est absurde.

Exercice VI.2.

Montrer la convergence et calculer la valeur de l'intégrale $\displaystyle\int_1^{+\infty}\frac{t-\lfloor t\rfloor}{t^2}dt$.

Exercice VI.3.

Soit f une fonction positive décroissante et continue par morceaux de $\mathbb{R}^+$ dans $\mathbb{R}$. Montrer que $\displaystyle\int_0^{+\infty} f(t)(\sin t)^2 dt$ converge si et seulement si $\displaystyle\int_0^{+\infty} f(t)dt$ converge.

Exercice VI.4.

Étudier la convergence de l'intégrale $\displaystyle\int_0^{+\infty} e^{-t^2|\sin t|}dt$.

VII Comportement à l'infini

Dans cette partie, on fait le lien entre le comportement d'une fonction à l'infini et la convergence de son intégrale.

Proposition VII.1.

Soit f une fonction de $\mathbb{R}^+$ dans $\mathbb{R}$ continue par morceaux telle que $\displaystyle\int_0^{+\infty} f(t)dt$ converge. S'il existe $\ell \in \mathbb{R}$ tel que $f(x) \xrightarrow[x\to+\infty]{} \ell$, alors $\ell = 0$.

Démonstration. Supposons que $\displaystyle\int_0^{+\infty} f(t)dt$ converge et qu'il existe $\ell \in \mathbb{R}$ tel que $f(x) \xrightarrow[x\to+\infty]{} \ell$. Supposons sans perte de généralité que $\ell \geq 0$ (quitte à remplacer f par $-f$). Soit $\mathrm{A} > 0$ tel que pour tout $x \geq \mathrm{A}$, $x \geq \dfrac{\ell}{2}$. On a alors pour tout $x \geq \mathrm{A}$,

$$\int_0^x f(t)dt = \int_0^{\mathrm{A}} f(t)dt + \int_{\mathrm{A}}^x f(t)dt \geq \int_0^{\mathrm{A}} f(t)dt + \frac{\ell}{2}(x-\mathrm{A}).$$

(cc) BY-NC-SA

Si $\ell > 0$, $x \longmapsto \frac{\ell}{2}(x - \mathrm{A})$ diverge vers $+\infty$ lorsque x tends vers $+\infty$. On en déduit qu'on a nécessairement $\ell = 0$. $\square$

Remarque. Cette propriété est aussi valable lorsque l'ensemble d'arrivée de f est $\mathbb{C}$ et que $\ell \in \mathbb{C}$. Le lecteur est encouragé à le montrer à titre d'exercice.

Exercice VII.2.

Soit une fonction de $\mathbb{R}^+$ dans $\mathbb{C}$ continument dérivable. On suppose que les intégrales $\int_0^{+\infty} f(t)dt$ et $\int_0^{+\infty} |f'(t)|\, dt$ convergent. Montrer que $f(x) \xrightarrow[x\to+\infty]{} 0$.

Exercice VII.3.

Soit une fonction de $\mathbb{R}^+$ dans $\mathbb{R}$ décroissante, telle que $\int_0^{+\infty} f(t)dt$ soit convergente. Montrer que $t \longmapsto tf(t)$ converge vers 0 lorsque t tend vers l'infini.

Exercice VII.4.

Soit $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$ une fonction uniformément continue sur $\mathbb{R}^+$ qui admet une intégrale généralisée sur $[0, +\infty[$. Montrer que $f(x) \xrightarrow[x\to+\infty]{} 0$.

Exercice VII.5.

Soit f une fonction de $\mathbb{R}^+$ dans $\mathbb{R}$ continument dérivable telle que les intégrales $\int_0^{+\infty} f(t)dt$ et $\int_0^{+\infty} (f'(t))^2 dt$ convergent. Montrer que $f(x) \xrightarrow[x\to+\infty]{} 0$.

(cc) BY-NC-SA

Correction de l'exercice I.8 :

1. On a pour tout $x \in]0,1]$,

$$\int_x^1 \frac{1}{t}dt = -\ln x \xrightarrow[x\to 0^+]{} +\infty.$$

On en déduit alors que $x \longmapsto \dfrac{1}{x}$ n'admet pas d'intégrale généralisée sur $]0,1]$.

2. Soit $\alpha \in]1,+\infty[$. On a pour tout $x \in [1,+\infty[$,

$$\int_1^x \frac{1}{t^\alpha}dt = \frac{1}{1-\alpha}\left(\frac{1}{x^{\alpha-1}}-1\right) \xrightarrow[x\to+\infty]{} \frac{1}{\alpha-1}.$$

On en déduit donc que $x \longmapsto \dfrac{1}{x^\alpha}$ admet bien une intégrale généralisée sur $]1,+\infty[$.

3. On a pour tout $x \in \left[0,\dfrac{\pi}{2}\right[$,

$$\int_0^x \tan t\, dt = \int_0^x \frac{\sin t}{\cos t}dt = \int_0^x -\frac{\cos' t}{\cos t}dt = -\ln\cos x + \ln\cos 0 = -\ln\cos x \xrightarrow[x\to\frac{\pi}{2}^-]{} +\infty.$$

On en déduit alors que $x \longmapsto \tan x$ n'admet pas d'intégrale généralisée sur $\left[0,\dfrac{\pi}{2}\right[$.

Correction de l'exercice II.4 :

→ (1) ⇒ (2) Supposons que f admet une intégrale généralisée sur $\mathbb{R}$. On en déduit donc que Les application $\mathrm{F}_d : x \longmapsto \displaystyle\int_0^x f(t)dt$ et $\mathrm{F}_g : x \longmapsto \displaystyle\int_x^0 f(t)dt$ convergent respectivement en $+\infty$ et $-\infty$. Soit $\varepsilon > 0$. En posant ℓ_d et ℓ_g les limites respectifs de F_d et F_g en $+\infty$ et $-\infty$, on peut considérer $\mathrm{A} > 0$ et $\mathrm{B} < 0$ tels que

$$\forall y > \mathrm{A}, |\mathrm{F}_d(y) - \ell_d| < \frac{\varepsilon}{2} \text{ et } \forall x < \mathrm{B},\ |\mathrm{F}_g(x) - \ell_g| < \frac{\varepsilon}{2}.$$

On en déduit donc que pour tout $y > \mathrm{A}$ et $x < \mathrm{B}$,

$$|\mathrm{F}_g(x) + \mathrm{F}_d(y) - (\ell_g + \ell_d)| \leq |\mathrm{F}_d(y) - \ell_d| + |\mathrm{F}_g(x) - \ell_g| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

L'inégalité ci-dessus se réécrit

$$\left|\int_x^y f(t)dt - \underbrace{(\ell_g + \ell_d)}_{\ell}\right| < \varepsilon,$$

ce qui est bien le résultat voulu.

→ (2) ⇒ (1) L'idée ici est d'utiliser le critère de Cauchy (Théorème II.3). Soit $\ell \in \mathbb{R}$ vérifiant la proposition (2). Soit $\varepsilon > 0$. $\mathrm{A} < 0$ et $\mathrm{B} > 0$ tels que pour tout $x < \mathrm{A}$ et $y > \mathrm{B}$, $|\mathrm{G}(x,y) - \ell| < \dfrac{\varepsilon}{2}$. Soit $y, y' \in]\mathrm{B} + \infty[$. On a

$$\left|\int_y^{y'} f(t)dt\right| = \left|\int_{\mathrm{A}-1}^{y} f(t)dt - \int_{\mathrm{A}-1}^{y'} f(t)dt\right| \leq \left|\int_{\mathrm{A}-1}^{y} f(t)dt - \ell\right| + \left|\int_{\mathrm{A}-1}^{y'} f(t)dt - \ell\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

On en déduit donc d'après le critère de Cauchy que f admet une intégrale généralisée sur $[0,+\infty[$. On peut montrer de la même manière que f admet également une intégrale généralisée sur $]-\infty,0]$. f admet donc bien une intégrale généralisée sur $\mathbb{R}$.

(cc) BY-NC-SA

Correction de l'exercice III.4 :

Remarquons d'abord que $g : t \longmapsto \frac{1}{t} - \left\lfloor \frac{1}{t} \right\rfloor$ admet bien une intégrale généralisée. En effet, g est continue par morceaux sur $]0,1]$, positive et majorée par la fonction positive constante égale à 1, qui est admet une intégrale généralisée sur $]0,1]$. Ceci implique donc grâce à II.2 que g admet bien une intégrale généralisée sur $]0,1]$.
On fait face à une intégrale avec une partie entière. On isole donc la partie de l'intégrale où est la partie entière, qui est une fonction constante par morceaux. Ensuite, on découpe cette intégrale sur les intervalles où la fonction à intégrer est constante. En particulier, la fonction $t \longmapsto \left\lfloor \frac{1}{t} \right\rfloor$ est constante sur les intervalles de la forme $\left] \frac{1}{k+1}, \frac{1}{k} \right]$. On a pour tout $x \in]0,1]$,

$$\begin{aligned}
\int_x^1 \frac{1}{t} - \left\lfloor \frac{1}{t} \right\rfloor dt &= -\ln x - \sum_{k=1}^{\left\lfloor \frac{1}{x} \right\rfloor - 1} \int_{\frac{1}{k+1}}^{\frac{1}{k}} \left\lfloor \frac{1}{t} \right\rfloor dt - \int_x^{\frac{1}{\left\lfloor \frac{1}{x} \right\rfloor}} \left\lfloor \frac{1}{t} \right\rfloor dt \\
&= -\ln x - \sum_{k=1}^{\left\lfloor \frac{1}{x} \right\rfloor - 1} \int_{\frac{1}{k+1}}^{\frac{1}{k}} k \, dt - \int_x^{\frac{1}{\left\lfloor \frac{1}{x} \right\rfloor}} \left\lfloor \frac{1}{x} \right\rfloor dt \\
&= -\ln x - \sum_{k=1}^{\left\lfloor \frac{1}{x} \right\rfloor - 1} \left(\frac{1}{k} - \frac{1}{k+1} \right) k - \left(\frac{1}{\left\lfloor \frac{1}{x} \right\rfloor} - x \right) \left\lfloor \frac{1}{x} \right\rfloor \\
&= -\ln x - \sum_{k=0}^{\left\lfloor \frac{1}{x} \right\rfloor - 1} \frac{1}{k+1} + x \left\lfloor \frac{1}{x} \right\rfloor.
\end{aligned}$$

On remarque tout d'abord que

$$x \left\lfloor \frac{1}{x} \right\rfloor = 1 - x \underbrace{\left(\frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \right)}_{\in [0,1]} \xrightarrow[x \to 0^+]{} 1.$$

Ensuite, on réarrange les termes de l'égalité ci-dessus pour obtenir, grâce à la continuité du logarithme en 1

$$\int_x^1 \frac{1}{t} - \left\lfloor \frac{1}{t} \right\rfloor dt = \ln \left\lfloor \frac{1}{x} \right\rfloor - \sum_{k=1}^{\left\lfloor \frac{1}{x} \right\rfloor} \frac{1}{k} + \underbrace{x \left\lfloor \frac{1}{x} \right\rfloor - \ln \left(x \left\lfloor \frac{1}{x} \right\rfloor \right)}_{\xrightarrow[x \to 0^+]{} 1}.$$

On en déduit alors qu'il suffit de trouver la limite (qui existe, étant donné que g admet bien une intégrale généralisée) de la fonction $x \longmapsto \ln \left\lfloor \frac{1}{x} \right\rfloor - \sum_{k=1}^{\left\lfloor \frac{1}{x} \right\rfloor} \frac{1}{k}$ lorsque x tend vers 0 à droite. Pour cela, on peut utiliser le lemme très important suivant.

Lemme VII.6.

Il existe une constante γ positive appelée constante gamma d'Euler, telle que

$$\sum_{k=1}^{n} \frac{1}{k} - \ln n \underset{n \to +\infty}{=} \gamma + o(1).$$

BY-NC-SA

Démonstration. Posons pour tout $n \in \mathbb{N}^*$, $u_n = \sum_{k=1}^{n} \frac{1}{k} - \ln n$. On a alors pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned} u_{n+1} - u_n = \frac{1}{n+1} - \ln\left(1 + \frac{1}{n}\right) &\underset{n\to+\infty}{=} \frac{1}{n+1} - \frac{1}{n} + \mathrm{O}\left(\frac{1}{n^2}\right) \\ &\underset{n\to+\infty}{=} \underbrace{-\frac{1}{n(n+1)}}_{=\mathrm{O}\left(\frac{1}{n^2}\right)} + \mathrm{O}\left(\frac{1}{n^2}\right) \underset{n\to+\infty}{=} \mathrm{O}\left(\frac{1}{n^2}\right). \end{aligned}$$

$\frac{1}{n^2}$ est le terme général d'une série absolument convergente, et donc $u_{n+1} - u_n$, aussi, ce qui signifie que la suite u_n est bien convergente, ce qui donne le résultat voulu.
Montrons désormais la (stricte) positiviste de γ et posons pour ceci pour tout $n \in \mathbb{N}^*$, $v_n = u_n - \frac{1}{n}$.
Il est clair que v_n converge vers γ et que $v_1 = 0$. Ainsi, il suffit de montrer la (stricte) croissance de v_n. Fixons alors $n \in \mathbb{N}^*$ et remarquons que

$$v_{n+1} - v_n = \frac{1}{n} - \ln\left(1 + \frac{1}{n}\right) > 0$$

étant donné que $\forall x > 0,\ x - \ln(x+1) > 0$. On conclue donc que $\gamma > 0$. □

L'application du lemme ci-dessus nous donne directement que la fonction $x \longmapsto \ln\left\lfloor \frac{1}{x} \right\rfloor - \sum_{k=1}^{\left\lfloor \frac{1}{x} \right\rfloor} \frac{1}{k}$ admet $-\gamma$ comme limite pour x tendant vers 0^+, et on en déduit finalement que $t \longmapsto \frac{1}{t} - \left\lfloor \frac{1}{t} \right\rfloor$ admet bien une intégrale généralisée sur $]0,1]$ égale à $1-\gamma$.
Remarquons au passage que ceci donne, comme conséquence directe, que $1 - \gamma > 0$ i.e. que $\gamma < 1$.
En effet, l'inégalité large est claire du fait que $g \geq 0$, tandis que la stricte provient, par exemple, du fait que $\int_0^1 g(t)dt \geq \int_{\frac{3}{5}}^{\frac{4}{5}} g(t)dt$ et que $g\big|_{\left[\frac{3}{5},\frac{4}{5}\right]}$ est continue et strictement positive, et donc d'intégrale strictement positive sur $\left[\frac{3}{5}, \frac{4}{5}\right]$.

Correction de l'exercice IV.3 :

Soit $\beta \in \mathbb{R}$. Si $\beta \leq 0$, alors $t \longmapsto |\ln t|^\beta$ est bornée sur l'intervalle borné $\left]0, \frac{1}{2}\right]$ et admet donc bien une intégrale généralisée sur $\left]0, \frac{1}{2}\right]$. On suppose maintenant que $\beta > 0$. On a alors pour tout $t \in \left]0, \frac{1}{2}\right]$

$$t^{\frac{1}{2}} |\ln t|^\beta = \left|2\beta t^{\frac{1}{2\beta}} \ln\left(t^{\frac{1}{2\beta}}\right)\right|^\beta \xrightarrow[t\to 0^+]{} 0.$$

On en déduit alors d'après la proposition IV.2 que $t \longmapsto |\ln t|^\beta$ admet bien une intégrale généralisée sur $\left]0, \frac{1}{2}\right]$ pour $\beta > 0$. Finalement, on conclut que $t \longmapsto |\ln t|^\beta$ admet bien une intégrale généralisée sur $\left]0, \frac{1}{2}\right]$ pour tout $\beta \in \mathbb{R}$.

Correction de l'exercice IV.4 :

→ On remarque tout d'abord que lorsque $\alpha > 1$ et $\beta \geq 0$, on peut facilement voir que l'intégrale converge, car le terme $(\ln t)^\beta$ ne fait que contribuer à la vitesse de convergence vers 0 de la fonction. En effet, on a $\frac{1}{t^\alpha (\ln t)^\beta} = \mathrm{O}\left(\frac{1}{t^\alpha}\right)$. On en déduit donc d'après les propositions II.2 et IV.1 que l'intégrale est convergente.

(cc) BY-NC-SA

→ Si $\alpha > 1$ et $\beta < 0$, alors on peut voir que la contribution de $(\ln t)^\beta$ est négligeable devant celle de t^α. En effet, on a en $+\infty$,

$$\frac{1}{t^\alpha(\ln t)^\beta} = \frac{1}{t^{\frac{1+\alpha}{2}}}\frac{(\ln t)^{-\beta}}{t^{\frac{\alpha-1}{2}}} = \frac{1}{t^{\frac{1+\alpha}{2}}}\underbrace{\left(\frac{\frac{2\beta}{1-\alpha}\ln t^{\frac{1-\alpha}{2\beta}}}{t^{\frac{1-\alpha}{2\beta}}}\right)^{-\beta}}_{\xrightarrow[t\to+\infty]{}0} = \mathrm{O}\left(\frac{1}{t^{\frac{\alpha+1}{2}}}\right).$$

Étant donné que $\frac{1+\alpha}{2} > 1$, on obtient bien en utilisant les propositions IV.1 et II.2 que l'intégrale est bien convergente. L'idée ici était d'écrire t^α sous la forme $t^{\alpha-\varepsilon}t^\varepsilon$, où ε est un nombre strictement positif tel qu'on ait $\alpha - \varepsilon > 1$, et d'ensuite contrôler le terme $(\ln t)^\beta$ en utilisant le terme t^ε.

→ Supposons maintenant que $\alpha < 1$. Si $\beta \leq 0$, on peut facilement voir que le terme $(\ln t)^\beta$ ne fait qu'accentuer la divergence vers l'infini de la fonction $\frac{1}{t^\alpha(\ln t)^\beta}$. En particulier, on a en $+\infty$,

$$\frac{1}{t^\alpha} = \mathrm{O}\left(\frac{1}{t^\alpha(\ln t)^\beta}\right).$$

On en déduit donc en utilisant les propositions IV.1 et II.2 que l'intégrale est divergente.

→ Si $\alpha < 1$ et $\beta > 0$, on peut considérer $\varepsilon > 0$ tel que $\alpha + \varepsilon < 1$. Dans ce cas, on a

$$t^{\alpha+\varepsilon}\frac{1}{t^\alpha(\ln t)^\beta} = \frac{t^\varepsilon}{(\ln t)^\beta} = \left(\frac{\varepsilon}{\beta}\frac{t^{\frac{\varepsilon}{\beta}}}{\ln(t^{\frac{\varepsilon}{\beta}})}\right)^\beta \xrightarrow[t\to+\infty]{} +\infty.$$

On peut alors déduire de la proposition IV.2 que l'intégrale est divergente.

→ Si $\alpha = 1$ et $\beta \neq 1$, alors on a pour tout $x \geq 2$,

$$\int_2^x \frac{1}{t(\ln t)^\beta}dt = \int_2^x \frac{\ln' t}{(\ln t)^\beta}dt = \frac{1}{1-\beta}(\ln x)^{1-\beta} - \frac{1}{1-\beta}(\ln 2)^{1-\beta}$$
$$\xrightarrow[x\to+\infty]{} \begin{cases} +\infty & \text{si } \beta < 1 \\ \frac{1}{\beta-1}(\ln 2)^{1-\beta} & \text{sinon.} \end{cases}$$

On en déduit donc que l'intégrale converge si et seulement si $\beta > 1$.

→ Si $\alpha = 1$ et $\beta = 1$, alors pour tout $x \geq 2$,

$$\int_2^x \frac{1}{t\ln t} = \int_2^x \frac{\ln' t}{\ln t}dt = \ln\ln t - \ln\ln 2 \xrightarrow[t\to+\infty]{} +\infty,$$

et donc l'intégrale diverge.

En conclusion, l'intégrale $\int_2^x \frac{1}{t^\alpha(\ln t)^\beta}dt$ converge si et seulement si $\alpha > 1$, ou $\alpha = 1$ et $\beta > 1$.

Correction de l'exercice IV.5 :

Pour tout $x \geq 0$, on a

$$\left|e^{-\lambda t^\alpha}f(t)\right| \leq e^{-|\mathrm{Re}\,\lambda|t^\alpha}|f(t)| = e^{-\frac{|\mathrm{Re}\,\lambda|}{2}t^\alpha}\underbrace{e^{-\frac{|\mathrm{Re}\,\lambda|}{2}t^\alpha}|f(t)|}_{\xrightarrow[t\to+\infty]{}0} \underset{t\to+\infty}{=} \mathrm{O}\left(e^{-\frac{|\mathrm{Re}\,\lambda|}{2}t^\alpha}\right) \underset{t\to+\infty}{=} \mathrm{O}\left(\frac{1}{t^2}\right).$$

(cc) BY-NC-SA

L'avant-dernière égalité est vraie car f est dominée par une fonction polynomiale. On en déduit finalement, en utilisant les propositions II.2 et IV.1 que l'intégrale est convergente.

Correction de l'exercice IV.6 :

Il suffit de faire le changement de variable $u = t - a$ et d'utiliser la proposition IV.1. L'intégrale converge si et seulement si $\alpha < 1$.

Correction de l'exercice IV.7 :

1. En 1^+, la fonction tend vers un réel fini. En effet,

$$\frac{1}{(\ln t)^{\ln t}} = \frac{1}{e^{(\ln t)(\ln \ln t)}} \underset{t \to 1^+}{\longrightarrow} 1.$$

La limite ci-dessus est obtenue par changement de variable et utilisation de la limite connue $t \ln t \underset{t \to 0^+}{\longrightarrow} 0$. Étudions maintenant la fonction en $+\infty$ a pour tout $t \geq 1$,

$$\frac{1}{(\ln t)^{\ln t}} = \frac{1}{e^{(\ln t)(\ln \ln t)}} = \frac{1}{t^{\ln \ln t}}.$$

On remarque que la puissance de t tend vers l'infini, et donc est supérieure à 2 pour t assez grand. En particulier, on a

$$t^2 \frac{1}{(\ln t)^{\ln t}} = \frac{1}{t^{-2+\ln \ln t}} \underset{t \to +\infty}{\longrightarrow} 0.$$

On déduit donc d'après la proposition IV.2 que l'intégrale est bien convergente.

2. En 0, $t \longmapsto \sin \dfrac{1}{t^2}$ est bornée. En $+\infty$, on voit facilement que $\sin \dfrac{1}{t^2} \underset{t \to +\infty}{\sim} \dfrac{1}{t^2}$ et donc en utilisant les propositions IV.1 et II.2, on voit que l'intégrale converge.

3. Utilisons la formule suivante. Pour tout $a, b \in \mathbb{R}$ tels que $ab \neq -1$, on a

$$\arctan a - \arctan b = \arctan \left(\frac{a-b}{1+ab} \right).$$

On a pour tout $t \geq 0$,

$$0 \leq \arctan(1+t^3) - \arctan(t^3) = \arctan \left(\frac{1}{1+t^3(1+t^3)} \right) \leq \frac{1}{1+t^3(1+t^3)} = \mathrm{O}\left(\frac{1}{t^6} \right).$$

On en déduit donc d'après les propositions IV.1 et II.2 que l'intégrale est convergente.

4. Étudions cette intégrale en chaque valeur singulière.
   - $\rightarrow$ En 0^+, on a $\dfrac{1}{\sqrt{x\,|x-1|\,|x-2|}} \sim \dfrac{1}{\sqrt{2x}}$, et la fonction de droite est intégrable au voisinage de 0.
   - $\rightarrow$ En 1, $\dfrac{1}{\sqrt{x\,|x-1|\,|x-2|}} \sim \dfrac{1}{\sqrt{|x-1|}}$, et la fonction de droite est intégrable au voisinage de 1.
   - $\rightarrow$ En 2, $\dfrac{1}{\sqrt{x\,|x-1|\,|x-2|}} \sim \dfrac{1}{\sqrt{2\,|x-2|}}$, et la fonction de droite est intégrable au voisinage de 2.
   - $\rightarrow$ En $+\infty$, $\dfrac{1}{\sqrt{x\,|x-1|\,|x-2|}} \sim \dfrac{1}{x^{\frac{3}{2}}}$, et la fonction de droite est intégrable en $+\infty$.

   On en déduit alors que l'intégrale est convergente.

(cc) BY-NC-SA

Correction de l'exercice IV.8 :

On commence d'abord par voir où sont les singularités pour cette fonction sur $\mathbb{R}$.

→ Si P admet une racine $a \in \mathbb{R}$ de multiplicité égale à 2, alors on peut écrire

$$\forall t \in \mathbb{R},\ \mathrm{P}(t) = \mathrm{C}(t-a)^2(t-b),$$

où $b \in \mathbb{R} \setminus \{a\}$ et $\mathrm{C} \neq 0$. On a alors au voisinage de a,

$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \sim \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|a-b|^{\frac{1}{2}}\,|t-a|}.$$

Cette fonction n'admet pas d'intégrale généralisée au voisinage de a, et donc l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}dt$ est divergente.

→ Si P admet une racine $a \in \mathbb{R}$ de multiplicité a égale à 3, alors on peut écrire

$$\forall t \in \mathbb{R},\ \mathrm{P}(t) = \mathrm{C}(t-a)^3,$$

où $\mathrm{C} \neq 0$. On a alors au voisinage de a,

$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \sim \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|t-a|^{\frac{3}{2}}}.$$

Cette fonction n'admet pas d'intégrale généralisée au voisinage de a, et donc l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}dt$ est divergente.

→ Si toutes les racines réelles de P sont simples, alors deux cas se présentent. Si P admet une seule racine réelle $a \in \mathbb{R}$, et dans ce cas, il existe Q un polynôme unitaire de degré 2 sans racines, et $\mathrm{C} \neq 0$ tels que

$$\forall t \in \mathbb{R},\ \mathrm{P}(t) = \mathrm{CQ}(t)(t-a).$$

$t \longmapsto \dfrac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}$ admet alors une singularité en a. On doit alors étudier le comportement de cette fonction en $+\infty$, $-\infty$ et a. On a en $+\infty$ et $-\infty$,

$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \sim \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|t|^{\frac{3}{2}}}.$$

La fonction de droite est intégrable en $+\infty$ et $-\infty$, et donc $t \longmapsto \dfrac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}$ est intégrable en $+\infty$ et $-\infty$. De plus, au voisinage de a, on a

$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \sim \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|\mathrm{Q}(a)|^{\frac{1}{2}}\,|t-a|^{\frac{1}{2}}}.$$

La fonction de droite est intégrable en a, ce qui nous permet de dire que $t \longmapsto \dfrac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}$ est également intégrable en a, et que finalement l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}dt$ est convergente.

Il reste le dernier cas à traiter où P admet trois racines distinctes $a, b, c \in \mathbb{R}$. Dans le cas, on peut affirmer l'existence de $\mathrm{C} \neq 0$ tel que

$$\forall t \in \mathbb{R},\ \mathrm{P}(t) = \mathrm{C}(t-a)(t-b)(t-c).$$

(cc) BY-NC-SA

On en déduit alors que P admet trois singularités en a, b et c, et qu'il faut donc étudier le comportement de la fonction $t \longmapsto \dfrac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}$ en $+\infty$, $-\infty$ et a, b, c. En $+\infty$ et $-\infty$, un raisonnement identique au précédent permet de dire que la fonction $t \longmapsto \dfrac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}$ est intégrable en $+\infty$ et $-\infty$. De plus, on a

$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \underset{t\to a}{\sim} \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|t-a|^{\frac{1}{2}}\,|a-b|^{\frac{1}{2}}\,|a-c|^{\frac{1}{2}}},$$
$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \underset{t\to b}{\sim} \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|t-b|^{\frac{1}{2}}\,|b-a|^{\frac{1}{2}}\,|b-c|^{\frac{1}{2}}},$$
$$\frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}} \underset{t\to c}{\sim} \frac{1}{|\mathrm{C}|^{\frac{1}{2}}\,|t-c|^{\frac{1}{2}}\,|c-a|^{\frac{1}{2}}\,|c-b|^{\frac{1}{2}}}.$$

Les trois fonctions ci-dessus sont intégrables respectivement en a, b et c. On en déduit donc que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{|\mathrm{P}(t)|^{\frac{1}{2}}}dt$ est convergente.

Correction de l'exercice IV.9 :

La fonction $t \longmapsto \dfrac{t^a-1}{\ln t}$ admet a priori deux singularités en 0 et 1. étudions donc le comportement de cette fonction en ces deux points. On a en 1^-,

$$\frac{t^a-1}{\ln t} = \frac{t^a-1}{t-1} \times \frac{t-1}{\ln t} \xrightarrow[t\to 1^-]{} \frac{a}{\ln'(1)} = a.$$

$t \longmapsto \dfrac{t^a-1}{\ln t}$ est donc (continue par morceaux et) bornée au voisinage de 1 et y est donc intégrable. En 0^+, lorsque $a > 0$ on a

$$\frac{t^a-1}{\ln t} \sim \frac{-1}{\ln t}.$$

D'après l'exercice IV.3, la fonction de droite est intégrable en 0, et donc, vu que cette dernière garde un signe constant, la fonction $t \longmapsto \dfrac{t^a-1}{\ln t}$ aussi.
Si $a = 0$, on peut également facilement que la fonction est intégrable sur $]0, 1]$ car elle est nulle. Enfin, il reste a traiter le cas $a < 0$. Dans ces cas, on a en 0^+,

$$\frac{t^a-1}{\ln t} \sim \frac{1}{t^{-a}\ln t}.$$

On peut facilement montrer que la fonction de droite est intégrable si et seulement si $a > -1$. La preuve de ce résultat est laissée comme exercice au lecteur. On en déduit alors finalement, toujours vu que cette dernière garde un signe constant et d'après la proposition II.2, que l'intégrale $\displaystyle\int_0^1 \frac{t^a-1}{\ln t}dt$ converge si et seulement si $a > -1$.

Correction de l'exercice V.3 :

La fonction $t \longmapsto \dfrac{\sin t}{t}$ admet une singularité en 0, il faut donc étudier son comportement en 0 et en $+\infty$. En 0^+, la fonction $t \longmapsto \dfrac{\sin t}{t}$ tend vers 1 et est donc bornée au voisinage de 0^+, ce qui fait qu'elle y est intégrable. Étudions maintenant le comportement de cette fonction en $+\infty$. On va utiliser la proposition V.2 pour montrer la convergence de l'intégrale. On va considérer f, g deux fonctions de $\mathbb{R}_+^*$ dans $\mathbb{R}$ telles

BY-NC-SA

que

$$\forall t \in \mathbb{R}_+^*,\ f'(t) = \sin t,\ f(t) = -\cos(t),\ g(t) = \frac{1}{t}.$$

La fonction $t \longmapsto f(t)g(t)$ admet une limite finie en $+\infty$ égale à 0, et de plus, on a pour tout $t \in \mathbb{R}_+^*$,

$$f(t)g'(t) = \frac{\cos t}{t^2}.$$

Cette fonction est intégrable en $+\infty$ car dominée par $t \longmapsto \frac{1}{t^2}$. de plus, f et g sont continument dérivables sur $[1, +\infty[$, et donc d'après la proposition V.2, $t \longmapsto f'(t)g(t)$ admet une intégrale généralisée sur $[1, +\infty[$, et finalement l'intégrale $\int_0^{+\infty} \frac{\sin t}{t} dt$ est convergente.

Correction de l'exercice V.4 :

On va une fois de plus faire une intégration par parties. Soit $\varepsilon > 0$ et $x \geq \varepsilon$. On a

$$\int_\varepsilon^x \frac{\sin t}{t} dt = \left[\frac{1-\cos t}{t}\right]_\varepsilon^x + \int_\varepsilon^t \frac{1-\cos t}{t^2} dt \tag{4}$$

$$= \left[\frac{1-\cos t}{t}\right]_\varepsilon^x + \int_\varepsilon^x \frac{2\left(\sin \frac{t}{2}\right)^2}{t^2} dt \tag{5}$$

$$\underset{u=\frac{t}{2}}{=} \underbrace{\left[\frac{1-\cos t}{t}\right]_\varepsilon^x}_{\mathrm{A}(\varepsilon,x)} + \underbrace{\int_{\frac{\varepsilon}{2}}^{\frac{x}{2}} \frac{(\sin u)^2}{u^2} du}_{\mathrm{B}(\varepsilon,x)} \tag{6}$$

On a vu dans l'exercice précédent que $t \longmapsto \frac{\sin t}{t}$ admet une intégrale généralisée sur $]0, +\infty[$. On en déduit donc que

$$\int_\varepsilon^x \frac{\sin t}{t} dt \xrightarrow[\varepsilon \to 0^+]{} \int_0^x \frac{\sin t}{t} dt \text{ et } \int_0^x \frac{\sin t}{t} dt \xrightarrow[x \to +\infty]{} \int_0^{+\infty} \frac{\sin t}{t} dt.$$

On a de plus,

$$\mathrm{A}(\varepsilon, x) = \underbrace{\frac{1-\cos x}{x}}_{\xrightarrow[x\to+\infty]{} 0} - \underbrace{\frac{1-\cos \varepsilon}{\varepsilon}}_{\xrightarrow[\varepsilon\to 0^+]{} 0} \xrightarrow[(\varepsilon,x)\to(0^+,+\infty)]{} 0$$

et $u \longmapsto \frac{(\sin u)^2}{u^2}$ est intégrable sur $]0, +\infty[$ (car dominée par $u \longmapsto u^{-2}$ en $+\infty$ et bornée par 1 en 0), et alors

$$\mathrm{B}(\varepsilon, x) = \int_{\frac{\varepsilon}{2}}^{\frac{x}{2}} \frac{(\sin u)^2}{u^2} dt \xrightarrow[(\varepsilon,x)\to(0^+,+\infty)]{} \int_0^{+\infty} \frac{(\sin u)^2}{u^2} dt.$$

Enfin, en faisant tendre ε vers 0^+ puis x vers $+\infty$ dans l'égalité 6, on obtient

$$\int_0^{+\infty} \frac{(\sin t)^2}{t^2} dt = \int_0^{+\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}.$$

(cc) BY-NC-SA

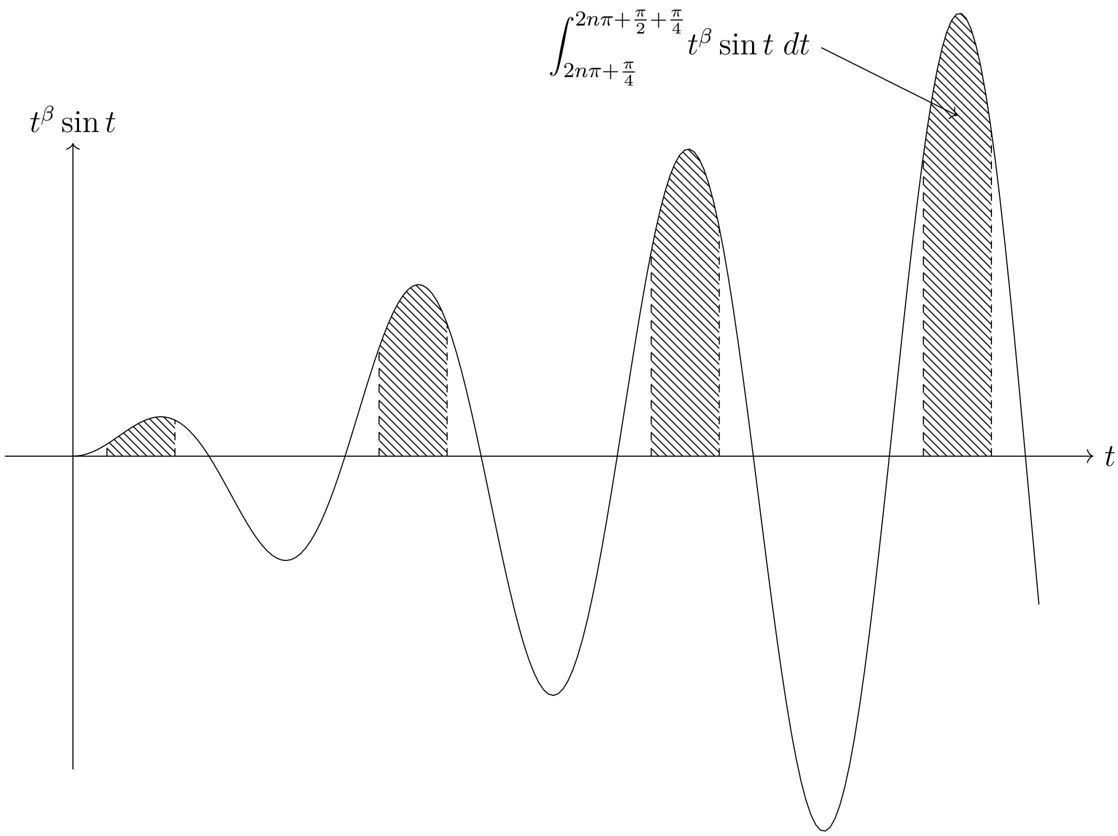
Correction de l'exercice V.5 :

Encore une fois, on pense ici à une intégration par parties. Pour faire converger l'intégrale, on pense poser pour tout $t \geq 1$, $f(t) = \dfrac{1}{t^\alpha}$ et $g'(t) = e^{it}$ (on précisera g un peu plus tard), et dériver f pour augmenter le degré du dénominateur et ensuite utiliser ce qu'on sait sur les intégrales de Riemann pour obtenir la convergence de l'intégrale via la proposition V.2. Pour que cette méthode fonctionne, if faut que $t \longmapsto f(t)g(t)$ converge vers une valeur finie en $+\infty$, et donc il est nécessaire d'avoir $\alpha > 0$. Supposons donc que $\alpha > 0$. Posons pour tout t, $g(t) = -ie^{it}$. On peut facilement voir que $t \longmapsto f(t)g(t)$ tend vers 0 en $+\infty$. De plus, pour tout $t \geq 1$,

$$f'(t)g(t) = \alpha\frac{ie^{it}}{t^{\alpha+1}} = \mathrm{O}\left(\frac{1}{t^{\alpha+1}}\right).$$

$t \longmapsto \dfrac{1}{t^{\alpha+1}}$ est intégrable (et donc admet une intégrale généralisée) sur $[1, +\infty[$ d'après la proposition IV.1, donc $f'(t)g(t)$ aussi d'après la proposition II.2, et alors d'après la proposition V.2, f et g étant continument dérivables sur $[1, +\infty[$, la fonction $t \longmapsto f(t)g'(t)$ (qui est égale à $t \longmapsto \dfrac{e^{it}}{t}$) admet une intégrale généralisée sur $[1, +\infty[$.

Il reste maintenant à traiter le cas $\alpha \leq 0$. Lorsque $\alpha = 0$, l'intégrale étudiée n'est clairement pas convergente. Supposons alors que $\alpha < 0$ et posons $\beta = -\alpha > 0$. On veut étudier l'intégrale $\displaystyle\int_1^{+\infty} t^\beta e^{it} dt$. Étudions la convergence de l'intégrale de la partie imaginaire de l'intégrande, qui est égale à $\displaystyle\int_1^{+\infty} t^\beta \sin t\, dt$. Les oscillations de cette fonction explosent en $+\infty$, donc notre intuition est de montrer qu'elle diverge en utilisant le critère de Cauchy (Théorème II.3).



(cc) BY-NC-SA

On a pour tout $n \geq 1$,

$$\begin{aligned}
\int_{2n\pi+\frac{\pi}{4}}^{2n\pi+\frac{\pi}{2}+\frac{\pi}{4}} t^{\beta}\sin t\, dt &\underset{(*)}{\geq} \frac{1}{\sqrt{2}}\int_{2n\pi+\frac{\pi}{4}}^{2n\pi+\frac{\pi}{2}+\frac{\pi}{4}} t^{\beta}dt \\
&= \frac{1}{(\beta+1)\sqrt{2}}\left(\left(2n\pi+\frac{\pi}{4}+\frac{\pi}{2}\right)^{\beta+1}-\left(2n\pi+\frac{\pi}{4}\right)^{\beta+1}\right) \\
&= \frac{\left(2n\pi+\frac{\pi}{4}\right)^{\beta+1}}{(\beta+1)\sqrt{2}}\left(\left(1+\frac{\pi}{2\left(2n\pi+\frac{\pi}{4}\right)}\right)^{\beta+1}-1\right) \\
&\underset{n\to+\infty}{=} \frac{\left(2n\pi+\frac{\pi}{4}\right)^{\beta+1}}{(\beta+1)\sqrt{2}}\left(\frac{(\beta+1)\pi}{2\left(2n\pi+\frac{\pi}{4}\right)}+o\left(\frac{1}{n}\right)\right) \\
&\underset{n\to+\infty}{=} \frac{\pi\left(2n\pi+\frac{\pi}{4}\right)^{\beta}}{2\sqrt{2}}+o\left(n^{\beta}\right) \xrightarrow[n\to+\infty]{} +\infty.
\end{aligned}$$

L'égalité $(*)$ est vraie car pour tout $t \in \left[2n\pi+\frac{\pi}{4}, 2n\pi+\frac{\pi}{2}+\frac{\pi}{4}\right]$, $\sin t \geq \frac{1}{\sqrt{2}}$. On en déduit par cette chaine d'inégalités que par exemple, pour tout $\varepsilon = 1$, pour tout $\mathrm{A} \geq 0$, il existe $x, y \in [1, +\infty[$ supérieurs à A tels que $\left|\int_x^y t^{\beta}\sin t\, dt\right| \geq 1$ (il suffit de prendre $x = 2n\pi+\frac{\pi}{4}$ et $y = 2n\pi+\frac{\pi}{2}+\frac{\pi}{4}$ pour n assez grand), et alors d'après le théorème II.3, on peut conclure que l'intégrale $\int_1^{+\infty} t^{\beta}\sin t$ est divergente, et par conséquent l'intégrale $\int_1^{+\infty} t^{\beta}e^{it}dt$ aussi. Finalement, on conclut que l'intégrale $\int_1^{+\infty}\frac{e^{it}}{t^{\alpha}}dt$ converge si et seulement si $\alpha > 0$.

Correction de l'exercice V.6 :

→ ($\Leftarrow$) Supposons que f admet une primitive périodique qu'on notera F. On peut alors affirmer que la fonction $t \longmapsto \frac{\mathrm{F}(t)}{t}$ converge vers 0 alors que t tend vers $+\infty$, car F est bornée sur $[1+\infty[$ (car continue periodique). De plus, en posant pour tout $t \geq 1$, $g(t) = \frac{1}{t}$, on voit que $g'(t)\mathrm{F}(t) = -\frac{\mathrm{F}(t)}{t^2}$. On en déduit donc que la fonction $t \longmapsto g'(t)\mathrm{F}(t)$ est dominée par $t \longmapsto \frac{1}{t^2}$ (car F est bornée) en $+\infty$, et est donc intégrable sur $[1, +\infty[$, et y admet par conséquent une intégrale généralisée. On en déduit alors d'après la proposition V.2, F et g étant continument dérivables sur $[1, +\infty[$ que $t \longmapsto \mathrm{F}'(t)g(t)$ (qui est égale à $t \longmapsto \frac{f(t)}{t}$) admet bien une intégrale généralisée sur $[1, +\infty[$.

→ ($\Rightarrow$) Montrons cette implication par contraposée. Soit T la plus petite période de f. Supposons qu'aucune primitive de F est périodique. On a alors pour tout $x \geq 1$, $\int_x^{x+\mathrm{T}} f(t)dt \neq 0$. Posons $\mathrm{C} \neq 0$ la valeur de cette intégrale (qui ne dépend pas de x car sa dérivée est nulle). On peut facilement voir que $t \longmapsto f(t) - \frac{\mathrm{C}}{\mathrm{T}}$ admet une primitive périodique. En effet, pour tout $x \geq 1$, si par exemple F est la primitive de $t \longmapsto f(t) - \frac{\mathrm{C}}{\mathrm{T}}$ qui s'annule en 1, on a

$$\mathrm{F}(x+\mathrm{T}) - \mathrm{F}(x) = \int_x^{x+\mathrm{T}} f(t) - \frac{\mathrm{C}}{\mathrm{T}}dt = 0,$$

(cc) BY-NC-SA

i.e. F est périodique. Maintenant, on peut écrire pour tout $t \geq 1$,

$$\frac{f(t)}{t} = \frac{f(t) - \frac{\mathrm{C}}{\mathrm{T}}}{t} + \frac{\mathrm{C}}{\mathrm{T}}\frac{1}{t}.$$

Dans le côté droit de l'égalité, le premier terme admet une intégrale généralisée sur $[1, +\infty[$ d'après le point précédent, car le numérateur est une fonction périodique qui admet une primitive periodique, et le second a une intégrale divergente sur $[1, +\infty[$, et donc la somme des deux n'admet pas d'intégrale généralisée, ce qui montre bien l'implication contraposée voulue.

Remarque. Attention, si le premier terme n'admettait pas d'intégrale généralisée sur $[1, +\infty[$, on aurait pas pu faire cette conclusion.

Correction de l'exercice VI.2 :

La fonction $t \longmapsto \dfrac{t - \lfloor t \rfloor}{t^2}$ est positive et dominée par $t \longmapsto \dfrac{1}{t^2}$ (car le numérateur est borné par 1). On en déduit donc directement d'après les propriétés IV.1 et II.2 que $\displaystyle\int_1^{+\infty} \frac{t - \lfloor t \rfloor}{t^2} dt$ est convergente. Calculons maintenant la valeur de cette intégrale. Soit $x \geq 1$, on a

$$\begin{aligned}
\int_1^x \frac{t - \lfloor t \rfloor}{t^2} dt &= \int_1^x \frac{1}{t} dt - \int_1^x \frac{\lfloor t \rfloor}{t^2} dt \\
&= \ln x - \sum_{k=1}^{\lfloor x \rfloor - 1} \int_k^{k+1} \frac{k}{t^2} dt - \int_{\lfloor x \rfloor}^x \frac{\lfloor x \rfloor}{t^2} dt \\
&= \ln x - \sum_{k=1}^{\lfloor x \rfloor - 1} k \left(\frac{1}{k} - \frac{1}{k+1} \right) + \lfloor x \rfloor \left(\frac{1}{x} - \frac{1}{\lfloor x \rfloor} \right) \\
&= \ln x - \sum_{k=1}^{\lfloor x \rfloor - 1} \frac{1}{k+1} + \frac{\lfloor x \rfloor}{x} - 1 \\
&= \underbrace{\ln\left(\frac{x}{\lfloor x \rfloor} \right)}_{\xrightarrow[x \to +\infty]{} 0} + \underbrace{\ln \lfloor x \rfloor - \sum_{k=1}^{\lfloor x \rfloor} \frac{1}{k}}_{\xrightarrow[x \to +\infty]{} -\gamma \text{ (lemme VII.6)}} + \underbrace{\frac{\lfloor x \rfloor}{x}}_{\xrightarrow[x \to +\infty]{} 1} \xrightarrow[x \to +\infty]{} 1 - \gamma.
\end{aligned}$$

On en déduit alors finalement que $\displaystyle\int_1^{+\infty} \frac{t - \lfloor t \rfloor}{t^2} dt = 1 - \gamma$.

Remarquons au passage que ce n'est pas une coïncidence qu'on retombe sur $1 - \gamma$, comme pour l'exercice III.4. En effet, les deux intégrales se retrouvent l'une de l'autre par un changement de variable $t = \dfrac{1}{u}$. Vérifions ceci et prenons $x > 1$, on a

$$\int_1^x \frac{t - \lfloor t \rfloor}{t^2} dt \underset{\substack{t = \frac{1}{u} \\ dt = \frac{-1}{u^2} du}}{=} \int_1^{\frac{1}{x}} \frac{\frac{1}{u} - \left\lfloor \frac{1}{u} \right\rfloor}{\frac{1}{u^2}} \cdot \frac{-1}{u^2} du = \int_{\frac{1}{x}}^1 \left(\frac{1}{u} - \left\lfloor \frac{1}{u} \right\rfloor \right) du$$

Ainsi, on voit bien que l'intégrale généralisée $\displaystyle\int_1^{+\infty} \frac{t - \lfloor t \rfloor}{t^2} dt$ est bien définie si, et seulement si, l'intégrale généralisée $\displaystyle\int_0^1 \left(\frac{1}{u} - \left\lfloor \frac{1}{u} \right\rfloor \right) du$ est bien définie et que, si c'est le cas, les deux sont bien égales.

Correction de l'exercice VI.3 :

→ (⇐) Si $\displaystyle\int_0^{+\infty} f(t)dt$ converge, alors étant donné que f est positive et continue par morceaux, et

(cc) BY-NC-SA

que pour tout $t \in \mathbb{R}^+$, $0 \leq f(t)(\sin t)^2 \leq f(t)$, alors nécessairement, d'après la proposition II.2, l'intégrale $\displaystyle\int_0^{+\infty} f(t)(\sin t)^2 dt$ converge.

→ (⇒) Supposons que $\displaystyle\int_0^{+\infty} f(t)(\sin t)^2 dt$ converge. On souhaite exploiter la décroissance de f en utilisant un argument de comparaison de série intégrale (comme utilisé pour montrer la convergence des sommes de termes du type $\dfrac{1}{n^\alpha}$). On utilise la découpe de la proposition VI.1 en posant $(a_n)_{n\in\mathbb{N}} = (n\pi)_{n\in\mathbb{N}}$. On a alors pour tout $n \in \mathbb{N}$,

$$\int_{n\pi}^{(n+1)\pi} f(t)(\sin t)^2 dt \geq f((n+1)\pi)\int_{n\pi}^{(n+1)\pi} (\sin t)^2 dt = \frac{\pi}{2} f((n+1)\pi) \geq 0.$$

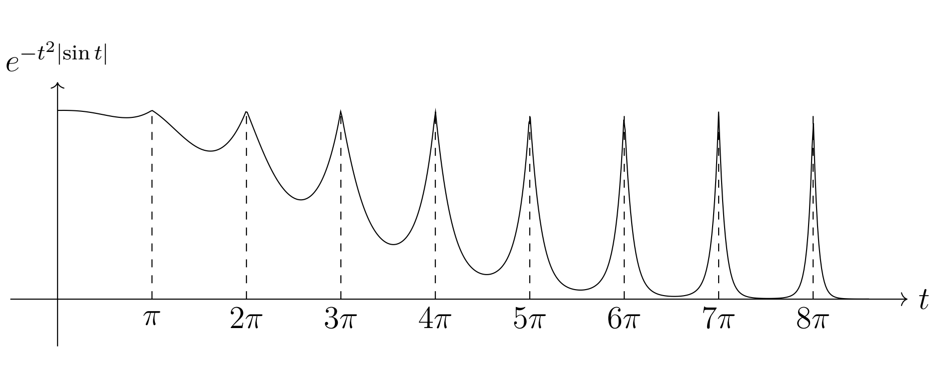
On sait d'après la proposition VI.1 que $\displaystyle\sum_{n=0}^{+\infty}\int_{n\pi}^{(n+1)\pi} f(t)(\sin t)^2 dt$ est convergente (car $t \longmapsto f(t)(\sin t)^2$ est positive et continue par morceaux, et son intégrale sur $[0, +\infty[$ converge). On en déduit alors que la série $\displaystyle\sum_{n=0}^{+\infty} f((n+1)\pi)$ est également convergente. On a de plus, f étant décroissante, pour tout $n \in \mathbb{N}$,

$$\int_{(n+1)\pi}^{(n+2)\pi} f(t)dt \leq \pi f((n+1)\pi),$$

et donc la série $\displaystyle\sum_{n=0}^{+\infty}\int_{(n+1)\pi}^{(n+2)\pi} f(t)dt$ est convergente, ce qui nous donne encore une fois d'après la propriété VI.1, étant donné que f est positive et continue par morceaux, que l'intégrale $\displaystyle\int_0^{+\infty} f(t)dt$ est convergente.

Correction de l'exercice VI.4 :

On commence par regarder la fonction à intégrer. On voit que ce qui pourrait faire diverger $\displaystyle\int_0^{+\infty} e^{-t^2|\sin t|}dt$ est que la puissance de l'exponentielle soit proche de 0 trop longtemps. On veut donc étudier les endroits où cette puissance est proche de 0. Ces endroits apparaissent périodiquement aux points $\{n\pi,\ n \in \mathbb{N}\}$.



On voit que la surface entre deux pics pics qui apparaissent sur la courbe de $t \longmapsto e^{-t^2|\sin t|}$ devient de plus en plus petite, plus on avance sur l'axe des abscisses. On veut donc voir si cette surface rétrécit suffisamment rapidement pour l'intégrale de cette fonction soit convergente. Découpons alors l'intégrale

(cc) BY-NC-SA

en morceaux, et étudions-les. On a pour tout $n \in \mathbb{N}^*$,

$$\begin{aligned}\int_{n\pi}^{(n+1)\pi} e^{-t^2|\sin t|}dt \leq \int_{n\pi}^{(n+1)\pi} e^{-n^2\pi^2|\sin t|}dt &= \int_0^{\pi} e^{-n^2\pi^2|\sin t|}dt \\ &\underset{(*)}{=} 2\int_0^{\frac{\pi}{2}} e^{-n^2\pi^2|\sin t|}dt \underset{|\sin t|\geq \frac{2}{\pi}|t|}{\leq} 2\int_0^{\frac{\pi}{2}} e^{-2n^2\pi t}dt \\ &= \frac{1}{n^2\pi}\left(1-e^{-n^2\pi^2}\right) = \mathrm{O}\left(\frac{1}{n^2}\right).\end{aligned}$$

L'égalité $(*)$ est vraie car la fonction $t \longmapsto e^{-n^2\pi^2|\sin t|}$ est symétrique par rapport à l'axe d'équation $x = \frac{\pi}{2}$.On en déduit donc que la série $\sum_{n=1}^{+\infty}\int_{n\pi}^{(n+1)\pi} e^{-t^2|\sin t|}dt$ est convergente (car la série $\sum_{n=1}^{+\infty}\frac{1}{n^2}$ est convergente) et qu'alors d'après la proposition VI.1, étant donné que la fonction $t \longmapsto e^{-t^2|\sin t|}$ est positive continue par morceaux sur $[0, +\infty[$ et y admet une intégrale généralisée, l'intégrale $\int_0^{+\infty} e^{-t^2|\sin t|}dt$ converge.

Correction de l'exercice VII.2 :

Par hypothèse, f' est intégrable sur $[0, +\infty[$, et donc f' y admet une intégrale généralisée, *i.e.* la fonction $x \longmapsto \int_0^x f'(t)dt$ admet une limite finie en $+\infty$, ce qui revient à dire que f admet une limite finie en $+\infty$. De plus, f admet une intégrale généralisée sur $[0, +\infty[$ et est continue par morceaux, ce qui nous permet de dire d'après la proposition VII.1 que $f(x) \xrightarrow[x\to+\infty]{} 0$.

Correction de l'exercice VII.3 :

Tout d'abord, on peut facilement montrer que f est positive. En effet, s'il existe $y \geq 0$ tel que $f(y) < 0$, alors pour tout $x \geq y$,

$$\int_y^x f(t)dt \leq (x-y)f(y) \xrightarrow[x\to+\infty]{} -\infty,$$

ce qui est impossible car f admet une intégrale généralisée sur $[0, +\infty[$. Montrons maintenant que $xf(x) \xrightarrow[x\to+\infty]{} 0$. Pour cela, on utilise le critère de Cauchy. Soit $\varepsilon > 0$. D'après le théorème II.3, étant donné que f est continue par morceaux sur $[0, +\infty[$, il existe $\mathrm{A} > 0$ tel que

$$\forall x, y \in [\mathrm{A}, +\infty[,\ \left|\int_y^x f(t)dt\right| \leq \varepsilon.$$

En substituant y par $\frac{x}{2}$ et en supposant que $x > 2\mathrm{A}$, on obtient

$$\left(x - \frac{x}{2}\right) f(x) \underset{\substack{\uparrow \\ f \text{ décroissante}}}{\leq} \int_{\frac{x}{2}}^x f(t)dt \leq \varepsilon,$$

ce qui donne $xf(x) \leq 2\varepsilon$. On en déduit donc que

$$\forall \varepsilon > 0,\ \exists \mathrm{A} > 0\ \forall x > 2\mathrm{A},\ 0 \leq xf(x) \leq \varepsilon,$$

c'est-à-dire que $xf(x) \xrightarrow[x\to+\infty]{} 0$.

Correction de l'exercice VII.4 :

L'hypothèse d'uniforme continuité de f impose que cette dernière ne peut pas varier trop brusquement.

(cc) BY-NC-SA

Maintenant, essayons de montrer que f tend vers 0 en $+\infty$. Supposons que ce n'est pas le cas, *i.e.*

$$\exists \varepsilon > 0,\ \forall \mathrm{A} > 0,\ \exists x > \mathrm{A},\ |f(x)| > \varepsilon.$$

Soit $\varepsilon > 0$ vérifiant la proposition ci-dessus. Ceci signifie que peu importe à quel point on va loin selon l'axe des abscisses, on va trouver des points où $|f|$ est supérieure ou égale à ε. On peut alors construire une suite (x_n) croissante à valeurs positives qui diverge vers $+\infty$ telle que pour tout $n \in \mathbb{N}^*$, $|f(x_n)| > \varepsilon$. Sans perte de généralité, quitte à remplacer f par $-f$, on suppose qu'une infinité de termes de la suite (x_n) vérifient $f(x_n) > 0$ (car tous les termes sont soit positifs ou négatifs, donc il existe une infinité de termes de signe positif ou une infinité de termes de signe négatif). Quitte donc à extraire de la suite (x_n), on suppose que pour tout $n \in \mathbb{N},\ f(x_n) > \varepsilon$. On va maintenant exploiter le fait que f ne varie pas trop brusquement pour trouver des intervalles au voisinage des termes de la suite (x_n) tels que f y est assez éloignée de 0 pour que la surface sous sa courbe sur cet intervalle soit non négligeable. Voici comment on va procéder. Soit $\eta > 0$ tel que pour

$$\forall x, y \in \mathbb{R}^+,\ |x-y| < \eta \implies |f(x) - f(y)| < \frac{\varepsilon}{2}.$$

On en déduit alors directement que pour tout $n \in \mathbb{N}^*$,

$$\int_{x_n-\eta}^{x_n+\eta} f(t)dt \geq 2\eta\left(f(x_n) - \frac{\varepsilon}{2}\right) \geq 2\eta\left(\varepsilon - \frac{\varepsilon}{2}\right) = \eta\varepsilon.$$

Ceci est en contradiction avec le critère de Cauchy (Théorème II.3). En effet, posons $\varepsilon' = \dfrac{\eta\varepsilon}{2}$. f est continue par morceaux sur $[0, +\infty[$. Si elle y admet une intégrale généralisée, alors d'après le critère de Cauchy, on peut affirmer qu'il existe $\mathrm{M} > 0$ tel que pour tout $x, y \in [\mathrm{M}, +\infty[$ tels que

$$\left|\int_x^y f(t)dt\right| < \frac{\eta\varepsilon}{2}.$$

On en déduit alors que pour $n \in \mathbb{N}^*$ assez grand pour qu'on ait $x_n > \mathrm{M} + \eta$,

$$\eta\varepsilon \leq \left|\int_{x_n-\eta}^{x_n+\eta} f(t)dt\right| \leq \frac{\eta\varepsilon}{2},$$

ce qui est absurde. On en déduit donc qu'on a effectivement $f(x) \xrightarrow[x\to+\infty]{} 0$.

Correction de l'exercice VII.5 :

On va montrer que f est uniformément continue et conclue via l'exercice précédent. On a pour tout $x, y \in \mathbb{R}^+$,

$$|f(x) - f(y)| = \left|\int_x^y f'(t)dt\right| \underset{\substack{\uparrow \\ \text{Cauchy-Schwarz}}}{\leq} \left|\int_x^y 1dt\right|^{\frac{1}{2}} \left|\int_x^y f'(t)^2dt\right|^{\frac{1}{2}} \leq \sqrt{|y-x|}\underbrace{\sqrt{\int_0^{+\infty} f'(t)^2dt}}_{\mathrm{C}}.$$

Soit $\varepsilon > 0$. Pour tout $x, y \in \mathbb{R}^+$, on a

$$|x-y| < \varepsilon^2 \implies |f(x) - f(y)| \leq \mathrm{C}\sqrt{\varepsilon^2} = \mathrm{C}\varepsilon.$$

On en déduit alors que f est uniformément continue, et on peut finalement conclure en utilisant l'exercice précédent.

(cc) BY-NC-SA

Chapitre 10

Familles sommables

Dans tout ce chapitre I désigne un ensemble non vide ($\mathrm{I} = \varnothing$ étant trivial, en utilisant la convention $\sum\limits_{i\in\varnothing} u_i = 0$). On adopte également les conventions et notations suivantes.

- → On note $\overline{\mathbb{R}^+} = \mathbb{R}^+ \cup \{+\infty\}$.
- → On convient que pour tout réel $x \in \overline{\mathbb{R}^+}$, dans $\overline{\mathbb{R}^+}$, $x + \infty = +\infty + x = +\infty$.
- → On convient aussi que pour tout $\mathrm{X} \subset \overline{\mathbb{R}^+}$, $\mathrm{X} = \varnothing$ si et seulement si $\sup \mathrm{X} = -\infty$, et si $+\infty \in \mathrm{X}$ alors $\sup \mathrm{X} = +\infty$.
- → Finalement, si $\sum\limits_{n\geq 0} u_n$ est une série à termes positifs non convergente, on posera $\sum\limits_{n=0}^{+\infty} u_n = +\infty$.

I Familles à termes positifs

Définition I.1.

Soit $(u_i)_{i\in\mathrm{I}}$ une famille de réels positifs. On dit que $(u_i)_{i\in\mathrm{I}}$ est sommable s'il existe $\mathrm{M} \geq 0$ tel que pour toute partie finie J de I,

$$\sum_{i\in\mathrm{J}} u_i \leq \mathrm{M}.$$

Dans ce cas, la somme de $(u_i)_{i\in\mathrm{I}}$, notée $\sum\limits_{i\in\mathrm{I}} u_i$, est $\sup\limits_{\substack{\mathrm{J}\subset\mathrm{I} \\ \mathrm{J}\text{ est finie}}} \left(\sum\limits_{i\in\mathrm{J}} u_i\right)$.

Convention.

- → Si $(u_i)_{i\in\mathrm{I}}$ n'est pas sommable, on convient que $\sum\limits_{i\in\mathrm{I}} u_i = +\infty$.
- → Soit $(u_i)_{i\in\mathrm{I}}$ et $(v_j)_{j\in\mathrm{J}}$ deux familles de termes réels positifs. On dit que, dans $\overline{\mathbb{R}^+}$ on a $\sum\limits_{i\in\mathrm{I}} u_i = \sum\limits_{i\in\mathrm{J}} v_i$, lorsque l'une des deux familles est sommable, si et seulement si l'autre l'est et que si elles le sont, alors leurs sommes sont égales.

Proposition I.2.

Soit $(u_i)_{i\in\mathrm{I}}$ une famille de réels positifs. Si $(u_i)_{i\in\mathrm{I}}$ est sommable, alors $\mathrm{K} := \{i \in \mathrm{I} \mid u_i > 0\}$ est au plus dénombrable.

Démonstration. Soit $n \in \mathbb{N}^*$. Supposons que $\mathrm{A}_{1/n} := \left\{ i \in \mathrm{I} \;\middle|\; u_i \geq \dfrac{1}{n} \right\}$ soit infini. Soit J une partie finie de $\mathrm{A}_{1/n}$ telle que $|\mathrm{J}| > n \sum\limits_{i\in\mathrm{I}} u_i$. Alors $\sum\limits_{i\in\mathrm{J}} u_i \geq \dfrac{|\mathrm{J}|}{n} > \sum\limits_{i\in\mathrm{I}} u_i$ ce qui est faux.

Donc pour tout $n \in \mathbb{N}^*$, $\mathrm{A}_{1/n}$ est fini, donc $\mathrm{K} = \bigcup\limits_{n\in\mathbb{N}^*} \mathrm{A}_{1/n}$ au plus dénombrable.

Dans toute la suite du chapitre, on suppose que I est au plus dénombrable.

(cc) BY-NC-SA

Proposition I.3.

Soient $(u_i)_{i\in\mathrm{I}}$ et $(v_i)_{i\in\mathrm{I}}$ deux familles de réels positifs.

1. Si $(v_i)_{i\in\mathrm{I}}$ est sommable et que pour tout $i\in\mathrm{I}$, $u_i\leq v_i$, alors $(u_i)_{i\in\mathrm{I}}$ est sommable et la somme vérifie $\sum\limits_{i\in\mathrm{I}}u_i\leq\sum\limits_{i\in\mathrm{I}}v_i$.
2. Si $(u_i)_{i\in\mathrm{I}}$ n'est pas sommable et que pour tout $i\in\mathrm{I}$, $u_i\leq v_i$, alors $(v_i)_{i\in\mathrm{I}}$ n'est pas sommable.
3. Soit $(\mathrm{J}_n)_{n\in\mathrm{N}}$ une suite croissante de parties finies de I telle que $\bigcup\limits_{n\in\mathrm{N}}\mathrm{J}_n=\mathrm{I}$. Alors $(u_i)_{i\in\mathrm{I}}$ est sommable si, et seulement si, la suite de terme général $\mathrm{S}_n:=\sum\limits_{i\in\mathrm{J}_n}u_i$ est majorée.

   Dans ce cas, $\mathrm{S}_n\xrightarrow[n\to+\infty]{}\sum\limits_{i\in\mathrm{I}}u_i$.
4. Soit $\lambda\geq 0$. Supposons que $(u_i)_{i\in\mathrm{I}}$ et $(v_i)_{i\in\mathrm{I}}$ soient sommables. Alors $(\lambda u_i)_{i\in\mathrm{I}}$ et $(u_i+v_i)_{i\in\mathrm{I}}$ sont sommables et
$$\sum_{i\in\mathrm{I}}\lambda u_i=\lambda\sum_{i\in\mathrm{I}}u_i\qquad\text{et}\qquad\sum_{i\in\mathrm{I}}(u_i+v_i)=\sum_{i\in\mathrm{I}}u_i+\sum_{i\in\mathrm{I}}v_i.$$

Démonstration.

1. Soit $\mathrm{J}\subset\mathrm{I}$ fini. On a alors
$$\sum_{j\in\mathrm{J}}u_j\leq\sum_{j\in\mathrm{J}}v_j\leq\sum_{i\in\mathrm{I}}v_i<+\infty.$$
Cette borne étant indépendante de J, on déduit que la famille $(u_i)_{i\in\mathrm{I}}$ est sommable et que de plus
$$\sum_{i\in\mathrm{I}}u_i\leq\sum_{i\in\mathrm{I}}v_i.$$
2. C'est simplement la contraposée du point précédent.
3. $(\Rightarrow)$ Pour tout $n\in\mathbb{N}$, $\mathrm{S}_n\leq\sum\limits_{i\in\mathrm{I}}u_i<+\infty$ et cette borne est indépendante de n.

   $(\Leftarrow)$ Soit J une partie finie de I. Il existe $\mathrm{N}\in\mathbb{N}$ tel que $\mathrm{J}\subset\mathrm{J}_\mathrm{N}$ donc en notant K un majorant de $(\mathrm{S}_n)_{n\in\mathbb{N}}$
$$\sum_{i\in\mathrm{J}}u_i\leq\mathrm{S}_\mathrm{N}\leq\mathrm{K},$$
donc $(u_i)_{i\in\mathrm{I}}$ est sommable, ce qui achève la démonstration de l'équivalence.

   Supposons maintenant que $(u_i)_{i\in\mathrm{I}}$ est sommable. Alors, pour tout $n\in\mathbb{N}$, $\mathrm{S}_n\leq\sum\limits_{i\in\mathrm{I}}u_i$. Croissante et majorée, (S_n) converge vers $\ell\leq\sum\limits_{i\in\mathrm{I}}u_i$. Par ailleurs, pour toute partie finie J de I, il existe $\mathrm{N}_\mathrm{J}\in\mathbb{N}$ tel que pour tout entier $n\geq\mathrm{N}_\mathrm{J}$, $\sum\limits_{i\in\mathrm{J}}u_i\leq\mathrm{S}_{\mathrm{N}_\mathrm{J}}\leq\mathrm{S}_n\leq\ell$. Donc $\sum\limits_{i\in\mathrm{I}}u_i=\sup\limits_{\substack{\mathrm{J}\subset\mathrm{I}\\ \mathrm{J}\text{ est finie}}}\left(\sum\limits_{i\in\mathrm{J}}u_i\right)\leq\ell$. Donc
$$\mathrm{S}_n\xrightarrow[n\to+\infty]{}\sum_{i\in\mathrm{I}}u_i.$$
4. Considérons une suite croissante $(\mathrm{J}_n)_{n\in\mathbb{N}}$ de parties finies de I telle que $\bigcup\limits_{n\in\mathbb{N}}\mathrm{J}_n=\mathrm{I}$ et posons, pour tout $n\in\mathbb{N}$, $\mathrm{S}_n=\sum\limits_{i\in\mathrm{J}_n}u_i$ et $\mathrm{S}'_n=\sum\limits_{i\in\mathrm{J}_n}v_i$. D'après la proposition I.3, $(\mathrm{S}_n)_{n\in\mathbb{N}}$ et $(\mathrm{S}'_n)_{n\in\mathbb{N}}$ sont

(cc) BY-NC-SA

croissantes et de limites respectives $\sum_{i\in I} u_i$ et $\sum_{i\in I} v_i$, donc $\sum_{i\in J_n}(u_i+v_i) = S_n + S'_n \xrightarrow[n\to+\infty]{} \sum_{i\in I} u_i + \sum_{i\in I} v_i$, donc $(S_n + S'_n)_{n\in\mathbb{N}}$ est majorée. D'après la proposition I.3, $(u_i+v_i)_{i\in I}$ est sommable, et sa somme vaut bien $\sum_{i\in I} u_i + \sum_{i\in I} v_i$.

Il est aisé de vérifier, au moyen de la définition, que $(\lambda u_i)_{i\in\mathbb{N}}$ est sommable.

Proposition I.4.

Soient $\sum_{n\geq 0} u_n$ une série à termes positifs, et $\sigma \in \mathfrak{S}(\mathbb{N})$.

Alors $\sum_{n\geq 0} u_n$ converge si, et seulement si, $\sum_{n\geq 0} u_{\sigma(n)}$ converge. Dans ce cas, $\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} u_{\sigma(n)}$.

Démonstration. ($\Rightarrow$) Soit $N \in \mathbb{N}$. Posons $M = \max(\sigma(0), \ldots, \sigma(N))$. Alors

$$\sum_{n=0}^{N} u_{\sigma(n)} \leq \sum_{m=0}^{M} u_m \leq \sum_{m=0}^{+\infty} u_m < +\infty.$$

$\sum_{n\geq 0} u_{\sigma(n)}$ étant à termes positifs, elle converge. Retenons de plus que $\sum_{n=0}^{+\infty} u_{\sigma(n)} \leq \sum_{n=0}^{+\infty} u_n$.

($\Leftarrow$) Posons, pour tout $n \in \mathbb{N}$, $v_n = u_{\sigma(n)}$. Alors, $v_{\sigma^{-1}(n)} = u_n$. Ainsi, $\sum_{n\geq 0} v_n$ converge, et $\sigma^{-1} \in \mathfrak{S}(\mathbb{N})$ donc d'après ($\Rightarrow$), $\sum_{n\geq 0} u_n$ converge, et retenons que $\sum_{n=0}^{+\infty} u_n \leq \sum_{n=0}^{+\infty} u_{\sigma(n)}$.

Donc, dans ce cas, $\sum_{n=0}^{+\infty} u_n = \sum_{n=0}^{+\infty} u_{\sigma(n)}$.

Théorème (Théorème de sommation par paquets) I.5.

Soient $(I_n)_{n\in\mathbb{N}}$ un partage de I (*i.e.* partition où l'ensemble vide est autorisé), et $(u_i)_{i\in I}$ une famille de réels positifs. $(u_i)_{i\in I}$ est sommable si, et seulement si, pour tout $n \in \mathbb{N}$, $(u_i)_{i\in I_n}$ est sommable et, en notant $U_n = \sum_{i\in I_n} u_i$, $\sum_{n\geq 0} U_n$ est convergente. Dans ce cas, $\sum_{i\in I} u_i = \sum_{n=0}^{+\infty} U_n$.

Démonstration.

→ ($\Leftarrow$) Soit J une partie finie de I. Chaque élément de J appartient à l'un des I_n. J étant fini, il existe donc $N \in \mathbb{N}$ tel que $J \subset I_0 \sqcup \cdots \sqcup I_N$. Il vient que $J = (J \cap I_0) \sqcup \cdots \sqcup (J \cap I_N)$. Ainsi

$$\sum_{i\in J} u_i = \sum_{n=0}^{N}\left(\sum_{i\in J\cap I_n} u_i\right) \leq \sum_{n=0}^{N} U_n \leq \sum_{n=0}^{+\infty} U_n < +\infty,$$

donc $(u_i)_{i\in I}$ est sommable et $\sum_{i\in I} u_i \leq \sum_{n=0}^{+\infty} U_n$.

→ ($\Rightarrow$) Soit $n \in \mathbb{N}$. Soit J une partie finie de I_n. Alors J est une partie finie de I, donc $\sum_{i\in J} u_i \leq \sum_{i\in I} u_i$ donc $(u_i)_{i\in I_n}$ est sommable.

(cc) BY-NC-SA

Soit $\varepsilon > 0$. Pour tout $n \in \mathbb{N}$, il existe une partie finie J_n de I_n telle que

$$\sum_{i \in \mathrm{J}_n} u_i \geq \mathrm{U}_n - \frac{\varepsilon}{2^n}.$$

Soit $\mathrm{N} \in \mathbb{N}$. Posons $\mathrm{J}'_\mathrm{N} = \mathrm{J}_0 \sqcup \cdots \sqcup \mathrm{J}_\mathrm{N}$. Ainsi,

$$\sum_{i \in \mathrm{I}} u_i \geq \sum_{i \in \mathrm{J}'_\mathrm{N}} u_i = \sum_{k=0}^{\mathrm{N}} \left(\sum_{i \in \mathrm{J}_k} u_i \right) \geq \sum_{k=0}^{\mathrm{N}} \mathrm{U}_k - \varepsilon \underbrace{\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^\mathrm{N}}\right)}_{\leq 2},$$

donc $2\varepsilon + \sum_{i \in \mathrm{I}} u_i \geq \sum_{k=0}^{\mathrm{N}} \mathrm{U}_k$. $\sum_{k \geq 0} \mathrm{U}_k$ est alors convergente et en passant à la limite $\varepsilon \longrightarrow 0^+$, il vient que $\sum_{i \in \mathrm{I}} u_i \geq \sum_{k=0}^{\mathrm{N}} \mathrm{U}_k$.

Finalement, en passant à la limite $\mathrm{N} \longrightarrow +\infty$, on déduit que $\sum_{i \in \mathrm{I}} u_i \geq \sum_{k=0}^{+\infty} \mathrm{U}_k$.

Remarquons que, dans le cas où $(u_i)_{i \in \mathrm{I}}$ est sommable, on a déjà montré l'égalité $\sum_{i \in \mathrm{I}} u_i = \sum_{n=0}^{+\infty} \mathrm{U}_n$.

Remarques.

1. Il est aisé de voir que l'égalité $\sum_{i \in \mathrm{I}} u_i = \sum_{n=0}^{+\infty} \mathrm{U}_n$ est toujours vraie, dans $\overline{\mathbb{R}^+}$, même dans le cas de non-sommabilité de $(u_i)_{i \in \mathrm{I}}$. En effet, supposons que $(u_i)_{i \in \mathrm{I}}$ n'est pas sommable. Par ce qui précède, on peut voir que
   - → Soit il existe $n \in \mathbb{N}$ tel que $(u_i)_{i \in \mathrm{I}_n}$ n'est pas sommable, auquel cas $\mathrm{U}_n = +\infty$ et donc $\sum_{k=0}^{+\infty} \mathrm{U}_k = +\infty$,
   - → Soit à ce que tous les $(u_i)_{i \in \mathrm{I}_n}$ soient sommables, mais que la série $\sum_{n \geq 0} \mathrm{U}_n$ soit divergente, auquel cas $\sum_{n \geq 0} \mathrm{U}_n = +\infty$.
2. On remarquera que si $\mathrm{I} = \mathbb{N}$ et $\mathrm{I}_k = \{k\}$, on obtient une identification, dont on usera souvent par abus de notation, entre série (à termes positifs) convergente (resp. non convergente) et famille sommable (resp. non sommable) sur $\mathbb{N}$. Plus exactement, la série $\sum_{n \geq 0} u_n$ est convergente si, et seulement si, la famille $\sum_{n \in \mathbb{N}} u_n$ est sommable. De plus, avec ou sans convergence de $\sum_{n \geq 0} u_n$, on a l'égalité $\sum_{n \in \mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n$ dans $\overline{\mathbb{R}^+}$.

Application. un cas très fréquent où le théorème ci-dessus s'avère utile est le suivant.
Soit $(a_{m,n})_{(m,n) \in \mathbb{N}^2} \in {\mathbb{R}^+}^{\mathbb{N}^2}$. On a dans $\overline{\mathbb{R}^+}$,

$$\sum_{n \in \mathbb{N}} \sum_{m \in \mathbb{N}} a_{m,n} = \sum_{m \in \mathbb{N}} \sum_{n \in \mathbb{N}} a_{m,n},$$

c'est-à-dire que si l'une des sommes ci-dessus est infinie, alors l'autre l'est également, et si l'une des sommes est finie, alors l'autre l'est également et les deux sommes sont égales. Le cas des sommes finies correspond au cas où $(a_{m,n})_{(m,n) \in \mathbb{N}^2} \in {\mathbb{R}^+}^{\mathbb{N}^2}$ est sommable.

(cc) BY-NC-SA

Démonstration. Posons $\mathrm{I}_m = \{(m,k),\ k \in \mathbb{N}\}$ et $\mathrm{I}'_n = \{(k,n),\ k \in \mathbb{N}\}$. On a alors dans $\overline{\mathbb{R}^+}$, en utilisant le théorème de sommation par paquets,

$$\sum_{n\in\mathbb{N}}\sum_{m\in\mathbb{N}} a_{m,n} = \sum_{n\in\mathbb{N}}\sum_{(m',n')\in\mathrm{I}'_n} a_{m',n'} = \sum_{(m,n)\in\mathbb{N}^2} a_{m,n} = \sum_{m\in\mathbb{N}}\sum_{(m',n')\in\mathrm{I}_m} a_{m',n'} = \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} a_{m,n}.$$

On rappelle encore une fois qu'il est bien de voir ces inégalités dans les ceux cas : si la famille $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable, alors le théorème de sommation par paquet nous permet d'affirmer que toutes les sommes ci-dessus sont finies et égales et lorsqu'elle n'est pas sommable, alors le premier point de la remarque précédente nous permet d'affirmer que toutes ces sommes sont infinies.
Par ailleurs, rappelons qu'on peut également appliquer le second point de la remarque précédente en écrivant

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} a_{m,n} = \sum_{n\in\mathbb{N}}\sum_{m\in\mathbb{N}} a_{m,n} = \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} a_{m,n} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_{m,n}. \tag{7}$$

II Familles de nombres complexes

Remarque préliminaire. Toutes les démonstrations sautées dans cette partie se font de manière exactement similaire à celles analogues de la partie précédente.

Définition II.1.

On dit que la famille $(z_i)_{i\in\mathrm{I}} \in \mathbb{C}^{\mathrm{I}}$ est sommable lorsque la famille $(|z_i|)_{i\in\mathrm{I}} \in {\mathbb{R}^+}^{\mathrm{I}}$ l'est.

Proposition II.2.

La famille $(z_i)_{i\in\mathrm{I}} \in \mathbb{C}^{\mathrm{I}}$ est sommable si, et seulement si, les quatre familles de termes à valeurs positives $(\mathrm{Re}(z_i)^+)_{i\in\mathrm{I}}$, $(\mathrm{Re}(z_i)^-)_{i\in\mathrm{I}}$, $(\mathrm{Im}(z_i)^+)_{i\in\mathrm{I}}$ et $(\mathrm{Im}(z_i)^-)_{i\in\mathrm{I}}$ le sont.
En particulier, les deux notions coïncident pour des familles à termes positifs

Démonstration. Il suffit de remarquer que pour tout $z \in \mathbb{C}$, on a

$$\left.\begin{matrix}\mathrm{Re}(z)^+\\ \mathrm{Re}(z)^-\\ \mathrm{Im}(z)^+\\ \mathrm{Im}(z)^-\end{matrix}\right\} \leq |z| \leq \mathrm{Re}(z)^+ + \mathrm{Re}(z)^- + \mathrm{Im}(z)^+ + \mathrm{Im}(z)^-.$$

Définition II.3.

Si la famille $(z_i)_{i\in\mathrm{I}} \in \mathbb{C}^{\mathrm{I}}$ est sommable, on définit alors sa somme par

$$\sum_{i\in\mathrm{I}} z_i = \sum_{i\in\mathrm{I}} \mathrm{Re}(z_i)^+ - \sum_{i\in\mathrm{I}} \mathrm{Re}(z_i)^- + i\sum_{i\in\mathrm{I}} \mathrm{Im}(z_i)^+ - i\sum_{i\in\mathrm{I}} \mathrm{Im}(z_i)^-.$$

Remarque. Remarquons que cette définition est cohérente car on a pour tout nombre complexe $z \in \mathbb{C}$,

$$|z| = \mathrm{Re}(z) + i\mathrm{Im}(z) = \mathrm{Re}(z)^+ - \mathrm{Re}(z)^- + i\mathrm{Im}(z)^+ - i\mathrm{Im}(z)^-.$$

BY-NC-SA

Proposition II.4.

Soit $\lambda \in \mathbb{C}$. Supposons que $(u_i)_{i\in\mathrm{I}}$ et $(v_i)_{i\in\mathrm{I}}$ soient sommables. Alors $(\lambda u_i)_{i\in\mathrm{I}}$ et $(u_i+v_i)_{i\in\mathrm{I}}$ sont sommables et

$$\sum_{i\in\mathrm{I}}\lambda u_i = \lambda\sum_{i\in\mathrm{I}}u_i \qquad \text{et} \qquad \sum_{i\in\mathrm{I}}(u_i+v_i) = \sum_{i\in\mathrm{I}}u_i + \sum_{i\in\mathrm{I}}v_i.$$

Démonstration. Il suffit d'utiliser la définition de la somme d'une famille de nombres complexes

$$\sum_{i\in\mathrm{I}}u_i = \sum_{i\in\mathrm{I}}\mathrm{Re}(u_i)^+ - \sum_{i\in\mathrm{I}}\mathrm{Re}(u_i)^- + i\sum_{i\in\mathrm{I}}\mathrm{Im}(u_i)^+ - i\sum_{i\in\mathrm{I}}\mathrm{Im}(u_i)^-.$$

et de montrer l'égalité pour chacune des sommes en utilisant la démonstration du point 4 de la proposition I.3.

Théorème II.5.

Soient $(\mathrm{I}_n)_{n\in\mathbb{N}}$ un partage de I (*i.e.* partition où l'ensemble vide est autorisé), et $(z_i)_{i\in\mathrm{I}}$ une famille sommable de complexes. Pour tout $n\in\mathbb{N}$, $(z_i)_{i\in\mathrm{I}_n}$ est sommable et, en notant $\mathrm{U}_n = \sum_{i\in\mathrm{I}_n}u_i$, $\sum_{n\geq 0}\mathrm{U}_n$ est absolument convergente. De plus, on a $\sum_{i\in\mathrm{I}}u_i = \sum_{n=0}^{+\infty}\mathrm{U}_n$.

Démonstration. Encore une fois, il suffit d'appliquer le théorème de sommation par paquets (I.5) aux familles de nombres réels positifs $\mathrm{Re}(z)^+, \mathrm{Re}(z)^-, \mathrm{Im}(z)^+$ et $\mathrm{Im}(z)^-$.

Remarques importantes.

→ Dans le cas précédent des familles positives, lorsqu'on considère une famille positive $(u_i)_{i\in\mathrm{I}}$, soit elle est sommable, et alors la somme $\sum_{i\in\mathrm{I}}u_i$ avait une valeur finie, ou alors, elle n'est pas sommable et donc on peut écrire sans ambiguïté $\sum_{i\in\mathrm{I}}u_i = +\infty$. Dans le cas des familles complexes, on ne peut plus faire cela sans ambiguïté. En effet, pour une famille de nombres complexes $(u_i)_{i\in\mathrm{I}}$, on a par définition,

$$\sum_{i\in\mathrm{I}}u_i = \underbrace{\sum_{i\in\mathrm{I}}\mathrm{Re}(u_i)^+}_{(1)} - \underbrace{\sum_{i\in\mathrm{I}}\mathrm{Re}(u_i)^-}_{(2)} + i\underbrace{\sum_{i\in\mathrm{I}}\mathrm{Im}(u_i)^+}_{(3)} - i\underbrace{\sum_{i\in\mathrm{I}}\mathrm{Im}(u_i)^-}_{(4)}.$$

Dans le cas par exemple d'une famille de nombres réels pas forcément positifs, les termes (3) et (4) sont nuls. Cette famille n'est pas sommable lorsque l'un les termes (1) ou (2) sont infinis. Dans le cas où les deux termes sont infinis, on obtient une expression de la forme $+\infty - \infty$ et donc ne peut pas définir de valeur pour $\sum_{i\in\mathrm{I}}u_i$.

→ Cette fois-ci, dans le cas des familles de nombres complexes, l'égalité $\sum_{n\geq 0}u_n = \sum_{n\in\mathbb{N}}u_n$ n'est plus toujours vraie et ni même a de sens sans sommabilité. Toutefois, on a bien équivalence entre **l'absolue convergence** de la série $\sum_{n\geq 0}u_n$ et la sommabilité de la famille $(u_n)_{n\in\mathbb{N}}$. De plus, mais cette fois uniquement lorsque la série de gauche est **absolument convergente** ou la famille de droite est sommable (on ne peut donc plus dire que la limite d'une série de nombres complexe et la somme de la famille associée sont égales dans $\overline{\mathbb{R}^+}$ comme on le faisait pour les familles positives), on a bien l'égalité $\sum_{n\geq 0}u_n = \sum_{n\in\mathbb{N}}u_n$.

On continuera donc à confondre série absolument convergente et famille sommable (qu'on appellera

CC BY-NC-SA

aussi série sommable) **mais uniquement après vérification de la sommabilité/convergence absolue**.

Un simple exemple où cette égalité est fausse est la famille $\left(\frac{(-1)^n}{n+1}\right)_{n\in\mathbb{N}}$. Cette famille n'est pas sommable, car en valeur absolue il s'agit des termes de la série harmonique, mais la série $\mathrm{S}_n = \sum_{n\geq 0} \frac{(-1)^n}{n+1}$ est convergente **mais n'est pas absolument convergente**. De plus, cette famille correspond au contrexemple du point précédent.

→ De même, la première remarque faite après le théorème de sommation par paquets n'est plus toujours valable. Rappelons cette remarque. pour toute famille $(u_i)_{i\in\mathrm{I}}$ de nombres positifs et tout partage $(\mathrm{I}_n)_{n\in\mathbb{N}}$ de I, l'égalité suivante est vraie dans $\overline{\mathbb{R}^+}$: $\sum_{n\geq 0} \mathrm{U}_n = \sum_{i\in\mathrm{I}} u_i$.

En effet, prenons le contre-exemple donné par la famille $(u_n)_{n\in\mathbb{N}} = ((-1)^n)_{n\in\mathbb{N}}$ et la partition $\mathrm{I}_n = \{2n, 2n+1\}$. Il est clair que la famille n'est pas sommable, alors que pour tout $n \in \mathbb{N}$, $\underbrace{\sum_{i\in\mathrm{I}_n} u_i}_{\mathrm{U}_n} = 0$ et alors $\sum_{n\geq 0} \mathrm{U}_n = 0$ ce qui clairement ne peut pas être égal à $\sum_{i\in\mathrm{I}} u_i$ car cette dernière n'est même pas définie.

III Applications.

1. Séries doubles

Soit $(z_{m,n})_{(m,n)\in\mathbb{N}^2} \in \mathbb{C}^{\mathbb{N}^2}$ une famille de nombres complexes sommable. Les propriétés suivantes sont vraies.

→ Pour tout $m \in \mathbb{N}$, la série $\sum_{n\geq 0} z_{m,n}$ est absolument convergente et converge vers une limite Z_m.

→ Pour tout $n \in \mathbb{N}$, la série $\sum_{m\geq 0} z_{m,n}$ est absolument convergente et converge vers une limite Z'_n.

→ Enfin, $\sum_{m\geq 0} \mathrm{Z}_m$ et $\sum_{n\geq 0} \mathrm{Z}'_n$ sont absolument convergentes et

$$\sum_{m\geq 0} \mathrm{Z}_m = \sum_{n\geq 0} \mathrm{Z}'_n = \sum_{(m,n)\in\mathbb{N}^2} z_{m,n}.$$

Ce résultat est une application directe du théorème de sommation par paquets pour les nombres complexes.

2. Produits

a) Produit de deux séries

Soit $(a_n)_{n\in\mathbb{N}}, (b_p)_{p\in\mathbb{N}}$ deux familles sommables. La famille $(u_{n,p})_{(n,p)\in\mathbb{N}^2} = (a_n b_p)_{(n,p)\in\mathbb{N}^2}$ est sommable, de somme

$$\sum_{(n,p)\in\mathbb{N}^2} u_{n,p} = \sum_{n=0}^{+\infty} a_n \sum_{p=0}^{+\infty} b_p.$$

(cc) BY-NC-SA

En effet, dans $\overline{\mathbb{R}^+}$, en utilisant la relation (7),

$$\sum_{(n,p)\in\mathbb{N}^2}|u_{n,p}| = \sum_{p=0}^{+\infty}\sum_{n=0}^{+\infty}|a_n b_p| = \sum_{p=0}^{+\infty}\left(|b_p|\underbrace{\sum_{n=0}^{+\infty}|a_n|}_{\text{constante}}\right) = \left(\sum_{n=0}^{+\infty}|a_n|\right)\left(\sum_{p=0}^{+\infty}|b_p|\right) < +\infty.$$

Ainsi $(u_{n,p})_{(n,p)\in\mathbb{N}^2}$ est sommable. On a donc le droit, par sommation par paquet, de refaire le même calcul et réutiliser la relation (7), sans valeurs absolues cette fois (tous les termes sont absolument convergents, donc il n'y a pas d'ambiguïté). Ainsi, on a bien le résultat

$$\sum_{(n,p)\in\mathbb{N}^2} u_{n,p} = \left(\sum_{n=0}^{+\infty}a_n\right)\left(\sum_{p=0}^{+\infty}b_p\right).$$

b) Produit de convolution (ou de Cauchy)

Soit $(a_m)_{m\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux familles de nombres complexes. On définit et note le produit de convolution (ou de Cauchy) $(c_p)_{p\in\mathbb{N}} = (a_m)_{m\in\mathbb{N}} * (b_n)_{n\in\mathbb{N}}$ de ces deux familles par

$$\forall p\in\mathbb{N},\ c_p = \sum_{m+n=p} a_m b_n = \sum_{m=0}^{p} a_m b_{p-m} = \sum_{n=0}^{p} a_{p-n} b_n.$$

Proposition III.1.

Si $\sum\limits_{m\geq 0} a_m$ et $\sum\limits_{n\geq 0} b_n$ sont absolument convergentes, alors $\sum\limits_{p\geq 0} c_p$ l'est aussi et on a l'égalité

$$\sum_{p=0}^{+\infty} c_p = \left(\sum_{m=0}^{+\infty}a_m\right)\left(\sum_{n=0}^{+\infty}b_n\right).$$

Démonstration. D'après ce qui précède, $(u_{m,n})_{(m,n)\in\mathbb{N}^2} = (a_m b_n)_{(m,n)\in\mathbb{N}^2}$ est sommable. Considérons la partition

$$\mathrm{I}_p = \{(m,n)\in\mathbb{N}^2,\ m+n=p\}.$$

On sait par le théorème de sommation par paquets qu'étant donné que $(\mathrm{I}_p)_{p\in\mathbb{N}}$ est un partage de $\mathbb{N}^2$, $(c_p)_{p\in\mathbb{N}} = \left(\sum\limits_{(m,n)\in\mathrm{I}_p} a_m b_n\right)_{p\in\mathbb{N}}$ est sommable et donc $\sum\limits_{p\geq 0} c_p$ est absolument convergente.

Finalement, toujours par le théorème de sommation par paquets, en plus du point précèdent,

$$\sum_{p=0}^{+\infty} c_p = \sum_{(m,n)\in\mathbb{N}^2} a_m b_n = \left(\sum_{m=0}^{+\infty}a_m\right)\left(\sum_{n=0}^{+\infty}b_n\right).$$

Remarque. Sous réserve d'avoir les bonnes hypothèses de convergence, le produit de Cauchy est particulièrement utile pour avoir une forme explicite du produit de deux séries entières, *i.e.* sous forme de polynôme de degré infini. En effet, pour exprimer le produit $\left(\sum\limits_{m=0}^{+\infty}a_m x^m\right)\left(\sum\limits_{n=0}^{+\infty}b_n x^n\right)$ sous forme de série entière lorsque les deux familles $(a_n x^n)$ et $(b_m x^m)$ sont sommables, il faut regrouper les termes de même degré de la forme $a_n b_m x^{n+m}$, d'où le fait de partitionner les termes par ensembles de la forme $\mathrm{I}_p = \{(m,n)\in\mathbb{N}^2,\ m+n=p\}$. On obtient dans ce cas l'égalité suivante

$$\left(\sum_{m=0}^{+\infty}a_m x^m\right)\left(\sum_{n=0}^{+\infty}b_n x^n\right) = \sum_{p=0}^{+\infty} c_p x^p.$$

(cc) BY-NC-SA

Chapitre 11.1

Espaces vectoriels normés et espaces métriques

I Espaces vectoriels normés

Dans toute cette partie, on considère $\mathbb{K}$ un corps égal à $\mathbb{R}$ ou $\mathbb{C}$ et E un $\mathbb{K}-$espace vectoriel.

Définition I.1.

Une norme sur E est une application $\mathrm{N} : \mathrm{E} \to \mathbb{R}^+$ vérifiant les trois conditions suivantes.

(N_1) $\forall x \in \mathrm{E},\ \mathrm{N}(x) = 0 \iff x = 0$ (séparation)

(N_2) $\forall (\lambda, x) \in \mathbb{K} \times \mathrm{E},\ \mathrm{N}(\lambda x) = |\lambda|\,\mathrm{N}(x)$ (homogénéité)

(N_3) $\forall (x, y) \in \mathrm{E}^2,\ \mathrm{N}(x+y) \leq \mathrm{N}(x) + \mathrm{N}(y)$ (sous-additivité)

Vocabulaire. une application $\mathrm{N} : \mathrm{E} \to \mathbb{R}^+$ vérifiant (N_2) et (N_3) est appelée semi-norme.

Proposition I.2.

1. Si $x \neq 0$, $\mathrm{N}(x) > 0$ et $\mathrm{N}\left(\dfrac{x}{\mathrm{N}(x)}\right) = 1$.
2. Si $\mathbb{K} = \mathbb{R}$ et $x \neq 0$, alors pour tout $\lambda \in \mathbb{K}$,
$$\mathrm{N}(\lambda x) = 1 \iff |\lambda| = \frac{1}{\mathrm{N}(x)} \iff \lambda \in \left\{-\frac{1}{\mathrm{N}(x)}, \frac{1}{\mathrm{N}(x)}\right\}.$$
L'espace $\mathbb{R}x$ contient donc deux vecteurs unitaires.
3. Pour tout $x \in \mathrm{E}$, $\mathrm{N}(x) = 1 \Longrightarrow \forall \theta \in \mathbb{R},\ \mathrm{N}(e^{i\theta}x) = 1$.
4. $\forall (x_1, \ldots, x_n) \in \mathrm{E}^n$, $\mathrm{N}\left(\displaystyle\sum_{i=1}^{n} x_i\right) \leq \displaystyle\sum_{i=1}^{n} \mathrm{N}(x_i)$.

Exemples. voici quelques normes usuelles pour $\mathrm{E} = \mathbb{K}^n$. On pose pour tout $x \in \mathrm{E}$, $x = (x_1, \ldots, x_n)$.

→ $\|x\|_\infty = \sup_{1 \leq k \leq n} |x_k|$

→ $\|x\|_1 = \displaystyle\sum_{k=1}^{n} |x_k|$

→ $\|x\|_2 = \left(\displaystyle\sum_{k=1}^{n} |x_k|^2\right)^{\frac{1}{2}}$

→ Pour tout $p \in \mathbb{N}^*$, $\|x\|_p = \left(\displaystyle\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}}$.

Remarque. La sous-additivité pour la norme $\|.\|_2$ n'est autre que l'inégalité de Minkowski.

Exercice I.3.

Montrer que pour tout $x \in \mathrm{E}$, $\|x\|_p \longrightarrow \|x\|_\infty$ lorsque p tend vers l'infini.

BY-NC-SA

Exemples. Pour $E = \mathcal{C}([a,b], \mathbb{C})$.

→ $\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|$

→ $\|f\|_1 = \int_a^b |f(x)|\, dx$

→ $\|f\|_2 = \left(\int_a^b |f(x)|^2\, dx\right)^{\frac{1}{2}}$

→ Pour $p \in \mathbb{N}^*$, $\|f\|_p = \left(\int_a^b |f(x)|^p\, dx\right)^{\frac{1}{p}}$.

Exemples. Quelques sous espaces normés de l'espace des suites

→ $E = \ell^\infty(\mathbb{N}, \mathbb{K}) = \{(u_n) \in \mathbb{K}^\mathbb{N},\ u_n \text{ est bornée}\}$, muni de $\|(u_n)\|_\infty = \sup\limits_{n\in\mathbb{N}} |u_n|$.

→ $E = \mathcal{C}_0(\mathbb{N}) = \{(u_n) \in \mathbb{K}^\mathbb{N},\ u_n \longrightarrow 0\}$ muni de $\|.\|_\infty$.

→ Pour $p \geq 1$, $\ell^p(\mathbb{N}, \mathbb{K}) = \{(u_n) \in \mathbb{K}^\mathbb{N},\ \sum\limits_{n=0}^{+\infty} |u_n|^p \text{ converge}\}$.

Remarque. Les inclusions suivantes sont vérifiées.

→ $\mathcal{C}_0(\mathbb{N}) \subset \ell^\infty(\mathbb{N}, \mathbb{K})$
En effet, si $u_n \xrightarrow[n\to+\infty]{} 0$, alors il existe $N \in \mathbb{N}$ pour lequel pour tout $n \geq N$, $|u_n| \leq 1$. Ceci donne pour tout $n \in \mathbb{N}$, $|u_n| \leq \max\{1, |u_0|, |u_1|, \dots, |u_N|\}$, donc $(u_n)_{n\in\mathbb{N}} \in \ell^\infty(\mathbb{N}, \mathbb{K})$.

→ $\ell^1(\mathbb{N}, \mathbb{K}) \subset \ell^2(\mathbb{N}, \mathbb{K})$
En effet, si on considère une suite $(u_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{N}, \mathbb{K})$ alors $u_n \longrightarrow 0$, ce qui implique qu'à partir d'un certain rang, $|u_n| \leq 1$ et alors $|u_n|^2 \leq |u_n|$. On trouve donc immédiatement que la convergence de la série $\sum\limits_{n=0}^{+\infty} |u_n|$ implique celle de $\sum\limits_{n=0}^{+\infty} |u_n|^2$, donc $(u_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}, \mathbb{K})$.

II Géométrie

Définition II.1.

Une partie A de E est dite convexe si pour tout $x, y \in A$ et $\lambda \in [0,1]$

$$\lambda x + (1-\lambda)y \in A.$$

Exercice II.2.

Montrer que tout sous-espace affine de E (*i.e.* les ensembles de la forme $a + F$ avec $a \in E$ et F sous espace vectoriel de E) est convexe.

Remarque. La boule unité fermée de E, $B_f(0,1) = \{x \in E,\ \ \|x\| \leq 1\}$ est convexe.
En effet, si $x, y \in B_f(0,1)$, alors pour tout $\lambda \in [0,1]$ on a

$$\|\lambda x + (1-\lambda)y\| \leq \lambda \|x\| + (1-\lambda) \|y\| \leq \lambda + 1 - \lambda = 1,$$

ce qui donne $\lambda x + (1-\lambda)y \in B_f(0,1)$.

Exercice II.3.

Soit $N : E \to \mathbb{R}$ qui vérifie (N_1) et (N_2). Soit $A = \{x \in E,\ N(x) \leq 1\}$. Montrer que si A est convexe, alors N vérifie (N_3).

BY-NC-SA

Exercice II.4.

Soit $\|.\|$ une norme sur E. Montrer que pour tous $x, y \in \mathrm{E} \setminus \{0\}$,

$$\|x-y\| \geq \frac{1}{2}\max(\|x\|, \|y\|)\left\|\frac{x}{\|x\|}-\frac{y}{\|y\|}\right\|.$$

Exercice II.5.

Une norme $\|.\|$ sur E est dite de somme stricte lorsque pour tous $x, y \in \mathrm{E}$

$$\|x+y\| = \|x\| + \|y\| \Longrightarrow (x, y) \text{ est liée.}$$

1. Parmi les normes suivantes, lesquelles sont de somme stricte ?
   (a) $\|.\|_p$ sur $\mathbb{K}^n$ avec $n \geq 2$ et $p \in \{1, 2, \infty\}$.
   (b) $\|.\|_p$ sur $\mathcal{C}([0,1], \mathbb{R})$ avec $p \in \{1, 2, \infty\}$.
2. Soit $\|.\|$ une norme sur E. Soit $x, y \in \mathrm{E}$. Montrer que si $\|x\| = \|y\| = 1$ et $\|x+y\| = 2$, alors $[x, y] \subset \mathrm{S}(0,1) = \{x \in \mathbb{E},\ \ \|x\| = 1\}$.
3. Montrer l'équivalence

$$\|.\| \text{ est stricte sur E} \Longleftrightarrow \mathrm{S}(0,1) \text{ ne contient pas de segment non trivial.}$$

III Espaces métriques

Définition III.1.

Soit X un ensemble. On dit qu'une application $d : \mathrm{X}^2 \to \mathbb{R}^+$ est une distance sur X lorsqu'elle vérifie les propriétés suivantes.

(D_1) $\forall (x, y) \in \mathrm{X}^2$, $d(x, y) = 0 \Leftrightarrow x = y$ (séparation)

(D_2) $\forall (x, y) \in \mathrm{X}^2$ $d(x, y) = d(y, x)$ (symétrie)

(D_3) $\forall (x, y, z) \in \mathrm{X}^3$ $d(x, y) \leq d(x, z) + d(z, y)$ (inégalité triangulaire)

Exemple 1. Soit $\|.\|$ une norme sur E. On appelle distance induite par $\|.\|$ sur E l'application $d : (x, y) \longmapsto \|x-y\|$.
La distance induite définit bien une distance. En effet, il est facile de montrer les implications suivantes :

→ $(\mathrm{N}_1) \Longrightarrow (\mathrm{D}_1)$

→ $(\mathrm{N}_2) \Longrightarrow (\mathrm{D}_2)$

→ $(\mathrm{N}_3) \Longrightarrow (\mathrm{D}_3)$

Exemple 2. On considère le cas où $\mathrm{X} = \ell^\infty(\mathbb{N}, \mathbb{K})$. L'application $d : ((x_n), (y_n)) \longrightarrow \sum_{n=0}^{+\infty} \frac{|x_n - y_n|}{2^n}$ définit bien une distance sur X.

BY-NC-SA

Proposition III.2.

Soit X une espace métrique et $d : X^2 \to \mathbb{R}^+$ une distance sur X. d vérifie les propriétés suivantes

1. $\forall n \in \mathbb{N}^* \setminus \{1\}$, $\forall (x_1, \ldots, x_n) \in X^n$, $d(x_1, x_n) \leq \sum_{i=1}^{n-1} d(x_i, x_{i+1})$.
2. $\forall (x, y, z) \in X^3$, $|d(x, y) - d(x, z)| \leq d(y, z)$.

Démonstration.

1. Il s'agit d'une simple récurrence. L'initialisation ($n = 2$) correspond à la propriété (D_3).
   Soit $n \in \mathbb{N}^* \setminus \{1\}$, supposons l'hypothèse de récurrence vérifiée :
   $$\forall n \in \mathbb{N}^* \setminus \{1\},\ \forall (x_1, \ldots, x_n) \in X^n,\ d(x_1, x_n) \leq \sum_{i=1}^{n-1} d(x_i, x_{i+1}).$$
   Soit $x_1, \ldots, x_{n+1} \in X$, on a
   $$d(x_1, x_{n+1}) \leq d(x_1, x_n) + d(x_n, x_{n+1}) \leq \sum_{i=1}^{n-1} d(x_i, x_{i+1}) + d(x_n, x_{n+1}) = \sum_{i=1}^{n} d(x_i, x_{i+1}).$$
   On en déduit donc immédiatement le résultat voulu.
2. Soit $(x, y, z) \in X^3$, il s'agit de montrer que
   (a) $d(x, y) - d(x, z) \leq d(y, z)$
   (b) $d(x, z) - d(x, y) \leq d(y, z)$
   Le point (a) est équivalent à $d(x, y) \leq d(x, z) + d(y, z)$, ce qui est vrai d'après (D_3). On peut dire la même chose du point (b).

IV Ensembles remarquables

Dans toute la suite X est espace métrique muni d'une distance d.

Définition IV.1.

Soit $(a, r) \in X \times]0, +\infty[$. On appelle boule ouverte de centre a et rayon r l'ensemble
$$B(a, r) = \{x \in X,\ d(x, a) < r\}.$$
De même, on appelle boule fermée de centre a et de rayon r l'ensemble
$$B_f(a, r) = \{x \in X,\ d(x, a) \leq r\}.$$
On appelle aussi sphère de centre a et de rayon r l'ensemble
$$S(a, r) = \{x \in X,\ d(x, a) = r\}.$$

Exemple. En espace vectoriel normé muni d'une norme $\|.\|$, lorsqu'on prend d égal à la distance induite par la norme, ces ensembles s'écrivent :

→ $B(a, r) = \{x \in X,\ \|x - a\| < r\}$

→ $B_f(a, r) = \{x \in X,\ \|x - a\| \leq r\}$

→ $S(a, r) = \{x \in X,\ \|x - a\| = r\}$

(cc) BY-NC-SA

Définition IV.2.

Soit $\mathrm{A} \subset \mathrm{X}$. On appelle distance induite sur A la restriction de d à A^2. Cette distance est notée d_{A}.

Notation. En substituant la distance d par d_{A} dans les définitions d'une boule et d'une sphère, on note pour $(a, r) \in \mathrm{A} \times \mathbb{R}^+$,

→ $\mathrm{B}_{\mathrm{A}}(a, r) = \{x \in \mathrm{A},\ d_{\mathrm{A}}(x, a) < r\} = \mathrm{B}(a, r) \cap \mathrm{A}$

→ $\mathrm{B}_{\mathrm{A},f}(a, r) = \{x \in \mathrm{A},\ d_{\mathrm{A}}(x, a) \leq r\} = \mathrm{B}_f(a, r) \cap \mathrm{A}$

→ $\mathrm{S}_{\mathrm{A}}(a, r) = \{x \in \mathrm{A},\ d_{\mathrm{A}}(x, a) = r\} = \mathrm{S}(a, r) \cap \mathrm{A}$

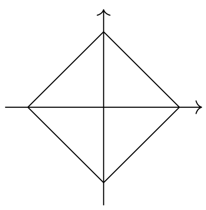
Exemples.

→ Pour $\mathrm{X} = \mathbb{R}$, $\mathrm{A} = [0, 1]$ et d la distance induite par la valeur absolue, $\mathrm{B}_{\mathrm{A}}(1, \frac{1}{2}) =]\frac{1}{2}, 1]$.

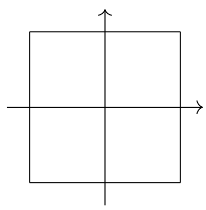
→ Pour $\mathrm{X} = \mathbb{C}$, $\mathrm{A} = \mathrm{S}(0, 1)$ et d la distance induite par le module, $\mathrm{B}_{\mathrm{A}}(i, 1) = \{e^{i\theta},\ \theta \in [\frac{\pi}{6}, \frac{5\pi}{6}]\}$.

Dessins en espace vectoriel normé : Voici des dessins de la boule unité $\mathrm{B}_f(0, 1)$ pour quelques exemples de X.

$\mathrm{X} = \mathbb{R}^2$ et d la distance induite par $\|.\|_1$.



$\mathrm{X} = \mathbb{R}^2$ et d la distance induite par $\|.\|_\infty$.



Proposition IV.3.

Lorsque X est un espace vectoriel normé muni d'une norme $\|.\|$ et de la distance induite par cette norme, on a les égalités suivantes pour tout $(a, r) \in \mathrm{X} \times \mathbb{R}_+^*$

1. $\mathrm{B}(a, r) = a + \mathrm{B}(0, r) = a + r\mathrm{B}(0, 1)$
2. $\mathrm{B}_f(a, r) = a + \mathrm{B}_f(0, r) = a + r\mathrm{B}_f(0, 1)$

Démonstration. Nous allons montrer le point (1). La démonstration du point 2 est identique : il suffit de remplacer les inégalités strictes par des inégalités larges.
Soit $(a, r) \in \mathrm{X} \times \mathbb{R}_+^*$, on a

$$\begin{aligned} x \in \mathrm{B}(a, r) &\iff \|x - a\| < r \\ &\iff \left\| \frac{1}{r}(x - a) \right\| < 1 \\ &\iff \frac{1}{r}(x - a) \in \mathrm{B}(0, 1) \\ &\iff x \in a + r\mathrm{B}(0, 1). \end{aligned}$$

(cc) BY-NC-SA

V Bornitude

Proposition V.1.

Soit A une partie de X. Les propriétés suivantes sont équivalentes.

1. $\{d(x, y),\ (x, y) \in \mathrm{A}^2\}$ est une partie bornée de $\mathbb{R}^+$.
2. Il existe une boule ouverte de X qui contient A.
3. Pour tout $a \in \mathrm{X}$, il existe r tel que $\mathrm{A} \subset \mathrm{B}(a, r)$.

Si A vérifie l'une de ces trois propriétés, on dit que A est borné et on note

$$\operatorname{diam}(\mathrm{A}) = \sup\{d(x, y),\ (x, y) \in \mathrm{A}^2\}.$$

Démonstration.

→ (1) $\Longrightarrow$ (2)
Soit $a \in \mathrm{A}$. L'ensemble $\mathrm{D} = \{d(x, y),\ (x, y) \in \mathrm{A}^2\}$ est une partie bornée de $\mathbb{R}$ donc admet une borne supérieure. On pose alors $r = 1 + \sup \mathrm{D}$. Pour tout $x \in \mathrm{A}$, $d(x, a) \leq \sup \mathrm{D} < r$.
On en déduit donc que $\mathrm{A} \subset \mathrm{B}(a, r)$.

→ (2) $\Longrightarrow$ (3)
Soit $\mathrm{B}(b, \varepsilon)$ une boule qui contient A. Soit $a \in \mathrm{X}$. On pose $r = \varepsilon + d(a, b)$.
Pour tout $x \in \mathrm{A}$, on a $d(x, b) \leq d(x, a) + d(a, b) < \varepsilon + d(a, b) = r$. Donc $\mathrm{A} \subset \mathrm{B}(a, r)$.

→ (3) $\Longrightarrow$ (1)
Soit $\mathrm{B}(a, r)$ une boule qui contient A. Pour tout $(x, y) \in \mathrm{A}^2$,

$$d(x, y) \leq d(x, a) + d(a, y) < 2r$$

L'ensemble D est alors majoré par $2r$.

Exercice V.2.

Soit E un espace vectoriel normé différent de $\{0\}$.

1. Montrer que pour tout $(a, r) \in \mathrm{E} \times \mathbb{R}^+$, $\operatorname{diam}(\mathrm{B}(a, r)) = \operatorname{diam}(\mathrm{B}_f(a, r))$.
2. Montrer que le rayon et le centre d'une boule ouverte ou fermée sont uniques.

Définition V.3.

Soit A une partie de X. On dit qu'une fonction $f : \mathrm{A} \to \mathrm{X}$ est bornée lorsque $f(\mathrm{A})$ est borné.

Notation. On désigne par $\mathcal{B}(\mathrm{A}, \mathrm{X})$ l'ensemble des fonctions bornées de A dans X.

Définition V.4.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans X. On dit que $(u_n)_{n\in\mathbb{N}}$ est bornée lorsque l'ensemble $\{u_n,\ n \in \mathbb{N}\}$ est borné.

BY-NC-SA

VI Suites dans un espace métrique et dans un espace vectoriel normé

Définition VI.1.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans X. On dit que (u_n) converge lorsque

$$\exists \ell \in \mathrm{X}\ \forall \varepsilon > 0\ \exists \mathrm{N} \in \mathbb{N}\ \forall n \geq \mathrm{N}\ d(u_n, \ell) \leq \varepsilon.$$

Proposition VI.2.

On suppose que X est un $\mathbb{K}-$espace vectoriel normé muni d'une norme $\|.\|$ et que d est la distance induite par cette norme.
Soit $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ deux suites à valeurs dans X.

1. Pour tout $(\ell, \ell') \in \mathbb{R}^2$, Si (u_n) converge vers ℓ et ℓ', alors $\ell = \ell'$.
2. pour tous $a, b, \ell, \ell' \in \mathbb{R}$, si $u_n \xrightarrow[n\to+\infty]{} \ell$ et $v_n \xrightarrow[n\to+\infty]{} \ell'$ alors $au_n + bv_n \xrightarrow[n\to+\infty]{} a\ell + b\ell'$.
3. Si $(\lambda_n)_{n\in\mathbb{N}}$ est une suite à valeurs dans $\mathbb{K}$ qui converge vers $\lambda \in \mathbb{K}$, et $(u_n)_{n\in\mathbb{N}}$ converge vers ℓ, alors la suite $(\lambda_n u_n)_{n\in\mathbb{N}}$ converge vers $\lambda\ell$.

VII Algèbres normées

Définition VII.1.

Soit $\mathcal{A}$ un ensemble. On dit que $(\mathcal{A}, +, ., *)$ est une $\mathbb{K}-$algèbre lorsque

→ $(\mathcal{A}, +, .)$ est un $\mathbb{K}-$espace vectoriel.

→ $*$ est une loi de composition interne.

→ $*$ est bilinéaire.

Vocabulaire.

→ Si $*$ est associative, on dit que $\mathcal{A}$ est une algèbre associative.

→ Si $*$ admet un élément neutre, on dit que $\mathcal{A}$ est une algèbre unitaire.

Définition VII.2.

Soit $\mathcal{A}$ une algèbre munie d'une norme $\|.\|$. On dit que $\|.\|$ est une norme d'algèbre lorsque pour tous $(a, b) \in \mathcal{A}^2$

$$\|a * b\| \leq \|a\| \, \|b\| \,.$$

Proposition VII.3.

Soit $\mathcal{A}$ une algèbre munie d'une norme d'algèbre $\|.\|$. Soit $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites à valeurs dans $\mathcal{A}$. L'implication suivante est vraie

$$a_n \xrightarrow[n\to+\infty]{} a \text{ et } b_n \xrightarrow[n\to+\infty]{} b \implies a_n * b_n \xrightarrow[n\to+\infty]{} a * b.$$

(cc) BY-NC-SA

Démonstration. Soit $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites vérifiant les hypothèses. On a

$$\begin{aligned}\|a_n * b_n - a * b\| &= \|a_n * b_n - a * b_n + a * b_n - a * b\| \\ &\leq \|(a_n - a) * b_n\| + \|a * (b_n - b)\| \\ &\leq \|a_n - a\|\,\|b_n\| + \|a\|\,\|b_n - b\| \xrightarrow[n\to+\infty]{} 0\end{aligned}$$

Exercice VII.4.

Soit $\mathcal{A}$ une algèbre munie d'une norme d'algèbre $\|.\|$. Soit $a \in \mathcal{A} \setminus \{0\}$ non nilpotent, montrer que la suite $\left(\|a^n\|^{\frac{1}{n}}\right)_{n\in\mathbb{N}}$ est convergente.

VIII Suites de fonctions

Dans toute cette partie, on considère A un ensemble et E un espace vectoriel normé muni d'une norme $\|.\|$.

Définition VIII.1.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de A dans E. On dit que la suite $(f_n)_{n\in\mathbb{N}}$ est simplement convergente lorsque pour tout $x \in \mathrm{A}$, la suite $(f_n(x))_{n\in\mathbb{N}}$ est convergente.

Vocabulaire. La fonction $f : x \longmapsto \lim_n f_n(x)$ est appelée limite simple de $(f_n)_{n\in\mathbb{N}}$.

Exemple. On considère ici le cas où $\mathrm{A} = [0,1]$, $\mathrm{E} = \mathbb{R}$ et $(f_n) = (x \longmapsto x^n)$. $(f_n)_{n\in\mathbb{N}}$ converge simplement vers f qui est définie par

$$f : x \longmapsto \begin{cases} 0 & \text{si } x \in [0,1[\\ 1 & \text{sinon} \end{cases}$$

Définition VIII.2.

Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de A dans E. On dit que $(f_n)_{n\in\mathbb{N}}$ est uniformément convergente lorsqu'il existe une fonction f de A dans E telle que

$$\sup_{x\in\mathrm{A}} \|f_n(x) - f(x)\| \xrightarrow[n\to+\infty]{} 0.$$

Remarque. Une formulation équivalente à la définition précédente est la suivante

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \forall x \in \mathrm{A},\ \ \|f_n(x) - f(x)\| \leq \varepsilon.$$

La convergence uniforme implique évidemment la convergence simple, mais la réciproque est fausse. En effet, il suffit de considérer la suite de fonctions $(x \longmapsto x^n)_{n\in\mathbb{N}}$. Cette suite de fonctions converge simplement vers la fonction f définie juste avant la définition précédente, qui vaut 0 sont $[0,1[$ et 1 en 1. Mais f ne converge pas uniformément vers f car pour tout $n \in \mathbb{N}$, $\sup_{x\in\mathrm{A}} \|f_n(x) - f(x)\| = 1$. Cette quantité ne peut donc pas tendre vers 0.

(cc) BY-NC-SA

Proposition VIII.3.

On considère le cas où $\mathrm{E} = \mathcal{C}([a,b],\mathbb{C})$, avec $a < b$. Soit $(f_n)_{n\in\mathbb{N}} \in \mathrm{E}^{\mathbb{N}}$.
Si (f_n) converge uniformément vers $f \in \mathcal{C}([a,b],\mathbb{C})$, alors on a

$$\|f_n - f\|_1 \xrightarrow[n\to+\infty]{} 0 \text{ et } \|f_n - f\|_2 \xrightarrow[n\to+\infty]{} 0.$$

Démonstration.
Pour tout $n \in \mathbb{N}$, on a

$$\|f_n - f\|_1 = \int_a^b |f_n(t) - f(t)|\, dt \leq (b-a)\, \|f_n - f\|_\infty \xrightarrow[n\to\infty]{} 0$$

et

$$\|f_n - f\|_2 = \sqrt{\int_a^b |f_n(t) - f(t)|^2\, dt} \leq \sqrt{(b-a)\, \|f_n - f\|_\infty^2} = \sqrt{b-a}\, \|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0.$$

(cc) BY-NC-SA

Correction de l'exercice I.3. :

Soit $x = (x_1, \dots, x_n) \in \mathrm{E}$. Si $\|x\|_\infty = 0$, on a $x = 0$ et alors pour tout $p \in \mathbb{N}^*$, $\|x\|_p = 0$. La proposition qu'on veut démontrer est donc vraie. Supposons donc à présent que $\|x\|_\infty > 0$. On pose $r = |\{i \in [\![1; n]\!],\ x_i = \|x\|_\infty\}|$ Quitte à permuter les x_i, on suppose sans perte de généralité que $x_1 = x_2 = \cdots = x_r = \|x\|_\infty$ et pour tout $k \geq r+1$, $|x_k| < \|x\|_\infty$.
On a alors pour tout $p \in \mathbb{N}^*$,

$$\|x\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} = \|x\|_\infty \left(r + \sum_{k=r+1}^n \left|\frac{x_k}{\|x\|_\infty}\right|^p\right)^{\frac{1}{p}}.$$

Pour tout $k \geq r+1$, $\left|\dfrac{x_k}{\|x\|_\infty}\right| \in [0, 1[$, donc $\left|\dfrac{x_k}{\|x\|_\infty}\right|^p \xrightarrow[p \to +\infty]{} 0$. On a alors

$$r + \sum_{k=r+1}^n \left|\frac{x_k}{\|x\|_\infty}\right|^p \xrightarrow[p \to +\infty]{} r.$$

Or

$$\|x\|_p = \|x\|_\infty \left(r + \sum_{k=r+1}^n \left|\frac{x_k}{\|x\|_\infty}\right|^p\right)^{\frac{1}{p}} \in \left[\|x\|_\infty, n^{\frac{1}{p}} \|x\|_\infty\right],$$

et finalement, par le lemme des gendarmes

$$\|x\|_p \xrightarrow[p \to +\infty]{} \|x\|_\infty.$$

Correction de l'exercice II.2. :

Soit F un sous-espace vectoriel de E et $a \in \mathrm{E}$. On pour $\mathrm{F}' = a + \mathrm{F}$. Soit $x, y \in \mathrm{F}'$ et $\lambda \in [0, 1]$. Par définition, il existe $u, v \in \mathrm{F}$ tels que $x = a + u$ et $y = a + v$. On a alors

$$\begin{aligned}\lambda x + (1-\lambda)y &= \lambda a + \lambda u + (1-\lambda)a + (1-\lambda)u \\ &= a + \underbrace{\lambda u + (1-\lambda)v}_{\in \mathrm{F}} \in \mathrm{F}'.\end{aligned}$$

Correction de l'exercice II.3. :

Supposons que A est convexe. Soit $x, y \in \mathrm{E} \setminus \{0\}$, $\dfrac{x}{\mathrm{N}(x)}$ et $\dfrac{y}{\mathrm{N}(y)}$ sont deux éléments de A, donc en posant $\lambda = \dfrac{\mathrm{N}(x)}{\mathrm{N}(x) + \mathrm{N}(y)} \in [0, 1]$ on a

$$\mathrm{N}\left(\lambda \frac{x}{\mathrm{N}(x)} + (1-\lambda)\frac{y}{\mathrm{N}(y)}\right) \leq 1.$$

Or

$$\begin{aligned}\mathrm{N}\left(\lambda \frac{x}{\mathrm{N}(x)} + (1-\lambda)\frac{y}{\mathrm{N}(y)}\right) &= \mathrm{N}\left(\frac{\mathrm{N}(x)}{\mathrm{N}(x)+\mathrm{N}(y)}\frac{x}{\mathrm{N}(x)} + \frac{\mathrm{N}(y)}{\mathrm{N}(x)+\mathrm{N}(y)}\frac{y}{\mathrm{N}(y)}\right) \\ &= \frac{\mathrm{N}(x+y)}{\mathrm{N}(x)+\mathrm{N}(y)}.\end{aligned}$$

On en déduit donc finalement que

$$\mathrm{N}(x+y) \leq \mathrm{N}(x) + \mathrm{N}(y).$$

CPGE paradise

(cc) BY-NC-SA

On a alors montré la propriété (N_3) pour N.

Correction de l'exercice II.4. :

Quitte à permuter x et y supposons que $\max(\|x\|, \|y\|) = \|x\|$. On a alors

$$\begin{aligned}
\frac{1}{2}\max(\|x\|, \|y\|)\left\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\right\| &= \frac{1}{2\|y\|}\|\|y\|x - \|x\|y\| \\
&= \frac{1}{2\|y\|}\|\|y\|x - \|y\|y + \|y\|y - \|x\|y\| \\
&\leq \frac{1}{2\|y\|}(\|\|y\|x - \|y\|y\| + \|\|y\|y - \|x\|y\|) \\
&= \frac{1}{2\|y\|}(\|y\|\|x-y\| + \|y\|\,|\|x\| - \|y\||) \\
&\leq \frac{1}{2\|y\|}(\|y\|\|x-y\| + \|y\|\|x-y\|) \\
&= \|x-y\|.
\end{aligned}$$

Correction de l'exercice II.5. :

1. Nous allons séparer les cas $p=1$, $p=2$ et $p=\infty$.
   (a) Mettons-nous sur $\mathbb{K}^n$.
   - → Si $p=\infty$, la norme $\|.\|_\infty$ n'est pas de somme stricte.
     En effet, il suffit de considérer le cas $n=2$ pour voir que ce n'est pas le cas :

     $$\|(1,0)+(1,1)\|_\infty = 2 = \|(1,0)\|_\infty + \|(1,1)\|_\infty$$

     mais $((1,0),(1,1))$ n'est pas liée.
   - → Si $p=1$, la norme $\|.\|_1$ n'est pas de somme stricte.
     En effet, il suffit de considérer le cas $n=2$ pour voir que ce n'est pas le cas :

     $$\|(1,0)+(0,1)\|_1 = 2 = \|(1,0)\|_1 + \|(0,1)\|_1$$

     mais $((1,0),(0,1))$ n'est pas liée.
   - → Si $p=2$, la norme $\|.\|_2$ est de somme stricte.
     En effet, si on considère $x=(x_1,\dots,x_n), y=(y_1,\dots,y_n) \in \mathbb{K}^n$,

     $$\begin{aligned}
     \|x+y\|_2 = \|x\|_2 + \|y\|_2 &\iff \|x+y\|_2^2 = (\|x\|_2 + \|y\|_2)^2 \\
     &\iff \|x\|_2^2 + \|y\|_2^2 + 2\sum_{k=1}^n x_k y_k = \|x\|_2^2 + \|y\|_2^2 + 2\|x\|\|y\| \\
     &\iff \sum_{k=1}^n x_k y_k = \sqrt{\left(\sum_{k=1}^n x_k^2\right)\left(\sum_{k=1}^n y_k^2\right)}.
     \end{aligned}$$

     La dernière ligne implique le cas d'égalité de l'inégalité de Cauchy-Schwarz (il suffit de majorer la somme des $x_k y_k$ par la somme des $|x_k y_k|$, qui est équivalent au fait que (x,y) soit liée.

(cc) BY-NC-SA

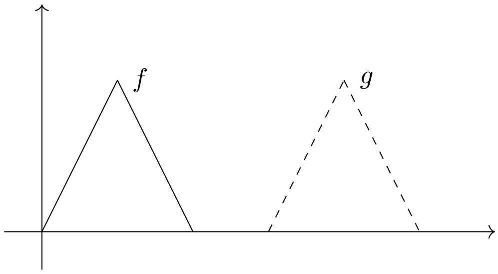
(b) Mettons-nous sur $\mathcal{C}([0,1],\mathbb{R})$.

→ Si $p=\infty$, la norme $\|.\|_\infty$ n'est pas de somme stricte.
En effet, il suffit de considérer les deux fonctions continues sur $[0,1]$ $f : t \longmapsto t$ et $g : t \longmapsto 1$.

$$\|f+g\|_\infty = 2 = \|f\|_\infty + \|g\|_\infty$$

alors que (f,g) n'est pas liée.

→ Si $p=1$, la norme $\|.\|_1$ n'est pas de somme stricte.
En effet, il suffit de prendre f et g respectivement décrites par les courbes ci-dessous.



Il est clair que

$$\|f+g\|_1 = \|f\|_1 + \|g\|_1$$

alors que (f,g) n'est pas liée.

→ Si $p=2$, la norme $\|.\|_2$ est de somme stricte.
En effet, si on considère $f,g \in \mathcal{C}([0,1],\mathbb{R})$,

$$\begin{aligned}\|f+g\|_2 = \|f\|_2 + \|g\|_2 &\iff \|f+g\|_2^2 = (\|f\|_2 + \|g\|_2)^2 \\ &\iff \|f\|_2^2 + \|g\|_2^2 + 2\int_0^1 f(t)g(t)dt = \|f\|_2^2 + \|g\|_2^2 + 2\|f\|\,\|g\| \\ &\iff \int_0^1 f(t)g(t)dt = \sqrt{\left(\int_0^1 f(t)^2 dt\right)\left(\int_0^1 g(t)^2 dt\right)}.\end{aligned}$$

La dernière ligne implique le cas d'égalité de l'inégalité de Cauchy-Schwarz (il suffit de majorer l'intégrale de fg par l'intégrale de $|fg|$), qui est équivalent au fait que (f,g) soit liée.

2. Soit $x,y \in \mathrm{E}$ vérifiant $\|x+y\| = 2$ et $\|x\| = \|y\| = 1$.
Soit $\lambda \in [0,1]$, on a

$$\underbrace{\|\lambda x + (1-\lambda)y\|}_{\leq 1} + \underbrace{\|(1-\lambda)x + \lambda y\|}_{\leq 1} \geq \|x+y\| = 2,$$

donc

$$\|\lambda x + (1-\lambda)y\| = \|(1-\lambda)x + \lambda y\| = 1$$

et finalement $[x,y] \subset \mathrm{S}(0,1)$.

3. $(\Rightarrow)$ Supposons que $\|.\|$ est stricte.
Soit $x,y \in \mathrm{S}(0,1)$. Supposons que $[x,y] \subset \mathrm{S}(0,1)$. On a

$$\left\|\frac{1}{2}x + \frac{1}{2}y\right\| = 1 = \left\|\frac{1}{2}x\right\| + \left\|\frac{1}{2}y\right\|$$

Donc (x,y) est liée et x et y sont de même norme. Lorsque $\mathbb{K} = \mathbb{R}$, deux cas se présentent.

→ $x = -y$ ce qui donne $\left\|\frac{1}{2}x + \frac{1}{2}y\right\| = 0$, chose qui n'est clairement pas possible.

→ $x = y$ et donc le segment $[x,y]$ est trivial.

(cc) BY-NC-SA

On en déduit alors immédiatement que $\mathrm{S}(0,1)$ ne contient aucun segment non trivial.
($\Leftarrow$) Nous allons procéder par contraposée. Supposons que $\|.\|$ est non stricte, *i.e.* il existe (x,y) libre telle que $\|x+y\| = \|x\|+\|y\|$.
On suppose sans perte de généralité de que $\|x\| \leq \|y\|$. On a

$$\begin{aligned}
2 \geq \left\| \frac{x}{\|x\|} + \frac{y}{\|y\|} \right\| &= \left\| \frac{x}{\|x\|} + \frac{y}{\|x\|} - \frac{y}{\|x\|} + \frac{y}{\|y\|} \right\| \\
&\geq \frac{1}{\|x\|}(\|x\|+\|y\|) - \|y\| \left(\frac{1}{\|x\|} - \frac{1}{\|y\|} \right) \\
&= 2,
\end{aligned}$$

donc $\left\| \frac{x}{\|x\|} + \frac{y}{\|y\|} \right\| = 2$ et $\frac{x}{\|x\|}, \frac{y}{\|y\|} \in \mathrm{S}(0,1)$. On en déduit d'après la question précédente que $\left[\frac{x}{\|x\|}, \frac{y}{\|y\|} \right] \subset \mathrm{S}(0,1)$. Ce segment est non trivial puisque (x,y) est libre.
$\mathrm{S}(0,1)$ contient donc bien un segment non trivial.

Correction de l'exercice V.2. :

1. Soit $(a,r) \in \mathrm{E} \times \mathbb{R}^+$. On sait déjà que $\mathrm{diam}(\mathrm{B}(a,r)) \leq 2r$ et $\mathrm{diam}(\mathrm{B}_f(a,r) \leq 2r$ (il suffit d'appliquer l'inégalité triangulaire).
   Comme $\mathrm{E} \neq \{0\}$, on peut considérer $u \in \mathrm{X}$ tel que $\|u\| = 1$. On pose alors
   $$x = a + ru \in \mathrm{B}_f(a,r) \text{ et } y = a - ru \in \mathrm{B}_f(a,r).$$
   On a donc $\|x-y\| = 2r$ ce qui donne $\mathrm{diam}(\mathrm{B}_f(a,r)) = 2r$.
   De même, on pose $(r_n)_{n\geq \mathrm{N}} = (r - \frac{1}{n})_{n \geq \mathrm{N}}$ avec N assez grand pour qu'on ait $r - \frac{1}{\mathrm{N}} \geq 0$ et
   $$x_n = a + r_n u \in \mathrm{B}(a,r) \text{ et } y_n = a - r_n u \in \mathrm{B}(a,r).$$
   On a alors $\|x_n - y_n\| = 2r_n \xrightarrow[n\to+\infty]{} 2r$. La suite $(\|x_n - y_n\|)_{n \geq \mathrm{N}}$ est à valeurs dans $\{\|x-y\|\,,\ (x,y) \in \mathrm{E}^2\}$ qui est majoré par $2r$, donc $\mathrm{diam}(\mathrm{B}(a,r)) = 2r$.

2. L'unicité du rayon d'une boule vient tout simplement de la question précédente. En effet, en espace vectoriel normé, on peut définir de manière unique le rayon d'une boule fermée ou ouverte comme le diamètre de la boule. Il reste donc à montrer l'unicité du centre. Nous allons le faire pour le cas d'une boule fermée. Le cas d'une boule ouverte peut être traité de la même manière.
   Soit $(a,r) \in \mathrm{E} \times \mathbb{R}^+$. On suppose que $b \in \mathrm{E}$ est aussi un centre de $\mathrm{B}_f(a,r)$ tel que $b \neq a$. On a alors pour tout $x \in \mathrm{B}_f(a,r)$, $\|x-b\| < r$ et $\|x-a\| < r$. Posons $v = \frac{a-b}{\|a-b\|}$. On a $a - rv \in \mathrm{B}_f(a,r)$, donc
   $$\begin{aligned}
   r \geq \|a - rv - b\| &= \left\| \left(\frac{r}{\|a-b\|} + 1 \right)(a-b) \right\| \\
   &= r + \|a-b\|\,,
   \end{aligned}$$
   ce qui donne $\|a-b\| = 0$ ce qui est absurde. On a donc bien l'unicité du centre.

Correction de l'exercice VII.4. :

On a pour tout $n \in \mathbb{N}^*$, $\|a^n\|^{\frac{1}{n}} \leq (\|a\|^n)^{\frac{1}{n}}$ et pour tout $n,m \in \mathbb{N}^*$, $\|a^{n+m}\| \leq \|a^n\|\,\|a^m\|$. Si on pose pour tout n, $v_n = \log\|a^n\|$, on peut affirmer grâce à cette inégalité que tout $m,n \in \mathbb{N}^*$, $v_{n+m} \leq v_n + v_m$.
Posons également pour tout n, $u_n = \log(\|a^n\|^{\frac{1}{n}})$. On a alors $(v_n)_{n\in\mathbb{N}}$ est sous additive et pour tout n,

(cc) BY-NC-SA

$u_n = \frac{v_n}{n}$. D'après le théorème de Fekete (hors programme, mais très classique)

$$u_n \xrightarrow[n\to+\infty]{} \inf_{k\in\mathbb{N}^*} u_k = \alpha$$

et finalement

$$\|a^n\|^{\frac{1}{n}} \xrightarrow[n\to+\infty]{} e^{\alpha}.$$

Remarque. ici, on suppose implicitement par convention que $e^{-\infty} = 0$.

(cc) BY-NC-SA

Chapitre 11.2

Ouverts et fermés

Dans tout le chapitre, on considère X un espace metrique muni d'une distance d.

I Ouverts

Définition I.1.

Soit O un sous ensemble de X. On dit que O est ouvert lorsque

$$\forall a \in \mathrm{O},\ \exists r > 0,\ \mathrm{B}(a, r) \subset \mathrm{O}.$$

Exemples.

→ $\varnothing$ et les intervalles ouverts sont des ouverts de $\mathbb{R}$.

→ $\mathbb{R}$ n'est pas ouvert dans $\mathbb{C}$.

Proposition I.2.

Soit I un ensemble et $(\mathrm{O}_i)_{i\in \mathrm{I}}$ est une famille d'ouverts de X.

1. $\varnothing$ et X sont des ouverts de X.
2. $\bigcup\limits_{i\in\mathrm{I}} \mathrm{O}_i$ est ouvert.
3. Si I est fini, alors $\bigcap\limits_{i\in\mathrm{I}} \mathrm{O}_i$ est ouvert.

Démonstration.

1. La proposition est évidente par définition.
2. Soit $a \in \bigcup\limits_{i\in\mathrm{I}} \mathrm{O}_i$. Il existe $i_0 \in \mathrm{I}$ tel que $a \in \mathrm{O}_{i_0}$. Il existe alors $r > 0$ tel que $\mathrm{B}(a, r) \subset \mathrm{O}_{i_0} \subset \bigcup\limits_{i\in\mathrm{I}} \mathrm{O}_i$. $\bigcup\limits_{i\in\mathrm{I}} \mathrm{O}_i$ est donc bien ouvert.
3. Supposons que I soit fini. Soit $a \in \bigcap\limits_{i\in\mathrm{I}} \mathrm{O}_i$. Pour tout $i \in \mathrm{I}$, il existe $r_i > 0$ tel que $\mathrm{B}(a, r_i) \subset \mathrm{O}_i$. I est fini, donc on peut définir $r = \min\{r_i\}_{i\in\mathrm{I}} > 0$. Par construction, on a bien $\mathrm{B}(a, r) \subset \bigcap\limits_{i\in\mathrm{I}} \mathrm{O}_i$. $\bigcap\limits_{i\in\mathrm{I}} \mathrm{O}_i$ est donc bien ouvert.

Attention : si I est infini, $\bigcap\limits_{i\in\mathrm{I}} \mathrm{O}_i$ n'est pas forcément ouvert. En effet, si on prend $\mathrm{I} = \mathbb{N}^*$ et pour tout $k \in \mathrm{I}$, $\mathrm{O}_k = \left]-\frac{1}{k}, \frac{1}{k}\right[$. On a $\bigcap\limits_{k\in\mathbb{N}^*} \mathrm{O}_k = \{0\}$ qui n'est pas ouvert.

Définition I.3.

Soit $a \in \mathrm{X}$. Une partie U de X est appelée voisinage de a s'il existe $r > 0$ tel que $\mathrm{B}(a, r) \subset \mathrm{U}$.

(cc) BY-NC-SA

Remarque. un ouvert est voisinage de tous ses points.

Exemple. Pour tout $(b, r) \in \mathrm{X} \times \mathbb{R}_+^*$, $\mathrm{B}(b, r)$ est un ouvert.
Montrons ce résultat. Soit $a \in \mathrm{B}(b, r)$. Posons $\varepsilon = r - d(a, b)$. On a alors pour tout $x \in \mathrm{B}(a, \varepsilon)$,

$$d(x, b) \leq d(x, a) + d(a, b) < \varepsilon + d(a, b) = r,$$

donc $\mathrm{B}(a, \varepsilon) \subset \mathrm{B}(b, r)$. On en déduit donc que $\mathrm{B}(b, r)$ est ouvert.

Proposition I.4.

Soit $a \in \mathrm{X}$. On note $\mathcal{V}(a)$ l'ensemble des voisinages de a.

1. Si $\mathrm{U} \in \mathcal{V}(a)$ et $\mathrm{U} \subset \mathrm{V}$, alors $\mathrm{V} \in \mathcal{V}(a)$.
2. L'intersection d'une famille finie d'éléments de $\mathcal{V}(a)$ appartient à $\mathcal{V}(a)$.
3. Si $a \neq b$, alors il existe $(\mathrm{U}, \mathrm{V}) \in \mathcal{V}(a) \times \mathcal{V}(b)$ tel que $\mathrm{U} \cap \mathrm{V} = \varnothing$.

Démonstration.

1. Soit U voisinage de a et V un ensemble contenant U. U étant un voisinage de a, il existe $r > 0$ tel que $\mathrm{B}(a, r) \subset \mathrm{U} \subset \mathrm{V}$. Donc V est aussi un voisinage de a.
2. Soit $p \in \mathbb{N}^*$ et $\mathrm{V}_1, \ldots, \mathrm{V}_p \in \mathcal{V}(a)$. Pour tout $i \in [\![1; p]\!]$, il existe $r_i > 0$ tel que $\mathrm{B}(a, r_i) \subset \mathrm{V}_i$. On pose $r = \min\{r_1, \ldots, r_p\} > 0$. On a alors $\mathrm{B}(a, r) \subset \bigcap_{i=1}^{p} \mathrm{V}_i$, donc $\bigcap_{i=1}^{p} \mathrm{V}_i \in \mathcal{V}(a)$.
3. Soit $r = \dfrac{d(a, b)}{10}$. Prenons $\mathrm{U} = \mathrm{B}(a, r)$ et $\mathrm{V} = \mathrm{B}(b, r)$. Si $x \in \mathrm{U} \cap \mathrm{V}$, alors $d(x, a) < r$ et $d(x, b) < r$. En sommant les deux inégalités, on obtient
$$d(a, b) \leq d(x, a) + d(x, b) < 2r = \frac{d(a, b)}{5}$$
ce qui est impossible. On en déduit donc que $\mathrm{U} \cap \mathrm{V} = \varnothing$.

II Fermés

Définition II.1.

Soit F une partie de X. On dit que F est fermé si $\mathrm{X} \setminus \mathrm{F}$ est ouvert.

Exemples.

→ Si $a \leq b$, alors $[a, b]$ est un fermé.

→ Pour tout $(b, r) \in \mathrm{X} \times \mathbb{R}^+$, $\mathrm{B}_f(b, r)$ est fermé.
En effet, si on prend $a \in \mathrm{X} \setminus \mathrm{B}_f(b, r)$, on peut considérer $\varepsilon = \frac{d(a,b)-r}{2} > 0$ (ε est non nul car sinon on aurait $a \in \mathrm{B}_f(b, r)$). Si $x \in \mathrm{B}(a, \varepsilon)$, alors $d(x, b) \geq d(a, b) - d(x, a) \geq \frac{d(a,b)+r}{2} > r$ donc $\mathrm{B}(a, \varepsilon) \subset \mathrm{X} \setminus \mathrm{B}_f(b, r)$ et alors $\mathrm{B}_f(b, r)$ est fermé.

→ Une réunion finie de points est un fermé.

Proposition II.2.

Soit F une partie de X. F est fermée si et seulement si pour toute suite $(x_n)_{n \in \mathbb{N}}$ à valeurs dans F, si (x_n) est convergente, alors sa limite est dans F.

(cc) BY-NC-SA

Démonstration.

($\Rightarrow$) Supposons que F soit fermé. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans F convergente, supposons que la limite ℓ de (x_n) est dans $X\setminus F$. $X\setminus F$ étant ouvert, il existe $\varepsilon>0$ tel que $B(\ell,\varepsilon)\subset X\setminus F$. Or par construction, il existe $N\in\mathbb{N}$ tel que pour tout $n\geq N$, $d(x_n,\ell)\leq\frac{\varepsilon}{2}$ et donc $x_n\in B(\ell,\varepsilon)\subset X\setminus F$ ce qui est absurde.

($\Leftarrow$) Réciproquement, supposons que toute suite à valeurs dans F convergente converge dans F. supposons de plus que F n'est pas fermé, *i.e.* $X\setminus F$ n'est pas ouvert. Il existe donc $a\in X\setminus F$ tel que pour tout $\varepsilon>0$, $B(a,\varepsilon)\not\subset X\setminus F$ *i.e.* $B(a,\varepsilon)\cap F\neq\varnothing$. Cette propriété étant vraie pour $\varepsilon=\dfrac{1}{n}$ et tout $n\in\mathbb{N}^*$, on peut définir une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans F qui vérifie pour tout $n\in\mathbb{N}^*$, $d(a,x_n)\leq\frac{1}{n}$. (x_n) est à valeurs dans F et converge vers a qui n'est pas dans F, ce qui est en contradiction avec les hypothèses.

Proposition II.3.

1. X et $\varnothing$ sont fermés dans X.
2. Soit I un ensemble. Si $(F_i)_{i\in I}$ est une famille de fermés de X, alors $\bigcap\limits_{i\in I}F_i$ est un fermé.
3. Si I est fini, alors $\bigcup\limits_{i\in I}F_i$ est un fermé.

Démonstration. pour montrer ces trois propriétés, il suffit de montrer que les complémentaires de ces ensembles sont ouverts en utilisant les propriétés vu sur les ouverts au début du chapitre.

Exercice II.4.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans X. Montrer que l'ensemble des valeurs d'adhérence de (u_n) est fermé.

Exercice II.5.

Soit $E=\ell^1(\mathbb{N},\mathbb{C})$ muni de $\|.\|_1$ et $N\in\mathbb{N}$. Montrer que $A=\{(u_n)\in E^{\mathbb{N}},\ u_0\neq 0,\dots,u_N\neq 0\}$ est ouvert.

Exercice II.6.

Soient $E=\mathcal{C}([0,1],\mathbb{R})$, $O=\{f\in E,\ \forall x\in[0,1],\ f(x)>0\}$ et $F=\{f\in E,\ \forall x\in[0,1],\ f(x)\geq 0\}$.

1. O est il ouvert pour la distance induite par $\|.\|_\infty$? $\|.\|_1$?
2. F est il fermé pour la distance induite par $\|.\|_\infty$? $\|.\|_1$?

III Topologie induite

Dans cette partie on considère A une partie de X et on considère d_A la restriction de d à A^2. (A,d_A) ainsi défini est un espace métrique, et pour tout $(a,r)\in A\times\mathbb{R}^+$, $B_{d_A}(a,r)=B(a,r)\cap A$.

(cc) BY-NC-SA

Proposition III.1.

1. Une partie O de A est ouverte pour d_A si et seulement s'il existe $\Omega \subset X$ ouvert tel que $O = A \cap \Omega$.
2. Une partie F de A est fermée pour d_A si et seulement s'il existe $G \subset X$ fermé tel que $F = A \cap G$.

Démonstration.

1. ($\Leftarrow$) Soit $a \in A \cap \Omega$. Puisque Ω est ouvert dans X, il existe $r > 0$ tel que $B(a, r) \subset \Omega$. On peut alors affirmer que $B_{d_A}(a, r) = B(a, r) \cap A \subset \Omega \cap A$. $O = A \cap \Omega$ est donc ouvert.
   ($\Rightarrow$) Soit $a \in O$. Il existe $r_a > 0$ tel que $B_{d_A}(a, r_a) \subset O$. On a donc

$$O = \bigcup_{a \in O} B_{d_A}(a, r_a) = \bigcup_{a \in O} (B(a, r_a) \cap A) = \underbrace{\left(\bigcup_{a \in O} B(a, r) \right)}_{\Omega} \cap A.$$

   Ω est une union d'ouverts, donc est un ouvert de X, ce qui nous permet de conclure.
2. Il suffit d'utiliser le résultat précédent en passant au complémentaire.

   ($\Leftarrow$) Si $F = A \cap G$, alors $A \setminus F = A \cap \underbrace{(X \setminus G)}_{\Omega}$. Ω est ouvert, donc en appliquant la propriété précédente, $A \setminus F$ est ouvert pour d_A et donc F est fermé pour d_A.
   ($\Rightarrow$) Si F est fermé pour la distance d_A, alors $A \setminus F$ est ouvert pour cette distance. D'après ce qui précède, il existe Ω un ouvert de X tel que $A \setminus F = \Omega \cap A$, ce qui donne $F = A \setminus (\Omega \cap A) = A \cap \underbrace{(X \setminus \Omega)}_{G}$.

   G est un fermé de X, ce qui nous permet de conclure.

Conséquences.

→ Tout ouvert de A est ouvert dans X si et seulement si A est ouvert dans X.

→ Tout fermé de A est fermé dans X si et seulement si A est fermé dans X.

IV Intérieur

Définition IV.1.

Soit A une partie de X. On dit que le point $a \in A$ est intérieur à A lorsqu'il existe $\varepsilon > 0$ tel que $B(a, \varepsilon) \subset A$ (*i.e.* $A \in \mathcal{V}(a)$). On note $\mathring{A}$ ou Int(A) l'ensemble des points intérieurs à A.

Exemples.

→ Dans $\mathbb{R}$, $\mathring{\mathbb{Q}} = \varnothing$ car $R \setminus \mathbb{Q}$ est dense dans $\mathbb{R}$.

→ Dans $\mathbb{C}$, $\mathring{\mathbb{R}} = \varnothing$.

Proposition IV.2.

Soit A une partie de X. $\mathring{A}$ est le plus grand ouvert inclus dans A.

Démonstration. Si $a \in \mathring{A}$, alors il existe $\varepsilon > 0$ tel que $B(a, \varepsilon) \subset A$. $B(a, \varepsilon)$ est ouvert, donc pour tout $b \in B(a, \varepsilon)$, il existe $\eta > 0$ tel que $B(b, \eta) \subset B(a, \varepsilon) \subset A$. b est donc un élément de $\mathring{A}$ et alors $B(a, \varepsilon) \subset \mathring{A}$. On peut donc conclure que $\mathring{A}$ est ouvert. Le fait que A soit le plus grand ouvert inclus dans A découle

(cc) BY-NC-SA

du fait que tout point appartenant à un ouvert inclus dans A est dans $\mathring{A}$ par définition de l'intérieur de A.

Exercice IV.3.

Soient A et B deux parties de X. Dans chaque cas, comparer les deux ensembles.

1. $\mathrm{Int}(A \cap B)$ et $\mathrm{Int}(A) \cap \mathrm{Int}(B)$.
2. $\mathrm{Int}(A \cup B)$ et $\mathrm{Int}(A) \cup \mathrm{Int}(B)$.

Exercice IV.4.

Soit E un espace vectoriel normé non réduit à $\{0\}$. Soit $(a, r) \in E \times \mathbb{R}_+^*$. Montrer que

$$\mathrm{Int}(B_f(a, r)) = B(a, r).$$

Attention : cette propriété n'est pas toujours vraie quand on est pas dans le cadre d'espaces vectoriels normés. Par exemple, pour $X = \mathbb{Z}$ et d la distance induite par la valeur absolue, on a

→ $B_f(0, 1) = \{-1, 0, 1\}$

→ $B(0, 1) = \{0\} \neq \mathrm{Int}(B_f(0, 1)) = \{-1, 0, 1\}$

V Adhérence

Définition V.1.

Soit $b \in X$. On dit que b est adhérent à A lorsque

$$\forall \varepsilon > 0,\ B(b, \varepsilon) \cap A \neq \varnothing.$$

On note $\overline{A}$ l'ensemble des points adhérents à A.

Proposition V.2.

Soit A une partie de X. Les propriétés suivantes sont vérifiées.

1. $\overline{A}$ est fermé.
2. $\overline{A}$ est le plus petit fermé contenant A.
3. $X \setminus \overline{A} = \mathrm{Int}(X \setminus A)$.

Démonstration. On a

$$\begin{aligned} x \notin \overline{A} &\iff \exists \varepsilon > 0,\ B(x, \varepsilon) \cap A = \varnothing \\ &\iff \exists \varepsilon > 0,\ B(x, \varepsilon) \subset X \setminus A \\ &\iff x \in \mathrm{Int}(X \setminus A). \end{aligned}$$

Ceci montre le point 3 et le point 1 car $\mathrm{Int}(X \setminus A)$ est ouvert. Montrons le deuxième point.
Soit B un fermé contenant A. $X \setminus A$ contient $X \setminus B$ qui est ouvert. $\mathrm{Int}(X \setminus A)$ étant le plus grand ouvert inclus dans $X \setminus A$, $X \setminus B \subset \mathrm{Int}(X \setminus A) = X \setminus \overline{A}$ et finalement $\overline{A} \subset B$. $\overline{A}$ est donc bien le plus petit fermé contenant A.

BY-NC-SA

Proposition V.3.

Soit $b \in \mathrm{X}$ et A une partie de X. L'équivalence suivante est vraie.

$$b \in \overline{\mathrm{A}} \iff \exists (a_n) \in \mathrm{A}^{\mathbb{N}},\ a_n \xrightarrow[n \to +\infty]{} b$$

Démonstration.
($\Leftarrow$) Soit $\varepsilon > 0$. Il existe $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $d(a_n, b) < \frac{\varepsilon}{2}$. (a_n) est à valeurs dans A donc $\mathrm{B}(b, \varepsilon) \cap \mathrm{A} \neq \varnothing$, donc $b \in \overline{\mathrm{A}}$.
($\Rightarrow$) Soit $b \in \overline{\mathrm{A}}$. On va construire la suite (a_n) par récurrence. On prend a_0 égal à n'importe quel élément de X. Soit $n \in \mathbb{N}$. Supposons $a_0, \ldots, a_n$ bien définis. b étant adhérent à A, on a $\mathrm{B}(b, \frac{1}{n+1}) \cap \mathrm{A} \neq \varnothing$. On peut donc prendre $a_{n+1} \in \mathrm{B}(b, \frac{1}{n+1}) \cap \mathrm{A}$. On a alors par construction,

$$\forall n \in \mathbb{N}^*,\ d(a_n, b) \leq \frac{1}{n} \text{ et } (a_n) \in \mathrm{A}^{\mathbb{N}}$$

et alors $a_n \xrightarrow[n \to +\infty]{} b$.

Corollaire V.4.

Soit $\mathrm{A} \subset \mathrm{X}$. L'équivalence suivante est vraie

$$\mathrm{A} \text{ fermé } \iff (\forall (a_n) \in \mathrm{A}^{\mathbb{N}},\ a_n \xrightarrow[n \to +\infty]{} \ell \implies \ell \in \mathrm{A})$$

Démonstration. d'après ce qui précède, on a

$$\mathrm{A} \text{ fermé } \iff \mathrm{A} = \overline{\mathrm{A}} \iff (\forall (a_n) \in \mathrm{A}^{\mathbb{N}},\ a_n \xrightarrow[n \to +\infty]{} \ell \implies \ell \in \mathrm{A})$$

Exercice V.5.

Trouver un ensemble $\mathrm{A} \subset \mathbb{R}$ tel que A, $\overline{\mathrm{A}}$, $\mathring{\mathrm{A}}$, $\overline{\mathring{\mathrm{A}}}$, $\mathring{\overline{\mathrm{A}}}$, $\overline{\mathring{\overline{\mathrm{A}}}}$ et $\mathring{\overline{\mathring{\mathrm{A}}}}$ soient deux à deux distincts.

Exercice V.6.

Soit E un espace vectoriel normé et $(a, r) \in \mathrm{E} \times \mathbb{R}_+^*$. Montrer que

$$\overline{\mathrm{B}(a, r)} = \mathrm{B}_f(a, r).$$

Exercice V.7.

Soit C une partie convexe d'un espace vectoriel normé E.

1. Montrer que $\overline{\mathrm{C}}$ et $\mathring{\mathrm{C}}$ sont convexes.
2. Soit $a \in \mathring{\mathrm{C}}$ et $b \in \overline{\mathrm{C}}$. Montrer que $[a, b[\subset \mathring{\mathrm{C}}$.

(cc) BY-NC-SA

Exercice V.8.

On pose $E = \mathcal{C}([0,1], \mathbb{R})$ et $F = \{f \in E,\ \forall x \in [0,1]\ f(x) \geq 0\}$.

1. Montrer que F est un fermé pour $\|.\|_\infty$ et $\|.\|_1$.
2. Donner $\mathring{F}$ pour $\|.\|_\infty$ et $\|.\|_1$.

VI Densité

Définition VI.1.

Soient A et B deux parties de X. On dit que B est dense dans A lorsque $A \subset \overline{B}$. De même, on dit que B est partout dense dans X lorsque $\overline{B} = X$.

Exemples.

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

→ Pour tout $n \in \mathbb{N}$, $\mathbb{Q}^n$ est dense dans $\mathbb{R}^n$ pour les normes usuelles ($\|.\|_1$, $\|.\|_2$, $\|.\|_\infty$, ...).

Proposition VI.2.

Soient A et B deux parties de X. Les propositions suivantes sont équivalentes.

1. B est dense dans A
2. Pour tout élément $a \in A$, il existe une suite (b_n) à valeurs dans B qui converge vers a.
3. Pour tout $a \in A$ et $\varepsilon > 0$, il existe $b \in B$ tel que $b \in B(x, \varepsilon)$.

Remarque. Ces caractérisations de la densité sont quelques fois plus commode à utiliser en exercice.

Proposition VI.3.

La propriété de densité est transitive, *i.e.* si A, B et C sont trois parties de X et si B et dense dans A et C est dense dans B alors C est dense dans A.

Exemple. Si E est un espace vectoriel normé et F est un sous-espace strict de E (*i.e.* $F \neq E$), alors $E \setminus F$ est dense dans E.
Montrons ce résultat. Soit $x \notin F$ et $a \in F$. Soit $\varepsilon > 0$. On a $a + \varepsilon \frac{x}{2\|x\|} \in F \setminus E$ (car sinon, on aurait $a + \varepsilon \frac{x}{2\|x\|} - a \in F$, *i.e.* $x \in F$). De plus $a + \varepsilon \frac{x}{2\|x\|} \in B(a, \varepsilon)$ et alors $B(a, \varepsilon) \cap (E \setminus F) \neq \varnothing$.
On a montré que $F \subset \overline{E \setminus F}$. On a aussi $E \setminus F \subset \overline{E \setminus F}$. On en déduit donc que $\overline{E \setminus F} = E$, *i.e.* $E \setminus F$ est dense dans E.

Remarque. le résultat montré ci-dessus peut être aussi formulé de la manière suivante :

$$\text{A est un sous espace vectoriel de E et } \mathring{A} \neq \varnothing \iff A = E$$

ou encore :

Si F est un sous espace stricte d'un espace vectoriel normé E, alors F est d'intérieur vide

Cette propriété se voit facilement en dimension 3 par exemple. Un sous-espace strict de $\mathbb{R}^3$ est soit un plan, soit une droite, ou alors un point qui dans tous les cas sont d'intérieur vide.

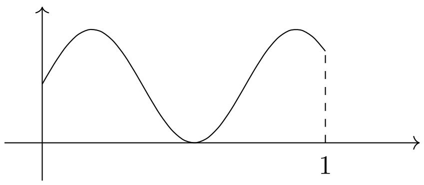
(cc) BY-NC-SA

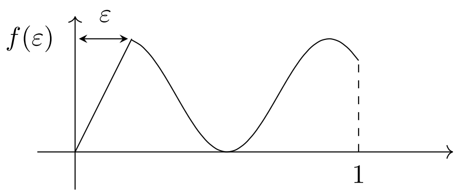
Conséquence. Si H est un hyperplan de E, alors soit H est fermé, soit H est dense dans E.
En effet, on a $H \subset \overline{H} \subset E$ donc soit $H = \overline{H}$, *i.e.* H est fermé, soit $\overline{H} = E$, *i.e.* H est dense dans E.

Exemple. l'hyperplan $H = \{f \in \mathcal{C}([0,1],\mathbb{R}), f(0) = 0\}$ n'est pas fermé pour la distance induite par $\|.\|_1$ car il est dense dans $\mathcal{C}([0,1],\mathbb{R})$.
En effet, si $f \in \mathcal{C}([0,1],\mathbb{R})$, on considère $g \in H$ la fonction définie par

$$g(x) = \begin{cases} \frac{x}{\varepsilon} f(\varepsilon) & \text{si } x \in [0,\varepsilon[\\ f(x) & \text{si } x \in [\varepsilon, 1] \end{cases}$$

Pour montrer que cet exemple est naturel, voici un dessin de la courbe de g à droite et celle de f à gauche





Montrons que g est proche de f. On a

$$\begin{aligned} \|f-g\|_1 &= \int_0^1 |f(x) - g(x)|\, dx \\ &= \int_0^{\varepsilon} |f(x) - g(x)|\, dx \leq 2\|f\|_\infty \varepsilon \end{aligned}$$

donc quitte à remplacer ε par $\dfrac{\varepsilon}{4\|f\|_\infty}$, on peut dire que $g \in B(f,\varepsilon)$ et donc H est dense dans $\mathcal{C}([0,1],\mathbb{R})$ pour la distance induite par $\|.\|_1$.

Exercice VI.4.

Soit G un sous groupe additif de $\mathbb{R}$. On note $A = \{x \in G,\ x > 0\}$.

1. Montrer que si $G \neq \{0\}$, alors A est non vide. On pose alors $a = \inf A$.
2. Montrer que si $a = 0$, alors G est dense dans $\mathbb{R}$.
3. Montrer que si $a > 0$, alors $G = a\mathbb{Z}$.
4. Application 1 : en utilisant ce qui précède, montrer que $\{\cos n,\ n \in \mathbb{N}\}$ est dense dans $[-1,1]$.
5. Application 2 : soit $f \in \mathcal{C}(\mathbb{R},\mathbb{R})$ admettant 1 et $\sqrt{2}$ comme périodes. Montrer que f est constante.

Définition VI.5.

On dit que X est un espace séparable si et seulement si il existe un sous ensemble A de X dense dans X et dénombrable.

Exercice VI.6.

Montrer que X est séparable si et seulement si il existe une famille dénombrable d'ouverts $(\Omega_i)_{i\in I}$ telle que pour tout ouvert O de X $\exists J \subset I$, $O = \bigcup_{j\in J} \Omega_j$.

(cc) BY-NC-SA

Définition VI.7.

Soit A une partie de X. Un point $a \in X$ est dit

→ isolé lorsque
$$\exists \varepsilon > 0,\ B(a, \varepsilon) \cap A = \{a\}.$$

→ d'accumulation lorsque
$$\forall \varepsilon > 0,\ (B(a, \varepsilon) \setminus \{a\}) \cap A \neq \varnothing.$$

L'ensemble des points d'accumulation et isolés de A sont respectivement notés A^{ac} et A^{i}.

Notation. On note A^{aci} l'ensemble des points d'accumulation de A qui sont dans A.

Proposition VI.8.

Soit A une partie de X.

1. $A^{aci} \cup A^{i} = A$.
2. $a \in A^{ac} \iff \forall \varepsilon > 0,\ B(a, \varepsilon) \cap A$ est infini.
3. $a \in A^{ac} \iff$ il existe une suite injective à valeurs dans A qui converge vers a.

Démonstration.

1. En regardant les définitions d'un point isolé et d'un point d'accumulation, on peut aisément remarquer que le fait qu'un point soit isolé est exactement la négation du fait qu'un point soit d'accumulation, donc naturellement, un point a de A est soit d'accumulation, soit isolé, d'où le fait que $A^{aci} \cup A^{i} = A$.
2. ($\Rightarrow$) Supposons que $a \in A^{ac}$. Supposons de plus que $\exists \varepsilon > 0$, $B(a, \varepsilon) \cap A$ est fini. On pose donc $(B(a, \varepsilon) \setminus \{a\}) \cap A = \{x_1, \ldots, x_p\}$ avec $p \in \mathbb{N}^*$. En posant $\tilde{\varepsilon} = \min\{d(a, x_i),\ i \in [\![1; p]\!]\}$, il est facile de remarquer que $(B(a, \frac{\tilde{\varepsilon}}{2}) \setminus \{a\}) \cap A = \varnothing$, ce qui est en contradiction avec le fait que a soit un point d'accumulation.
   ($\Leftarrow$) Si pour tout $\varepsilon > 0$ $B(a, \varepsilon) \cap A$ est infini, alors d'une manière évidente $(B(a, \varepsilon) \cap A) \setminus \{a\} = (B(a, \varepsilon) \setminus \{a\}) \cap A$ est non vide, donc a est un point d'accumulation.
3. ($\Rightarrow$) Supposons que $a \in A^{ac}$, *i.e.* $\forall \varepsilon > 0$, $B(a, \varepsilon) \cap A$ est infini.
   Construisons une suite (x_n) qui vérifie la propriété voulue par récurrence. On définit x_0 comme un point pris au hasard dans A différent de a.
   Soit $n \in \mathbb{N}$, supposons que x_n est bien défini. On prend alors x_{n+1} dans $B(a, \frac{1}{2} \min(\frac{1}{n+2}, d(a, x_n))) \cap A$ (c'est possible car cet ensemble est non vide). Par construction (x_n) est injective et pour tout $n \in \mathbb{N}^*$ $d(x_n, a) \leq \frac{1}{n+1}$ et donc $x_n \xrightarrow[n \to +\infty]{} a$, d'où le résultat voulu.
   ($\Leftarrow$) Supposons l'existence d'une telle suite qu'on note (x_n). Soit $\varepsilon > 0$. Supposons que $B(a, \varepsilon) \cap A$ est fini. Il existe alors $p \in \mathbb{N}^*$ tel que $(B(a, \varepsilon) \setminus \{a\}) = \{x_1, \ldots, x_p\}$. En posant $\tilde{\varepsilon} = \min\{d(a, x_i),\ i \in [\![1; p]\!]\}$, il est facile de remarquer que $(B(a, \frac{\tilde{\varepsilon}}{2}) \setminus \{a\}) \cap A = \varnothing$, ce qui est absurde car pour n assez grand, x_n est dans $(B(a, \frac{\tilde{\varepsilon}}{2}) \setminus \{a\}) \cap A$.

Proposition VI.9.

En dimension finie, tout ensemble A borné infini admet un point d'accumulation, *i.e.* $A^{ac} \neq \varnothing$.

Démonstration. Il suffit de considérer une suite à valeurs dans cet ensemble injective (cette suite existe, car l'ensemble est infini) et utiliser le fait que cette suite est bornée, le théorème de Bolzano-Weierstraß et le point 3 de la proposition VI.8.

(cc) BY-NC-SA

Correction de l'exercice II.4. :

Soit $a \in \mathrm{X}$. Le fait que a ne soit pas une valeur d'adhérence de (u_n) est équivalent au fait qu'il existe $\varepsilon > 0$ tel que $\mathrm{B}(a,\varepsilon) \cap \{u_n,\ n \in \mathbb{N}\}$ soit fini. Soit $b \in \mathrm{B}(a,\varepsilon)$. Posons $\eta = \frac{\varepsilon - d(a,b)}{2}$. La boule $\mathrm{B}(b,\eta)$ est incluse dans $\mathrm{B}(a,\varepsilon)$ et donc $\mathrm{B}(b,\eta) \cap \{u_n,\ n \in \mathbb{N}\}$ est fini, car inclus dans $\mathrm{B}(a,\varepsilon) \cap \{u_n,\ n \in \mathbb{N}\}$. On peut affirmer alors que $b \in \mathrm{X} \setminus \mathrm{Adh}(u_n)$, et alors $\mathrm{B}(a,\varepsilon) \subset \mathrm{X} \setminus \mathrm{Adh}(u_n)$ ce qui donne finalement que $\mathrm{X} \setminus \mathrm{Adh}(u_n)$ soit ouvert, *i.e.* $\mathrm{Adh}(u_n)$ est fermé.

Correction de l'exercice II.5. :

Soit $(u_n) \in \mathrm{A}$. Posons $\varepsilon = \min\{|u_0|, \ldots, |u_\mathrm{N}|\}$. Montrons que $\mathrm{B}((u_n), \frac{\varepsilon}{2}) \subset \mathrm{A}$. Soit $(v_n) \in \mathrm{B}((u_n), \frac{\varepsilon}{2})$. Pour tout $i \in [\![0;\mathrm{N}]\!]$, on a

$$|u_i - v_i| \leq \sum_{k=0}^{+\infty} |u_k - v_k| < \frac{\varepsilon}{2}$$

donc

$$0 < |u_i| - \frac{\varepsilon}{2} \leq |v_i|$$

ce qui nous donne que pour tout $i \in [\![0;\mathrm{N}]\!]$, $v_i \neq 0$ et alors $(v_n) \in \mathrm{A}$, donc $\mathrm{B}((u_n), \frac{\varepsilon}{2}) \subset \mathrm{A}$. A est donc ouvert.

Correction de l'exercice II.6. :

On se place dans l'espace vectoriel normé $\mathrm{E} = \mathcal{C}([0,1], \mathbb{R})$.

1. Étudions $\mathrm{O} = \{f \in \mathrm{E},\ \forall x \in [0,1],\ f(x) > 0\}$.

   (a) O est ouvert pour la distance induite par $\|.\|_\infty$.
   En effet, si on considère $f \in \mathrm{O}$, alors on a $\varepsilon = \min_{x \in [0,1]} f(x) > 0$.
   Soit $g \in \mathrm{B}(f, \frac{\varepsilon}{2})$, alors

   $$\|g - f\|_\infty \leq \frac{\varepsilon}{2}$$

   ce qui donne

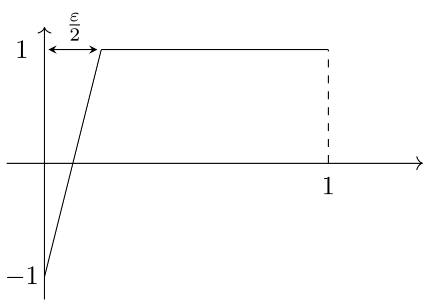
   $$\forall x \in [0,1],\quad |g(x)| \geq |f(x)| - \frac{\varepsilon}{2} \geq \frac{\varepsilon}{2} > 0,$$

   donc $g \in \mathrm{O}$ et alors $\mathrm{B}(f, \frac{\varepsilon}{2}) \subset \mathrm{O}$. D'où le fait que O soit ouvert pour $\|.\|_\infty$.

   (b) O n'est pas ouvert pour la distance induite par $\|.\|_1$.
   Exhibons un contrexemple qui permet d'affirmer cela. Soit $\varepsilon > 0$. considérons la fonction f partout égale à 1 sur $[0,1]$, et g une fonction de E définie de la manière suivante :

   $$g(x) = \begin{cases} \frac{4x}{\varepsilon} - 1 & \text{si } x \in [0, \frac{\varepsilon}{2}[\\ 1 & \text{si } x \in [\frac{\varepsilon}{2}, 1] \end{cases}$$

   Pour rendre ce choix plus parlant, voici un dessin de la fonction g :



BY-NC-SA

On a alors

$$\|f-g\|_1 = \int_0^1 |f(x)-g(x)|\,dx = \int_0^{\frac{\varepsilon}{2}} \left|\frac{4x}{\varepsilon} - 2\right| = \left[2x - \frac{2x^2}{\varepsilon}\right]_0^{\frac{\varepsilon}{2}} = \frac{\varepsilon}{2},$$

donc $g \in \mathrm{B}(f,\varepsilon)$ mais $g \notin \mathrm{O}$. On en déduit que O n'est pas ouvert pour la distance induite par $\|.\|_1$.

2. Étudions $\mathrm{F} = \{f \in \mathrm{E},\ \forall x \in [0,1],\ f(x) \geq 0\}$.

(a) F est fermé pour la distance induite par $\|.\|_\infty$.
Pour montrer cela, on va montrer que $\mathrm{E} \setminus \mathrm{F} = \{f \in \mathrm{E},\ \exists x\ [0,1],\ f(x) < 0\}$ est ouvert.
Soit $f \in \mathrm{E}\setminus\mathrm{F}$. Soit $a \in [0,1]$ tel que $f(a) < 0$. On pose $\varepsilon = \frac{|f(a)|}{2}$. Montrons que $\mathrm{B}(f,\varepsilon) \subset \mathrm{E}\setminus\mathrm{F}$.
soit $g \in \mathrm{B}(f,\varepsilon)$, on a

$$\|g-f\|_\infty < \varepsilon$$

et donc

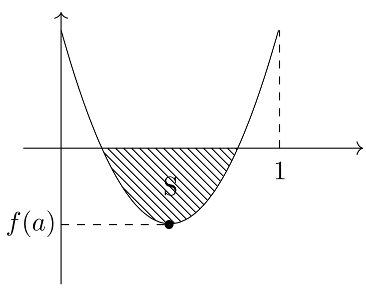
$$g(a) \leq f(a) + \varepsilon < 0.$$

On a alors $g \in \mathrm{B}(f,\varepsilon)$ et donc $\mathrm{B}(f,\varepsilon) \subset \mathrm{E} \setminus \mathrm{F}$. On en déduit finalement que $\mathrm{E} \setminus \mathrm{F}$ est ouvert, *i.e.* F est fermé pour la distance induite par $\|.\|_\infty$.

(b) F est fermé pour la distance induite par $\|.\|_1$.
On va procéder de la même manière qu'avant. Soit $f \in \mathrm{E} \setminus \mathrm{F}$. Soit $a \in [0,1]$ tel que $f(a) < 0$. Par continuité de f, il exsite $\eta > 0$ tel que pour tout $x \in [a-\eta, a+\eta]$, $f(x) < \frac{f(a)}{2}$. Donc pour tout $g \in \mathrm{F}$,

$$\|f-g\|_1 = \int_0^1 |f(x)-g(x)|\,dx \geq \int_{a-\eta}^{a+\eta} g(x) - f(x)dx \geq -\int_{a-\eta}^{a+\eta} f(x)dx \geq -\eta f(a).$$

En prenant donc $\varepsilon = -\eta f(a)$, on a $\mathrm{B}(f,\varepsilon) \subset \mathrm{E} \setminus \mathrm{F}$ et donc $\mathrm{E} \setminus \mathrm{F}$ est ouvert, *i.e.* F est fermé pour la distance induite par $\|.\|_1$.
Pour comprendre un peu ce qui se passe, considérons le cas où f est décrite par la courbe ci-dessous :



La distance en norme 1 entre f et une fonction partout positive ou nulle va être au moins égale à la surface S hachurée sur la figure. Ici, on a minoré cette surface par $-\eta f(a)$ ce qui nous a permis de trouver une boule incluse dans $\mathrm{E} \setminus \mathrm{F}$ de centre f.

Correction de l'exercice IV.3. :

1. On a $\mathrm{Int}(\mathrm{A} \cap \mathrm{B}) = \mathrm{Int}(\mathrm{A}) \cap \mathrm{Int}(\mathrm{B})$.
Montrons ce résultat. Soit $x \in \mathrm{X}$.

($\subset$) Supposons que $x \in \mathrm{Int}(\mathrm{A} \cap \mathrm{B})$. On dispose alors de $r > 0$ tel que $\mathrm{B}(x,r) \subset \mathrm{A} \cap \mathrm{B}$, *i.e.* $\mathrm{B}(x,r) \subset \mathrm{A}$ et $\mathrm{B}(x,r) \subset \mathrm{B}$ et donc $x \in \mathrm{Int}(\mathrm{A})$ et $x \in \mathrm{Int}(\mathrm{B})$. On a alors $x \in \mathrm{Int}(\mathrm{A}) \cap \mathrm{Int}(\mathrm{B})$

($\supset$) Supposons que $x \in \mathrm{Int}(\mathrm{A}) \cap \mathrm{Int}(\mathrm{B})$. On dispose alors de $r_\mathrm{A} > 0$ et $r_\mathrm{B} > 0$ tels que $\mathrm{B}(x, r_\mathrm{A}) \subset \mathrm{A}$ et $\mathrm{B}(x, r_\mathrm{B}) \subset \mathrm{B}$. En posant $r = \min\{r_\mathrm{A}, r_\mathrm{B}\} > 0$, on a $\mathrm{B}(x,r) \subset \mathrm{A} \cap \mathrm{B}$ et finalement $x \in \mathrm{Int}(\mathrm{A} \cap \mathrm{B})$.

BY-NC-SA

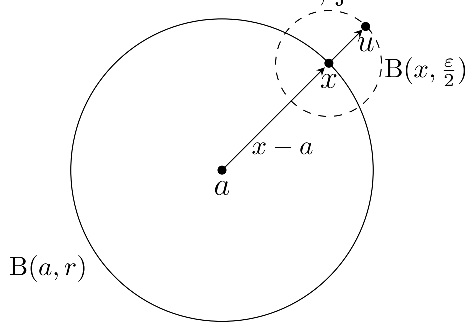
2. On a $\mathrm{Int}(\mathrm{A}) \cup \mathrm{Int}(\mathrm{B}) \subset \mathrm{Int}(\mathrm{A} \cup \mathrm{B})$.
   Pour montrer ce résultat, on va commencer par supposer que $x \in \mathrm{Int}(\mathrm{A}) \cup \mathrm{Int}(\mathrm{B})$. Sans perte de généralité, on suppose que $x \in \mathrm{Int}(\mathrm{A})$, on dispose alors de $r > 0$ tel que $\mathrm{B}(x, r) \subset \mathrm{A} \subset \mathrm{A} \cup \mathrm{B}$, donc $x \in \mathrm{Int}(\mathrm{A} \cup \mathrm{B})$.
   L'inclusion réciproque est fausse. Pour le voir, il suffit de prendre $\mathrm{A} = \mathbb{Q}$ et $\mathrm{B} = \mathbb{R} \setminus \mathbb{Q}$. On a $\mathrm{Int}(\mathbb{Q}) = \mathrm{Int}(\mathbb{R} \setminus \mathbb{Q}) = \varnothing$ et $\mathrm{Int}((\mathbb{R} \setminus \mathbb{Q}) \cup \mathbb{Q}) = \mathbb{R} \not\subset \mathrm{Int}(\mathbb{Q}) \cup \mathrm{Int}(\mathbb{R} \setminus \mathbb{Q}) = \varnothing$.

Correction de l'exercice IV.4. :

On a $\mathrm{B}(a, r) \subset \mathrm{B}_f(a, r)$ et est ouvert, donc $\mathrm{B}(a, r) \subset \mathrm{Int}(\mathrm{B}_f(a, r)) \subset \mathrm{B}_f(a, r)$. Montrons l'inclusion réciproque. Pour cela, nous allons montrer que $\mathrm{Int}(\mathrm{B}_f(a, r)) \setminus \mathrm{B}(a, r) = \varnothing$. Supposons le contraire, on dispose alors de $x \in \mathrm{Int}(\mathrm{B}_f(a, r)) \setminus \mathrm{B}(a, r) \subset \mathrm{B}_f(a, r) \setminus \mathrm{B}(a, r)$. On a alors nécessairement $\|x - a\| = r$. Soit $\varepsilon > 0$ tel que $\mathrm{B}(x, \varepsilon) \subset \mathrm{B}_f(a, r)$, posons $u = x + \varepsilon \frac{x-a}{2\|x-a\|}$. On a $u \in \mathrm{B}(x, \varepsilon)$, mais

$$\|a - u\| = \left\| a - x + \varepsilon \frac{x - a}{\|x - a\|} \right\| = \|x - a\| \left(1 + \frac{\varepsilon}{2\,\|x - a\|} \right) = \|x - a\| + \frac{\varepsilon}{2} = r + \frac{\varepsilon}{2} > r,$$

ce qui est en contradiction avec le fait que $\mathrm{B}(x, \varepsilon) \subset \mathrm{B}_f(a, r)$. Ceci donne finalement, on a $\mathrm{Int}(\mathrm{B}_f(a, r)) = \mathrm{B}(a, r)$. Pour comprendre l'intuition derrière le choix de u, jetons un coup d'oeil au dessin ci-dessous.



On souhaite sortir de la boule $\mathrm{B}(a, r)$, donc à partir de x, on va se déplacer vers la direction opposée à l'origine, d'où la présence du vecteur directeur unitaire $\frac{x-a}{\|x-a\|}$, mais on veut rester dans $\mathrm{B}(x, \varepsilon)$, on va donc multiplier ce vecteur "déplacement" par $\frac{\varepsilon}{2}$.

Correction de l'exercice V.5. :

Il suffit de prendre $\mathrm{A} = [0, 1[\cup]1, 2[\cup (]3, 4[\cap \mathbb{Q}) \cup \{5\}$.
En effet, on a

→ $\mathring{\mathrm{A}} =]0, 1[\cup]1, 2[$

→ $\overline{\mathring{\mathrm{A}}} = [0, 2]$

→ $\overline{\mathrm{A}} = [0, 2] \cup [3, 4] \cup \{5\}$

→ $\mathring{\overline{\mathrm{A}}} =]0, 2[\cup]3, 4[$

→ $\overline{\mathring{\overline{\mathrm{A}}}} = [0, 2] \cup [3, 4]$

→ $\mathring{\overline{\mathring{\mathrm{A}}}} =]0, 2[$

Tous ces ensembles sont bien deux à deux distincts.

Correction de l'exercice V.6. :

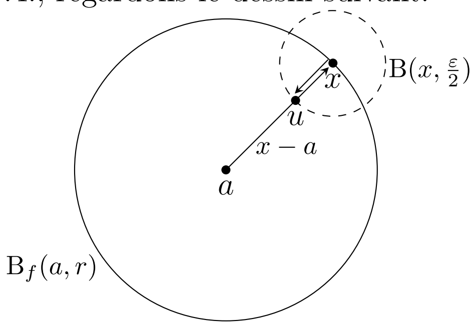
On va procéder d'une manière similaire à l'exercice IV.4. $\overline{\mathrm{B}(a, r)}$ étant le plus petit fermé qui contient $\mathrm{B}(a, r)$, on a $\mathrm{B}(a, r) \subset \overline{\mathrm{B}(a, r)} \subset \mathrm{B}_f(a, r)$. Il reste donc à montrer l'inclusion réciproque.

(cc) BY-NC-SA

Soit $x \in \mathrm{B}_f(a,r) \setminus \mathrm{B}(a,r)$. soit $\varepsilon > 0$, posons $u = x + \varepsilon \frac{a-x}{2\|a-x\|}$. On a

$$\|a-u\| = \left\| a - x - \varepsilon \frac{a-x}{2\|a-x\|} \right\| = \|a-x\| \left(1 - \frac{\varepsilon}{2\|a-x\|}\right) = \|a-x\| - \frac{\varepsilon}{2} = r - \frac{\varepsilon}{2} < r.$$

On suppose bien sûr ici que ε est suffisamment petit pour que $r - \frac{\varepsilon}{2}$ soit positif. On a donc $u \in \mathrm{B}(a,r)$ et de même, on a $\|x-u\| = \frac{\varepsilon}{2}$, et alors $u \in \mathrm{B}(x,\varepsilon)$ ce qui montre enfin que $\mathrm{B}_f(a,r) \subset \overline{\mathrm{B}(a,r)}$ et finalement que $\mathrm{B}_f(a,r) = \overline{\mathrm{B}(a,r)}$.
Pour faire le lien avec l'exercice IV.4., regardons le dessin suivant.



Ici, au lieu de partir dans la direction du vecteur $x-a$, on va vers la direction opposée puisqu'on veut rester dans la boule $\mathrm{B}(a,r)$. Mais on veut aussi être dans la boule $\mathrm{B}(x,\varepsilon)$, on multiplie alors pour cela le vecteur directeur $\frac{a-x}{\|a-x\|}$ par $\frac{\varepsilon}{2}$.

Correction de l'exercice V.7. :

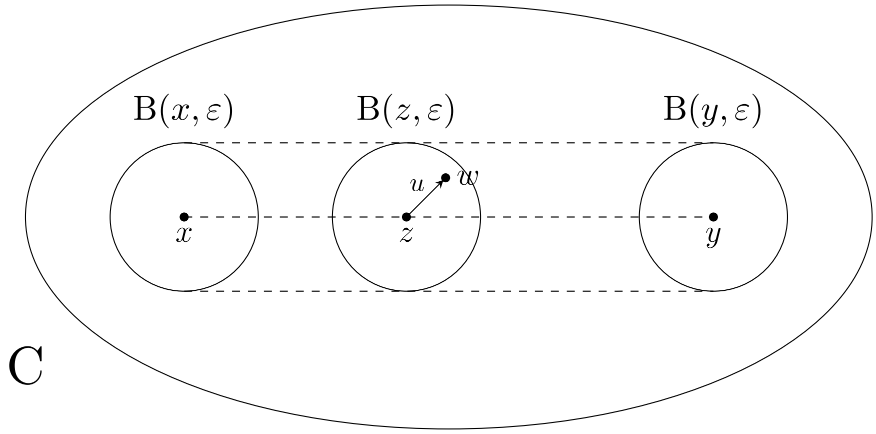
1. Montrons que $\overline{\mathrm{C}}$ et $\mathring{\mathrm{C}}$ sont convexes.

   → Convexité de $\overline{\mathrm{C}}$
   Soit $x, y \in \overline{\mathrm{C}}$. On dispose de deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ à valeurs dans C telles que

   $$x_n \xrightarrow[n\to+\infty]{} x \text{ et } y_n \xrightarrow[n\to+\infty]{} y.$$

   Soit $\lambda \in [0,1]$, on a $\lambda x_n + (1-\lambda) y_n \xrightarrow[n\to+\infty]{} \lambda x + (1-\lambda) y$, donc $\lambda x + (1-\lambda) y \in \overline{\mathrm{C}}$ et finalement $\overline{\mathrm{C}}$ est convexe.

   → Convexité de $\mathring{\mathrm{C}}$
   Soit $x, y \in \mathring{\mathrm{C}}$. On dispose de $\varepsilon > 0$ tel que $\mathrm{B}(x,\varepsilon) \subset \mathrm{C}$ et $\mathrm{B}(y,\varepsilon) \subset \mathrm{C}$ (quitte à prendre le minimum des rayons des deux boules). Soit $\lambda \in [0,1]$, montrons que $z = \lambda x + (1-\lambda) y \in \mathring{\mathrm{C}}$. Pour cela, regardons d'abord le dessin suivant.



   Le dessin ci-dessus suggère de montrer que $\mathrm{B}(z,\varepsilon) \subset \mathrm{C}$. Soit $w \in \mathrm{B}(z,\varepsilon)$. Posons $u = w - z$. On a alors $w = z + u = \lambda(x+u) + (1-\lambda)(y+u)$, et

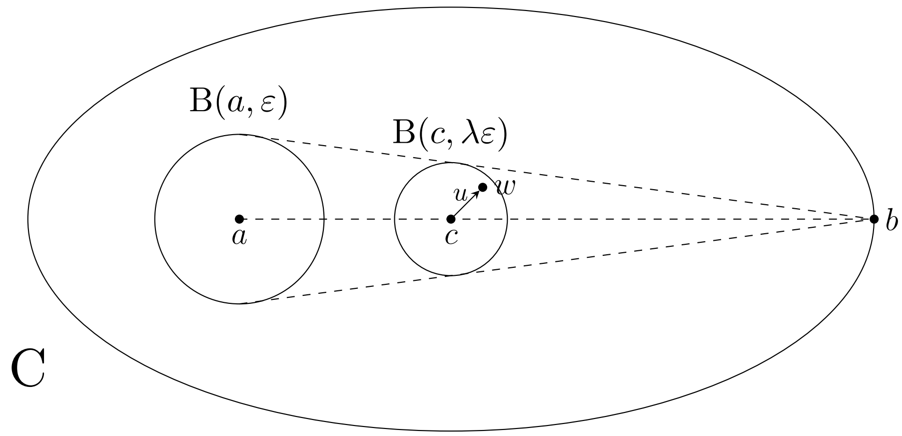
   $$\|x+u-x\| = \|w-z\| < \varepsilon \text{ et } \|y+u-y\| = \|w-z\| < \varepsilon,$$

(cc) BY-NC-SA

ce qui donne $y+u \in \mathrm{B}(y,\varepsilon) \subset \mathrm{C}$ et $x+u \in \mathrm{B}(x,\varepsilon) \subset \mathrm{C}$. Or w est combinaison linéaire convexe de $x+u$ et $y+u$, donc par convexité de C, $w \in \mathrm{C}$ et alors $z \in \mathring{\mathrm{C}}$ et finalement $\mathring{\mathrm{C}}$ est convexe.

2. Cette question est similaire à la question précédente. Soit $\lambda \in [0,1]$. On pose $c = \lambda a + (1-\lambda)b$. $a \in \mathring{\mathrm{C}}$, donc on dispose de $\varepsilon > 0$ tel que $\mathrm{B}(a,\varepsilon) \subset \mathrm{C}$. Si $\lambda = 0$, il n'y a rien à montrer puisque la question revient à montrer que $\varnothing \subset \mathrm{C}$, on considère dorénavant donc que $\lambda \in]0,1]$.

   → Partons du cas $b \in \mathrm{C}$. Observons le dessin suivant.



Le rayon de la boule du milieu peut être trouvé en utilisant des théorèmes de géométrie élémentaire (Thalès par exemple). Ce dessin suggère de montrer que $\mathrm{B}(c,\lambda\varepsilon) \subset \mathrm{C}$.
Soit $w \in \mathrm{B}(c,\lambda\varepsilon)$, on pose $u = w - c$. On a alors $w = c + u = \lambda(a + \frac{u}{\lambda}) + (1-\lambda)b$. Or

$$\left\| a + \frac{u}{\lambda} - a \right\| = \left\| \frac{w-c}{\lambda} \right\| < \varepsilon$$

donc $a + \frac{u}{\lambda} \in \mathrm{B}(a,\varepsilon) \subset \mathrm{C}$ et $b \in \mathrm{C}$ et w est combinaison convexe de $a + \frac{u}{\lambda}$ et b, ce qui nous permet de dire par convexité de C que $w \in \mathrm{C}$ et alors $\mathrm{B}(c,\lambda\varepsilon) \subset \mathrm{C}$ et finalement $[a,b[\subset \mathring{\mathrm{C}}$.

   → Passons au cas plus général où $b \in \overline{\mathrm{C}}$.
On reprend les notations du cas où $b \in \mathrm{C}$, et on considère une suite (b_n) à valeurs dans C qui tend vers b. Posons pour tout $n \in \mathbb{N}$, $c_n = \lambda a + (1-\lambda)b_n$. La démonstration du cas $b \in \mathrm{C}$ nous permet d'affirmer que $\mathrm{B}(c_n,\lambda\varepsilon) \subset \mathrm{C}$. On a alors

$$\begin{aligned}\|w - c_n\| &= \|w - c + c - c_n\| \\ &\leq \|w-c\| + \|c - c_n\| \xrightarrow[n\to+\infty]{} \|w-c\| < \lambda\varepsilon.\end{aligned}$$

Ceci nous permet d'affirmer l'existence de $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $\|w - c_n\| < \lambda\varepsilon$ et donc $w \in \mathrm{B}(c_n,\lambda\varepsilon) \subset \mathrm{C}$. On a alors $\mathrm{B}(c,\lambda\varepsilon) \subset \mathrm{C}$ et $c \in \mathring{\mathrm{C}}$. Finalement on conclut que $[a,b[\subset \mathring{\mathrm{C}}$.
L'intuition derrière le passage au cas où $b \in \overline{\mathrm{C}}$ était de prendre n assez grand pour que la boule $\mathrm{B}(c_n,\lambda\varepsilon) \subset \mathrm{C}$ soit quasi confondue avec $\mathrm{B}(c,\lambda\varepsilon)$ pour dire que tous les points de $\mathrm{B}(c,\lambda\varepsilon)$ sont dans C.

Correction de l'exercice V.8. :

1. Cette question a été déjà faite dans la correction de l'exercice II.6.

2. Trouvons $\mathring{\mathrm{F}}$ pour $\|.\|_1$ et $\|.\|_\infty$.

   → Pour $\|.\|_\infty$.
On pose $\mathrm{O} = \{f \in \mathrm{E},\ \forall x \in \mathrm{E},\ f(x) > 0\}$. On a d'après l'exercice 3, O est ouvert et $\mathrm{O} \subset \mathrm{F}$, donc $\mathrm{O} \subset \mathring{\mathrm{F}}$. Il reste donc à voir si $(\mathrm{F} \setminus \mathrm{O}) \cap \mathring{\mathrm{F}} = \varnothing$.

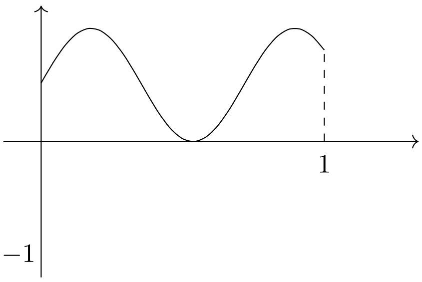
CC BY-NC-SA

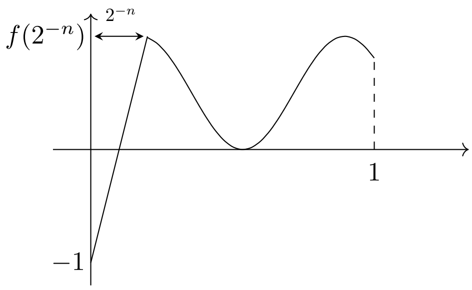
Soit $f \in \mathrm{F} \setminus \mathrm{O}$ et soit $\varepsilon > 0$. Le fait que $f \in \mathrm{F} \setminus \mathrm{O}$ implique $\exists a \in [0,1]$, $f(a) = 0$. En considérant $g = f - \frac{\varepsilon}{2}$, on peut voir immédiatement que $g \notin \mathrm{F}$ (car $g(a) = -\frac{\varepsilon}{2} < 0$) et $g \in \mathrm{B}(f, \varepsilon)$. On en déduit que $f \notin \mathring{\mathrm{F}}$ et finalement $\mathring{\mathrm{F}} = \mathrm{O}$.

→ Pour $\|.\|_1$.
Soit $f \in \mathrm{F}$. On considère la suite de fonctions $(g_n)_{n\in\mathbb{N}}$ définie sur $[0,1]$ pour tout $n \in \mathbb{N}$ par

$$g(x) = \begin{cases} \frac{x}{2^{-n}}(f(2^{-n}) + 1) - 1 & \text{si } x \in [0, 2^{-n}[\\ f(x) & \text{si } x \in [2^{-n}, 1]. \end{cases}$$

Voici un dessin d'un terme de cette suite de fonctions à droite et de la courbe de f à gauche.





Bien sûr, pour tout $n \in \mathbb{N}$, g_n est bien une fonction continue sur $[0,1]$. Il s'agit de la même idée que la question 1 de l'exercice II.6 où on a montré que O n'est pas ouvert, sauf que la fonction considérée est f au lieu d'une fonction qui est identiquement égale à 1 à laquelle on a ajouté une petite perturbation au voisinage de 0, qui fait sortir g_n de F sans trop s'éloigner de f. On a pour tout $n \in \mathbb{N}$, $g_n(0) = -1$, donc (g_n) est valeurs dans $\mathrm{E} \setminus \mathrm{F}$. De plus, on a

$$\begin{aligned} \|f - g_n\|_1 &= \int_0^1 |f(x) - g_n(x)|\, dx = \int_0^{2^{-n}} |f(x) - g_n(x)|\, dx \\ &\leq 2^{-n} \left(\|f\|_\infty + \|g_n\|_\infty\right) = 2^{-n}(\|f\|_\infty + \max\{1, f(2^{-n})\}) \\ &\leq 2^{-n}(\|f\|_\infty + \max\{1, \|f\|_\infty\}) \xrightarrow[n\to+\infty]{} 0, \end{aligned}$$

ce qui donne pour tout $\varepsilon > 0$, $\mathrm{B}(f, \varepsilon) \not\subset \mathrm{F}$. On conclut que $\mathring{\mathrm{F}} = \varnothing$.

Correction de l'exercice VI.4. :

1. Supposons que $\mathrm{G} \neq \{0\}$ et que A est vide. Alors nécessairement, on dispose de $x \in \mathrm{G}$ tel que $x < 0$. Or $-x \in \mathrm{G}$ et $-x > 0$ donc $-x \in \mathrm{A}$ ce qui est absurde.
2. Supposons que $a = 0$. Soit $x \in \mathbb{R}$. Étant donné que $\inf \mathrm{A} = 0$, A contient des éléments aussi petits qu'on veut. Soit $\varepsilon > 0$, on dispose donc de $\theta \in \mathrm{A}$ tel que $\theta < \varepsilon$. Or, pour tout $n \in \mathbb{Z}$, $n\theta \in \mathrm{G}$. On a alors

$$\left| x - \left\lfloor \frac{x}{\theta} \right\rfloor \theta \right| = \left(\frac{x}{\theta} - \left\lfloor \frac{x}{\theta} \right\rfloor \right) \theta \leq \theta < \varepsilon$$

et $\left\lfloor \frac{x}{\theta} \right\rfloor \theta \in \mathrm{G}$, donc on a trouvé un élément de G aussi proche qu'on veut de x, ce qui signifie que G est dense dans $\mathbb{R}$.
3. Supposons que $a > 0$.
   → Montrons d'abord que $a \in \mathrm{G}$.
   Supposons par l'absurde que $a \notin \mathrm{G}$. On dispose alors de $b, c \in \mathrm{A}$ tels que $b > c$ et $|b - a| \leq \frac{a}{4}$ et $|c - a| \leq \frac{a}{4}$. On a donc

$$b - c = |b - a + a - c| \leq |b - a| + |a - c| \leq \frac{a}{2}.$$

(cc) BY-NC-SA

Or $b-c \in \mathrm{A}$, ce qui implique
$$a < b-c \leq \frac{a}{2}.$$
Ceci est absurde, donc $a \in \mathrm{G}$.

→ Montrons ensuite que $\mathrm{G} = a\mathbb{Z}$.
Par stabilité par addition et soustraction, on a évidemment $a\mathbb{Z} \subset \mathrm{G}$.
Supposons que $\mathrm{G} \neq a\mathbb{Z}$. On dispose donc de $x \in \mathrm{G} \setminus a\mathbb{Z}$. Quitte à passer à l'opposé, on suppose que $x > 0$. On a alors
$$0 < x - \left\lfloor \frac{x}{a} \right\rfloor a = a\left(\frac{x}{a} - \left\lfloor \frac{x}{a} \right\rfloor\right) < a,$$
ce qui donne $x - \left\lfloor \frac{x}{a} \right\rfloor a \in \mathrm{A}$ et $x - \left\lfloor \frac{x}{a} \right\rfloor a < a$ ce qui est absurde. On en déduit donc que $\mathrm{G} = a\mathbb{Z}$.

4. Pour commencer, remarquons que par parité du cos, $\{\cos(n), n \in \mathbb{N}\} = \{\cos(n), n \in \mathbb{Z}\}$.

→ Montrons d'abord que $\mathrm{G} = \mathbb{Z} + 2\pi\mathbb{Z}$ est dense dans $\mathbb{R}$.
G est un sous groupe additif de $\mathbb{R}$. Supposons qu'il existe $a \in \mathbb{R}^*$ tel que $\mathrm{G} = a\mathbb{Z}$. Il existe alors $n, m \in \mathbb{Z}^*$ tel que $1 = an$ et $2\pi = am$, donc, $\frac{2\pi}{1} = \frac{an}{am}$, *i.e.* $\pi = \frac{n}{2m}$ ce qui implique que π est rationnel, absurde. On en déduit que d'après les questions précédentes, G est dense dans $\mathbb{R}$.

→ Montrons à présent que $\{\cos(n), n \in \mathbb{N}\}$ est dense dans $\mathbb{R}$.
On a
$$\mathrm{B} = \{\cos(n), n \in \mathbb{N}\} = \{\cos(n), n \in \mathbb{Z}\} = \{\cos(n), n \in \mathrm{G}\}.$$
Soit $x \in [-1, 1]$ et y un antécédent de x par cos. G étant dense dans $\mathbb{R}$, on dispose d'une suite (y_n) à valeurs dans G qui tend vers y. Ceci implique que la suite $(\cos(y_n))$ est valeurs dans B et tend vers x par continuité du cos. On en déduit donc que $\{\cos(n), n \in \mathbb{N}\}$ est bien dense dans $[-1, 1]$.

5. Pour commencer, remarquons que $f(\mathbb{Z} + \sqrt{2}\mathbb{Z}) = \{f(1)\}$ par périodicité de f. $\sqrt{2}$ étant irrationnel, on peut montrer comme on l'a fait pour $\mathbb{Z} + 2\pi\mathbb{Z}$ que $\mathrm{H} = \mathbb{Z} + \sqrt{2}\mathbb{Z}$ est dense dans $\mathbb{R}$. Soit $x \in \mathbb{R}$, on dispose d'une suite (x_n) à valeurs dans H telle que (x_n) converge vers x. On a donc par continuité de f
$$f(1) = f(x_n) \xrightarrow[n \to +\infty]{} f(x).$$
On en déduit donc que f est constante égale à $f(1)$.

Correction de l'exercice VI.6. :

(⇐) Supposons que pour tout i, $\Omega_i \neq \varnothing$ et considérons pour tout i, a_i un élément de Ω_i.
Posons $\mathrm{A} = \{a_i,\ i \in \mathrm{I}\}$, c'est un ensemble dénombrable. Montrons que A est dense dans X. Soit $x \in \mathrm{X}$ et $\varepsilon > 0$. Par hypothèse, il existe $\mathrm{J} \subset \mathrm{I}$ tel que $\mathrm{B}(x, \varepsilon) = \bigcup_{i \in \mathrm{J}} \Omega_i$, il suffit donc de prendre $i \in \mathrm{J}$ pour voir que a_i est dans $\mathrm{B}(x, \varepsilon) \cap \mathrm{A}$. D'où la densité de A dans X.

(⇒) Soit $\mathrm{A} = \{a_i,\ i \in \mathbb{N}\}$ une partie dénombrable dense de X. Considérons la famille d'ouverts $(\Omega_{n,m})_{n,m \in \mathbb{N}} = (\mathrm{B}(a_n, \frac{1}{m+1}))_{n,m \in \mathbb{N}}$ et montrons qu'elle vérifie la propriété voulue.
Soit O un ouvert de X. Considérons l'ensemble $\mathrm{J} = \{(n, m) \in \mathbb{N}^2,\ \mathrm{B}(a_n, \frac{1}{m+1}) \subset \mathrm{O}\}$.
On a $\bigcup_{(n,m) \in \mathrm{J}} \Omega_{n,m} \subset \mathrm{O}$. Montrons l'inclusion réciproque. Soit $x \in \mathrm{O}$. Comme A est dense dans X, il existe $a_n \in \mathrm{A}$ (avec $n \in \mathbb{N}$) et $m \in \mathbb{N}$ arbitrairement grand tel que $x \in \mathrm{B}(a_n, \frac{1}{m+1})$. On prend alors m assez grand pour qu'on ait $\mathrm{B}(a_n, \frac{1}{m+1}) \subset \mathrm{O}$, *i.e.* $(n, m) \in \mathrm{J}$ ce qui nous donne $x \in \bigcup_{(m,n) \in \mathrm{J}} \Omega_{n,m}$.
On a donc bien obtenu le résultat voulu.

(cc) BY-NC-SA

Chapitre 11.3

Limites et continuité

Dans tout ce chapitre, on considère X et Y deux espaces metriques respectivement munis des distances d et δ, A une partie de X, $a \in \overline{\mathrm{A}}$ et f une fonction de A dans Y.

I Limites

Définition I.1.

On dit que f admet une limite en a selon A lorsqu'il exsite $\ell \in \mathrm{Y}$ tel que

$$\forall \mathrm{V} \in \mathcal{V}(\ell),\ \exists \mathrm{U} \in \mathcal{V}(a),\ f(\mathrm{U} \cap \mathrm{A}) \subset \mathrm{V}\ (*)$$

Remarques.

→ une formulation équivalente à la définition ci-dessous est la suivante.
f admet une limite en a selon A s'il existe $\ell \in \mathrm{Y}$ tel que

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \mathrm{A},\ d(a,x) < \eta \Longrightarrow \delta(f(x),a),\ell) < \varepsilon.$$

→ La propriété $(*)$ est équivalente à

$$\forall \mathrm{V} \in \mathcal{V}(\ell),\ f^{-1}(\mathrm{V}) \cap \mathrm{A} \in \{\mathrm{U} \cap \mathrm{A},\ \mathrm{U} \in \mathcal{V}(a)\}.$$

Proposition I.2.

1. Si f tends vers ℓ et ℓ', alors $\ell = \ell'$.
2. Si f admet une limite ℓ en a selon A et $a \in \mathrm{A}$, alors nécessairement $f(a) = \ell$.
3. Si f admet une limite en a selon A, elle au bornée au voisinage de a (dans A).

Démonstration.

1. Soit $\ell, \ell' \in \mathrm{Y}$ telles que f tend vers ℓ et ℓ' en a selon A.Supposons que $\ell \neq \ell'$. Appliquons la définition de limite avec $\varepsilon = \frac{\delta(\ell,\ell')}{10}$. Il existe $\eta, \eta' > 0$ tel que
   → $\forall x \in \mathrm{A},\ d(a,x) < \eta \Longrightarrow \delta(f(x),\ell) < \frac{\delta(\ell,\ell')}{10}$
   → $\forall x \in \mathrm{A},\ d(a,x) < \eta' \Longrightarrow \delta(f(x),\ell') < \frac{\delta(\ell,\ell')}{10}$
   On a alors pour tout $x \in \mathrm{A}$ tel que $d(x,a) < \min(\eta,\eta')$
   $$\delta(\ell,\ell') < \delta(f(x),\ell) + \delta(f(x),\ell') < \frac{\delta(\ell,\ell')}{5},$$
   ce qui est absurde et finalement $\ell = \ell'$.
2. Si a admet une limite en a selon A, alors $\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \mathrm{A},\ \ d(x,a) < \eta \Longrightarrow \delta(f(x),\ell) < \varepsilon$. Or pour tout $\eta > 0$, $d(a,a) = 0 < \eta$, donc pour tout $\varepsilon > 0$, $d(f(a),\ell) < \varepsilon$ ce qui donne directement que $d(f(a),\ell) = 0$ et finalement $f(a) = \ell$.
3. Il suffit d'appliquer la définition avec $\varepsilon = 1$. En effet, il existe $\eta > 0$ tel que pour tout $x \in \mathrm{A}$, $d(x,a) < \eta \implies \delta(f(x),\ell) < 1$. Ceci implique que $f(\mathrm{A} \cap \mathrm{B}(a,\eta)) \subset \mathrm{B}(\ell,1)$ qui est borné. $\mathrm{B}(a,\eta)$ est un voisinage de a, donc f est bien bornée au voisinage de a dans A.

(cc) BY-NC-SA

Proposition I.3.

Les propositions suivantes sont équivalentes

1. f possède une limite en a selon A.
2. Pour toute suite $(x_n) \in \mathrm{A}^{\mathbb{N}}$ qui converge vers a, la suite $(f(x_n))$ converge.

Démonstration.
(1) $\Rightarrow$ (2) Soit $(x_n) \in \mathrm{A}^{\mathbb{N}}$ tel que $x_n \xrightarrow[n\to+\infty]{} a$.
Soit $\varepsilon > 0$. Il existe $\eta > 0$ tel que $\forall x \in \mathrm{B}(a, \eta) \cap \mathrm{A}$, $\delta(f(x), \ell) < \varepsilon$.
(x_n) converge vers a, il existe donc $\mathrm{N} \in \mathbb{N}$ tel que $\forall n \geq \mathrm{N}$, $d(x_n, a) < \eta$ et alors $\delta(f(x_n), \ell) < \varepsilon$.
On en déduit donc que $(f(x_n))$ converge vers ℓ.

(2) $\Rightarrow$ (1)

→ Montrons tout d'abord que toutes les suites de la forme $(f(x_n))$ avec (x_n) une suite à valeurs dans A convergeant vers a convergent vers une même limite ℓ.
Soit (x_n) et (y_n) deux suites à valeurs dans A convergeant vers a. À partir de ces deux suites, on construit la suite (z_n) définie de la manière suivante : $\forall n \in \mathbb{N}$, $\begin{cases} z_{2n} = x_n. \\ z_{2n+1} = y_n \end{cases}$
On a alors bien évidemment $z_n \xrightarrow[n\to+\infty]{} a$ et donc par hypothèse $f(z_n)$ converge vers une limite qu'on notera ℓ. En considérant les deux suites extraites $(f(z_{2n+1})) = (f(y_n))$ et $(f(z_{2n})) = (f(x_n))$, on en déduit immédiatement que $(f(x_n))$ et $(f(y_n))$ admettent la même limite que $(f(z_n))$.

→ Il reste à montrer maintenant que la limite de f en a selon A est bien ℓ.
Supposons que f ne converge pas vers ℓ, *i.e.*

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x \in \mathrm{B}(a, \eta) \cap \mathrm{A},\ \delta(f(x), \ell) \geq \varepsilon.$$

Cette proposition peut être reformulée de la manière suivante

$$\exists \varepsilon > 0,\ \forall n \in \mathbb{N},\ \exists x \in \mathrm{B}\left(a, \frac{1}{n+1}\right) \cap \mathrm{A},\ \delta(f(x), \ell) \geq \varepsilon.$$

On peut alors construire une suite (x_n) à valeurs dans A telle que pour tout $n \in \mathbb{N}$, $d(x_n, a) < \frac{1}{n+1}$ et $\delta(f(x_n), \ell) \geq \varepsilon$.
En résumé, on a

- (x_n) est une suite à valeurs dans A qui converge vers a.
- Pour tout n, $\delta(f(x_n), \ell) \geq \varepsilon$ donc $(f(x_n))$ ne peut pas tendre vers ℓ.

ce qui est absurde par hypothèse.

Proposition I.4.

Soit B une partie de Y et Z un espace métrique. Si

→ $a \in \overline{\mathrm{A}}$ et f est une fonction de A dans Y qui admet une limite b en a selon A.

→ $b \in \overline{\mathrm{B}}$ et g est une fonction de B dans Z qui admet une limite ℓ en b selon B.

→ $f(\mathrm{A}) \subset \mathrm{B}$.

Alors $g \circ f$ admet la limite ℓ en a selon A.

Démonstration. soit $\mathrm{W} \in \mathcal{V}(\ell)$.

→ Il existe $\mathrm{V} \in \mathcal{V}(b)$ tel que $g(\mathrm{V} \cap \mathrm{B}) \subset \mathrm{W}$.

→ Il existe $\mathrm{U} \in \mathcal{V}(a)$ tel que $f(\mathrm{U} \cap \mathrm{A}) \subset \mathrm{V}$.

BY-NC-SA

On a aussi $f(\mathrm{U}\cap\mathrm{A})\subset f(\mathrm{A})\subset\mathrm{B}$, donc $f(\mathrm{U}\cap\mathrm{A})\subset\mathrm{V}\cap\mathrm{B}$ et alors $g\circ f(\mathrm{U}\cap\mathrm{A})\subset\mathrm{W}$.
En résumé, on a montré que

$$\forall \mathrm{W}\in\mathcal{V}(\ell),\ \exists \mathrm{U}\in\mathcal{V}(a),\ g\circ f(\mathrm{U}\cap\mathrm{A})\subset\mathrm{W}.$$

Attention : la condition $f(\mathrm{A})\subset\mathrm{B}$ est nécessaire.
En effet, si on prend $\mathrm{A}=\mathbb{R}^*$, $\mathrm{B}=\mathbb{R}^*$, $\mathrm{Y}=\mathbb{R}$, $\mathrm{Z}=\mathbb{R}$ tous munis de la distance induite par la valeur absolue. On considère le cas où

→ f est une fonction de A dans Y, définie par $\forall x\neq 0,\ f(x)=x\sin\left(\frac{1}{x}\right)$.

→ g est une fonction de $\mathbb{R}$ dans Z, définie par $\forall x\in\mathbb{R},\ g(x)=1_{x=0}$.

On a alors

→ f admet une limite 0 en 0 selon $\mathrm{A}=\mathbb{R}^*$.

→ g admet une limite 0 en 0 selon $\mathrm{B}=\mathbb{R}^*$.

→ $f(\mathrm{A})=f(\mathbb{R}^*)$ contient 0 donc $f(\mathrm{A})\not\subset\mathbb{R}^*=\mathrm{B}$.

mais $g\circ f$ n'admet pas de limite en 0 selon $\mathbb{R}^*$. En effet, si on considère les suites à valeurs dans $\mathbb{R}^*$, $(u_n)_{n\in\mathbb{N}}=\left(\frac{1}{2\pi n}\right)$ et $(v_n)_{n\in\mathbb{N}}=\left(\frac{1}{2\pi n+\frac{\pi}{2}}\right)$, on a

→ $g\circ f(u_n)\xrightarrow[n\to+\infty]{}1$.

→ $g\circ f(v_n)\xrightarrow[n\to+\infty]{}0$.

On en déduit donc que $g\circ f$ n'admet pas de limite en 0.

Exercice I.5.

Soit $f\in\mathcal{C}([-1,1],\mathbb{R})$ telle que

$$\frac{f(x)-f(\frac{x}{2})}{x}\xrightarrow[x\to 0]{}\ell\in\mathbb{R}.$$

Montrer que f est dérivable en 0.

Exercice I.6.

Soit $f\in\mathcal{C}(\mathbb{R},\mathbb{R})$ telle que

$$f(x+1)-f(x)\xrightarrow[x\to+\infty]{}\mathrm{L}\in\mathbb{R}.$$

Montrer que $x\longmapsto\dfrac{f(x)}{x}$ admet une limite en $+\infty$.

II Continuité

Dans cette partie, on considère que l'ensemble de départ de f est X et que $a\in\mathrm{X}$. Nous nous allons donc dire "f admet une limite ℓ en a" au lieu de "f admet une limite ℓ en a selon X" lorsqu'il n'y a pas ambigüité.

BY-NC-SA

1. Continuité locale

Proposition II.1.

Soit f une fonction de X dans Y. Les propositions suivantes sont équivalentes.

1. f admet une limite en a.
2. $\forall \mathrm{V} \in \mathcal{V}(f(a)),\ \exists \mathrm{U} \in \mathcal{V}(a),\ f(\mathrm{U}) \subset \mathrm{V}$.
3. $\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \mathrm{X},\ d(a, x) < \eta \implies \delta(f(x), f(a)) < \varepsilon$.

Lorsque f vérifie l'une de ces propriétés, on dit que f est continue en a.

Démonstration.
(1) $\Rightarrow$ (2) Si f admet une limite en a selon X, alors nécessairement, d'après ce qui précède, puisque $a \in \mathrm{X}$, cette limite est égale à $f(a)$. Le point (2) est tout simplement l'application de la définition de limite pour $\ell = f(a)$.
(2) $\Rightarrow$ (3) soit $\varepsilon > 0$. On considère le voisinage de $f(a)$, $\mathrm{B}(f(a), \varepsilon)$. Par hypothèse, il existe $\mathrm{U} \in \mathcal{V}(a)$ tel que $f(\mathrm{U}) \subset \mathrm{B}(f(a), \varepsilon)$. U étant un voisinage de a, il existe $\eta > 0$ tel que $\mathrm{B}(a, \eta) \subset \mathrm{U}$. et alors on a $f(\mathrm{B}(a, \eta)) \subset \mathrm{U}$. En résumé,

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ f(\mathrm{B}(a, \eta)) \subset \mathrm{B}(f(a), \varepsilon)$$

i.e.

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ d(x, a) < \eta \implies \delta(f(x), a) < \varepsilon.$$

(3) $\Rightarrow$ (1) Il s'agit tout simplement de l'une des formulations du fait que f admet une limite $f(a)$ en a selon X.

Proposition II.2.

Les propositions suivantes sont équivalentes.

1. f est continue en a.
2. Pour toute suite $(x_n) \in \mathrm{X}^{\mathbb{N}}$, $x_n \xrightarrow[n \to +\infty]{} a \implies (f(x_n))$ converge.

Remarques.
La composition de deux fonctions continues est continue.
Lorsque Y est un espace vectoriel normé muni d'une norme $\|.\|$ et δ la distance induite par $\|.\|$, alors les opérations conservent la continuité :

- $\rightarrow$ La somme de deux fonctions continues est continue.
- $\rightarrow$ Lorsque $\|.\|$ est une norme d'algèbre, le produit de deux fonctions continues est continu.

Exercice II.3.

Soit a, b deux réels tels que $a < b$. On dit qu'une fonction $f : [a, b] \longrightarrow \mathbb{R}$ est réglée lorsque f admet une limite à droite et à gauche en tout point de $[a, b]$. Par exemple, les fonctions monotones sont réglées.
Supposons que f soit réglée. Montrer que l'ensemble des points de discontinuité de f est dénombrable.

2. Continuité globale

Pour commencer, on va énoncer un résultat qui sera utile par la suite.

(cc) BY-NC-SA

Lemme II.4.

Les propositions suivantes sont équivalentes

1. f est continue en a.
2. $\forall \mathrm{V} \in \mathcal{V}(f(a)),\ f^{-1}(\mathrm{V}) \in \mathcal{V}(a)$.

Démonstration.
(1) $\Rightarrow$ (2) Si f est continue en a, alors par définition,

$$\forall \mathrm{V} \in \mathcal{V}(f(a)),\ \exists \mathrm{U} \in \mathcal{V}(a),\ f(\mathrm{U}) \subset \mathrm{V}$$

et donc

$$\forall \mathrm{V} \in \mathcal{V}(f(a)),\ \exists \mathrm{U} \in \mathcal{V}(a),\ f^{-1}(f(\mathrm{U})) \subset f^{-1}(\mathrm{V}).$$

Or $\mathrm{U} \subset f^{-1}(f(\mathrm{U})) \subset f^{-1}(\mathrm{V})$ et finalement pour tout $\mathrm{V} \in \mathcal{V}(f(a))$, on a $f^{-1}(\mathrm{V}) \in \mathcal{V}(a)$ car contient U un voisinage de a.

(2) $\Rightarrow$ (1) Soit $\mathrm{V} \in \mathcal{V}(f(a))$, posons $\mathrm{U} = f^{-1}(\mathrm{V}) \in \mathcal{V}(a)$. On a $f(\mathrm{U}) = f(f^{-1}(\mathrm{V})) \subset \mathrm{V}$, d'où la continuité de f en a.

Proposition II.5.

Les propositions suivantes sont équivalentes.

1. f est continue en tout point de X.
2. Pour tout ouvert Ω de Y, $f^{-1}(\Omega)$ est un ouvert de X.
3. pour tout fermé F de Y, $f^{-1}(\mathrm{F})$ est un fermé de X.

Démonstration.
(1) $\Rightarrow$ (2) Soit Ω un ouvert de Y et $x \in f^{-1}(\Omega)$. Ω est un ouvert contenant $f(x)$, donc $\Omega \in \mathcal{V}(f(x))$, alors par continuité de f et le lemme précédent, $f^{-1}(\Omega) \in \mathcal{V}(x)$. Il existe alors $\varepsilon > 0$ tel que $\mathrm{B}(x, \varepsilon) \subset f^{-1}(\Omega)$ et finalement $f^{-1}(\Omega)$ est ouvert.

(2) $\Rightarrow$ (1) Soit $\varepsilon > 0$ et $b \in \mathrm{X}$. Posons $\Omega = \mathrm{B}(f(b), \varepsilon)$. Par hypothèse, $f^{-1}(\Omega)$ est ouvert et donc il existe $\eta > 0$ tel que $\mathrm{B}(b, \varepsilon) \subset f^{-1}(\Omega)$, *i.e.* $f(\mathrm{B}(b, \eta)) \subset \mathrm{B}(f(a), \varepsilon)$, ce qui donne directement la continuité de f en b, donc en tout point de X.

(2) $\Rightarrow$ (3) Soit F un fermé de Y. $\mathrm{Y} \setminus \mathrm{F}$ est ouvert de Y, donc $f^{-1}(\mathrm{Y} \setminus \mathrm{F}) = \mathrm{X} \setminus f^{-1}(\mathrm{F})$ est ouvert et finalement $f^{-1}(\mathrm{F})$ est fermé.

(3) $\Rightarrow$ (2) Le sens réciproque peut être fait de la même manière et est laissé comme exercice au lecteur.

Applications.

→ Si f est une fonction continue de X dans Y, alors pour tout $b \in \mathrm{Y}$, $f^{-1}(\{b\})$ est fermé car $\{b\}$ est fermé.

→ Si de plus $\mathrm{Y} = \mathbb{R}$, alors $\{x \in \mathrm{X},\ f(x) < 0\}$ est ouvert car est égal à $f^{-1}(]-\infty, 0[)$ et $]-\infty, 0[$ est ouvert.

→ Si Y est un espace vectoriel normé et g une fonction continue de X dans Y, alors l'ensemble où f coïncide avec g est un fermé car il est égal à

$$\{x \in \mathrm{X}\ f(x) = g(x)\} = \{x \in \mathrm{X}\ (f-g)(x) = 0\} = (f-g)^{-1}(\{0\}).$$

Le fait que $f - g$ continue et $\{0\}$ est fermé permet de conclure.

(cc) BY-NC-SA

Exemple. (principe des prolongements des identités)
Soit E un espace vectoriel normé et $f, g \in \mathcal{C}(\mathrm{X}, \mathrm{E})$. Si f et g coïncident sur une partie A dense dans X, alors $f = g$.
En effet, si B est l'ensemble où f coïncide avec g, alors d'après le troisième point de la proposition précédente, B est fermé. On a alors $\mathrm{A} \subset \mathrm{B}$ et donc $\mathrm{X} = \overline{\mathrm{A}} \subset \overline{\mathrm{B}} = \mathrm{B}$ car B est fermé. On obtient donc que $\mathrm{B} = \mathrm{X}$ et que finalement $f = g$.

Exercice II.6.

Soit f une fonction de X dans Y. Montrer que les propositions suivantes sont équivalentes.

1. f est continue en tout point de X.
2. Pour toute partie A de X, $f(\overline{\mathrm{A}}) \subset \overline{f(\mathrm{A})}$.

Exercice II.7.

Soit $f : (\mathbb{R}, +) \longrightarrow (\mathbb{R}, +)$ un morphisme de groupe. Montrer que les propriétés suivantes sont équivalentes.

1. f est continue.
2. $\exists \lambda \in \mathbb{R}, \forall x \in \mathbb{R},\ f(x) = \lambda x$.
3. f est continue en un point.
4. f est continue en 0.
5. f est bornée au voisinage de 0.

Exercice II.8.

Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ un morphisme de corps continu. Montrer que $f = \mathrm{Id}$.

Exercice II.9.

Soit $f : \mathbb{C} \longrightarrow \mathbb{C}$ un morphisme de corps continu. Montrer que $f = \mathrm{Id}$ ou $f = (z \longmapsto \overline{z})$.

3. Homéomorphismes

Définition II.10.

Soit f une application de X dans Y. Si

- $\rightarrow$ f est continue.
- $\rightarrow$ f est bijective.
- $\rightarrow$ f^{-1} est continue.

alors on dit que f est un homéomorphisme de X dans Y.

Exemples.

- $\rightarrow$ Soit E un $\mathbb{K}-$espace vectoriel normé et $(\lambda, a) \in \mathbb{K}^* \times \mathrm{E}$. L'application $h_{\lambda,a} : \begin{cases} \mathrm{E} \longrightarrow \mathrm{E} \\ x \longmapsto \lambda x + a \end{cases}$ est un homéomorphisme.
- $\rightarrow$ La fonction tangente est un homéomorphisme de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ dans $\mathbb{R}$.

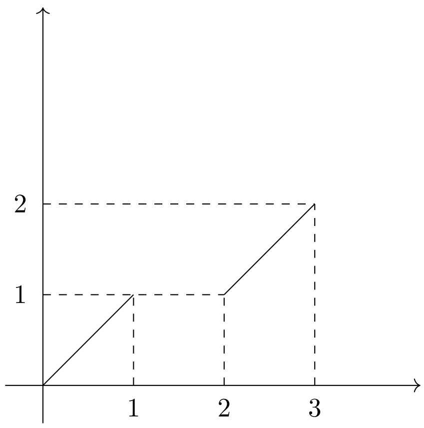
BY-NC-SA

Proposition II.11.

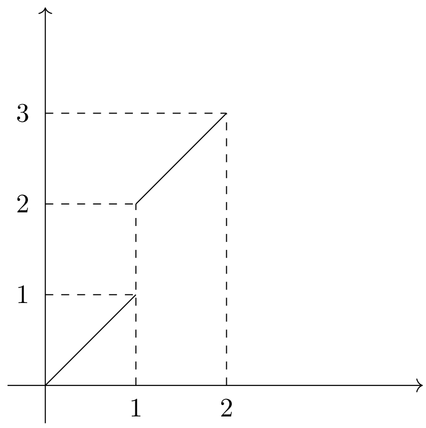
Soit I et J deux intervalles de $\mathbb{R}$. Si $f : \mathrm{I} \longrightarrow \mathrm{J}$ est une bijection continue, alors f^{-1} est continue.

Attention : cette propriété n'est plus vraie lorsque I et J ne sont pas des intervalles. En effet, si on considère la fonction de $[0,1[\cup[2,3]$ dans $[0,2]$ $f : x \longmapsto \begin{cases} x & \text{si } 0 \leq x < 1 \\ x-1 & \text{si } 2 \leq x \leq 3 \end{cases}$, cette fonction est continue bijective mais sa bijection réciproque n'est pas continue.

Courbe de f



Courbe de f^{-1}



Proposition II.12.

Soient Ω et Ω' deux ouverts de $\mathbb{R}$. Si $f : \Omega \longrightarrow \Omega'$ est une bijection continue, alors f^{-1} est continue.

4. Applications lipschitziennes

Définition II.13.

Soit f une application de X dans Y. On dit que f est lipschitzienne ou $k-$lipschitzienne s'il existe $k \in \mathbb{R}^+$ tel que

$$\forall (x,y) \in \mathrm{X}^2,\ \delta(f(x), f(y)) \leq k d(x,y).$$

Exemples.

→ Soit E un espace vectoriel normé muni d'une norme $\|.\|$. l'application $x \longmapsto \|x\|$ est $1-$lipschitzienne. En effet, pour tout $x, y \in \mathrm{E}$, on a

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

→ Soit $a \in \mathrm{X}$. L'application $x \longmapsto d(a,x)$ est $1-$lipschitzienne. En effet, pour tout $x, y \in \mathrm{X}$

$$|d(a,x) - d(a,y)| \leq d(x,y).$$

(cc) BY-NC-SA

Exercice II.14.

Soit I un intervalle de $\mathbb{R}$ et $f \in \mathcal{C}^1(\mathrm{I}, \mathbb{C})$. Montrer que les deux propriétés suivantes sont équivalentes.

1. f est lipschitzienne.
2. f' est bornée.

Vocabulaire. lorsque f est lipschitzienne, on appelle rapport de lipschitzianité le nombre

$$k = \inf\{\alpha \in \mathbb{R}^+,\ f \text{ est } \alpha\text{−lipschitzienne}\}.$$

Exercice II.15.

Soit A une partie non vide X. Posons pour tout $x \in \mathrm{X}$,

$$d(x, \mathrm{A}) = \inf_{a \in \mathrm{A}} d(x, a).$$

1. Montrer que pour tout $x \in \mathrm{X}$, $d(x, \mathrm{A}) = 0 \iff x \in \overline{\mathrm{A}}$.
2. Montrer que $x \longmapsto d(x, \mathrm{A})$ est 1−lipschitzienne.
3. Montrer que si A et B sont deux fermés disjoints non vides de X, alors il existe deux ouverts U et V de X tels que $\mathrm{A} \subset \mathrm{U}$, $\mathrm{B} \subset \mathrm{V}$, et $\mathrm{U} \cap \mathrm{V} = \varnothing$.

Exercice II.16.

Soit $k \in \mathbb{R}^+$, I un ensemble, et $(f_i)_{i \in \mathrm{I}}$ une famille de fonctions k−lipschitziennes. Supposons qu'il existe $a \in \mathrm{X}$ tel que la famille $(f_i(a))_{i \in \mathrm{I}}$ est majorée.

Montrer que $f : \begin{cases} \mathrm{X} & \longrightarrow \mathbb{R} \\ x & \longmapsto \sup\limits_{i \in \mathrm{I}} f_i(x) \end{cases}$ est correctement définie et est k−lipschitzienne.

Exercice II.17.

Soit A une partie de X et $f : \mathrm{A} \longrightarrow \mathbb{R}$ une application k−lipschitzienne. Posons pour tout $x \in \mathrm{A}$

$$g(x) = \inf_{y \in \mathrm{A}} \left(f(y) + kd(x, y)\right)$$

Montrer que g est k−lipschitzienne et que $g_{|\mathrm{A}} = f$.

5. Uniforme continuité

Définition II.18.

Soit f une application de X dans Y. On dit que f est uniformément continue lorsque

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall (x, y) \in \mathrm{X}^2,\ d(x, y) < \eta \Longrightarrow \delta(f(x), f(y)) < \varepsilon.$$

Exemples.

(cc) BY-NC-SA

→ Si $f \in \mathcal{C}([a,b],\mathbb{C})$, alors f est uniformément continue d'après le théorème de Heine.

→ Si f est une application de X dans Y $k-$lipschitzienne, alors f est uniformément continue. La démonstration de ce résultat est laissée comme exercice au lecteur.

Attention : Il n'y a pas de réciproque pour le point précédent. Une fonction uniformément continue n'est pas nécessairement lipschitzienne.
En effet, il suffit de considérer la fonction de $[0,1]$ dans $[0,1]$ $f : x \longmapsto \sqrt{x}$. C'est une fonction continue sur un segment, donc elle est uniformément continue d'après le théorème de Heine.
Supposons que f est lipschitzienne. On dispose alors de $k \geq 0$ tel que pour tout $x \in]0,1]$,

$$\left|\frac{f(x)-f(0)}{x-0}\right| \leq k \ i.e. \ \frac{1}{\sqrt{x}} \leq k,$$

ce qui est absurde.

Théorème (Théorème de Heine) II.19.

Une fonction f continue sur un segment $[a,b]$ est uniformément continue.

Démonstration. formulons tout d'abord la négation de l'uniforme continuité

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x, y \in [a,b],\ \ |x-y| < \eta \text{ et } |f(x)-f(y)| \geq \varepsilon.$$

Supposons donc cette proposition vérifiée. On dispose alors de deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ telles que pour tout $n \in \mathbb{N}$, $|x_n - y_n| \leq \frac{1}{n}$ $(*)$ et $|f(x_n) - f(y_n)| \geq \varepsilon$. (x_n) et (y_n) étant à valeurs dans un fermé borné de $\mathbb{R}$, le théorème de Bolzano-Weierstrass nous permet d'affirmer l'existence d'une extractrice φ telle que $(x_{\varphi(n)})$ converge. De même, on dispose d'une extractrice ψ telle que $(y_{\varphi o \psi(n)})$ soit convergente. L'inégalité $(*)$ impose que $(x_{\varphi o \psi(n)})$ et $(y_{\varphi o \psi(n)})$ admettent la même limite qu'on note ℓ.
On a alors

$$\left|f(x_{\varphi o \psi(n)}) - f(y_{\varphi o \psi(n)})\right| \xrightarrow[n\to+\infty]{} |f(\ell)-f(\ell)| = 0.$$

Or pour tout $n \in \mathbb{N}$, $\left|f(x_{\varphi o \psi(n)}) - f(y_{\varphi o \psi(n)})\right| \geq \varepsilon > 0$, ce qui absurde, d'où l'uniforme continuité de f sur $[a,b]$.

Exercice II.20.

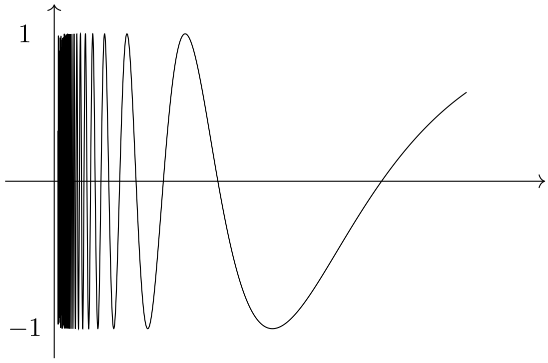
Soit A une partie de X et $f : \mathrm{A} \longrightarrow \mathbb{R}$ une application uniformément continue. Montrer que f admet une limite en tout point de $\overline{\mathrm{A}}$.

Exercice II.21.

Soit E un espace vectoriel normé et $f : \mathrm{B}_f(0,1) \longrightarrow \mathbb{R}$ une application uniformément continue. Montrer que f est bornée.

Attention : une application continue bornée n'est pas nécessairement uniformément continue.
En effet, considérons la fonction de $]0,1]$ dans $[-1,1]$, $f : x \longmapsto \cos(\frac{1}{x})$.

CC BY-NC-SA



Cette fonction est $\mathcal{C}^\infty$ sur $]0,1]$, mais elle n'est pas uniformément continue. Montrons ce résultat.
On va tout d'abord énoncer la négation de l'uniforme continuité.

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists (x,y) \in \mathrm{X}^2,\ d(x,y) < \eta \text{ et } \delta(f(x), f(y)) \geq \varepsilon$$

i.e., en adaptant cette proposition à notre problème,

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists (x,y) \in]0,1]^2,\ \ |x-y| < \eta \text{ et } |f(x) - f(y)| \geq \varepsilon.$$

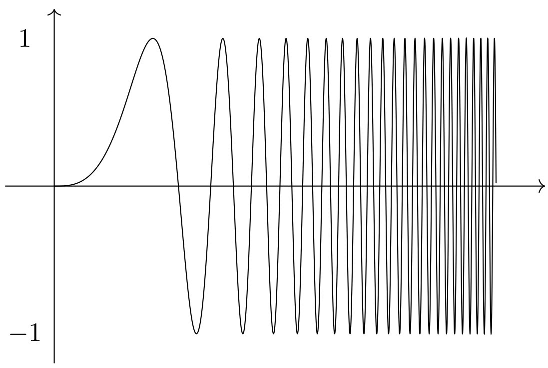
Considérons les suites $(x_n)_{n\in\mathbb{N}^*} = (\frac{1}{2n\pi})_{n\in\mathbb{N}^*}$ et $(y_n)_{n\in\mathbb{N}^*} = (\frac{1}{(2n+1)\pi})_{n\in\mathbb{N}^*}$. On a

$$|x_n - y_n| = \left| \frac{1}{2n\pi} - \frac{1}{(2n+1)\pi} \right| \xrightarrow[n\to+\infty]{} 0,$$

mais pour tout $n \in \mathbb{N}^*$

$$|f(x_n) - f(y_n)| = 1,$$

donc en prenant $\varepsilon = \frac{1}{2}$, on a pour tout $\eta > 0$, on peut prendre n assez grand pour qu'on ait $|x_n - y_n| < \eta$ et $|f(x_n) - f(y_n)| = 1 \geq \varepsilon$ ce qui montre bien la négation de l'uniforme continuité.
L'intuition derrière ce contrexemple est que f oscille très rapidement au voisinage de 0 (voir figure ci-dessus), ce qui fait qu'on ne peut plus contrôler l'écart entre l'image de deux points assez proches.
Il existe une infinité d'autres contrexemples, on peut par exemple prendre la fonction de $\mathbb{R}$ dans $[-1,1]$ $x \longmapsto \sin(x^2)$



on peut prouver de la même manière la non uniforme continuité en prenant les suites $(x_n)_{n\in\mathbb{N}} = \left(\sqrt{2n\pi}\right)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}} = \left(\sqrt{(2n+\frac{1}{2})\pi}\right)_{n\in\mathbb{N}}$. Ici encore, on voit que la fonction oscille très vite au voisinage de $+\infty$ ce qui empêche la fonction d'être uniformément continue.

(cc) BY-NC-SA

6. Extensions du théorème de Heine

Exercice II.22.

Soit $\mathrm{T} \in \mathbb{R}^{+}$ et $\mathrm{E} = \mathcal{C}_{\mathrm{T}}(\mathbb{R}, \mathbb{C})$ l'ensemble des fonctions continues $\mathrm{T}-$périodiques de $\mathbb{R}$ dans $\mathbb{C}$. Montrer que toute fonction de E est uniformément continue.

Exercice II.23.

Soit $\mathrm{E} = \{f \in \mathcal{C}(\mathbb{R}, \mathbb{C}),\ f(x) \xrightarrow[x \to \pm\infty]{} 0\}$. Montrer que toute fonction de E est uniformément continue.

(cc) BY-NC-SA

Correction de l'exercice I.5. :

Soit $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $x \in \mathrm{B}(0, \eta)$

$$\left| \frac{f(x) - f\left(\frac{x}{2}\right)}{x} - \ell \right| < \varepsilon.$$

Or pour tout $k \in \mathbb{N}$, si $x \in \mathrm{B}(0, \eta)$, alors $\frac{x}{2^k} \in \mathrm{B}(0, \eta)$. On a alors pour tout $k \in \mathbb{N}$,

$$\left| \frac{f(\frac{x}{2^k}) - f(\frac{x}{2^{k+1}})}{\frac{x}{2^k}} - \ell \right| < \varepsilon$$

i.e.

$$\left| \frac{f(\frac{x}{2^k}) - f(\frac{x}{2^{k+1}})}{x} - \frac{\ell}{2^k} \right| < \frac{\varepsilon}{2^k}.$$

En sommant toutes ces inégalités pour $k \in [\![0; n-1]\!]$ et utilisant l'inégalité triangulaire, on obtient que pour tout $n \in \mathbb{N}$ et $x \in \mathrm{B}(0, \eta)$,

$$\left| \frac{f(x) - f(\frac{x}{2^n})}{x} - \ell \sum_{k=0}^{n-1} \frac{1}{2^k} \right| \leq \sum_{k=0}^{n-1} \left| \frac{f(\frac{x}{2^{k+1}}) - f(\frac{x}{2^k})}{x} - \frac{\ell}{2^k} \right| < \varepsilon \sum_{k=0}^{n-1} \frac{1}{2^k}.$$

En faisant tendre n vers l'infini, par continuité de f, l'inégalité ci-dessus devient

$$\left| \frac{f(x) - f(0)}{x} - 2\ell \right| \leq 2\varepsilon,$$

ce qui montre que f est dérivable en 0 et que $f'(0) = 2\ell$.

Correction de l'exercice I.6. :

Soit $\varepsilon > 0$, $\mathrm{A} > 1$ tel que pour tout $x \geq \mathrm{A} - 1$,

$$\mathrm{L} - \varepsilon \leq f(x+1) - f(x) \leq \mathrm{L} + \varepsilon.$$

Posons pour tout x, $n(x)$ l'unique entier vérifiant $\mathrm{A} \leq x - n(x) < \mathrm{A} + 1$. On a alors

$$\mathrm{L} - \varepsilon \leq f(x - (n(x) - 1) + 1) - f(x - (n(x) - 1)) \leq \mathrm{L} + \varepsilon$$

$$\mathrm{L} - \varepsilon \leq f(x - (n(x) - 2) + 1) - f(x - (n(x) - 2)) \leq \mathrm{L} + \varepsilon$$

$$\vdots$$

$$\mathrm{L} - \varepsilon \leq f((x - 1 + 1) - f(x - 1) \leq \mathrm{L} + \varepsilon.$$

En sommant toutes les inégalités ci-dessus, on obtient

$$n(x)(\mathrm{L} - \varepsilon) \leq f(x) - f(x - n(x)) \leq n(x)(\mathrm{L} + \varepsilon)$$

et on a $x - n(x)$ est toujours dans $[\mathrm{A}, \mathrm{A} + 1]$ où f est bornée car continue, donc

$$\underbrace{\frac{n(x)}{x}}_{\xrightarrow[x \to +\infty]{} 1} (\mathrm{L} - \varepsilon) + \underbrace{\frac{f(x - n(x))}{x}}_{\xrightarrow[x \to +\infty]{} 0} \leq \frac{f(x)}{x} \leq \underbrace{\frac{n(x)}{x}}_{\xrightarrow[x \to +\infty]{} 1} (\mathrm{L} + \varepsilon) + \underbrace{\frac{f(x - n(x))}{x}}_{\xrightarrow[x \to +\infty]{} 0}.$$

(cc) BY-NC-SA

On peut donc choisir B > A de sorte que pour tout $x > \mathrm{B}$,

$$\left|\frac{f(x-n(x))}{x}\right| \leq \varepsilon \text{ et } \mathrm{L}-2\varepsilon \leq \frac{n(x)}{x}(\mathrm{L}+\varepsilon) \leq \mathrm{L}+2\varepsilon,$$

et alors pour tout $x > \mathrm{B}$

$$\mathrm{L}-3\varepsilon \leq \frac{f(x)}{x} \leq \mathrm{L}+3\varepsilon.$$

On a donc bien montré que $x \longmapsto \frac{f(x)}{x}$ admet une limite en $+\infty$ et que

$$\frac{f(x)}{x} \xrightarrow[x\to+\infty]{} \mathrm{L}.$$

Correction de l'exercice II.3. :

Posons pour tout $x \in [a,b]$, $f(x^+)$ et $f(x^-)$ respectivement la limite à droite et à gauche en x. Définissons les ensembles suivants.

- $\rightarrow$ C est l'ensemble des points de discontinuité de f.
- $\rightarrow$ $\mathrm{A} = \{x \in [a,b],\ \ |f(x^+) - f(x^-)| > 0\}$.
- $\rightarrow$ Pour tout $p \in \mathbb{N}^*$, $\mathrm{A}_p = \left\{x \in [a,b],\ \ |f(x^+) - f(x^-)| \geq \frac{1}{p}\right\}$.
- $\rightarrow$ $\mathrm{B} = \{x \in [a,b],\ f(x^-) = f(x^+) \text{ et } f(x^+) \neq f(x)\}$.
- $\rightarrow$ Pour tout $p \in \mathbb{N}^*$ $\mathrm{B}_p = \left\{x \in \mathrm{B},\ \ |f(x^+) - f(x)| \geq \frac{1}{p}\right\}$.

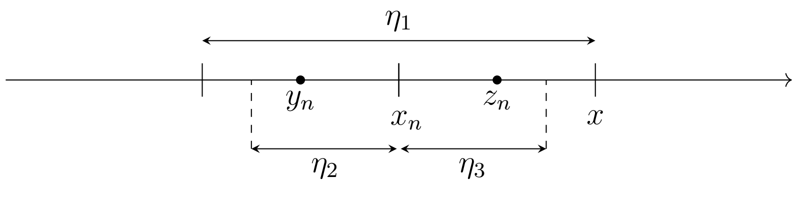
On a alors

$$\mathrm{C} = \mathrm{B} \cup \mathrm{A} = \bigcup_{p\in\mathbb{N}^*} \mathrm{B}_p \cup \bigcup_{p\in\mathbb{N}^*} \mathrm{A}_p.$$

On va montrer que pour tout $p \in \mathbb{N}^*$, A_p et B_p sont finis. Une fois cela fait, on peut facilement conclure car on aura montré que A et B sont union dénombrable d'ensembles finis et donc qu'ils sont dénombrables. Supposons par l'absurde qu'il existe $p \in \mathbb{N}^*$ tel que A_p est infini. A_p est infini borné en dimension finie, il possède donc un point d'accumulation qu'on note x. On dispose alors d'une suite $(x_n)_{n\in\mathbb{N}}$ injective à valeurs dans A qui converge vers x. Quitte à extraire de la suite (x_n) on suppose sans perte de généralité qu'elle est croissante.

- $\rightarrow$ Soit $\eta_1 > 0$ tel que pour tout $\alpha \in]x-\eta_1, x[$, $|f(\alpha) - f(x^-)| < \frac{1}{4p}$.
- $\rightarrow$ Soit $n \in \mathbb{N}$ tel que $x_n \in]x-\eta_1, x[$. On suppose aussi sans perte de généralité que $x \neq a$.
- $\rightarrow$ Soit $\eta_2 > 0$ tel que pour tout $\alpha \in]x_n-\eta_2, x_n[$, $|f(\alpha) - f(x_n^-)| < \frac{1}{4p}$.
- $\rightarrow$ Soit $\eta_3 > 0$ tel que pour tout $\alpha \in]x_n, x_n+\eta_3[$, $|f(x_n^+) - f(\alpha)| < \frac{1}{4p}$.
  Quitte à remplacer η_2 par $\min(|x_n - x|, \eta_2)$ et η_3 par $\min(|x_n - x + \eta_1|, \eta_3)$, on suppose que $\eta_2 \leq |x_n - x + \eta_1|$ et que $\eta_3 \leq |x_n - x|$ pour qu'on ait $]x_n-\eta_2, x_n[\subset]x-\eta_1, x[$ et $]x_n, x_n+\eta_3[\subset]x-\eta_1, x[$.
- $\rightarrow$ Soit $y_n, z_n \in [a,b]$ tels que $y_n \in]x_n-\eta_2, x_n[$ et $y_n \in]x_n, x_n+\eta_3[$.

Pour que tout cela soit plus clair, observons tous les paramètres qu'on vient de définir sur le dessin suivant :



On a

$$\begin{aligned}|f(y_n) - f(z_n)| &\geq \left|f(x_n^-) - f(x_n^+)\right| - \left|f(x_n^+) - f(z_n)\right| - \left|f(y_n) - f(x_n^-)\right| \\ &> \frac{1}{p} - \frac{1}{4p} - \frac{1}{4p} = \frac{1}{2p}\end{aligned}$$

(cc) BY-NC-SA

et

$$|f(y_n)-f(z_n)| \leq \left|f(y_n)-f(x^-)\right| + \left|f(z_n)-f(x^-)\right| < \frac{1}{4p}+\frac{1}{4p}=\frac{1}{2p},$$

ce qui est absurde.
On va procéder de la même manière pour les ensembles B_p. Supposons par l'absurde qu'il existe $p \in \mathbb{N}^*$ tel que B_p est infini. B_p est infini borné en dimension finie, il admet donc un point d'accumulation qu'on note x. On dispose alors d'une suite $(x_n)_{n\in\mathbb{N}}$ injective à valeurs dans B_p qu'on suppose sans perte de généralité, quitte à extraire, croissante.

→ Soit $\eta_1 > 0$ tel que pour tout $\alpha \in]x-\eta_1, x[$, $|f(\alpha)-f(x^-)| < \frac{1}{4p}$

→ Soit $n \in \mathbb{N}$ tel que $x_n \in]x-\eta_1, x[$.

→ Soit $\eta_2 > 0$ tel que pour tout $\alpha \in]x_n-\eta_2, x_n+\eta_2[$, $|f(\alpha)-f(x_n^-)| < \frac{1}{2p}$
Quitte à remplacer η_2 par $\min(|x_n-x|, |x_n-x+\eta|, \eta_2)$, on suppose que $]x_n-\eta_2, x_n+\eta_2[\subset]x-\eta_1, x[$.

→ Soit $y_n \in]x_n-\eta_2, x_n+\eta_2[$.

On a alors

$$\left|f(x_n)-f(x^-)\right| < \frac{1}{4p}$$

et

$$\begin{aligned}\left|f(x_n)-f(x^-)\right| &\geq |f(x_n)-f(y_n)| - \left|f(y_n)-f(x^-)\right| \\ &\geq \left|f(x_n)-f(x_n^-)\right| - \left|f(x_n^-)-f(y_n)\right| - \left|f(y_n)-f(x^-)\right| \\ &> \frac{1}{p}-\frac{1}{2p}-\frac{1}{4p}=\frac{1}{4p},\end{aligned}$$

ce qui est absurde. On a donc bien obtenu le résultat voulu.
L'intuition derrière le fait que A_p et B_p soient finis est que lorsqu'on se rapproche par exemple par la gauche d'un point x où f admet une limite à gauche, on ne peut pas rencontrer un nombre infini de points de discontinuité de taille $\frac{1}{p}$ car sinon la fonction fera des pas trop grands pour tendre vers $f(x^-)$ à gauche de x.
Remarque. ici, x n'est pas nécessairement dans A. On a juste eu besoin du fait que f admet une limite à gauche de x.

Correction de l'exercice II.6. :

(1) ⇒ (2) Supposons que f est continue. Soit $x \in \overline{\mathrm{A}}$. On dispose d'une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans A qui converge vers x. On a alors par continuité de f, la suite $(f(x_n))_{n\in\mathbb{N}}$ est à valeurs dans $f(\mathrm{A})$ converge vers $f(x)$, *i.e.* $f(x) \in \overline{f(\mathrm{A})}$ et on a finalement $f(\overline{\mathrm{A}}) \subset \overline{f(\mathrm{A})}$

(2) ⇒ (1) Supposons que pour tout $\mathrm{A} \subset \mathrm{X}$, $f(\overline{\mathrm{A}}) \subset \overline{f(\mathrm{A})}$. Soit F un fermé de Y. Posons $\mathrm{A} = f^{-1}(\mathrm{F})$ et montrons que A est fermé. Une fois cela fait, on peut aisément conclure que f est continue d'après une propriété du cours.
On a $f(\overline{\mathrm{A}}) \subset \overline{f(\mathrm{A})} = \overline{f(f^{-1}(\mathrm{F}))} \subset \overline{\mathrm{F}} = \mathrm{F}$, et donc $\overline{\mathrm{A}} \subset f^{-1}(\mathrm{F}) = \mathrm{A}$ *i.e.* A est fermé.

Correction de l'exercice II.7. :

(1) ⇒ (2) On a pour tout $n \in \mathbb{Z}$ et $x \in \mathbb{R}$, $f(nx) = nf(x)$. Lorsque $n \neq 0$, en substituant x par $\frac{x}{n}$, $f(x) = nf(\frac{x}{n})$, *i.e.* $f(\frac{x}{n}) = \frac{1}{n}f(x)$.
On a alors pour tout $(p,q) \in \mathbb{Z}\times\mathbb{N}^*$, $f(\frac{p}{q}x) = \frac{p}{q}f(x)$, et alors pour tout $\alpha \in \mathbb{Q}$, $f(\alpha x) = \alpha f(x)$.
Par densité de $\mathbb{Q}$ dans $\mathbb{R}$, on dispose d'une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{Q}$ qui converge vers x. On utilise alors cette suite pour avoir le résultat suivant : $f(x_n) = x_n f(1) \xrightarrow[n\to+\infty]{} xf(1)$ et f continue donc $f(x_n) \xrightarrow[n\to+\infty]{} f(x)$. En posant $\lambda = f(1)$, on obtient que pour tout $x \in \mathbb{R}$, $f(x) = \lambda x$.

(cc) BY-NC-SA

(2) $\Rightarrow$ (3) Le résultat est évident car toutes les homothéties de $\mathbb{R}$ sont continues.

(3) $\Rightarrow$ (4) Idem.

(4) $\Rightarrow$ (5) Il suffit d'appliquer la définition de la continuité en 0 avec $\varepsilon = 1$.

(5) $\Rightarrow$ (1) Soit $\alpha > 0$ tel que f est bornée sur $[-\alpha, \alpha]$ et posons $\mathrm{M} = \sup_{x\in[-\alpha,\alpha]} |f(x)|$. Soit $\varepsilon > 0$. On considère $n \in \mathbb{N}^*$ tel que $\frac{\mathrm{M}}{n} < \varepsilon$. On a alors

$$f\left(\left[-\frac{\alpha}{n}, \frac{\alpha}{n}\right]\right) = \frac{1}{n} f\left([-\alpha, \alpha]\right) \subset \left[-\frac{\mathrm{M}}{n}, \frac{\mathrm{M}}{n}\right] \subset [-\varepsilon, \varepsilon],$$

ce qui nous donne la continuité de f en 0. La continuité de f en tout point $x \in \mathbb{R}$ vient du fait que

$$f(y) = \underbrace{f(y-x)}_{\xrightarrow[y\to x]{}0} + f(x) \xrightarrow[y\to x]{} f(x)$$

par continuité de f en 0.

Correction de l'exercice II.8. :

D'après l'exercice précédent, on dispose de $\lambda \in \mathbb{R}$, tel que $f = x \longmapsto \lambda x$. Or, f étant un morphisme de corps, on a $f(1) = 1$, *i.e.* $\lambda = 1$ et alors $f = \mathrm{Id}$.

Correction de l'exercice II.9. :

En utilisant une méthode analogue à celle utilisée pour les deux exercices précédents, on peut montrer que pour tout nombre complexe $a+ib$, $f(a+ib) = a + f(i)b$. Il reste donc à déterminer la valeur de $f(i)$ pour déterminer f. On a $f(i)^2 = f(i^2) = f(-1) = -1$ et alors $f(i) \in \{i, -i\}$. On peut donc immédiatement en déduire que $f = \mathrm{Id}$ ou $f = (z \longmapsto \overline{z})$.

Correction de l'exercice II.14. :

(1) $\Rightarrow$ (2) Posons $k \geq 0$ le rapport de lipschitziannité de f. On a alors pour tout $a \in \mathrm{I}$,

$$k \geq \left|\frac{f(x) - f(a)}{x - a}\right| \xrightarrow[x\to a]{} |f'(a)|$$

i.e. f' est bornée par k.

(2) $\Rightarrow$ (1) Supposons que f' est bornée par k. Utilisons le théorème des accroissements finis. On a pour tout $x, y \in \mathrm{I}$ différents, il existe $c \in \mathrm{I}$ tel que

$$|f(x) - f(y)| = |f'(c)|\,|x-y| \leq k\,|x-y|\,.$$

Correction de l'exercice II.15. :

1. On a

$$d(x, \mathrm{A}) = 0 \iff \exists (a_n) \in \mathrm{A}^{\mathbb{N}},\ d(x, a_n) \xrightarrow[n\to+\infty]{} 0 \iff x \in \overline{\mathrm{A}}.$$

2. Il s'agit de l'inégalité triangulaire. Soit $x, y \in \mathrm{X}$.
   On a pour tout $a \in \mathrm{A}$, $d(y, \mathrm{A}) \leq d(a, y) \leq d(x, a) + d(x, y)$ et alors en passant à la borne inférieure

BY-NC-SA

sur a des deux côtés, on obtient

$$d(y,\mathrm{A}) \leq d(x,\mathrm{A}) + d(x,y). \tag{8}$$

De même, on a pour tout $a \in \mathrm{A}$, $d(x,\mathrm{A}) \leq d(x,a) \leq d(y,a) + d(x,y)$ et alors en passant à la borne inférieure des deux côtés, on obtient

$$d(x,\mathrm{A}) \leq d(y,\mathrm{A}) + d(x,y). \tag{9}$$

En combinant les inégalités 8 et 9, on obtient

$$-d(x,y) \leq d(x,\mathrm{A}) - d(y,\mathrm{A}) \leq d(x,y)$$

i.e.

$$|d(x,\mathrm{A}) - d(y,\mathrm{A})| \leq d(x,y).$$

3. Considèrons l'application $f : \begin{cases} \mathrm{X} & \longrightarrow \mathbb{R} \\ x & \longmapsto d(x,\mathrm{B}) - d(x,\mathrm{A}). \end{cases}$

   → Si $x \in \mathrm{A}$, on a $x \notin \mathrm{B} = \overline{\mathrm{B}}$, et alors $f(x) = d(x,\mathrm{B}) > 0$. On a donc $\mathrm{A} \subset \{x \in \mathrm{X},\ f(x) > 0\} = f^{-1}(\mathbb{R}_+^*) = \mathrm{U}$.

   → Si $x \in \mathrm{B}$, on a $x \notin \mathrm{A} = \overline{\mathrm{A}}$, et alors $f(x) = -d(x,\mathrm{A}) < 0$. On a donc $\mathrm{B} \subset \{x \in \mathrm{X},\ f(x) < 0\} = f^{-1}(\mathbb{R}_-^*) = \mathrm{V}$.

   U et V sont les images réciproques des ouverts $\mathbb{R}_+^*$ et $\mathbb{R}_-^*$ par une fonction continue (car lipschitzienne) et donc sont ouverts. De plus on a $\mathrm{U} \cap \mathrm{V} = \varnothing$, on a alors bien obtenu le résultat voulu.

Correction de l'exercice II.16. :

Montrons tout d'abord que f est bien définie, *i.e.* pour tout $i \in \mathrm{I}$ et $x \in \mathbb{R}$, la famille $(f_i(x))_{i\in\mathrm{I}}$ est majorée. Soit M un majorant de $(f_i(a))_{i\in\mathrm{I}}$, on a pour tout $x \in \mathbb{R}$ et $i \in \mathrm{I}$,

$$f_i(x) - f_i(a) \leq k\,|x-a|$$

et alors

$$f_i(x) \leq k\,|x-a| + \mathrm{M},$$

ce qui montre bien que f est définie.
Montrons ensuite que f est $k-$lipschitzienne. On a pour $i \in \mathrm{I}$ et $x, y \in \mathbb{R}$,

$$|f_i(x) - f_i(y)| \leq k\,|x-y|$$

i.e.

$$f_i(y) - k\,|x-y| \leq f_i(x) \leq f_i(y) + k\,|x-y|\,.$$

On peut donc passer à la borne supérieure en i de chaque côté de l'inégalité pour obtenir que

$$f(y) - k\,|x-y| \leq f(x) \leq f(y) + k\,|x-y|$$

i.e.

$$|f(x) - f(y)| \leq k\,|x-y|\,.$$

Correction de l'exercice II.17. :

→ Montrons que g est bien définie et est $k-$lipschitzienne.
Posons pour tout $y \in \mathrm{X}$, $\varphi_y : x \longmapsto f(y) + kd(x,y)$. Toutes les fonctions $-\varphi_y$ sont $k-$lipschitziennes

CC BY-NC-SA

et par lipschitzianité de $-\varphi_y$ pour tout $y \in \mathrm{X}$,

$$-\varphi_y(0) = -(f(y) + kd(0, y)) \leq -f(0).$$

La famille $(\varphi_y(0))_{y\in\mathrm{X}}$ est donc majorée, et alors d'après l'exercice précédent, $h = \sup_{y\in\mathrm{X}}(-\varphi_y)$ est bien définie et est $k-$lipschitzienne et donc de même pour $g = -h$.

→ Montrons que $g_{|\mathrm{A}} = f$.
On a pour tout $x, y \in \mathrm{A}$, $f(y) + kd(x, y) \geq f(x)$ (car f est $k-$lipschitzienne). Ce minorant est atteint pour $y = x$, donc $\inf\limits_{y\in\mathrm{X}} f(y) + kd(x, y) = f(x)$.

Correction de l'exercice II.20. :

Soit $a \in \overline{\mathrm{A}}$ et $(x_n)_{n\in\mathbb{N}} \in \mathrm{A}^{\mathbb{N}}$ telle que $x_n \xrightarrow[n\to+\infty]{} a$. On va montrer que la suite $(f(x_n))_{,\in\mathbb{N}}$ est de Cauchy.
Soit $\varepsilon > 0$. Par uniforme continuité de f, on peut considérer $\eta > 0$ tel que pour tout $x, y \in \mathrm{A}$,

$$d(x, y) < \eta \Longrightarrow |f(x) - f(y)| < \varepsilon.$$

$(x_n)_{n\in\mathbb{N}}$ étant convergente et donc de Cauchy, on peut considérer $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n, m \geq \mathrm{N}$, $d(x_n, x_m) < \eta$.
On a alors pour tout $n, m > \mathrm{N}$, $|f(x_n) - f(x_m)| < \varepsilon$. $(f(x_n))_{n\in\mathbb{N}}$ est donc de Cauchy, *i.e.* convergente. f admet donc bien une limite en tout point de $\overline{\mathrm{A}}$.

Correction de l'exercice II.21. :

Supposons sans perte de généralité que $f(0) = 0$. Par uniforme continuité de f, on peut considérer $\eta > 0$ tel que pour tout $x, y \in \mathrm{B}_f(0, 1)$, $\|x - y\| < \eta \Longrightarrow |f(x) - f(y)| < 1$.
Soit $p \in \mathbb{N}^*$ tel que $\frac{1}{p} \leq \eta$. On a pour tout $x \in \mathrm{B}_f(0, 1)$,

$$|f(x)| = |f(x) - f(0)| \leq \sum_{k=0}^{p-1} \left| f\left((k + 1)\frac{x}{p}\right) - f\left(k\frac{x}{p}\right) \right|.$$

Or, on a pour tout $k \in [\![0; p - 1]\!]$, $\left\|(k + 1)\frac{x}{p} - k\frac{x}{p}\right\| = \frac{\|x\|}{p} < \frac{1}{p} < \eta$ et alors

$$|f(x)| \leq \sum_{k=0}^{p-1} \left| f\left((k + 1)\frac{x}{p}\right) - f\left(k\frac{x}{p}\right) \right| \leq p$$

donc f est bornée sur $\mathrm{B}_f(0, 1)$.

Correction de l'exercice II.22. :

f est une fonction continue sur le segment $[-\mathrm{T}, 2\mathrm{T}]$ et y est donc uniformément continue. Montrons maintenant que f est uniformément continue sur $\mathbb{R}$.
Soit $\varepsilon > 0$. Soit $\eta > 0$ tel que pour tout $x, y \in [-\mathrm{T}, 2\mathrm{T}]$,

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

Soit $x, y \in \mathbb{R}$ tel que $|x - y| < \eta$. Quitte à remplacer η par $\min(\eta, \mathrm{T})$, on suppose que $\eta \leq \mathrm{T}$. On dispose alors de $n \in \mathbb{Z}$ tel que $x - n\mathrm{T} \in [-\mathrm{T}, 2\mathrm{T}]$ et $y - n\mathrm{T} \in [-\mathrm{T}, 2\mathrm{T}]$. On a alors $|(x - n\mathrm{T}) - (y - n\mathrm{T})| < \eta$ et donc

$$|f(x) - f(y)| = |f(x - n\mathrm{T}) - f(y - n\mathrm{T})| < \varepsilon,$$

d'où l'uniforme continuité de f.
Remarque. on a choisi de regarder l'uniforme continuité de f dans l'intervalle $[-\mathrm{T}, 2\mathrm{T}]$ et non pas $[0, \mathrm{T}]$

(cc) BY-NC-SA

pour éviter le cas pathologique où x se retrouve sur les bords de l'intervalle $[0, \mathrm{T}]$ et y en dehors de cet intervalle, ce qui nous empêcherait de conclure car on ne pourra pas utiliser l'uniforme continuité de f sur $[0, \mathrm{T}]$.

Correction de l'exercice II.23. :

Soit $\varepsilon > 0$ et $\mathrm{A} > 0$ tel que pour tout $x \in \mathbb{R}$, $|x| \geq \mathrm{A} \Longrightarrow |f(x)| < \varepsilon$.
f est uniformément continue sur $[-\mathrm{A}, \mathrm{A}]$ car continue, donc il existe $\eta > 0$ tel que $\eta < \mathrm{A}$ et pour tout $x, y \in [-\mathrm{A}, \mathrm{A}]$, $|x - y| < \eta \Longrightarrow |f(x) - f(y)| < \varepsilon$.
Soit $x, y \in \mathbb{R}$, supposons que $|x - y| < \eta$, trois cas de présentent :

- $\rightarrow$ $x, y \in [-\mathrm{A}, \mathrm{A}]$, alors on a directement que $|f(x) - f(y)| < \varepsilon$.
- $\rightarrow$ $x, y \in]\mathrm{A}, +\infty[$ ou $x, y \in]-\infty, \mathrm{A}[$, alors on a $|f(x) - f(y)| \leq |f(x)| + |f(y)| < 2\varepsilon$.
- $\rightarrow$ $x \in [-\mathrm{A}, \mathrm{A}]$ et $y \notin [-\mathrm{A}, \mathrm{A}]$. On suppose sans perte de généralité que $x < \mathrm{A} < y$. On a alors $|f(x) - f(y)| \leq |f(x) - f(\mathrm{A})| + |f(y) - f(\mathrm{A})| \leq 3\varepsilon$.

Ceci nous donne l'uniforme continuité de f sur $\mathbb{R}$.

(cc) BY-NC-SA

Chapitre 11.4

Comparaison de normes et espaces produits

Dans tout ce chapitre, $\mathbb{K}$ est un corps égal à $\mathbb{R}$ ou $\mathbb{C}$, E est un $\mathbb{K}-$espace vectoriel et n est un entier naturel strictement positif.

I Comparaison de normes

Définition I.1.

Soient N_1 et N_2 deux normes sur E. On dit que N_1 est plus fine que N_2 (ou domine N_2) s'il existe $\mathrm{C} > 0$ tel que pour tout $x \in \mathrm{E}$, $\mathrm{N}_2(x) \leq \mathrm{C} \cdot \mathrm{N}_1(x)$.

Proposition I.2.

Si N_1 est plus fine que N_2, alors

1. Pour tout $(x_n) \in \mathrm{E}^{\mathbb{N}}$, $x_n \xrightarrow[n\to+\infty]{} \ell$ pour $\mathrm{N}_1 \Longrightarrow x_n \xrightarrow[n\to+\infty]{} \ell$ pour N_2.
2. Si F est une partie fermée de E pour N_2, alors elle l'est pour N_1.
3. Si O est une partie ouverte de E pour N_2, alors elle l'est pour N_1.

Démonstration.
Supposons que N_1 est plus fine que N_2. soit $\mathrm{C} > 0$ tel que pour tout $x \in \mathrm{E}$, $\mathrm{N}_2(x) \leq \mathrm{C} \cdot \mathrm{N}_1(x)$.

1. Soit (x_n) une suite à valeurs dans E qui converge vers ℓ pour N_1. On a alors
$$\mathrm{N}_2(x_n - \ell) \leq \mathrm{C} \cdot \mathrm{N}_1(x_n - \ell) \xrightarrow[n\to+\infty]{} 0,$$
d'où le résultat voulu.
2. Soit F une partie fermée de E pour N_2. Soit (x_n) une suite à valeurs dans F convergente pour N_1 vers $x \in \mathrm{E}$. (x_n) converge alors pour N_2 vers x. F étant fermé pour N_2, on a $x \in \mathrm{F}$ et alors F est fermé pour N_1.
3. Si O est ouvert pour N_2, alors $\mathrm{E} \setminus \mathrm{O}$ est fermé pour N_1 donc pour N_2 et finalement O est ouvert pour N_1.

Définition I.3.

On dit que deux normes N_1 et N_2 sont équivalentes lorsque N_1 est plus fine que N_2 et N_2 est plus fine que N_1. Dans ce cas, on note $\mathrm{N}_1 \sim \mathrm{N}_2$.

(cc) BY-NC-SA

Proposition I.4.

Soit N_1 et N_2 deux normes équivalentes.

1. Les suites convergentes pour N_1 sont exactement les suites convergentes pour N_2.
2. Les ouverts (resp. fermés) pour N_1 sont exactement les ouverts (resp. fermés) pour N_2.
3. Les fonctions continues pour N_1 sont exactement les fonctions continues pour N_2.
4. Les application lipschitziennes pour N_1 sont exactement les applications lipschitziennes pour N_2, mais le rapport de lipschitzianité ne reste plus le même quand on passe d'une norme à l'autre.

Exemple. Dans le cas où $E = \mathbb{K}^n$, les normes $\|.\|_\infty$, $\|.\|_1$ et $\|.\|_2$ sont équivalentes. En effet, on a pour tout $x \in E$,

→ $\frac{1}{n}\|x\|_1 \le \|x\|_\infty \le \|x\|_1$, donc $\|.\|_1 \sim \|.\|_\infty$.

→ $\|x\|_2 \le \|x\|_1 \le \sqrt{n}\,\|x\|_2$, donc $\|.\|_2 \sim \|.\|_1$.

→ Par transitivité, ces trois normes sont deux à deux équivalentes.

Proposition I.5.

Lorsque E est de dimension finie, alors toutes les normes sur E sont équivalentes.

Démonstration. la démonstration de ce résultat sera faite dans une partie ultérieure de ce cours, car nous n'avons pas encore exposé les outils nécessaires pour le montrer.

Exemple 1. Dans le cas où $E = \mathcal{C}([a,b],\mathbb{R})$ avec $a < b$, $\|.\|_2$ est plus fine que $\|.\|_1$ et $\|.\|_\infty$ est plus fine que $\|.\|_2$, mais parmi ces trois normes, il n'y en a aucune qui est équivalente à une autre.

→ On a pour tout $f \in \mathcal{C}([a,b],\mathbb{R})$, $\|f\|_1 \le \sqrt{b-a}\,\|f\|_2 \le (b-a)\,\|f\|_\infty$.

→ Les inégalités dans le sens contraire sont toutes fausses en général. En effet, si on considère la suite de fonctions $(f_n)_{n\in\mathbb{N}} = (x \longmapsto x^n)_{n\in\mathbb{N}}$ de $[0,1]$ dans $\mathbb{R}$, alors on a $f_n \xrightarrow[n\to+\infty]{} 0$ pour $\|.\|_1$ et $\|.\|_2$, mais $\|f_n\|_\infty = 1$ pour tout n, donc (f_n) ne tend pas vers 0 pour $\|.\|_\infty$.

Exemple 2. On se place dans le cas où $E = \mathcal{C}^1([0,1],\mathbb{R})$ et on considère la norme $\|.\|$ telle que pour tout $f \in E$, $\|f\| = |f(0)| + \displaystyle\int_0^1 |f'(x)|\,dx$.

→ $\|.\|$ est plus fine que $\|.\|_\infty$.
En effet, on a pour tout $f \in E$ et $x \in [a,b]$,

$$|f(x)| = \left|f(0) + \int_0^x f'(t)dt\right| \le |f(0)| + \int_0^x |f'(t)|\,dt \le |f(0)| + \int_0^1 |f'(t)|\,dt = \|f\|,$$

ce qui entraine $\|f\|_\infty \le \|f\|$.

→ $\|.\|_\infty$ n'est pas plus fine que $\|.\|$.
En effet, si on considère la suite de fonctions $(f_n)_{n\in\mathbb{N}} = (x \longmapsto \frac{1}{n}\sin(n\pi x))$, on trouve

- $\|f_n\|_\infty = \frac{1}{n} \xrightarrow[n\to+\infty]{} 0$
- Pour tout n, $\|f_n\| = \displaystyle\int_0^1 |\sin(n\pi t)|\,dt \ge \int_0^1 \sin(n\pi t)^2 dt \ge \frac{1}{2}$ et donc $\|f_n\|$ ne tend pas vers 0.

BY-NC-SA

Exercice I.6.

On reprend les notations de l'exemple précédent. Soit $f \in \mathrm{E}$. Montrer que

$$\int_0^1 f(t)\,|\sin(nt)|\,dt \xrightarrow[n\to+\infty]{} \frac{2}{\pi}\int_0^1 f(t)dt.$$

Exercice I.7.

On pose $\mathrm{E} = \mathbb{R}_n[\mathrm{X}]$. Comparer les normes suivantes sur E, telles que pour tout $\mathrm{P} = \sum_{k=0}^{n} a_k \mathrm{X}^k \in \mathrm{E}$,

→ $\|\mathrm{P}\|_0 = \max(|a_0|, \dots, |a_n|)$,

→ $\|\mathrm{P}\|_1 = \sum_{k=0}^{n} |a_k|$,

→ $\|\mathrm{P}\|_\infty = \sup_{x\in[0,1]} |\mathrm{P}(x)|$.

II Espaces produits

Soit $p \in \mathbb{N}^*$. On considère p espaces métriques $\mathrm{E}_1, \dots, \mathrm{E}_p$ munis respectivement des distances $d_1, \dots, d_p$ et on pose $\mathrm{E} = \mathrm{E}_1 \times \cdots \times \mathrm{E}_p$. On définit pour tout $i \in [\![1;p]\!]$ la $i-$ème projection canonique

$$\pi_i : \begin{cases} \mathrm{E}_1 \times \cdots \times \mathrm{E}_p & \longrightarrow \mathrm{E}_i \\ (x_1, \dots, x_p) & \longmapsto x_i. \end{cases}$$

Remarque préliminaire. Pour tout $i \in [\![1;p]\!]$, si A_i est une partie de E_i, alors

$$\pi_i^{-1}(\mathrm{A}_i) = \mathrm{E}_1 \times \cdots \times \mathrm{E}_{i-1} \times \mathrm{A}_i \times \mathrm{E}_{i+1} \times \cdots \times \mathrm{E}_p.$$

1. Métrique produit

On munit E de la distance d définie par

$$\forall (x_1, \dots, x_p), (y_1, \dots, y_p) \in \mathrm{E}, \quad \max_{i\in[\![1;p]\!]} d_i(x_i, y_i).$$

Remarque. Pour tout $x = (x_1, \dots, x_p) \in \mathrm{E}$ et $\varepsilon > 0$

$$\mathrm{B}_d(x,\varepsilon) = \{(y_1,\dots,y_p) \in \mathrm{E},\ \forall i \in [\![1;p]\!],\ d_i(x_i,y_i) < \varepsilon\} = \mathrm{B}_{d_1}(x_1,\varepsilon) \times \cdots \times \mathrm{B}_{d_p}(x_p,\varepsilon).$$

Vocabulaire. On appelle pavé ouvert tout produit de la forme $\omega_1 \times \cdots \times \omega_p$ tel que pour tout $i \in [\![1;p]\!]$, ω_i un ouvert de E_i.

Proposition II.1.

1. Tout pavé ouvert est ouvert dans E.
2. Tout ouvert de E est réunion de pavés ouverts.

(cc) BY-NC-SA

Démonstration.

1. Soit $\Omega = \omega_1 \times \cdots \times \omega_p$ un pavé ouvert. soit $x = (x_1, \ldots, x_p) \in \Omega$. Pour tout $i \in [\![1;p]\!]$, il existe $\varepsilon_i > 0$ tel que $\mathrm{B}_{d_i}(x_i, \varepsilon_i) \subset \omega_i$. On pose alors $\varepsilon = \min_{i \in [\![1;p]\!]} \varepsilon_i > 0$.
   Le choix de ε nous permet d'affirmer que $\mathrm{B}(x, \varepsilon) \subset \mathrm{B}_{d_1}(x_1, \varepsilon_1) \times \cdots \times \mathrm{B}_{d_p}(x_p, \varepsilon_p) \subset \Omega$ donc Ω est ouvert.
2. Soit Ω un ouvert de E. Pour tout $a \in \mathrm{O}$, il existe $\varepsilon_a > 0$ tel que $\mathrm{B}(a, \varepsilon_a) \subset \Omega$.
   On peut alors écrire $\Omega = \bigcup_{a \in \Omega} \mathrm{B}(a, \varepsilon_a)$. Une boule ouverte de E est par définition un pavé ouvert, donc Ω est bien une union de pavés ouverts.

Proposition II.2.

Soit $\mathrm{F}_1, \ldots, \mathrm{F}_p$ des parties de $\mathrm{E}_1, \ldots, \mathrm{E}_p$. Si pour tout $i \in [\![1;p]\!]$, F_i est fermé dans E_i, alors $\mathrm{F} = \mathrm{F}_1 \times \cdots \times \mathrm{F}_p$ est fermé dans E.

Démonstration. On commence par remarquer que $\mathrm{F} = \mathrm{F}_1 \times \cdots \times \mathrm{F}_p = \bigcap_{i=1}^{p} \mathrm{G}_i$ avec pour tout $i \in [\![1;p]\!]$,

$$\mathrm{G}_i = \mathrm{E}_1 \times \cdots \times \mathrm{E}_{i-1} \times \mathrm{F}_i \times \mathrm{E}_{i+1} \times \cdots \times \mathrm{E}_p.$$

Or on a pour tout $i \in [\![1;p]\!]$,

$$\mathrm{E} \setminus \mathrm{G}_i = \mathrm{E}_1 \times \cdots \times \mathrm{E}_{i-1} \times (\mathrm{E}_i \setminus \mathrm{F}_i) \times \mathrm{E}_{i+1} \times \cdots \times \mathrm{E}_p.$$

$\mathrm{E} \setminus \mathrm{G}_i$ est un produit de d'ouverts, il est donc ouvert et alors G_i est fermé. Finalement, on en déduit que F est intersection de fermés, il est donc fermé.

Exemple. Mettons nous dans le cas où $\mathrm{E} = \mathbb{R}^p$.

→ La distance associée à $\mathbb{R}^p$ est la distance induite par $\|.\|_\infty$.

→ Pour tout $x = (x_1, \ldots, x_p) \in \mathbb{R}^p$ et $\varepsilon > 0$, $\mathrm{B}(x, \varepsilon) =]x_1 - \varepsilon, x_1 + \varepsilon[\times \cdots \times]x_p - \varepsilon, x_p + \varepsilon[$.

→ Les pavés ouverts s'écrivent sous forme de $\omega_1 \times \cdots \times \omega_p$ avec pour tout $i \in [\![1;p]\!]$, ω_i un ouvert de $\mathbb{R}$.

→ Tout ouvert de $\mathbb{R}^p$ est réunion d'ensembles de la forme $]x_1 - \varepsilon, x_1 + \varepsilon[\times \cdots \times]x_p - \varepsilon, x_p + \varepsilon[$.

Proposition II.3.

1. Pour tout $k \in [\![1;p]\!]$, π_k est 1−lipschitzienne.
2. Pour tout $k \in [\![1;p]\!]$, si Ω est un ouvert de E, $\pi_k(\Omega)$ est ouvert.

Démonstration. soit $k \in [\![1;p]\!]$.

1. Il suffit d'écrire la définition de la 1−lipschitzianité de π_k. En effet, on a pour tous $x, y \in \mathrm{E}$, $d_k(\pi_k(x), \pi_k(y)) = d_k(x_k, y_k) \leq d(x, y)$.
2. Soit Ω un ouvert de E. Soit $a = (a_1, \ldots, a_p) \in \Omega$ et $\varepsilon > 0$ tel que $\mathrm{B}(a, \varepsilon) \subset \Omega$. On a alors $\mathrm{B}_{d_k}(a_k, \varepsilon) = \pi_k(\mathrm{B}(a, \varepsilon)) \subset \pi_k(\Omega)$, donc $\pi_k(\Omega)$ est ouvert.

Attention : si F est un fermé de E, sa projection $\pi_k(\mathrm{F})$ avec $k \in [\![1;p]\!]$ n'est pas forcément un fermé. En effet, dans le cas où $\mathrm{E} = \mathbb{R}^2$, on peut considérer $\mathrm{F} = \{(x, y) \in \mathbb{R}^2,\ xy = 1\}$ qui est fermé car l'image réciproque du fermé $\{1\}$ par l'application continue $(x, y) \longmapsto xy$, mais $\pi_1(\mathrm{F}) = \mathbb{R}^*$ n'est pas fermé.

(cc) BY-NC-SA

2. Fonctions à valeurs dans un espace produit

Proposition II.4.

Soit $(x_n)_{n\in\mathbb{N}} = (x_{n,1}, \dots, x_{n,p}) \in \mathrm{E}^{\mathbb{N}}$ et $x = (x_1, \dots, x_p) \in \mathrm{E}$. Les propositions suivantes sont équivalentes.

1. $x_n \xrightarrow[n\to+\infty]{} x$
2. $\forall i \in [\![1;p]\!]$, $x_{n,i} \xrightarrow[n\to+\infty]{} x_i$

Démonstration.
(1) $\Rightarrow$ (2) Si $x_n \xrightarrow[n\to+\infty]{} x$, alors pour tout $i \in [\![1;p]\!]$, par continuité de π_i (π_i est $1-$lipschitzienne donc continue)

$$x_{n,i} = \pi_i(x_n) \xrightarrow[n\to+\infty]{} \pi_i(x) = x_i.$$

(2) $\Rightarrow$ (1) Il suffit d'écrire la définition de la distance

$$d(x_n, x) = \max_{i\in[\![1;p]\!]} d_i(x_{n,i}, x_i) \xrightarrow[n\to+\infty]{} 0.$$

Proposition II.5.

Soit A un ensemble, f une fonction de A dans E et $a \in \overline{\mathrm{A}}$. On pose pour tout $x \in \mathrm{A}$, $f(x) = (f_1(x), \dots, f_p(x))$ tel que pour tout $i \in [\![1;p]\!]$, f_i une fonction de A dans E_i. On a alors équivalence entre les points suivants.

1. f admet une limite $\ell = (\ell_1, \dots, \ell_p)$ en a.
2. Pour tout $i \in [\![1;p]\!]$, f_i admet la limite ℓ_i en a.

Démonstration. la démonstration se fait d'une manière identique à celle pour les suites et est donc laissée comme exercice au lecteur.

Proposition II.6.

On reprend les notations précédentes et on considère cette fois-ci que $a \in \mathrm{A}$. la propriété précédente implique l'équivalence entre les deux points suivants.

1. f est continue en a.
2. Pour tout $i \in [\![1;p]\!]$, f_i est continue en a_i.

3. Fonctions d'un espace produit

Dans cette partie, on considère Y un ensemble et pour tout $f \in \mathrm{Y}^{\mathrm{E}}$, $i \in [\![1;p]\!]$ et $a \in \mathrm{E}$, on désigne par $i-$ème application partielle de f en a l'application

$$f_{a,i} : \begin{cases} \mathrm{E}_i & \longrightarrow \mathrm{Y} \\ x & \longmapsto f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p). \end{cases}$$

(cc) BY-NC-SA

Proposition II.7.

Soit $f : \mathrm{E} \longrightarrow \mathrm{Y}$ et $a \in \mathrm{E}$ une application. Si f est continue, alors ses toutes ses applications partielles en a sont continues.

Démonstration. considérons pour tout $i \in [\![1;p]\!]$, l'application

$$j_{a,i} : \begin{cases} \mathrm{E}_i & \longrightarrow \mathrm{E} \\ x & \longmapsto (a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p). \end{cases}$$

Cette application est une isométrie, elle est donc continue. Or on a pour tout $i \in [\![1;p]\!]$, $f_{a,i} = f \circ j_{a,i}$. $f_{a,i}$ est composition de fonctions continues, elle est donc continue.

Définition II.8.

On considère le cas particulier où pour tout $i \in [\![1;p]\!]$, E_i est un intervalle ouvert de $\mathbb{R}$ et Y est un espace vectoriel normé. Pour tout $i \in [\![1;p]\!]$, on appelle $i-$ème dérivée partielle de f lorsqu'elle existe, l'application suivante

$$\frac{\partial f}{\partial x_i} : \begin{cases} \mathrm{E} & \longrightarrow \mathrm{Y} \\ x & \longmapsto \lim\limits_{t_i \to x_i} \dfrac{f_{x,i}(t_i) - f_{x,i}(x_i)}{t_i - x_i} \end{cases}$$

Cette dernière associe un réel x à la dérivée de la $i-$ème application partielle de f en x au point x_i.

Vocabulaire. une fonction de E vers Y continue, qui admet des dérivées partielles continues, est dite dans $\mathcal{C}^1$.

Attention : la continuité des applications partielles de f n'implique pas la continuité de f. En effet, il suffit de considérer l'application suivante.

$$f : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon.} \end{cases} \end{cases}$$

Les applications partielles de f en tout point de R^2 sont continues sur $\mathbb{R}$, mais f n'est pas continue en $(0,0)$, car en considérant la suite qui tend vers 0, $((\frac{1}{n}, \frac{1}{n}))_{n \in \mathbb{N}^*}$ on voit que pour tout $n \in \mathbb{N}^*$ $f(\frac{1}{n}, \frac{1}{n}) = \frac{1}{2}$ et ne tend donc pas vers 0, ce qui fait que f n'est pas continue en $(0,0)$.

Exercice II.9.

Soit f une application de $\mathbb{R}^2$ dans $\mathbb{R}$ qui admet des applications partielles continues en tout point de $\mathbb{R}^2$. Montrer que f est limite simple d'une suite de fonctions continues.

4. Polynômes à plusieurs indeterminées

Dans cette partie, pour tout $\alpha \in \mathbb{N}^n$, on note $\alpha_1, \dots, \alpha_n$ ses coordonnées. On appellera tout élément $\alpha \in \mathbb{N}^n$ un multi-indice.

BY-NC-SA

Définition II.10.

Soit $\mathbb{K}$ un corps. On pose $\mathbb{K}[X_1, \ldots, X_n]$ l'ensemble des sommes de la forme

$$P(X_1, \ldots, X_n) = \sum_{(\alpha_1,\ldots,\alpha_n)\in A} a_{\alpha_1,\ldots,\alpha_n} X_1^{\alpha_1} \ldots X_n^{\alpha_n},$$

avec A une partie finie de $\mathbb{N}^n$.

Dans toute la suite de cette partie, on fixe $P \in \mathbb{K}[X_1, \ldots, X_n]$.

Remarque. en première année, les polynômes ont étés vus comme des suites stationnaires en 0. De la même manière, on peut voir les polynômes de $\mathbb{K}[X_1, \ldots, X_n]$ comme des suites à valeurs dans $\mathbb{K}$ indexées par $\mathbb{N}^n$ et dont le nombre de termes non nuls est fini. Ici par exemple, on se limite dans la somme aux indices $(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ tels que $a_{\alpha_1,\ldots,\alpha_n} \neq 0$, mais sachant qu'il y a un nombre fini de coefficients non nuls, on aurait pu écrire

$$P(X_1, \ldots, X_n) = \sum_{(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}^n} a_{\alpha_1,\ldots,\alpha_n} X_1^{\alpha_1} \ldots X_n^{\alpha_n}.$$

Vocabulaire. Soit $(\alpha_1, \ldots, \alpha_n) \in A$.

- → $a_{\alpha_1,\ldots,\alpha_n}$ est un coefficient de P.
- → $X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ est un monôme.
- → Le degré total du monôme $X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ est $\alpha_1 + \cdots + \alpha_n$.
- → Pour tout $i \in [\![1; n]\!]$, le degré partiel du monôme $X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ en X_i est α_i.
- → Le degré de P est défini par $\deg P = \max_{\alpha\in A'}(\alpha_1 + \cdots + \alpha_n)$ avec A' l'ensemble des multi-indices α tels que $a_\alpha \neq 0$.

Définition II.11.

Soit $d \in \mathbb{N}$. On dit que P est homogène de degré d si

$$\forall(\alpha_1, \ldots, \alpha_n) \in A,\ a_{\alpha_1,\ldots,\alpha_n} \neq 0 \Longrightarrow \alpha_1 + \cdots + \alpha_n = d.$$

L'ensemble des polynômes homogènes de degré d est noté $H_d(X_1, \ldots, X_n)$.

Proposition II.12.

Soit $P = \sum_{(\alpha_1,\ldots,\alpha_n)\in A} a_{\alpha_1,\ldots,\alpha_n} X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ et $Q = \sum_{(\alpha_1,\ldots,\alpha_n)\in B} b_{\alpha_1,\ldots,\alpha_n} X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ deux éléments de $\mathbb{K}[X_1, \ldots, X_n]$. On pose $C = A \cup B$.

1. $P + Q = \sum_{(\alpha_1,\ldots,\alpha_n)\in C} (a_{\alpha_1,\ldots,\alpha_n} + b_{\alpha_1,\ldots,\alpha_n}) X_1^{\alpha_1} \ldots X_n^{\alpha_n}$.
2. $P.Q = \sum_{(\alpha,\beta)\in A\times B} a_\alpha b_\beta X_1^{\alpha_1+\beta_1} \ldots X_n^{\alpha_n+\beta_n}$.

(cc) BY-NC-SA

Proposition II.13.

Soit d et d' deux éléments de $\mathbb{N}$.

1. Si $\mathrm{P} \in \mathrm{H}_d(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ et $\mathrm{Q} \in \mathrm{H}_{d'}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$, alors $\mathrm{PQ} \in \mathrm{H}_{d+d'}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$.
2. $\mathbb{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n] = \bigoplus\limits_{k\in\mathbb{N}} \mathrm{H}_k(\mathrm{X}_1, \ldots, \mathrm{X}_n)$.
3. $\mathbb{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$ est une $\mathbb{K}-$algèbre commutative unitaire intègre.

Démonstration.

1. Soit P et Q deux éléments appartenant respectivement à $\mathrm{H}_d(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ et $\mathrm{H}_{d'}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$. Soit A (resp. B) l'ensemble des multi-indices α (resp. β) tels que $a_\alpha \neq 0$ (resp. $b_\beta \neq 0$). On a alors
$$\mathrm{P} = \sum_{(\alpha_1,\ldots,\alpha_n)\in \mathrm{A}} a_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n} \text{ et } \mathrm{Q} = \sum_{(\alpha_1,\ldots,\alpha_n)\in \mathrm{B}} b_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n}.$$
Ceci donne
$$\mathrm{PQ} = \sum_{(\alpha,\beta)\in \mathrm{A}\times\mathrm{B}} a_\alpha b_\beta \mathrm{X}_1^{\alpha_1+\beta_1} \ldots \mathrm{X}_n^{\alpha_n+\beta_n}$$
et pour tout $(\alpha, \beta) \in \mathrm{A} \times \mathrm{B}$, $\alpha_1 + \beta_1 + \cdots + \alpha_n + \beta_n = d + d'$, donc $\mathrm{PQ} \in \mathrm{H}_{d+d'}(\mathrm{X}_1, \ldots, \mathrm{X}_n)$.

2. Soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$. Il suffit de diviser la somme de P selon le degré de chaque monôme. En effet, on a
$$\mathrm{P} = \sum_{(\alpha_1,\ldots,\alpha_n)\in \mathrm{A}} a_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n} = \sum_{d=0}^{\deg \mathrm{P}} \underbrace{\sum_{\substack{(\alpha_1,\ldots,\alpha_n)\in \mathrm{A}\\ \alpha_1+\cdots+\alpha_n = d}} a_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n}}_{\in \mathrm{H}_d(\mathrm{X}_1,\ldots,\mathrm{X}_n)}.$$
L'unicité de cette écriture découle du fait qu'un polynôme nul est nécessairement à coefficients nuls, on a donc bien $\mathrm{P} \in \bigoplus\limits_{k\in\mathbb{N}} \mathrm{H}_k(\mathrm{X}_1, \ldots, \mathrm{X}_n)$, d'où le résultat voulu.

3. Le fait que $\mathbb{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$ soit une algèbre commutative ne présente pas de difficulté majeure, nous allons donc uniquement montrer que cette algèbre est intègre.
Soit $\mathrm{P}, \mathrm{Q} \in \mathbb{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n] \setminus \{0\}$. On pose
$$\mathrm{P} = \sum_{(\alpha_1,\ldots,\alpha_n)\in \mathrm{A}} a_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n} \text{ et } \mathrm{Q} = \sum_{(\alpha_1,\ldots,\alpha_n)\in \mathrm{B}} b_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n}.$$
Posons aussi
$$\mathrm{PQ} = \sum_{(\alpha_1,\ldots,\alpha_n)\in \mathrm{C}} c_{\alpha_1,\ldots,\alpha_n} \mathrm{X}_1^{\alpha_1} \ldots \mathrm{X}_n^{\alpha_n}.$$
Considérons $\alpha, \beta \in \mathbb{N}^n$ maximaux pour l'ordre lexicographique tels que $a_\alpha \neq 0$ et $b_\beta \neq 0$.
Le coefficient devant $\mathrm{X}^{\alpha_1+\beta_1} \ldots \mathrm{X}^{\alpha_n+\beta_n}$ est $a_\alpha b_\beta = c_{\alpha+\beta}$ et $\alpha + \beta$ est le multi-indice maximal pour l'ordre lexicographique vérifiant $c_{\alpha+\beta} \neq 0$. PQ admet donc un coefficient non nul ce qui nous permet de déduire qu'il est non nul, d'où l'intégrité de $\mathbb{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$.

(cc) BY-NC-SA

Définition II.14.

Soit $\mathrm{P} = \sum_{(\alpha_1,\dots,\alpha_n)\in \mathrm{A}} a_{\alpha_1,\dots,\alpha_n} \mathrm{X}_1^{\alpha_1}\dots \mathrm{X}_n^{\alpha_n}$ un élément de $\mathbb{K}[\mathrm{X}_1,\dots,\mathrm{X}_n]$.
On appelle fonction polynôme associée à P la fonction

$$\tilde{\mathrm{P}} : \begin{cases} \mathbb{K}^n & \longrightarrow \mathbb{K} \\ (x_1,\dots,x_p) & \longmapsto \sum_{(\alpha_1,\dots,\alpha_n)\in \mathrm{A}} a_{\alpha_1,\dots,\alpha_n} x_1^{\alpha_1}\dots x_n^{\alpha_n}. \end{cases}$$

Proposition II.15.

Lorsque $\mathbb{K}=\mathbb{R}$ ou $\mathbb{K}=\mathbb{C}$, pour tout $\mathrm{P}\in\mathbb{K}[\mathrm{X}_1,\dots,\mathrm{X}_n]$, $\tilde{\mathrm{P}}$ est continue.

Démonstration. pour tout $i\in[\![1;p]\!]$, $\tilde{\mathrm{X}}_i=\pi_i$, donc $\tilde{\mathrm{X}}_i$ est continue. Pour tout $\mathrm{P}\in\mathbb{K}[\mathrm{X}_1,\dots,\mathrm{X}_n]$, la fonction $\tilde{\mathrm{P}}$ est produit et combinaison linéaire des fonctions continues $\tilde{\mathrm{X}}_i$ avec $i\in[\![1;n]\!]$, elle est donc continue.

Proposition II.16.

Soient $\mathrm{A}_1,\dots,\mathrm{A}_n$ des parties infinies de $\mathbb{K}$. Soit $\mathrm{P}\in\mathbb{K}[\mathrm{X}_1,\dots,\mathrm{X}_n]$. Si $\tilde{\mathrm{P}}$ s'annule sur $\mathrm{A}_1\times\dots\times\mathrm{A}_n$, alors P est nul.

Démonstration. Montrons ce résultat par récurrence sur n.
Lorsque $n=1$, P est un polynôme à une seule indéterminée qui s'annule en un nombre infini de points, il est donc nul.
Soit $n\geq 2$. Soit $\mathrm{P} = \sum_{(\alpha_1,\dots,\alpha_n)\in \mathrm{A}} a_{\alpha_1,\dots,\alpha_n} \mathrm{X}_1^{\alpha_1}\dots \mathrm{X}_n^{\alpha_n}\in\mathbb{K}[\mathrm{X}_1,\dots,\mathrm{X}_n]$. Réécrivons P de la manière suivante

$$\mathrm{P}=\sum_{k=0}^{\mathrm{N}}\mathrm{Q}_k(\mathrm{X}_1,\dots,\mathrm{X}_{n-1})\mathrm{X}_n^k$$

avec $\mathrm{N}\in\mathbb{N}$ et pour tout $k\in[\![1;\mathrm{N}]\!]$, $\mathrm{Q}_k\in\mathbb{K}[\mathrm{X}_1,\dots,\mathrm{X}_{n-1}]$.
Soit $(a_1,\dots,a_{n-1})\in\mathrm{A}_1\times\dots\times\mathrm{A}_{n-1}$. Pour tout $a_n\in\mathrm{A}_n$,

$$\tilde{\mathrm{P}}(a_1,\dots,a_{n-1},a_n)=\sum_{k=0}^{\mathrm{N}}\mathrm{Q}_k(a_1,\dots,a_{n-1})a_n^k=0.$$

Le polynôme $\mathrm{R}(\mathrm{Y})=\sum_{k=0}^{\mathrm{N}}\mathrm{Q}_k(a_1,\dots,a_{n-1})\mathrm{Y}^k$ est nul sur A_n et possède donc une infinité de racines. Ce polynôme est donc nul. On peut alors affirmer que pour tout $k\in[\![1;\mathrm{N}]\!]$ et $(a_1,\dots,a_{n-1})\in\mathrm{A}_1\times\dots\times\mathrm{A}_{n-1}$, $\mathrm{Q}_k(a_1,\dots,a_{n-1})=0$. Q_k s'annule donc sur $\mathrm{A}_1\times\dots\times\mathrm{A}_{n-1}$ et alors par hypothèse de récurrence, pour tout $k\in[\![1;\mathrm{N}]\!]$, $\mathrm{Q}_k=0$ et finalement $\mathrm{P}=0$.

Exercice II.17.

Montrer que $\mathrm{GL}_n(\mathbb{C})=\{\mathrm{M}\in\mathcal{M}_n(\mathbb{C}),\ \det\mathrm{M}\neq 0\}$ est ouvert.

(cc) BY-NC-SA

Exercice II.18.

Soit $P \in \mathbb{C}[X_1, \dots, X_n] \setminus \{0\}$. Montrer que $Z(P) = \{(x_1, \dots, x_n) \in \mathbb{C}^n,\ \tilde{P}(x_1, \dots, x_n) = 0\}$ est fermé d'intérieur vide.

Exercice II.19.

Soit $P \in \mathbb{C}[X, Y]$ non constant. Montrer que $Z(P)$ est infini.

(cc) BY-NC-SA

Correction de l'exercice I.6 :

On commence par démontrer le lemme suivant.

Lemme II.20.

Soit f une fonction continue et g une fonction de signe constant sur $[a,b]$ non identiquement nulle. Il existe $c \in [a,b]$ tel que

$$\int_a^b f(t)g(t)dt = f(c)\int_a^b g(t)dt.$$

Démonstration. On suppose sans perte de généralité que g est positive. On a alors $\displaystyle\int_a^b g(t)dt > 0$. Soit $\alpha, \beta \in [a,b]$ tels que $f(\alpha) = \inf\limits_{t\in[a,b]} f(t)$ et $f(\beta) = \sup\limits_{t\in[a,b]} f(t)$. On a

$$\int_a^b f(\alpha)g(t)dt \leq \int_a^b f(t)g(t)dt \leq \int_a^b f(\beta)g(t)dt,$$

et alors

$$f(\alpha) \leq \frac{\displaystyle\int_a^b f(t)g(t)dt}{\displaystyle\int_a^b g(t)dt} \leq f(\beta). \tag{10}$$

f étant continue, le théorème des valeurs intermédiaires nous permet d'affirmer l'existence d'un réel $c \in [a,b]$ tel que

$$\frac{\displaystyle\int_a^b f(t)g(t)dt}{\displaystyle\int_a^b g(t)dt} = f(c),$$

d'où le résultat voulu.

Remarque. L'inégalité 10 est en général fausse lorsque g n'est pas de signe constant.

Montrons à présent que l'intégrale de l'exercice converge vers la quantité voulue.
Soit $n \in \mathbb{N}^*$, on a

$$\int_0^1 f(t)\left|\sin(nt)\right| dt = \sum_{k=0}^{\lfloor \frac{n}{2\pi}\rfloor -1} \int_{\frac{2k\pi}{n}}^{\frac{(2k+1)\pi}{n}} f(t)\sin(nt)dt - \int_{\frac{(2k+1)\pi}{n}}^{\frac{(2k+2)\pi}{n}} f(t)\sin(nt)dt + \int_{\lfloor \frac{n}{2\pi}\rfloor \frac{2\pi}{n}}^1 f(t)\left|\sin(nt)\right| dt.$$

En appliquant le lemme précédent sur les deux premières intégrales ci-dessus, on peut affirmer, pour tout $k \in [\![0; \lfloor \frac{n}{2\pi}\rfloor - 1]\!]$ l'existence de $c_{k,n} \in [\frac{2k\pi}{n}, \frac{(2k+1)\pi}{n}]$ et $c'_{k,n} \in [\frac{(2k+1)\pi}{n}, \frac{(2k+2)\pi}{n}]$ tels que

$$\int_{\frac{2k\pi}{n}}^{\frac{(2k+1)\pi}{n}} f(t)\sin(nt)dt = f(c_{k,n})\int_{\frac{2k\pi}{n}}^{\frac{(2k+1)\pi}{n}} \sin(nt)dt = \frac{2}{n}f(c_{k,n})$$

et

$$\int_{\frac{(2k+1)\pi}{n}}^{\frac{(2k+2)\pi}{n}} f(t)\sin(nt) = f(c'_{k,n})\int_{\frac{(2k+1)\pi}{n}}^{\frac{(2k+2)\pi}{n}} \sin(nt) = -\frac{2}{n}f(c'_{k,n}).$$

(cc) BY-NC-SA

En posant $\tilde{c}_{k,n} = c_{\frac{k}{2},n}$ si k pair et $\tilde{c}_{k,n} = c'_{\frac{k-1}{2},n}$ sinon, on obtient

$$\begin{aligned}\int_0^1 f(t)\left|\sin(nt)\right| dt &= \frac{2}{n}\sum_{k=0}^{\lfloor \frac{n}{2\pi}\rfloor -1} f(c_{k,n}) + \frac{2}{n}\sum_{k=0}^{\lfloor \frac{n}{2\pi}\rfloor -1} f(c'_{k,n}) + \int_{\lfloor \frac{n}{2\pi}\rfloor \frac{2\pi}{n}}^1 f(t)\left|\sin(nt)\right| dt\\ &= \frac{2}{\pi}\left(\sum_{k=0}^{2\lfloor \frac{n}{2\pi}\rfloor -1} \frac{\pi}{n} f(\tilde{c}_{k,n})\right) + \int_{\lfloor \frac{n}{2\pi}\rfloor \frac{2\pi}{n}}^1 f(t)\left|\sin(nt)\right| dt.\end{aligned}$$

La somme dans la ligne ci-dessus est une somme de Riemann sur $[0,1]$, elle converge donc vers $\displaystyle\int_0^1 f(t)dt$. De plus, $t \longmapsto f(t)\left|\sin(nt)\right|$ est bornée uniformément (sa borne supérieure ne dépend pas de n) sur $[0,1]$ et $\lfloor \frac{n}{2\pi}\rfloor \frac{2\pi}{n} \xrightarrow[n\to+\infty]{} 1$, donc

$$\int_{\lfloor \frac{n}{2\pi}\rfloor \frac{2\pi}{n}}^1 f(t)\left|\sin(nt)\right| dt \xrightarrow[n\to+\infty]{} 0.$$

On en déduit finalement que

$$\int_0^1 f(t)\left|\sin(nt)\right| \xrightarrow[n\to+\infty]{} \frac{2}{\pi}\int_0^1 f(t)dt.$$

<u>Correction de l'exercice I.7 :</u>

Pour tout $\mathrm{P} = \displaystyle\sum_{k=0}^{n} a_k \mathrm{X}^k \in \mathrm{E}$, on a

$$\left\|\mathrm{P}\right\|_0 = \max(\left|a_0\right|, \ldots, \left|a_n\right|) \geq \frac{1}{n}\sum_{k=0}^{n}\left|a_k\right| \geq \frac{1}{n}\max(\left|a_0\right|, \ldots, \left|a_n\right|) = \frac{1}{n}\left\|\mathrm{P}\right\|_0,$$

donc

$$\left\|\mathrm{P}\right\|_0 \geq \frac{1}{n}\left\|\mathrm{P}\right\|_1 \geq \frac{1}{n}\left\|\mathrm{P}\right\|_0$$

et alors $\left\|.\right\|_0 \sim \left\|.\right\|_1$. On a aussi

$$\left\|\mathrm{P}\right\|_\infty \leq \sum_{k=0}^{n}\left|a_k\right| = \left\|\mathrm{P}\right\|_1$$

donc $\left\|.\right\|_1$ est plus fine que $\left\|.\right\|_\infty$, mais $\left\|.\right\|_\infty$ n'est pas toujours plus fine que $\left\|.\right\|_1$.
En effet, si on considère la suite de polynômes $\mathrm{P}_n = \displaystyle\sum_{k=1}^{n}(-1)^k k \mathrm{X}^{k-1}$, on a pour tout $x \in [0,1]$,

$$\begin{aligned}\tilde{\mathrm{P}}_n(x) &= \left(\sum_{k=0}^{n} - x^k\right)' = \left(x\frac{1-(-x)^n}{1+x}\right)'\\ &= -x\frac{1-(-x)^n}{(1+x)^2} + \frac{1-(-x)^n}{1+x} - \frac{n(-x)^n}{1+x}\end{aligned}$$

et alors pour tout $x \in [0,1]$,

$$\left|\tilde{\mathrm{P}_n}(x)\right| \leq \left|x\frac{1-(-x)^n}{(1+x)^2}\right| + \left|\frac{1-(-x)^n}{1+x}\right| + \left|\frac{n(-x)^n}{1+x}\right| \leq 2+2+n = n+4$$

et donc $\left\|\mathrm{P}_n\right\|_\infty \leq n+4$.
On a aussi $\left\|\mathrm{P}_n\right\|_1 = \displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$, donc s'il existe $\mathrm{C} > 0$ tel que pour tout $n \in \mathbb{N}$,
$\left\|\mathrm{P}_n\right\|_1 \leq \mathrm{C}\left\|\mathrm{P}_n\right\|_\infty$, alors on aura $\dfrac{n(n+1)}{2} \leq \mathrm{C}(n+4)$ ce qui est faux pour n assez grand, d'où le résultat

(cc) BY-NC-SA

voulu.

Correction de l'exercice II.9 :

On va approcher f par une suite de fonctions de première application partielle affine par morceaux. Soit $n \in \mathbb{N}^*$. On subdivise $\mathbb{R}$ en la famille d'intervalles $([\frac{k}{n}, \frac{k+1}{n}[)_{k\in\mathbb{Z}}$. Pour tout $n \in \mathbb{N}^*$, on définit f_n de la manière suivante : pour tout $(x, y) \in \mathbb{R}^2$ et $k \in \mathbb{Z}$, si $x \in [\frac{k}{n}, \frac{k+1}{n}[$, on pose $x = \frac{k+\varepsilon_{n,x}}{n}$ avec $\varepsilon_{n,x} = nx - \lfloor nx \rfloor \in [0, 1[$. On a alors

$$f_n(x, y) = f\left(\frac{k}{n}, y\right) + \varepsilon_{n,x}\left(f\left(\frac{k+1}{n}, y\right) - f\left(\frac{k}{n}, y\right)\right).$$

Montrons que pour tout $n \in \mathbb{N}^*$, f_n est continue. Soit $n \in \mathbb{N}$ et $x \in \mathbb{R}$, soit $k \in \mathbb{Z}$ tel que $x \in [\frac{k}{n}, \frac{k+1}{n}[$.

→ Cas $x \in]\frac{k}{n}, \frac{k+1}{n}[$.
$]\frac{k}{n}, \frac{k+1}{n}[$ étant ouvert, on peut supposer que tous les couples considérés sont dans $]\frac{k}{n}, \frac{k+1}{n}[\times\mathbb{R}$. On pose alors pour tout $u \in]\frac{k}{n}, \frac{k+1}{n}[$, $u = \frac{k+\varepsilon_{n,u}}{n}$ avec $\varepsilon_{n,u} \in [0, 1[$. On en déduit que lorsque $(u, v) \longrightarrow (x, y)$, $\varepsilon_{n,u} \longrightarrow \varepsilon_{n,x}$ et par continuité de la seconde application partielle de f,

$$\begin{aligned} f_n(u, v) &= f\left(\frac{k}{n}, v\right) + \varepsilon_{n,u}\left(f\left(\frac{k+1}{n}, v\right) - f\left(\frac{k}{n}, v\right)\right) \\ &\xrightarrow[(u,v)\to(x,y)]{} f\left(\frac{k}{n}, y\right) + \varepsilon_{n,x}\left(f\left(\frac{k+1}{n}, y\right) - f\left(\frac{k}{n}, y\right)\right) = f(x, y), \end{aligned}$$

donc f_n est continue en (x, y)

→ Cas $x = \frac{k}{n}$.
Étudions séparément la continuité à droite et à gauche.

- Continuité pour u à droite de x.
Lorsque $(u, v) \in [\frac{k}{n}, \frac{k+1}{n}[\times\mathbb{R}$, si $(u, v) \longrightarrow (\frac{k}{n}, y)$, alors $\varepsilon_{n,u} \longrightarrow 0$ et $f\left(\frac{k+1}{n}, v\right) - f\left(\frac{k}{n}, v\right)$ reste bornée. On peut donc affirmer que

$$f_n(u, v) = f\left(\frac{k}{n}, v\right) + \varepsilon_{n,u}\left(f\left(\frac{k+1}{n}, v\right) - f\left(\frac{k}{n}, v\right)\right) \xrightarrow[(u,v)\to(x,y)]{} f\left(\frac{k}{n}, y\right),$$

d'où la continuité de f_n à droite en x.

- Continuité pour u à gauche de $\frac{k}{n}$.
Lorsque $(u, v) \in]\frac{k-1}{n}, \frac{k}{n}] \times \mathbb{R}$, si $(u, v) \longrightarrow (\frac{k}{n}, y)$, alors $\varepsilon_{n,u} \longrightarrow 1$ et

$$f_n(u, v) = f\left(\frac{k-1}{n}, v\right) + \varepsilon_{n,u}\left(f\left(\frac{k}{n}, v\right) - f\left(\frac{k-1}{n}, v\right)\right) \xrightarrow[(u,v)\to(x,y)]{} f\left(\frac{k}{n}, y\right),$$

d'où la continuité à gauche en x.

On en déduit donc que pour tout n, f_n est continue sur $\mathbb{R}^2$. Il reste maintenant à montrer que f_n converge simplement vers f.
Soit $(x, y) \in \mathbb{R}^2$. On a $\varepsilon_{n,x} \in [0, 1[$ et le terme $f\left(\frac{k+1}{n}, y\right) - f\left(\frac{k}{n}, y\right) \xrightarrow[n\to+\infty]{} 0$, car la première application de f est continue. On a alors

$$f_n(x, y) = f\left(\frac{k}{n}, y\right) + \varepsilon_{n,x}\left(f\left(\frac{k+1}{n}, y\right) - f\left(\frac{k}{n}, y\right)\right) \xrightarrow[n\to+\infty]{} f(x, y),$$

d'où le résultat voulu.

(cc) BY-NC-SA

Correction de l'exercice II.17 :

On a pour tout $\mathrm{M} = (m_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{C})$,

$$\det \mathrm{M} = \sum_{\sigma\in \mathrm{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} m_{i,\sigma(i)}.$$

L'application det est alors polynomiale et donc continue. Or $\mathrm{GL}_n(\mathbb{C}) = \det^{-1}(\mathbb{R}^*)$. Il s'agit de l'image réciproque d'un ouvert par une application continue, c'est donc un ouvert.

Correction de l'exercice II.18 :

On a $\mathrm{Z(P)} = \tilde{\mathrm{P}}^{-1}(\{0\})$. Il s'agit de l'image réciproque d'un fermé par une application continue (car polynomiale), c'est donc un fermé.
Supposons que Z(P) n'est pas d'intérieur vide. Il existe alors $\varepsilon > 0$ et $x = (x_1, \dots, x_n) \in \mathrm{Z(P)}$ tel que $\mathrm{B}(x, \varepsilon) \subset \mathrm{Z(P)}$. P s'annule alors sur $\mathrm{B}(x,\varepsilon) = \mathrm{B}_{d_1}(x_1, \varepsilon) \times \dots \times \mathrm{B}_{d_n}(x_n, \varepsilon)$. Pour tout $i \in [\![1; n]\!]$, $\mathrm{B}_{d_i}(x_i, \varepsilon)$ est infinie. P s'annule sur le produit d'ensembles infinis, il est donc nul ce qui est absurde. On en déduit donc que Z(P) est bien fermé d'intérieur vide.

Correction de l'exercice II.19 :

On peut écrire P de la manière suivante

$$\mathrm{P(X, Y)} = \sum_{k=0}^{\mathrm{N}} \mathrm{Q}_k(\mathrm{Y})\mathrm{X}^k$$

avec $\mathrm{Q_N} \neq 0$.

- → Cas N = 0.
  Lorsque N = 0, alors Q_0 est nécessairement non nul et admet au moins une racine qu'on note a. On a alors pour tout $x \in \mathbb{C}$, $\tilde{\mathrm{P}}(x, a) = 0$ ce qui implique que $\mathbb{C} \times \{a\} \subset \mathrm{Z(P)}$. Z(P) est donc infini.
- → Cas $\mathrm{N} \geq 1$.
  $\mathrm{A} = \{a \in \mathbb{C}, \mathrm{Q_N}(a) \neq 0\}$ est infini. Soit $a \in \mathrm{A}$. $x \longmapsto \tilde{\mathrm{P}}(x, a)$ est une fonction polynomiale de degré N, elle admet donc au moins une racine b et donc $(a, b) \in \mathrm{Z(P)}$. A étant infini, on peut construire une infinité de couples d'éléments distincts de Z(P). Z(P) est alors bien infini.

(cc) BY-NC-SA

Chapitre 11.5

Compacité

Dans ce chapitre, on considère X un espace métrique muni d'une distance d. Tous les ensembles considérés sont, lorsque cela est nécessaire, non vides.

I Valeurs d'adhérence

Proposition I.1.

Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans X et $a \in X$. Les propriétés suivantes sont équivalentes.

1. $\forall \varepsilon > 0,\ \forall N \in \mathbb{N},\ \exists n \geq N,\ x_n \in B(a, \varepsilon)$.
2. Pour tout $\varepsilon > 0$, l'ensemble des indices n tels que $d(x_n, a) < \varepsilon$ est infini.
3. Il existe une extractrice φ tell que $x_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$.

Lorsque ces propriétés sont vérifiées, on dit que a est une valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$.

Démonstration.

→ (1) ⇒ (2)
La proposition (1) implique que pour tout $\varepsilon > 0$, l'ensemble des indices $n \in \mathbb{N}$ telle que $d(a, x_n) < \varepsilon$ est non borné. Il est donc infini.

→ (2) ⇒ (3)
Construisons par récurrence une extractrice φ qui vérifie pour tout $n \in \mathbb{N}$, $d(a, x_{\varphi(n)}) < \frac{1}{n+1}$. Pour $n = 0$, l'ensemble $A_0 = \{n \in \mathbb{N},\ d(a, x_n) < 1\}$ est infini donc non vide. On pose alors $\varphi(0) = m$ avec $m \in A$.
Soit $n \in \mathbb{N}$, supposons que $\varphi(0), \varphi(1), \ldots, \varphi(n)$ sont correctement définis.
L'ensemble $A_n = \{k \in \mathbb{N},\ d(a, x_k) < \frac{1}{n+2}\}$ est infini. On dispose donc de $r \in A_n$ tel que pour tout $i \in [\![1; n]\!]$, $r > \varphi(i)$. On pose alors $\varphi(n+1) = r$, ce qui donne bien $d(a, x_{\varphi(n+1)}) < \frac{1}{n+2}$.

→ (3) ⇒ (1)
Soit $\varepsilon > 0$ et $N \in \mathbb{N}$. $x_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$, donc il existe $n' \geq N$ tel que $x_{\varphi(n')} \in B(a, \varepsilon)$, il suffit alors de voir que $n = \varphi(n')$ vérifie la propriété voulue.

Notation. Pour toute suite $(x_n) \in X^{\mathbb{N}}$, on note $\mathrm{Adh}((x_n)_{n\in\mathbb{N}})$ l'ensemble des valeurs d'adhérence de (x_n). Pour simplifier, et lorsqu'il n'y a pas ambiguïté, on notera $\mathrm{Adh}((x_n)_{n\in\mathbb{N}})$ comme $\mathrm{Adh}(x_n)$.

Proposition I.2.

Pour toute suite $(x_n) \in X^{\mathbb{N}}$, on a $\mathrm{Adh}(x_n) = \bigcap_{p\in\mathbb{N}} \overline{\{x_n,\ n \geq p\}}$.

Démonstration. Posons pour tout $p \in \mathbb{N}$, $A_p = \{x_n,\ n \geq p\}$.

→ (⊂) Soit $a \in \mathrm{Adh}(x_n)$. Pour tout $p \in \mathbb{N}$ et tout $\varepsilon > 0$ il existe $n \geq p$ tel que $x_n \in B(a, \varepsilon)$, *i.e.* $A_p \cap B(a, \varepsilon) \neq \varnothing$.
On peut alors affirmer que pour tout $p \in \mathbb{N}$, $a \in \overline{A_p}$, *i.e.* $a \in \bigcap_{p\in\mathbb{N}} \overline{A_p}$.

(cc) BY-NC-SA

→ ($\supset$) Soit $a \in \bigcap_{p\in\mathbb{N}} \overline{\mathrm{A}_p}$. Soit $\varepsilon > 0$ et $\mathrm{N} \in \mathbb{N}$. $a \in \overline{\mathrm{A}_\mathrm{N}}$, donc pour tout $\varepsilon > 0$, il existe $n \geq \mathrm{N}$ tel que $x_n \in \mathrm{B}(a, \varepsilon)$, *i.e.* $a \in \mathrm{Adh}(x_n)$.

II Définitions et propriétés structurelles

Définition II.1.

L'espace métrique (X, d) est dit compact lorsque toute suite à valeurs dans X admet une valeur d'adhérence dans X.
Une partie A de X est dite compacte si elle l'est pour la distance induite d_A.

Exemples.

→ Pour tout $a, b \in \mathbb{R}$ tel que $a \leq b$, $[a, b]$ est une partie compacte de $\mathbb{R}$.

→ $\mathbb{R}$ n'est pas compact pour la distance induite par la valeur absolue. En effet, si on considère la suite $(u_n)_{n\in\mathbb{N}}$ définie par $\forall n \in \mathbb{N}$, $u_n = n$, on peut voir que toute extractrice de u_n tend vers l'infini et donc (u_n) n'admet pas de valeur d'adhérence.

→ $\overline{\mathbb{R}}$ est compact pour la distance $d : (x, y) \longmapsto |\arctan(x) - \arctan(y)|$ (avec $\arctan(+\infty) = \frac{\pi}{2}$ et $\arctan(-\infty) = -\frac{\pi}{2}$).

Proposition II.2.

1. Si A est une partie compacte de X, alors A est fermée et bornée.
2. Soit Y une partie de X et A une partie de Y. Si Y est compact, alors

$$\text{A est compact} \iff \text{A est fermé.}$$

Démonstration.

1. Soit A une partie compacte de X.

   → Montrons que A est fermé.
   Soit $a \in \overline{\mathrm{A}}$. Par définition, il existe une suite $(a_n) \in \mathrm{A}^\mathbb{N}$ qui converge vers A. A est compact, il existe donc une extractrice φ telle que $(a_{\varphi(n)})$ converge vers $b \in \mathrm{A}$. Par unicité de la limite, on a $a = b \in \mathrm{A}$ et finalement $\mathrm{A} = \overline{\mathrm{A}}$, *i.e.* A est fermé.

   → Montrons que A est borné.
   Si A est non borné, alors par définition il existe une suite $(a_n) \in \mathrm{A}^\mathbb{N}$ et $a \in \mathrm{A}$ tel que $d(a, a_n) \xrightarrow[n\to+\infty]{} +\infty$. Aucune suite extraite ne peut alors converger car pour toute extractrice φ et $x \in \mathrm{A}$, $d(x, a_{\varphi(n)}) \geq d(a, a_{\varphi(n)}) - d(x, a) \xrightarrow[n\to+\infty]{} +\infty$.

2. ($\Rightarrow$) Le sens de gauche à droite est une simple conséquence de la propriété précédente.

   ($\Leftarrow$) Supposons que A est fermé. Soit $(a_n) \in \mathrm{A}^\mathbb{N}$. Y est compact, il existe donc une extractrice φ et un élément x de Y tel que $a_{\varphi(n)} \xrightarrow[n\to+\infty]{} x$. A étant fermé, on a nécessairement $x \in \mathrm{A}$ et donc A est compact.

(cc) BY-NC-SA

III Produit d'espaces compacts

Proposition III.1.

Soit $p \in \mathbb{N}^*$ et $X_1, \ldots, X_p$ des compacts munis des distances $d_1, \ldots, d_p$. L'espace produit $X_1 \times \cdots \times X_p$ est compact.

Démonstration. Procédons par récurrence.

→ Le cas $p = 1$ est évident.

→ Soit $p \in \mathbb{N}^*$. Supposons que p vérifie la propriété voulue. Soit $X_1, \ldots, X_{p+1}$ $p+1$ compacts et soit $(x_n)_{n\in\mathbb{N}} = ((x_{1,n}, \ldots, x_{p+1,n}))_{n\in\mathbb{N}}$.
L'espace produit $X_1 \times \cdots \times X_p$ étant compact par hypothèse, il existe une extractrice φ telle que $(y_{\varphi(n)}) = (x_{1,\varphi(n)}, \ldots, x_{p,\varphi(n)})_{n\in\mathbb{N}}$ converge vers $(a_1, \ldots, a_p) \in X_1 \times \cdots \times X_p$. X_{p+1} est compact, il existe donc une extractrice ψ telle que la suite $(x_{p+1,\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ converge vers $a_{p+1} \in X_{p+1}$.
On a alors

$$\begin{aligned} d(x_{\varphi\circ\psi(n)}, (a_1, \ldots, a_{p+1})) &= \max_{i\in[\![1;p+1]\!]} d_i(x_{i,\varphi\circ\psi(n)}, a_i) \\ &\leq d(y_{\varphi\circ\psi(n)}, (a_1, \ldots, a_p)) + d_{p+1}(x_{p+1,\varphi\circ\psi(n)}, a_{p+1}) \xrightarrow[n\to+\infty]{} 0. \end{aligned}$$

La suite $(x_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ converge donc bien vers $(a_1, \ldots, a_{p+1}) \in X_1 \times \cdots \times X_{p+1}$.
L'espace $X_1 \times \cdots \times X_{p+1}$ est donc bien compact.

Proposition III.2.

Soit $n \in \mathbb{N}^*$ et A une partie de $\mathbb{R}^n$ munie de $\|.\|_\infty$. A est compact si et seulement s'il est fermé et borné.

Démonstration.

(⇒) Cette implication a déjà été montrée auparavant.

(⇐) Supposons que A est fermée et bornée. A étant borné, il existe $M \geq 0$ tel que pour tout $a \in A$, $\|a\|_\infty \leq M$. On a alors $A \subset [-M, M]^n$. $[-M, M]$ étant compact, $[-M, M]^n$ est un produit de compacts de $\mathbb{R}$, c'est donc un compact. A est fermé borné et est inclus dans un compact, il est donc compact d'après ce qui précède.

Conséquence. Par équivalence des normes en dimension finie, si on munit $\mathbb{R}^n$ d'une norme quelconque $\|.\|$, l'équivalence ci-dessus reste vraie.

Attention : Cette équivalence est fausse en dimension infinie en général. En effet, pour le voir, il suffit de considérer l'exemple suivant.

On considère l'espace vectoriel normé $E = \mathcal{C}([0, 2\pi], \mathbb{R})$ muni de la norme $\|.\|_2$ définie par

$$\forall f \in E, \ \ \|f\|_2 = \left(\int_0^{2\pi} f(t)^2 dt\right)^{\frac{1}{2}}.$$

On considère la suite à valeurs dans E $(f_n)_{n\in\mathbb{N}} = (x \longmapsto \sin(nx))_{n\in\mathbb{N}}$.
On a pour tout $n \in \mathbb{N}$,

$$\|f_n\|_2^2 = \int_0^{2\pi} \sin(nt)^2 dt = \int_0^{2\pi} \frac{1 - \cos(2nt)}{2} = \pi.$$

Donc (f_n) est une suite à valeurs dans $B_f(0, \pi)$ qui est une partie fermée bornée de E. et pour tout

(cc) BY-NC-SA

$n, m \in \mathbb{N}$ tels que $m \neq m$,

$$\|f_n - f_m\|_2^2 = \int_0^{2\pi} (\sin(nt) - \sin(mt))^2 dt = \int_0^{2\pi} \sin(nt)^2 + \sin(mt)^2 - 2\sin(nt)\sin(mt) dt = 2\pi.$$

On en déduit que pour toute extractrice φ et tout $n \in \mathbb{N}$, $\left\|f_{\varphi(n)} - f_{\varphi(n+1)}\right\|_2 = \sqrt{2\pi}$ et donc $(f_{\varphi(n)})$ ne peut pas converger.
En résumé, on a trouvé une suite (f_n) à valeurs dans un fermé borné dont aucune suite extraite ne converge, donc $\mathrm{B}_f(0, \pi)$ est fermé borné mais n'est pas compact.

Définition III.3.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite à valeurs dans X. On dit que (u_n) est de Cauchy si

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \forall m \geq \mathrm{N},\ d(u_n, u_m) < \varepsilon.$$

Remarque. Toute suite convergente est de Cauchy. Montrons ce résultat. Soit $\varepsilon > 0$ si $(u_n) \in \mathrm{X}^{\mathbb{N}}$ est convergente vers une limite $u \in \mathrm{X}$, il suffit de prendre $\mathrm{N} \in \mathbb{N}$ assez grand pour qu'on ait pour tout $n \geq \mathrm{N}$, $d(u_n, u) < \frac{\varepsilon}{2}$. On a alors pour tout $n, m \geq \mathrm{N}$, $d(u_n, u_m) \leq d(u_n, u) + d(u_m, u) < \varepsilon$.

Attention : la réciproque est fausse. En effet, si on considère l'espace vectoriel normé $\mathbb{R}[\mathrm{X}]$ muni de $\|.\|_1$, et la suite $(\mathrm{P}_n) \in \mathbb{R}[\mathrm{X}]^{\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N},\ \mathrm{P}_n(\mathrm{X}) = \sum_{k=0}^{n} \frac{1}{2^k} \mathrm{X}^k.$$

Montrons que cette suite de Cauchy. Soit $\varepsilon > 0$, on a pour tout $m, n \in \mathbb{N}$, si $n > m$, alors

$$\|\mathrm{P}_n - \mathrm{P}_m\|_1 = \sum_{k=m+1}^{n} \frac{1}{2^k} \leq \sum_{k=m+1}^{+\infty} \frac{1}{2^k} = \frac{1}{2^m}.$$

Il existe donc $\mathrm{N} \in \mathbb{N}$ assez grand tel que pour tout $n, m > \mathrm{N}$, $\frac{1}{2^m} < \varepsilon$ et alors $\|\mathrm{P}_n - \mathrm{P}_m\|_1 < \varepsilon$, donc (P_n) est de Cauchy.
En revanche, cette suite n'est pas convergente. En effet, pour tout polynôme Q de degré m, pour tout $n > m$,

$$\|\mathrm{P}_n - \mathrm{Q}\| \geq \sum_{k=m+1}^{n} \frac{1}{2^k} \xrightarrow[n\to+\infty]{} \frac{1}{2^m} \neq 0.$$

(P_n) ne peut donc pas converger vers Q et alors (P_n) n'est pas convergente.

Définition III.4.

On dit que X est complet lorsque toute suite à valeurs dans X de Cauchy converge.

IV Suites dans un compact

Proposition IV.1.

Si X est compact, alors il est complet.

Démonstration. Supposons que X est compact. Soit $(u_n) \in \mathrm{X}^{\mathbb{N}}$ une suite de Cauchy. Soit φ une extractrice et $u \in \mathrm{X}$ tels que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} u$. Soit $\varepsilon > 0$ et $\mathrm{N} \in \mathbb{N}$ tel que pour tout $m, n \geq \mathrm{N}$ $d(u_n, u_m) < \frac{\varepsilon}{2}$.

BY-NC-SA

Soit $N' \in \mathbb{N}$ tel pour tout $n \geq N'$ $d(u_{\varphi(n)}, u) < \frac{\varepsilon}{2}$. On a alors

$$\forall n \geq \max(N, N'),\ d(u_n, u) \leq d(u_n, u_{\varphi(n)}) + d(u_{\varphi(n)}, u) < \varepsilon.$$

(u_n) est alors bien convergente et donc X est complet.

Remarque. La démonstration ci-dessus nous dévoile un résultat utile : peu importe l'espace X, une suite $(u_n) \in X^{\mathbb{N}}$ de Cauchy qui admet une valeur d'adhérence est convergente. Plus généralement, une suite de Cauchy admet au plus une seule valeur d'adhérence.

Attention : La réciproque de ce résultat est fausse. En effet, $\mathbb{R}$ est complet, mais n'est pas compact. Montrons que $\mathbb{R}$ est complet. Si $(u_n) \in \mathbb{R}^{\mathbb{N}}$ est de Cauchy, alors il existe $N \in \mathbb{N}$ tel que pour tous $m, n \geq N$, $|u_n - u_m| < \varepsilon$. En prenant $\varepsilon = 1$, on voit que pour tout $n \geq N$, $|u_n - u_N| < 1$. On en déduit donc immédiatement que (u_n) est bornée par $\max(|u_0|, |u_1|, \ldots, |u_{N-1}|, |u_N|)+1$. On peut donc appliquer Bolzano-Weierstrass pour affirmer que (u_n) admet une valeur d'adhérence. (u_n) est une suite de Cauchy qui admet une valeur d'adhérence, elle est donc convergente et donc $\mathbb{R}$ est complet. Pour montrer que $\mathbb{R}$ n'est pas compact, il suffit de voir que la suite $(u_n) = (n)_{n\in\mathbb{N}}$ n'a pas de valeur d'adhérence.

Proposition IV.2.

On suppose que X est compact. Soit $(u_n) \in X^{\mathbb{N}}$. On a alors

$$(u_n) \text{ converge } \iff (u_n) \text{ admet au plus une valeur d'adhérence.}$$

Démonstration.

($\Rightarrow$) Cette implication ne présente pas de difficulté et est laissée comme exercice au lecteur.

($\Leftarrow$) Par compacité de X, (u_n) admet au moins une valeur d'adhérence, donc (u_n) admet exactement une valeur d'adhérence qu'on note u. Supposons que (u_n) ne converge pas vers u. Il existe donc $\varepsilon > 0$ tel que $A_\varepsilon = \{n \in \mathbb{N},\ d(u_n, u) \geq \varepsilon\}$ est infini. On considère alors φ une extractrice telle que $\varphi(\mathbb{N}) \subset A_\varepsilon$. Par construction, on a pour tout $n \in \mathbb{N}$, $d(u_{\varphi(n)}, u) \geq \varepsilon$. Mais $(u_{\varphi(n)})$ est à valeurs dans le compact X, donc il existe $x \in X$ et ψ une extractrice telle que $u_{\varphi\circ\psi(n)} \xrightarrow[n\to+\infty]{} x$, mais $\varepsilon \leq d(u_{\varphi\circ\psi(n)}, u) \xrightarrow[n\to+\infty]{} d(x, u)$, donc $x \neq u$. On a trouvé une valeur d'adhérence de (u_n) différente de u ce qui est absurde.

Exercice IV.3.

Soit $d \in \mathbb{N}^*$ et $(u_n) \in (\mathbb{R}^d)^{\mathbb{N}}$ vérifiant les propriétés suivantes.

1. (u_n) est bornée.
2. (u_n) possède un nombre fini de valeurs d'adhérences.
3. $u_{n+1} - u_n \xrightarrow[n\to+\infty]{} 0$.

Montrer que (u_n) est convergente.

Exercice IV.4.

Supposons que X est compact. Soit (u_n) une suite à valeurs dans X. Montrer que

$$d(u_p, \mathrm{Adh}(u_n)) \xrightarrow[p\to+\infty]{} 0.$$

(cc) BY-NC-SA

Exercice IV.5.

Supposons que X est compact. Soit $f \in \mathcal{C}(\mathrm{X}, \mathrm{X})$ et (u_n) une suite définie par

$$\begin{cases} u_0 \in \mathrm{X} \\ \forall n \in \mathbb{N},\ u_{n+1} = f(u_n). \end{cases}$$

On pose $\mathrm{A} = \mathrm{Adh}(u_n)$. Montrer que $f(\mathrm{A}) = \mathrm{A}$.

Exercice IV.6.

Soit Y un espace métrique muni d'une distance δ. On pose pour tout $f \in \mathrm{Y}^{\mathrm{X}}$,

$$\Gamma_f = \{(x, f(x)), x \in \mathrm{X}\} \subset \mathrm{X} \times \mathrm{Y}.$$

1. Montrer que pour tout $f \in \mathcal{C}(\mathrm{X}, \mathrm{Y})$, Γ_f est fermé dans $\mathrm{X} \times \mathrm{Y}$.
2. Donner un exemple de X, Y et f vérifiant f discontinue et Γ_f fermé.
3. On suppose que Y est compact. Montrer que si Γ_f est fermé, alors f est continue.

V Continuité uniforme

Théorème (Théorème de Heine) V.1.

Soit Y un espace métrique muni d'une distance δ et $f \in \mathcal{C}(\mathrm{X}, \mathrm{Y})$. Si X est compact, alors f est uniformément continue.

Démonstration. On commence par écrire la définition de l'uniforme continuité de f

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x, y \in \mathrm{X},\ d(x, y) < \eta \Longrightarrow \delta(f(x), f(y)) < \varepsilon.$$

Supposons que la négation de cette proposition est vraie, *i.e.*

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x, y \in \mathrm{X},\ d(x, y) < \eta \text{ et } \delta(f(x), f(y)) \geq \varepsilon.$$

Il existe donc un certain $\varepsilon > 0$ pour lequel il existe deux suites (x_n) et (y_n) à valeurs dans X vérifiant

$$\forall n \in \mathbb{N},\ d(x_n, y_n) < \frac{1}{n+1} \text{ et } \delta(f(x_n), f(y_n)) \geq \varepsilon.$$

X étant compact, il existe une extractrice φ telle que $(x_{\varphi(n)})$ est convergente. De même, il existe une extractrice ψ telle que $(y_{\varphi\circ\psi(n)})$ soit convergente. Or, on a pour tout $n \in \mathbb{N}$,

$$d(x_{\varphi\circ\psi(n)}, y_{\varphi\circ\psi(n)}) < \frac{1}{\varphi\circ\psi(n) + 1} \xrightarrow[n\to+\infty]{} 0.$$

$(x_{\varphi\circ\psi(n)})$ et $(y_{\varphi\circ\psi(n)})$ ont donc la même limite. On a alors par continuité de f,

$$\varepsilon \leq \delta(f(x_{\varphi\circ\psi(n)}), f(y_{\varphi\circ\psi(n)})) \xrightarrow[n\to+\infty]{} 0.$$

ce qui est absurde. On en déduit donc que f est uniformément continue.

(cc) BY-NC-SA

Exercice V.2.

Soit $f \in \mathcal{C}([0,1]^2, \mathbb{R})$, montrer que $\varphi : \begin{cases} [0,1] & \longrightarrow \mathbb{R} \\ x & \longmapsto \sup\limits_{y\in[0,1]} f(x,y) \end{cases}$ est continue.

VI Optimisation sur un compact

Proposition VI.1.

Soit Y un espace métrique muni d'une distance δ et $f \in \mathcal{C}(\mathrm{X}, \mathrm{Y})$.

1. Si X est compact, alors $f(\mathrm{X})$ est compact.
2. Si $\mathrm{X} \neq \varnothing$ et $\mathrm{Y} = \mathbb{R}$, alors f est bornée et atteint ses bornes.

Démonstration.

1. Soit $(y_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $f(\mathrm{X})$. Il existe une suite $(x_n) \in \mathrm{X}^{\mathbb{N}}$ telle que pour tout $n \in \mathbb{N}$, $f(x_n) = y_n$. Comme X est compact, il existe une extractrice φ telle que $(x_{\varphi(n)})$ converge vers un élément x de X. Par continuité de f, on a

$$y_{\varphi(n)} = f(x_{\varphi(n)}) \xrightarrow[n\to+\infty]{} f(x) \in f(\mathrm{X}).$$

   On en déduit donc que $f(\mathrm{X})$ est compact.

2. D'après le point précédent, $f(\mathrm{X})$ est compact donc c'est un fermé borné, cet ensemble contient alors sa borne supérieure et sa borne inférieure car ce sont les limites de suites à valeurs dans $f(\mathrm{X})$. On en déduit qu'il existe $a, b \in \mathrm{X}$ tel que $f(a) = \sup f(\mathrm{X})$ et $f(b) = \inf f(\mathrm{X})$. f est donc bornée et atteint ses bornes.

Corollaire VI.2.

Si X est compact, F fermé dans X et $f \in \mathcal{C}(\mathrm{X}, \mathrm{Y})$, alors $f(\mathrm{F})$ est fermé dans Y.

Démonstration. F est un fermé dans un compact, c'est donc un compact et alors $f(\mathrm{F})$ est compact et finalement c'est un fermé dans Y.

Exercice VI.3.

On suppose que X est compact. Soit $f \in \mathcal{C}(\mathrm{X}, \mathrm{Y})$ bijective. Montrer que f^{-1} est continue.

Exercice VI.4.

Soit $f, g \in \mathcal{C}([0,1],[0,1])$ tels que $f \circ g = g \circ f$. Montrer qu'il existe $a \in [0,1]$ tel que $f(a) = g(a)$.

BY-NC-SA

Exercice VI.5.

Soit E un espace vectoriel normé muni d'une norme $\|.\|$. Soit A et B deux parties de E. On pose

$$\mathrm{A}+\mathrm{B}=\{a+b,\ (a,b)\in\mathrm{A}\times\mathrm{B}\}\,.$$

Dire lesquelles des propositions suivantes sont vraies.

1. Si A ou B est ouvert, alors $\mathrm{A}+\mathrm{B}$ est ouvert.
2. Si A et B sont compacts, alors $\mathrm{A}+\mathrm{B}$ est compact.
3. Si A est compact et B est fermé, alors $\mathrm{A}+\mathrm{B}$ est fermé.
4. Si A et B sont fermés alors $\mathrm{A}+\mathrm{B}$ est fermé.

Exercice VI.6.

On suppose que X est compact non vide. Soit $f\in\mathrm{X}^{\mathrm{X}}$. telle que

$$\forall(x,y)\in\mathrm{X}^2,\ x\neq y\Longrightarrow d(f(x),f(y))<d(x,y).$$

1. Montrer que f admet un unique point fixe ℓ.
2. Considérons la suite $(u_n)\in\mathrm{X}^{\mathbb{N}}$ définie par $\begin{cases}u_0\in\mathrm{X}\\ \forall n\in\mathbb{N},\ u_{n+1}=f(u_n).\end{cases}$

   Montrer que $u_n\xrightarrow[n\to+\infty]{}\ell$.

VII Compléments

1. Recouvrement fini d'un compact

Proposition VII.1.

On suppose que X est compact. Pour tout $\varepsilon>0$, il existe $p\in\mathbb{N}^*$ et $a_1,\dots,a_p$ tels que

$$\mathrm{X}=\bigcup_{k=1}^{p}\mathrm{B}(a_k,\varepsilon).$$

On dit que les boules de l'union ci-dessus "recouvrent" le compact X. On appelle la famille d'ensembles $(\mathrm{B}(a_k,\varepsilon))_{k\in[\![1;p]\!]}$ recouvrement fini (resp. infini) de X lorsque cette famille est finie (resp. infinie).

Démonstration. Supposons que

$$\exists\varepsilon>0,\ \forall p\in\mathbb{N}^*,\ \forall(a_1,\dots,a_p)\in\mathrm{X}^p,\ \bigcup_{k=1}^{p}\mathrm{B}(a_k,\varepsilon)\subsetneq\mathrm{X}.$$

Soit $\varepsilon>0$ vérifiant la propriété ci-dessus. Construisons la suite $(a_n)_{n\in\mathbb{N}}\in\mathrm{X}^{\mathbb{N}}$ de la manière suivante : on prend tout d'abord $a_0\in\mathrm{X}$ au hasard. Soit $n\in\mathbb{N}$ tel que a_n est bien défini. $\mathrm{A}_n=\mathrm{X}\setminus\bigcup_{k=0}^{n}\mathrm{B}(a_k,\varepsilon)$ est non vide, on peut donc prendre a_{n+1} dans A_n. Par construction, on a pour tout $n\neq m$, $d(a_n,a_m)\geq\varepsilon$ et donc pour toute extractrice φ, pour tout $n\in\mathbb{N}$, $d(a_{\varphi(n)},a_{\varphi(n+1)})\geq\varepsilon$. $(a_{\varphi(n)})$ ne peut donc pas converger et alors X n'est pas compact, ce qui est absurde.

(cc) BY-NC-SA

Remarque. Il existe une sorte de réciproque à ce résultat qu'on appelle propriété de Borel-Lebesgue. L'énoncé de cette propriété est le suivant.

Une partie A d'un espace métrique est compacte si et seulement si de tout recouvrement (potentiellement infini) de A, on peut extraire un recouvrement fini.

Le lecteur curieux est encouragé à se renseigner sur la démonstration de ce résultat quelques fois utile aux oraux.

Exercice VII.2.

On suppose que X est compact.

1. Soit $\varepsilon > 0$. Montrer qu'il existe $\mathrm{N} \in \mathbb{N}^*$ tel que
$$\exists (x_1, \ldots, x_\mathrm{N}) \in \mathrm{X}^\mathrm{N},\ \forall i \neq j,\ d(x_i, x_j) \geq \varepsilon \tag{$*$}$$
et tel que toute famille de points $y_1, \ldots, y_n$ vérifiant la propriété $(*)$ est de cardinal inférieur à N (*i.e.* $n \leq \mathrm{N}$).
2. Soit $f : \mathrm{X} \longrightarrow \mathrm{X}$ une isométrie. Montrer que f est surjective.

2. Fermés emboités

Proposition VII.3.

On suppose que X est compact. Pour toute suite de fermés non vides décroissante pour l'inclusion $(\mathrm{F}_n)_{n\in\mathrm{N}} \in \mathcal{P}(\mathrm{X})^\mathbb{N}$, $\bigcap\limits_{n\in\mathbb{N}} \mathrm{F}_n \neq \varnothing$.

Démonstration. Soit (x_n) une suite à valeurs dans X telle que pour tout $n \in \mathbb{N}$, $x_n \in \mathrm{F}_n$ (cette suite existe car tous les F_n sont non vides). La suite (F_n) étant décroissante, on a pour tout $n \in \mathbb{N}$ et tout $i \leq n$, $x_n \in \mathrm{F}_i$.
X étant compact, il existe une extractrice φ telle que $(x_{\varphi(n)})$ converge vers un élément x de X. On a alors pour tout $p \in \mathbb{N}$, la suite $(x_{\varphi(n)})$ est à valeurs dans F_p à partir d'un certain rang donc x est limite d'une suite à valeurs dans F_p, *i.e.* $x \in \overline{\mathrm{F}_p} = \mathrm{F}_p$. On a donc pour tout $p \in \mathbb{N}$, $x \in \mathrm{F}_p$ et alors $x \in \bigcap\limits_{n\in\mathbb{N}} \mathrm{F}_n$ et finalement $\bigcap\limits_{n\in\mathbb{N}} \mathrm{F}_n \neq \varnothing$.

Corollaire VII.4.

On suppose que X est compact. Soit (f_n) une suite de fonctions à valeurs dans $\mathcal{C}(\mathrm{X}, \mathbb{R})$ et $f \in \mathcal{C}(\mathrm{X}, \mathbb{R})$. Si pour tout $x \in \mathrm{X}$, la suite $(f_n(x))_{n\in\mathbb{N}}$ est croissante et converge vers $f(x)$, alors (f_n) converge uniformément vers f.

Démonstration. Soit $\varepsilon > 0$. Posons pour tout $n \in \mathbb{N}$, $\mathrm{F}_n = \{x \in \mathrm{X},\ |f_n(x) - f(x)| \geq \varepsilon\}$. Remarquons tout d'abord que pour tout $n \in \mathbb{N}$, F_n est l'image réciproque d'un fermé par une fonction continue, c'est donc un fermé dans X. De plus, pour tout $n \in \mathbb{N}$, si $x \in \mathrm{F}_{n+1}$, alors $|f_{n+1}(x) - f(x)| \geq \varepsilon$, *i.e.* $f(x) - \varepsilon \geq f_{n+1}(x)$ mais $f_{n+1}(x) \geq f_n(x)$, donc $f(x) - \varepsilon \geq f_n(x)$ et finalement $x \in \mathrm{F}_n$. La suite (F_n) est donc décroissante.

→ S'il existe $n_\varepsilon \in \mathbb{N}$ tel que $\mathrm{F}_{n_\varepsilon} = \varnothing$, alors pour tout $n \geq n_\varepsilon$, $\mathrm{F}_n = \varnothing$ et alors $\forall n \geq n_\varepsilon,\ \forall x \in \mathrm{X},\ |f_n(x) - f(x)| < \varepsilon$.

→ Sinon, toutes les hypothèses de la propriété précédente sont vérifiées, on peut donc affirmer que

(cc) BY-NC-SA

$\bigcap_{n\in\mathbb{N}} F_n \neq \varnothing$. Soit $x \in \bigcap_{n\in\mathbb{N}} F_n$. On a pour tout $n \in \mathbb{N}$, $|f_n(x) - f(x)| \geq \varepsilon$, donc $(f_n(x))$ ne converge pas vers $f(x)$ ce qui est absurde.

On peut donc affirmer l'existence de $n_\varepsilon \in \mathbb{N}$ tel que pour tout $n \geq n_\varepsilon$ et pour tout $x \in X$, $|f_n(x) - f(x)| < \varepsilon$, *i.e.* (f_n) converge uniformément vers f.

Exercice VII.5.

Montrer que $[0,1]^{\mathbb{N}}$ muni de la distance $d : ((x_n),(y_n)) \longmapsto \sum_{k=0}^{+\infty} \frac{|x_k - y_k|}{2^k}$ est compact.

Exercice VII.6.

Soit $\sum z_n$ une série complexe absolument convergente. On note

$$K = \left\{ z \in \mathbb{C},\ \exists A \in \mathcal{P}(\mathbb{N}),\ z = \sum_{n\in A} z_n \right\}.$$

En utilisant l'exercice précédent, montrer que K est compact.

(cc) BY-NC-SA

Correction de l'exercice IV.3. :

Nous allons montrer que (u_n) est à valeurs dans un compact et admet au plus une valeur d'adhérence. Une fois cela fait, on peut directement appliquer la propriété précédente pour affirmer que (u_n) converge. Posons $\mathrm{A} = \{a_1, \ldots, a_p\}$ l'ensemble des valeurs d'adhérence de (u_n) avec $p \in \mathbb{N}^*$ (on peut dire que $p \geq 1$ car (u_n) est à valeurs dans un compact non vide).

Supposons que $|\mathrm{A}| > 1$. Posons $\varepsilon = \frac{1}{3} \min\limits_{i,j \in [\![1;p]\!], i<j} \|a_i - a_j\| > 0$. Pour tout $i \in [\![1;p]\!]$, l'ensemble

$$\mathrm{C} = \left\{ n \in \mathbb{N},\ u_n \in \mathrm{X} \setminus \bigcup_{i=1}^{p} \mathrm{B}(a_i, \varepsilon) \right\}$$

est fini et donc borné.

Soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $n \notin \mathrm{C}$ et $\|u_{n+1} - u_n\| < \varepsilon$. Soit $n_1 \geq \mathrm{N}$ tel que $u_{n_1} \in \mathrm{B}(a_1, \varepsilon)$, $n_2 \geq n_1$ tel que $u_{n_2} \notin \mathrm{B}(a_1, \varepsilon)$ et $k = \max\{\ell \in [\![n_1; n_2]\!],\ u_\ell \in \mathrm{B}(a_1, \varepsilon)\}$. On a par construction $u_k \in \mathrm{B}(a_1, \varepsilon)$ et $\exists i \neq 1\ u_{k+1} \in \mathrm{B}(a_i, \varepsilon)$. On a donc

$$\begin{aligned}\|u_k - u_{k+1}\| &= \|u_k - a_1 + a_1 - a_i + a_i - u_{k+1}\| \\ &\geq \|a_1 - a_i\| - \|u_k - a_1\| - \|u_{k+1} - a_i\| \\ &\geq 3\varepsilon - \varepsilon - \varepsilon = \varepsilon,\end{aligned}$$

ce qui est absurde car $k \geq \mathrm{N}$, donc $|\mathrm{A}| \leq 1$.

Considérons M un majorant de $(\|u_n\|)$. (u_n) est à valeurs dans $\mathrm{B}_f(0, \mathrm{M}+1)$ qui est un fermé borné dans un $\mathbb{R}-$espace vectoriel de dimension finie. C'est donc un compact, ce qui nous permet de conclure.

Correction de l'exercice IV.4. :

Posons pour tout $p \in \mathbb{N}$, $\varepsilon_p = d(u_p, \mathrm{Adh}(u_n))$. Supposons que (ε_p) ne tend pas vers 0. Soit $\delta > 0$ tel que l'ensemble $\mathrm{A} = \{p \in \mathbb{N},\ \varepsilon_p > \delta\}$ est infini et (u_n) est à valeurs dans un compact, on peut donc considérer une extractrice φ à valeurs dans A telle que $(u_{\varphi(p)})$ converge vers un certain $b \in \mathrm{Adh}(u_n)$. Pour tout $p \in \mathbb{N}$, $d(u_{\varphi(p)}, \mathrm{Adh}(u_n)) > \delta$ et alors $\delta < d(u_{\varphi(p)}, b) \xrightarrow[p \to +\infty]{} 0$ ce qui est absurde.

Correction de l'exercice IV.5. :

Montrons successivement les inclusions dans les deux sens.

- → $f(\mathrm{A}) \subset \mathrm{A}$
  Soit $a \in \mathrm{A}$. Il existe une extractrice φ telle que $u_{\varphi(n)} \xrightarrow[n \to +\infty]{} a$. On a alors $u_{\varphi(n)+1} = f(u_{\varphi(n)}) \xrightarrow[n \to +\infty]{} f(a)$. $f(a)$ est donc une valeur d'adhérence de (u_n), *i.e.* $f(a) \in \mathrm{A}$. On conclut donc que $f(\mathrm{A}) \subset \mathrm{A}$.
- → $\mathrm{A} \subset f(\mathrm{A})$
  On reprend les notations du point précédent.
  On considère l'extractrice $\varphi' : n \longmapsto \varphi(n) - 1$. Soit ψ une extractrice telle que $(u_{\varphi' \circ \psi(n)})_{n \geq 1}$ converge vers un certain $b \in \mathrm{A}$. On a alors $f(u_{\varphi' \circ \psi(n)}) = f(u_{\varphi \circ \psi(n)-1}) = u_{\varphi \circ \psi(n)}$. De plus, on a $f(u_{\varphi' \circ \psi(n)}) \xrightarrow[n \to +\infty]{} f(b)$ et $u_{\varphi \circ \psi(n)} \xrightarrow[n \to +\infty]{} a$. On a donc finalement $a = f(b) \in f(\mathrm{A})$ et alors $\mathrm{A} \subset f(\mathrm{A})$.

Correction de l'exercice IV.6. :

1. Soit $(w_n) = ((x_n, f(x_n)))_{n \in \mathbb{N}}$ une suite convergente à valeurs dans Γ_f. On pose (x, y) la limite de cette suite. On a $x_n \xrightarrow[n \to +\infty]{} x$, donc $f(x_n) \xrightarrow[n \to +\infty]{} f(x)$. Mais on a aussi $f(x_n) \xrightarrow[n \to +\infty]{} y$, donc $(x, y) = (x, f(x)) \in \Gamma_f$. La limite de (w_n) est dans Γ_f ce qui nous permet d'affirmer que Γ_f est fermé.

(cc) BY-NC-SA

2. Considérons la fonction f définie par

$$f : \begin{cases} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto \begin{cases} \frac{1}{x} \text{ si } x \neq 0 \\ 0 \text{ sinon.} \end{cases} \end{cases}$$

Posons $g : (x, y) \longmapsto xy - 1$. On a alors

$$\Gamma_f = g^{-1}(\{0\}) \cup \{(0, 0)\}.$$

$g^{-1}(\{0\})$ est l'image réciproque d'un fermé par une fonction continue, c'est donc un fermé. $\{(0, 0)\}$ étant aussi un fermé, on en déduit que Γ_f est fermé dans $\mathrm{X} \times \mathrm{Y}$.

3. Soit $x \in \mathrm{X}$. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans X qui converge vers x. Considérons la suite $(w_n)_{n\in\mathbb{N}}$ à valeurs dans Γ_f définie par $(w_n) = (x_n, f(x_n))_{n\in\mathbb{N}}$. Par hypothèse, Y est compact, il existe donc une extractrice φ telle que $(f(x_{\varphi(n)}))$ converge dans Y et donc $(w_{\varphi(n)})_{n\in\mathbb{N}}$ converge dans Γ_f car c'est un fermé, *i.e.* il existe $y \in \mathrm{X}$ tel que $w_{\varphi(n)} \xrightarrow[n\to+\infty]{} (y, f(y))$. On a alors $x_n \xrightarrow[n\to+\infty]{} y$ et $f(x_{\varphi(n)}) \xrightarrow[n\to+\infty]{} f(y)$. Or $x_{\varphi(n)} \xrightarrow[n\to+\infty]{} x$, donc $y = x$ et $f(x_{\varphi(n)}) \xrightarrow[n\to+\infty]{} f(x)$. On a donc montré que $(f(x_n))$ admet une unique valeur d'adhérence $f(x)$ et est à valeurs dans le compact Y, et par conséquent converge vers $f(x)$. D'où la continuité de f.

Correction de l'exercice V.2. :

Soit $x \in [0, 1]$ et $\varepsilon > 0$. f est continue sur le compact $[0, 1]^2$, donc d'après le théorème de Heine, f est uniformément continue. On dispose donc de $\eta > 0$ tel que

$$\forall (x, y), (x', y') \in [0, 1]^2, \ \ \|(x, y) - (x', y')\| < \eta \Longrightarrow |f(x, y) - f(x', y')| < \varepsilon.$$

Soit x' tel que $|x - x'| < \eta$, on a alors pour tout $y \in [0, 1]$, $\|(x, y) - (x', y)\| = |x - x'| < \eta$ et donc

$$|f(x, y) - f(x', y)| < \varepsilon$$

i.e.

$$f(x', y) - \varepsilon < f(x, y) < f(x', y) + \varepsilon.$$

En passant à la borne supérieure de chaque côté de l'inégalité, on obtient

$$\varphi(x') - \varepsilon \leq \varphi(x) \leq \varphi(x') + \varepsilon$$

i.e.

$$|\varphi(x) - \varphi(x')| \leq \varepsilon,$$

d'où la continuité de φ.

Correction de l'exercice VI.3. :

→ **Méthode 1 :** Soit F un fermé de X. X étant compact, on peut affirmer que F est compact. f est continue, donc $f(\mathrm{F})$ est compact donc fermé. Or on a $(f^{-1})^{-1}(\mathrm{F}) = f(\mathrm{F})$, donc $(f^{-1})^{-1}(\mathrm{F})$ est fermé. La caractérisation des fonctions continues par image réciproque des fermés nous permet de conclure.

→ **Méthode 2 :** On va utiliser l'exercice IV.6. f est continue, donc Γ_f est fermé dans $\mathrm{X} \times \mathrm{Y}$. On considère l'application

$$i : \begin{cases} \mathrm{X} \times \mathrm{Y} & \longrightarrow \mathrm{Y} \times \mathrm{X} \\ (x, y) & \longmapsto (y, x). \end{cases}$$

(cc) BY-NC-SA

On a $\Gamma_{f^{-1}} = i(\Gamma_f)$ et i envoie les fermés sur des fermés (démonstration laissée en exercice), donc $\Gamma_{f^{-1}}$ est fermé dans $\mathrm{Y} \times \mathrm{X}$. X étant compact, l'exercice IV.6. nous donne que f^{-1} est continue.

→ **Méthode 3 :** Considérons (y_n) une suite à valeurs dans Y qui converge vers b. On pose $a = f^{-1}(b)$ et $(x_n) = (f^{-1}(y_n))_{n\in\mathbb{N}}$. Pour montrer que f^{-1} est continue, il faut montrer que $x_n \xrightarrow[n\to+\infty]{} a$.
(x_n) est à valeurs dans le compact X, on peut donc considérer a' une valeur d'adhérence de (x_n). Il existe une extractrice φ telle que $x_{\varphi(a)} \xrightarrow[n\to+\infty]{} a'$. On a par continuité de f, $f(x_{\varphi(n)}) \xrightarrow[n\to+\infty]{} f(a')$, mais on a aussi $f(x_{\varphi(n)}) = y_n \xrightarrow[n\to+\infty]{} b$, donc $f(a') = b$, *i.e.* $a' = f^{-1}(b) = a$. On a alors montré que (x_n) admet une unique valeur d'adhérence qui est a. (x_n) étant à valeurs dans un compact, on peut déduire que $x_n \xrightarrow[n\to+\infty]{} a$, d'où le résultat voulu.

Correction de l'exercice VI.4. :

Posons $\mathrm{A} = \{x \in [0,1],\ g(x) = x\}$.

→ A est non vide.
En effet, en considérant $h = g - \mathrm{Id}$, on a $h(0) = g(0) \geq 0$ et $h(1) = g(1) - 1 \leq 0$. h étant continue, le théorème des valeurs intermédiaires nous permet d'affirmer qu'il existe $a \in [0,1]$ tel que $h(a) = 0$, *i.e.* $g(a) = a$. A est donc non vide

→ A est fermé, car il est égal à $h^{-1}(\{0\})$, qui est l'image réciproque d'un fermé par une fonction continue. A est un fermé inclus dans le compact $[0,1]$, il est donc compact et admet alors un minimum $\alpha = \min \mathrm{A}$ et un maximum $\beta = \max \mathrm{A}$ tel que $g(\alpha) = \alpha$ et $g(\beta) = \beta$. Or, $f(g(\alpha)) = f(\alpha) = g(f(\alpha))$. $f(\alpha)$ est donc un point fixe de g, *i.e.* $f(\alpha) \in \mathrm{A}$. De même, on a $f(\beta) \in \mathrm{A}$.
Posons $w = f - g$. On a $w(\alpha) = f(\alpha) - g(\alpha) = f(\alpha) - \alpha \geq 0$ et $w(\beta) = f(\beta) - g(\beta) = f(\beta) - \beta \leq 0$. w étant continue, le théorème des valeurs intermédiaires nous permet donc d'affirmer qu'il existe $a \in [\alpha, \beta]$ tel que $w(a) = 0$, *i.e.* $f(a) = g(a)$.

Correction de l'exercice VI.5. :

1. Cette proposition est vraie.
Soit $a \in \mathrm{A}$, $b \in \mathrm{B}$, et $\varepsilon > 0$ tel que $\mathrm{B}(a,\varepsilon) \subset \mathrm{A}$ ou $\mathrm{B}(b,\varepsilon) \subset \mathrm{B}$. Soit $x \in \mathrm{B}(a+b,\varepsilon)$. On peut écrire x comme $x = a + b + u$ tel que $\|u\| \leq \varepsilon$. On a alors $a + u \in \mathrm{B}(a,\varepsilon) \subset \mathrm{A}$ ou $b + u \in \mathrm{B}(b,\varepsilon)$, donc $x \in \mathrm{A}+\mathrm{B}$ et finalement $\mathrm{B}(a+b,\varepsilon) \subset \mathrm{A}+\mathrm{B}$. On en déduit donc que $\mathrm{A}+\mathrm{B}$ est ouvert.
2. Cette proposition est vraie.
Soit $(w_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathrm{A}+\mathrm{B}$. On peut écrire $(w_n) = (a_n + b_n)_{n\in\mathbb{N}}$ avec (a_n) et (b_n) deux suites à valeurs respectivement dans A et B. A est compact, il existe donc une extractrice φ et $a \in \mathrm{A}$ tel que $a_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$. De même B étant compact et $(b_{\varphi(n)})$ étant à valeurs dans B, il existe une extractrice ψ et $b \in \mathrm{B}$ tel que $b_{\psi\circ\varphi(n)} \xrightarrow[n\to+\infty]{} b$. On a alors $a_{\psi\circ\varphi(n)} + b_{\psi\circ\varphi(n)} \xrightarrow[n\to+\infty]{} a + b$, d'où la compacité de $\mathrm{A}+\mathrm{B}$.
3. Cette proposition est vraie.
Soit $(c_n)_{n\in\mathbb{N}} = (a_n + b_n)$ une suite à valeurs dans $\mathrm{A}+\mathrm{B}$ qui converge vers c. A étant compact, on peut considérer une extractrice φ telle que $a_{\varphi(n)} \xrightarrow[n\to+\infty]{} a \in \mathrm{A}$. On a alors $b_{\varphi(n)} = c_{\varphi(n)} - a_{\varphi(n)} \xrightarrow[n\to+\infty]{} c - a$. B est fermé et $b_{\varphi(n)}$ est à valeurs dans B, ce qui nous permet de dire que $c - a \in \mathrm{B}$. On en déduit que $c = \underbrace{a}_{\in \mathrm{A}} + \underbrace{c-a}_{\in \mathrm{B}} \in \mathrm{A}+\mathrm{B}$, donc $\mathrm{A}+\mathrm{B}$ est fermé.
4. Cette proposition est fausse.
En effet, il suffit de considérer le cas $\mathrm{A} = \mathbb{Z}$ et $\mathrm{B} = \sqrt{2}\mathbb{Z}$. $\mathrm{A}+\mathrm{B} = \mathbb{Z} + \sqrt{2}\mathbb{Z}$ est un sous groupe de $\mathbb{R}$, donc d'après l'exercice VI.4. du chapitre 11.2, il est soit égal à $a\mathbb{Z}$ avec $a \in \mathbb{R}$ ou alors il est dense dans $\mathbb{R}$. Le premier cas n'est pas possible (il est facile de montrer que cela implique que $\sqrt{2}$ est rationnel), donc $\mathbb{Z} + \sqrt{2}\mathbb{Z}$ est dense dans $\mathbb{R}$. Pour résumer, on a
   → $\mathrm{A} = \mathbb{Z}$ fermé dans $\mathbb{R}$.

BY-NC-SA

→ $\mathrm{B} = \sqrt{2}\mathbb{Z}$ fermé dans $\mathbb{R}$.

→ $\overline{\mathrm{A}+\mathrm{B}} = \overline{\mathrm{Z}+\sqrt{2}\mathbb{Z}} = \mathbb{R} \neq \mathrm{A}+\mathrm{B}$, donc $\mathrm{A}+\mathrm{B}$ n'est pas fermé dans $\mathbb{R}$.

Correction de l'exercice VI.6. :

1. Existence et unicité du point fixe.

→ Existence
On considère la fonction $\varphi : x \longmapsto d(x, f(x))$.

- Montrons que φ est continue.
Soit $x, y \in \mathrm{X}$. On a

$$\begin{aligned}|\varphi(x)-\varphi(y)| &\leq |d(x,f(x))-d(y,f(x))| + |d(y,f(x))-d(y,f(y))| \\ &\leq d(x,y)+d(f(x),f(y)) \\ &\leq 2d(x,y).\end{aligned}$$

φ est lipschitzienne et donc continue.

- Montrons que φ atteint 0.
X est compact donc φ y atteint son minimum, *i.e.* il existe $a \in \mathrm{X}$ tel que $\varphi(\ell) = \inf_{x\in\mathrm{X}} \varphi(x)$.
Supposons que $\ell \neq f(\ell)$, *i.e.* $\varphi(\ell) \neq 0$.
On a alors

$$\varphi(f(\ell)) = d(f(\ell), f(f(\ell))) < d(\ell, f(\ell)) = \varphi(\ell),$$

ce qui est absurde par construction. On en déduit donc que $\varphi(\ell) = 0$, *i.e.* $f(\ell) = \ell$ et donc que f admet un point fixe.

→ Unicité
Soit a et a' deux points fixes de f qu'on suppose distincts. On a $d(a,a') = d(f(a), f(a')) < d(a,a')$ ce qui est absurde, d'où l'unicité du point fixe.

2. Considérons la suite $(\varepsilon_n) = (d(u_n, a))_{n\in\mathbb{N}}$. On a pour tout $n \in \mathbb{N}$,

$$\varepsilon_{n+1} = d(u_{n+1}, a) = d(f(u_n), f(a)) \leq d(u_n, a) = \varepsilon_n.$$

(ε_n) est donc décroissante minorée par 0, elle est alors convergente. Notons α sa limite.

→ Si $\alpha = 0$, alors $u_n \xrightarrow[n\to+\infty]{} \ell$ ce qui est bien le résultat voulu.

→ Supposons que $\alpha > 0$.
X étant compact, on peut considérer a une valeur d'adhérence de u_n. Soit φ une extractrice telle que $u_{\varphi(n)} \xrightarrow[n\to+\infty]{} a$. Par continuité de $u \longmapsto d(\ell, u)$, $d(u_{\varphi(n)}, \ell) \xrightarrow[n\to+\infty]{} d(a,\ell)$, donc $d(a,\ell) = \alpha$.
Or, $u_{\varphi(n)+1} = f(u_{\varphi(n)}) \xrightarrow[n\to+\infty]{} f(a)$, donc $f(a)$ est aussi une valeur d'adhérence de (u_n). Par un raisonnement similaire à ce qui précède, $f(a)$ étant une valeur d'adhérence de (u_n), on peut affirmer que $d(f(a), \ell) = \alpha$. On a alors

$$\alpha = d(f(a), \ell) = d(f(a), f(\ell)) < d(a, \ell) = \alpha,$$

ce qui est absurde, donc a bien $\alpha = 0$, d'où le résultat voulu. La dernière inégalité provient du fait que $a \neq \ell$ car $f(a) = f(\ell)$ entrainerait $\alpha = 0$ ce qui est faux par hypothèse.

Correction de l'exercice VII.2. :

1. Posons

$$\mathrm{A} = \{n \in \mathbb{N},\ \exists (y_1, \ldots, y_n) \in \mathrm{X}^n \text{ vérifiant } (*)\}.$$

(cc) BY-NC-SA

Il est facile de voir que $1 \in \mathrm{A}$. En utilisant la caractérisation d'un compact par recouvrement fini d'ouverts (proposition VII.1), on dispose de $p \in \mathbb{N}^*$ tel que

$$\mathrm{X} = \bigcup_{k=1}^{p} \mathrm{B}\left(a_k, \frac{\varepsilon}{3}\right).$$

Soit $n > p$ et $y_1, \dots, y_n \in \mathrm{X}^n$. Par le principe des tiroirs, il existe $(k, i, j) \in \mathbb{N}^{*3}$, tel que $i \neq j$, $y_i \in \mathrm{B}\left(a_k, \frac{\varepsilon}{3}\right)$ et $y_j \in \mathrm{B}\left(a_k, \frac{\varepsilon}{3}\right)$. On a alors

$$d(y_i, y_j) \leq d(y_i, a_k) + d(a_k, y_j) \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3} < \varepsilon.$$

A est donc une partie de $\mathbb{N}$ majorée par p, elle admet donc un maximum qu'on note N. N vérifie bien la propriété voulue.

2. **Rappel :** $f : \mathrm{X} \longrightarrow \mathrm{X}$ est une isométrie si $\forall (x, y) \in \mathrm{X}^2,\ d(f(x), f(y)) = d(x, y)$.
On pose $\mathrm{N} = \max \mathrm{A}$ et on considère $x_1, \dots, x_{\mathrm{N}}$ vérifiant $(*)$. Posons pour tout $i \in [\![1; \mathrm{N}]\!]$, $f(x_i) = y_i$. f étant une isométrie, on a pour tout $i \neq j$, $d(y_i, y_j) \geq d(x_i, x_j) \geq \varepsilon$, *i.e.* $y_1, \dots, y_{\mathrm{N}}$ vérifie $(*)$.

$\rightarrow$ Montrons que $\mathrm{X} = \bigcup\limits_{k=1}^{\mathrm{N}} \mathrm{B}(y_k, \varepsilon)$.

Supposons que $\mathrm{X} \setminus \bigcup\limits_{k=1}^{\mathrm{N}} \mathrm{B}(y_k, \varepsilon) \neq \varnothing$. On dispose alors de $z \in \mathrm{X} \setminus \bigcup\limits_{k=1}^{\mathrm{N}} \mathrm{B}(y_k, \varepsilon)$. Par construction, $(y_1, \dots, y_{\mathrm{N}}, z)$ est une famille de cardinal $\mathrm{N}+1$ qui vérifie $(*)$ ce qui est absurde. On en déduit donc qu'on a bien $\mathrm{X} = \bigcup\limits_{k=1}^{\mathrm{N}} \mathrm{B}(y_k, \varepsilon)$.

$\rightarrow$ Montrons finalement que f est surjective.

On a $\{y_1, \dots, y_{\mathrm{N}}\} \subset f(\mathrm{X})$ et $\mathrm{X} = \bigcup\limits_{k=1}^{\mathrm{N}} \mathrm{B}(y_k, \varepsilon)$, donc pour tout $z \in \mathrm{X}$,

$$d(z, f(\mathrm{X})) \leq d(z, \{y_1, \dots, y_{\mathrm{N}}\}) \leq \varepsilon.$$

Cette inégalité est valable pour tout ε, ce qui nous donne que pour tout $z \in \mathrm{X}$, $d(z, f(\mathrm{X})) = 0$, *i.e.* $z \in \overline{f(\mathrm{X})}$. X est compact et f est continue, car c'est une isométrie, on a donc que $f(\mathrm{X})$ est compact, donc $\overline{f(\mathrm{X})} = f(\mathrm{X})$.
En résumé, on a $\mathrm{X} \subset \overline{f(\mathrm{X})} = f(\mathrm{X})$ ce qui nous donne que $f(\mathrm{X}) = \mathrm{X}$, *i.e.* f est surjective.

Correction de l'exercice VII.5. :

Soit $((x_{k,n})_{n\in\mathbb{N}})_{k\in\mathbb{N}}$ une suite à valeurs dans $[0,1]^{\mathbb{N}}$.
On va définir par récurrence une suite d'extractrices $(\varphi_n)_{n\in\mathbb{N}}$ et une suite $(x_n)_{n\in\mathbb{N}}$ de la manière suivante.

- $\rightarrow$ φ_0 est une extractrice telle que $x_{\varphi_0(k),0} \xrightarrow[k\to+\infty]{} x_0 \in [0,1]$. Cette extractrice existe car la suite $(x_{k,0})_{k\in\mathbb{N}}$ est à valeurs dans le compact $[0,1]$.
- $\rightarrow$ Soit $n \in \mathbb{N}$, on suppose que φ_n est bien définie, on définit alors φ_{n+1} comme une extractrice telle que la suite $x_{\varphi_0\circ\cdots\circ\varphi_n\circ\varphi_{n+1}(k),n+1} \xrightarrow[n\to+\infty]{} x_{n+1} \in [0,1]$. Encore une fois, cette extractrice existe car la suite $(x_{\varphi_0\circ\cdots\circ\varphi_n(k),n+1})_{k\in\mathbb{N}}$ est à valeurs dans le compact $[0,1]$.

Considérons à présent l'application

$$\varphi : \begin{cases} \mathbb{N} & \longrightarrow \mathbb{N} \\ n & \longmapsto \varphi_0 \circ \cdots \circ \varphi_n(n). \end{cases}$$

Montrons qu'il s'agit bien d'une extractrice. Pour cela, il faut montrer que φ est strictement croissante.

(cc) BY-NC-SA

Soit $\ell, k \in \mathbb{N}$ tel que $\ell < k$. On a

$$\varphi(k) = \varphi_0 \circ \cdots \circ \varphi_\ell \circ \underbrace{\varphi_{\ell+1} \circ \cdots \circ \varphi_k}_{\psi}(k).$$

ψ est une extractrice car c'est une composition d'extractrices, on a alors $\psi(k) \geq k > \ell$. De plus $\varphi_0 \circ \cdots \circ \varphi_\ell$ est strictement croissante car c'est une composition de fonctions strictement croissantes, on a alors

$$\varphi(k) = \varphi_0 \circ \cdots \circ \varphi_\ell \circ \psi(k) \geq \varphi_0 \circ \cdots \circ \varphi_\ell(k) > \varphi_0 \circ \cdots \circ \varphi_\ell(\ell) = \varphi(\ell),$$

donc φ est bien une extractrice.
Montrons à présent que pour la distance d, on a $(x_{\varphi(k),n})_{n\in\mathbb{N}} \xrightarrow[k\to+\infty]{} (x_n)_{n\in\mathbb{N}}$.
Soit $\varepsilon > 0$. On a pour tout $n \in \mathbb{N}$, $x_{\varphi(k),n} \xrightarrow[k\to+\infty]{} x_n$, *i.e.*

$$\forall n \in \mathbb{N},\ \exists \mathrm{K}_n \in \mathbb{N},\ \forall k \geq \mathrm{K}_n,\ \left|x_{\varphi(k),n} - x_n\right| \leq \varepsilon.$$

Soit $\mathrm{N} \in \mathbb{N}$ tel que $\dfrac{1}{2^{\mathrm{N}}} \leq \varepsilon$ et soit $\mathrm{K} = \max\{\mathrm{K}_1, \ldots, \mathrm{K}_{\mathrm{N}}\}$. On a pour tout $k \geq \mathrm{K}$,

$$\begin{aligned}
d((x_{\varphi(k),n})_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}}) &= \sum_{n=0}^{\infty} \frac{\left|x_{\varphi(k),n} - x_n\right|}{2^n} \\
&= \sum_{n=0}^{\mathrm{N}-1} \frac{\left|x_{\varphi(k),n} - x_n\right|}{2^n} + \sum_{n=\mathrm{N}}^{\infty} \frac{\left|x_{\varphi(k),n} - x_n\right|}{2^n} \\
&\leq \sum_{n=0}^{\mathrm{N}-1} \frac{\varepsilon}{2^n} + \sum_{n=\mathrm{N}}^{\infty} \frac{2}{2^n} \\
&\leq 2\varepsilon + \frac{4}{2^{\mathrm{N}}} \leq 6\varepsilon.
\end{aligned}$$

On a donc trouvé une suite extraite de $((x_{k,n})_{n\in\mathbb{N}})_{k\in\mathbb{N}}$ qui converge vers une suite $(x_n)_{n\in\mathbb{N}}$ dans $[0,1]^{\mathbb{N}}$ pour la distance d, ce qui montre bien que $[0,1]^{\mathbb{N}}$ est compact pour cette distance.

Correction de l'exercice VII.6. :

On pose $\mathrm{X}_c = \{0,1\}^{\mathbb{N}}$. Une démonstration quasi identique à celle de l'exercice précédent permet d'affirmer que X_c muni de la distance d vue dans cet exercice est compact.
Pour toute partie A de $\mathbb{N}$, on considère la fonction

$$\mathbb{1}_{\mathrm{A}} : x \longmapsto \begin{cases} 1 \text{ si } x \in \mathrm{A} \\ 0 \text{ sinon.} \end{cases}$$

On considère également l'application

$$f : \begin{cases} \mathrm{X}_c & \longrightarrow \mathrm{K} \\ (\varepsilon_n) & \longmapsto \displaystyle\sum_{n=0}^{\infty} \varepsilon_n z_n. \end{cases}$$

Cette application est définie correctement et est surjective (pour la surjectivité, il suffit d'utiliser les suites de la forme $(\varepsilon_n) = (\mathbb{1}_{\mathrm{A}}(n))_{n\in\mathbb{N}}$).
Si on montre que f est continue, alors $f(\mathrm{X}_c) = \mathrm{K}$ est compact car il s'agit de l'image d'un compact par une application continue. Montrons donc que f est continue.

(cc) BY-NC-SA

Soit $\varepsilon > 0$ et $(\varepsilon_n)_{n\in\mathbb{N}} \in \mathrm{X}_c$. Pour tout $\eta > 0$ et $(\gamma_n) \in \mathrm{X}_c$, on a

$$d((\varepsilon_n), (\gamma_n)) < \eta \Longrightarrow \forall n \in \mathbb{N},\ \ |\varepsilon_n - \gamma_n| < 2^n\eta,$$

donc pour tout $\mathrm{N} \in \mathbb{N}$, lorsque $\eta < \dfrac{1}{2^{\mathrm{N}}}$, on a pour tout $n \in [\![0;\mathrm{N}]\!]$, $|\varepsilon_n - \gamma_n| < 2^n\eta < 1$, *i.e.* $\varepsilon_n = \gamma_n$.
On a alors pour tout $(\gamma_n)_{n\in\mathbb{N}}$ et $\mathrm{N} \in \mathbb{N}$,

$$d((\varepsilon_n), (\gamma_n)) < \frac{1}{2^{\mathrm{N}}} \Longrightarrow |f((\varepsilon_n)) - f((\gamma_n))| \leq \left|\sum_{k=\mathrm{N}+1}^{\infty} (\varepsilon_k - \gamma_k)z_k\right| \leq \sum_{k=\mathrm{N}+1}^{\infty} |z_k|\,.$$

En prenant $\mathrm{N} \in \mathbb{N}$ tel que $\displaystyle\sum_{k=\mathrm{N}+1}^{\infty} |z_k| < \varepsilon$, on obtient

$$d((\varepsilon_n), (\gamma_n)) < \frac{1}{2^{\mathrm{N}}} \Longrightarrow |f((\varepsilon_n)) - f((\gamma_n))| \leq \sum_{k=\mathrm{N}+1}^{\infty} |z_k| \leq \varepsilon,$$

ce qui montre la continuité de f et par conséquent la compacité de K.

(cc) BY-NC-SA

Chapitre 11.6

Applications linéaires continues

I Quelques propriétés

Dans ce chapitre, on considère deux $\mathbb{K}-$espaces vectoriels de dimension quelconque E et F munis respectivement des normes $\|.\|_{\mathrm{E}}$ et $\|.\|_{\mathrm{F}}$, où $\mathbb{K}$ est égal à $\mathbb{R}$ ou $\mathbb{C}$.

Proposition I.1.

Soit $u \in \mathcal{L}(\mathrm{E},\mathrm{F})$. On a

$$u \text{ est continue } \iff \exists \mathrm{C} > 0,\ \forall x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} \leq \mathrm{C}\,\|x\|_{\mathrm{E}}\,.$$

Démonstration.

($\Leftarrow$) On a pour tout $y \in \mathrm{E}$,

$$\|u(x) - u(y)\| = \|u(x-y)\| \leq \mathrm{C}\,\|x-y\| \xrightarrow[x\to y]{} 0,$$

donc u est bien continue.

($\Rightarrow$) Pour cette implication, on va uniquement utiliser le fait que u est continue en 0.
Montrons que u est bornée sur $\mathrm{B}_f(0,1)$. u étant continue en 0, il existe $r > 0$ tel que

$$\forall x \in \mathrm{B}_f(0,r),\ \|u(x)\|_{\mathrm{F}} \leq 1.$$

Ici, on a utilisé la définition de la continuité en 0 en prenant $\varepsilon = 1$. Soit $x \in \mathrm{B}_f(0,1)$, on a $rx \in \mathrm{B}_f(0,r)$ et alors $\|u(rx)\|_{\mathrm{F}} \leq 1$. On en déduit donc que

$$\forall x \in \mathrm{B}_f(0,1),\ \|u(x)\|_{\mathrm{F}} \leq \frac{1}{r}.$$

On notant $\mathrm{C} = \frac{1}{r}$, on a pour tout $x \in \mathrm{E}\setminus\{0\}$, $\left\|u\left(\frac{x}{\|x\|_{\mathrm{E}}}\right)\right\|_{\mathrm{F}} \leq \mathrm{C}$, *i.e.*

$$\forall x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} \leq \mathrm{C}\,\|x\|_{\mathrm{E}}\,,$$

d'où le résultat voulu.

Proposition I.2.

Soit $u \in \mathcal{L}(\mathrm{E},\mathrm{F})$. les propositions suivantes sont équivalentes.

1. u est continue.
2. u est continue en 0.
3. u est continue en un point de E.
4. u est lipschitzienne.
5. u est bornée sur $\mathrm{S}(0,1)$.
6. u est bornée sur $\mathrm{B}_f(0,1)$.
7. u est bornée sur une boule non réduite à un point.

BY-NC-SA

Démonstration.

→ Les implications (1) ⇒ (2) ⇒ (3) sont évidentes.

→ (3) ⇒ (4) Si u est continue en un point $a \in \mathrm{E}$, alors $u(x+a) \xrightarrow[x \to 0]{} u(a)$, donc

$$u(x) = u(x+a) - u(a) \xrightarrow[x \to 0]{} 0.$$

u est donc continue en 0 et donc d'après la démonstration de la proposition I.1., u est lipschitzienne.

→ (4) ⇒ (5) si u est lipschitzienne de facteur de lipschitzianité $k \geq 0$, alors on a pour tout $x \in \mathrm{S}(0,1)$,

$$\|u(x)\|_{\mathrm{F}} = \|u(x) - u(0)\|_{\mathrm{F}} \leq k \|x\|_{\mathrm{E}} = k,$$

donc u est bornée sur $\mathrm{S}(0,1)$.

→ (5) ⇒ (6) Soit M un majorant de $x \longmapsto \|u(x)\|_{\mathrm{F}}$ sur $\mathrm{S}(0,1)$. On a pour tout $x \in \mathrm{B}_f(0,1) \setminus \{0\}$,

$$\|u(x)\|_{\mathrm{F}} = \|x\|_{\mathrm{E}} \left\| u\left(\frac{x}{\|x\|_{\mathrm{E}}}\right) \right\|_{\mathrm{F}} \leq \mathrm{M} \|x\|_{\mathrm{E}} \leq \mathrm{M},$$

donc u est bien bornée sur $\mathrm{B}_f(0,1)$.

→ (6) ⇒ (7) Cette implication est évidente.

→ (7) ⇒ (1) Soit $(a,r) \in \mathrm{E} \times \mathbb{R}_+^*$ tel que u est bornée sur $\mathrm{B}_f(a,r)$ et M un majorant de $x \longmapsto \|u(x)\|_{\mathrm{F}}$ sur $\mathrm{B}_f(a,r)$. On a pour tout $x \in \mathrm{B}_f(0,1)$,

$$r \|u(x)\|_{\mathrm{F}} - \|u(a)\|_{\mathrm{F}} \leq \|u(a+rx)\|_f \leq \mathrm{M},$$

et alors

$$\|u(x)\|_{\mathrm{F}} \leq \frac{\mathrm{M} + \|u(a)\|_{\mathrm{F}}}{r} := \mathrm{C}.$$

On en déduit que pour tout $x \in \mathrm{E} \setminus \{0\}$,

$$\|u(x)\|_{\mathrm{F}} = \|x\|_{\mathrm{E}} \left\| u\left(\frac{x}{\|x\|_{\mathrm{E}}}\right) \right\| \leq \mathrm{C} \|x\|_{\mathrm{E}}.$$

La proposition I.1. nous permet donc d'affirmer le fait que u est continue.

Proposition I.3.

Soit $u \in \mathcal{L}(\mathrm{E},\mathrm{F})$ continue. Les trois nombres suivants sont égaux.

1. $\sup\limits_{x \in \mathrm{B}_f(0,1)} \|u(x)\|_{\mathrm{F}}$
2. $\sup\limits_{x \in \mathrm{S}(0,1)} \|u(x)\|_{\mathrm{F}}$
3. $\inf \{\mathrm{C} \geq 0,\ \forall x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} \leq \mathrm{C} \|x\|_{\mathrm{E}}\}$

De plus, lorsque E est de dimension finie, les bornes inférieures (2) et (1) sont atteintes alors que (3) est atteinte peu importe la dimension de E. On appelle ce nombre norme d'opérateur de u et on le note $|||u|||$.

Démonstration.

→ (1) = (2)

- On a $\mathrm{S}(0,1) \subset \mathrm{B}_f(0,1)$, donc (1) ≥ (2).

(cc) BY-NC-SA

- On a pour tout $x \in \mathrm{B}_f(0,1) \setminus \{0\}$,

$$\|u(x)\|_{\mathrm{F}} = \|x\|_{\mathrm{E}} \left\| u\left(\frac{x}{\|x\|_{\mathrm{E}}}\right) \right\| \leq \|x\|_{\mathrm{E}} \sup_{y \in \mathrm{S}(0,1)} \|u(y)\|_{\mathrm{F}} \leq \sup_{x \in \mathrm{S}(0,1)} \|u(x)\|_{\mathrm{F}}.$$

On en déduit donc que (1) $\leq$ (2) et que finalement (1) $=$ (2).

$\to$ (2) $=$ (3) On a

$$\begin{aligned}
(3) &= \inf \{\mathrm{C} \geq 0,\ \forall x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} \leq \mathrm{C}\,\|x\|_{\mathrm{E}}\} \\
&= \inf \left\{\mathrm{C} \geq 0,\ \forall x \in \mathrm{E} \setminus \{0\},\ \left\| u\left(\frac{x}{\|x\|_{\mathrm{E}}}\right) \right\|_{\mathrm{F}} \leq \mathrm{C}\right\} \\
&= \inf \{\mathrm{C} \geq 0,\ \forall x \in \mathrm{S}(0,1),\ \|u(x)\|_{\mathrm{F}} \leq \mathrm{C}\} \\
&= \sup_{x \in \mathrm{S}(0,1)} \|u(x)\|_{\mathrm{F}} = (2).
\end{aligned}$$

Or, $(2) \in \{\mathrm{C} \geq 0,\ \forall x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} \leq \mathrm{C}\,\|x\|_{\mathrm{E}}\}$, donc la borne inférieure (3) est bien atteinte.

Notation. On note $\mathcal{L}_c(\mathrm{E}, \mathrm{F})$ l'ensemble des applications linéaires continues de E dans F.

Proposition I.4.

$\|\!|\cdot|\!\|$ est une norme sur $\mathcal{L}_c(\mathrm{E}, \mathrm{F})$.

Démonstration. La démonstration est laissée comme exercice au lecteur.

Proposition I.5.

Soit G un espace vectoriel normé, $u \in \mathcal{L}_c(\mathrm{E}, \mathrm{F})$ et $v \in \mathcal{L}_c(\mathrm{F}, \mathrm{G})$. On a

$$|\!|\!| v \circ u |\!|\!| \leq |\!|\!| v |\!|\!| \times |\!|\!| u |\!|\!|.$$

Démonstration. On a pour tout $x \in \mathrm{E}$,

$$\|v \circ u(x)\|_{\mathrm{G}} \leq |\!|\!| v |\!|\!| \, \|u(x)\|_{\mathrm{F}} \leq |\!|\!| v |\!|\!| \times |\!|\!| u |\!|\!| \, \|x\|_{\mathrm{E}}.$$

On peut alors dire par définition de la norme d'opérateur que $|\!|\!| v \circ u |\!|\!| \leq |\!|\!| v |\!|\!| \times |\!|\!| u |\!|\!|$.

Remarque. en pratique, pour une application linéaire $u \in \mathcal{L}_c(\mathrm{E}, \mathrm{F})$ donnée, on peut calculer $|\!|\!| u |\!|\!|$ en utilisant l'égalité suivante :

$$|\!|\!| u |\!|\!| = \sup_{x \in \mathrm{E} \setminus \{0\}} \frac{\|u(x)\|_{\mathrm{F}}}{\|x\|_{\mathrm{E}}}.$$

Application. On peut utiliser cette nouvelle définition de la continuité d'une application linéaire pour énoncer une nouvelle formulation des résultats de comparaison de normes.
Soit N_1 et N_2 deux normes sur l'espace vectoriel normé E. On a

$$\begin{aligned}
\mathrm{N}_1 \text{ plus fine que } \mathrm{N}_2 &\iff \exists \mathrm{C} > 0,\ \forall x \in \mathrm{E},\ \mathrm{N}_2(x) \leq \mathrm{C} \cdot \mathrm{N}_1(x) \\
&\iff \text{L'application linéaire } \mathrm{I}d : \begin{cases} (\mathrm{E}, \mathrm{N}_1) & \longrightarrow (\mathrm{E}, \mathrm{N}_2) \\ \quad x & \longmapsto x \end{cases} \text{ est continue.} \\
&\iff \text{Pour tout } \Omega \in \mathcal{P}(\mathrm{E}),\ \Omega \text{ est ouvert dans } (\mathrm{E}, \mathrm{N}_2) \Longrightarrow \mathrm{Id}^{-1}(\Omega) \text{ est ouvert dans } (\mathrm{E}, \mathrm{N}_1).
\end{aligned}$$

(cc) BY-NC-SA

II Exercices

Exercice II.1.

1. On pose $E = \mathcal{C}([0,1], \mathbb{R})$, $a \in [0,1]$ et $u \in \mathcal{L}(E, \mathbb{R})$ définie par $\forall f \in E,\ u(f) = f(a)$. Étudier la continuité de u pour les normes $\|.\|_\infty$, $\|.\|_1$ et $\|.\|_2$.
2. On pose $F = \mathcal{C}^1([0,1], \mathbb{R})$ muni de la norme $\|.\|$ définie par $\forall f \in F$, $\|f\| = \|f\|_\infty + \|f'\|_\infty$. On considère $v \in \mathcal{L}(F, E)$ définie par $\forall f \in F,\ v(f) = f'$. Étudier la continuité de v pour $\|.\|$ et $\|.\|_\infty$ lorsque E est muni de $\|.\|_\infty$.
3. Calculer $\|\|u\|\|$ pour $\|.\|_\infty$.
4. Soit $u \in \mathcal{L}(E, \mathbb{R})$ définie par

$$\forall f \in E,\ u(f) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} f\left(\frac{1}{2^n}\right).$$

Montrer que u est bien définie, linéaire et continue pour $\|.\|_\infty$ mais que $\|\|u\|\|$ n'est pas atteinte.

Exercice II.2.

Soit $u \in \mathcal{L}(E, F)$. Montrer que les deux propositions suivantes sont équivalentes.

1. u est continue.
2. $\{x \in E,\ \|u(x)\|_F = 1\}$ est fermé dans E.

Exercice II.3.

On suppose que $\mathbb{K} = \mathbb{R}$. Soit u un morphisme additif de E dans F borné sur $B_f(0,1)$. Montrer que u est une application linéaire continue.

Exercice II.4.

Soit $u \in \mathcal{L}(E, \mathbb{R})$. Montrer que les propositions suivantes sont équivalentes.

1. u est continue.
2. $\operatorname{Ker} u$ est fermé.

Exercice II.5.

Soit X un espace métrique compact. On considère l'espace vectoriel $E = \mathcal{C}(X, \mathbb{R})$ muni de $\|.\|_\infty$.

1. Soit u une forme linéaire de E dans $\mathbb{R}$. On suppose que

$$\forall f \in E,\ f \geq 0 \Longrightarrow u(f) \geq 0.$$

Montrer que u est continue et trouver $\|\|u\|\|$.
2. Soit $\chi : E \longrightarrow \mathbb{R}$ un morphisme d'anneau unitaire et linéaire, montrer que χ est continu.
3. Montrer qu'il existe $x_0 \in X$ tel que pour tout $f \in E$, $\chi(f) = f(x_0)$.

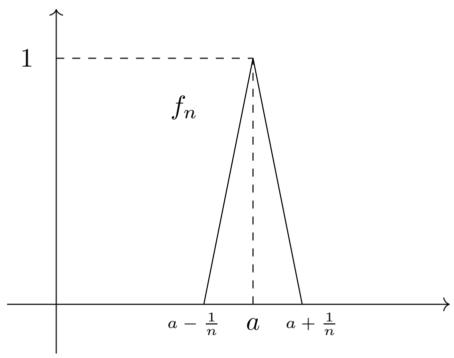
BY-NC-SA

Correction de l'exercice II.1. :

1. Étudions la continuité de u pour les 3 normes de l'énoncé.
   - $\to$ Continuité pour $\|.\|_\infty$.
     On a pour tout $f \in \mathrm{E}$, $|u(f)| = |f(a)| \leq \|f\|_\infty$, donc u est continue pour $\|.\|_\infty$.
   - $\to$ Continuité pour $\|.\|_1$.
     On considère la suite de fonctions (f_n) définie par

     $$\forall n \in \mathbb{N}^*,\ \forall x \in [0,1],\ f_n(x) = \begin{cases} n\left(x - a + \frac{1}{n}\right) & \text{si } x \in \left[a - \frac{1}{n}, a\right] \\ n\left(-x + a + \frac{1}{n}\right) & \text{si } x \in \left[a, a + \frac{1}{n}\right] \\ 0 & \text{sinon.} \end{cases}$$

     Pour comprendre l'intuition derrière cette construction, voici un dessin d'un terme de la suite (f_n).

     On a pour tout $n_0 \in \mathbb{N}^*$ assez grand pour qu'on ait $\left[a - \frac{1}{n}, a + \frac{1}{n}\right] \subset [0,1]$ et $n \geq n_0$,

     $$\frac{|u(f)|}{\|f\|_1} = \frac{|f(a)|}{1/n} = n \xrightarrow[n \to +\infty]{} +\infty.$$

     L'application

     $$\begin{cases} \mathrm{E} \setminus \{0\} & \longrightarrow \mathbb{R} \\ f & \longmapsto \dfrac{u(f)}{\|f\|_1} \end{cases}$$

     est alors non bornée, et alors u n'est pas continue pour $\|.\|_1$.
   - $\to$ Continuité pour $\|.\|_2$.
     On pose $(g_n) = (\sqrt{f_n})_{n \in \mathbb{N}^*}$. On a pour tout $n \geq n_0$,

     $$\frac{|u(f)|}{\|f\|_2} = \frac{1}{\sqrt{1/n}} = \sqrt{n} \xrightarrow[n \to +\infty]{} +\infty,$$

     ce qui nous permet comme avant de conclure que u n'est pas continue pour $\|.\|_2$. La non-continuité pour $\|.\|_2$ aurait aussi pu être démontrée en utilisant le fait que $\|.\|_2$ est plus fine que $\|.\|_1$.
2. Étudions la continuité de v pour les deux normes de l'énoncé.
   - $\to$ Continuité pour $\|.\|$.
     On a pour tout $f \in \mathcal{C}^1([0,1], \mathbb{R})$, $\|v(f)\|_\infty = \|f'\|_\infty \leq \|f\|_\infty + \|f'\|_\infty = \|f\|$. On en déduit donc que u est continue pour $\|.\|$.
   - $\to$ Continuité pour $\|.\|_\infty$.
     On considère la suite (f_n) de fonctions à valeurs dans $\mathcal{C}([0,1], \mathbb{R})$ définie par

     $$\forall n \in \mathbb{N},\ \forall x \in [0,1],\ f_n(x) = e^{nx}.$$

(cc) BY-NC-SA

On a alors pour tout $n \in \mathbb{N}$,
$$\frac{\|v(f_n)\|_\infty}{\|f_n\|_\infty} = n \xrightarrow[n\to+\infty]{} +\infty,$$
ce qui permet de dire que u n'est pas continue pour $\|.\|_\infty$.

3. On a pour tout $f \in \mathrm{E}$, $|u(f)| \leq \|f\|_\infty$, donc $\|\!|u|\!\| \leq 1$. De plus, si on considère f la fonction constante égale à 1 sur $[0,1]$, on a $|u(f)| = \|f\|_\infty = 1$, ce qui nous permet d'affirmer que $\|\!|u|\!\| = 1$.

4. $\to$ Montrons que u est bien définie et linéaire.
Soit $f \in \mathrm{E}$. f est continue donc bornée sur le segment $[0,1]$. Soit M un majorant de $|f|$. On a pour tout $n \in \mathbb{N}$,
$$\left|\frac{(-1)^n}{2^n} f\left(\frac{1}{2^n}\right)\right| \leq \frac{\mathrm{M}}{2^n}.$$
Le terme de gauche est le terme général d'une série convergence, donc la série $\sum \frac{(-1)^n}{2^n} f\left(\frac{1}{2^n}\right)$ est absolument convergente, ce qui nous permet d'affirmer que u est bien définie.
On a aussi pour tout $f, g \in \mathrm{E}$,
$$\begin{aligned} u(f+g) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n}\left(f\left(\frac{1}{2^n}\right) + g\left(\frac{1}{2^n}\right)\right) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} f\left(\frac{1}{2^n}\right) + \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} g\left(\frac{1}{2^n}\right) \\ &= u(f) + u(g), \end{aligned}$$
donc u est bien linéaire.

$\to$ Montrons que u est continue.
On a pour tout $f \in \mathrm{E}$,
$$|u(f)| = \left|\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} f\left(\frac{1}{2^n}\right)\right| \leq \sum_{n=0}^{\infty} \frac{1}{2^n} \|f\|_\infty = 2\|f\|_\infty,$$
donc u est bien continue et $\|\!|u|\!\| \leq 2$.

$\to$ Montrons que $\|\!|u|\!\| = 2$.
Soit $(f_\mathrm{N})_{\mathrm{N}\in\mathbb{N}}$ une suite à valeurs dans $\mathrm{B}_f(0,1)$ telle que pour tout $\mathrm{N} \in \mathbb{N}$, $\forall n \leq \mathrm{N}$, $f_\mathrm{N}\left(\frac{1}{2^n}\right) = (-1)^n$ et $\|f_\mathrm{N}\|_\infty \leq 1$. On a alors pour tout $\mathrm{N} \in \mathbb{N}$,
$$\begin{aligned} |u(f_\mathrm{N})| &= \left|\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} f_\mathrm{N}\left(\frac{1}{2^n}\right)\right| = \left|\sum_{n=0}^{\mathrm{N}} \frac{1}{2^n} + \sum_{n=\mathrm{N}+1}^{\infty} \frac{(-1)^n}{2^n} f_\mathrm{N}\left(\frac{1}{2^n}\right)\right| \\ &= \left|2 - \frac{1}{2^\mathrm{N}} + \sum_{n=\mathrm{N}+1}^{\infty} \frac{(-1)^n}{2^n} f_\mathrm{N}\left(\frac{1}{2^n}\right)\right| \\ &\geq 2 - \frac{1}{2^\mathrm{N}} - \sum_{n=\mathrm{N}+1}^{\infty} \frac{1}{2^n} \|f_\mathrm{N}\|_\infty \\ &\geq 2 - \frac{1}{2^{\mathrm{N}-1}} \xrightarrow[\mathrm{N}\to\infty]{} 2. \end{aligned}$$
Par caractérisation séquentielle de la borne supérieure, on a $\sup\limits_{f\in\mathrm{B}_f(0,1)} |u(f)| = 2$ et alors $\|\!|u|\!\| = 2$.

$\to$ Montrons finalement que $\|\!|u|\!\|$ n'est pas atteinte.
Supposons le contraire. On dispose alors de $f \in \mathrm{E}$ telle que $\|f\|_\infty = 1$ et $|u(f)| = 2$.

(cc) BY-NC-SA

En supposant sans perte de généralité que $u(f) = 2$, on a que

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} f\left(\frac{1}{2^n}\right) = 2 = \sum_{n=0}^{\infty} \frac{1}{2^n},$$

et alors

$$\sum_{n=0}^{\infty} \frac{1}{2^n} \underbrace{\left(1 - (-1)^n f\left(\frac{1}{2^n}\right)\right)}_{\geq 0} = 0.$$

Il s'agit d'une somme de nombres positifs qui est nulle, donc tous les termes de la somme sont nuls. On a alors pour tout $n \in \mathbb{N}$, $f\left(\frac{1}{2^n}\right) = (-1)^n$. f n'admet pas de limite en 0 et n'est donc pas continue en 0, ce qui nous permet d'affirmer que $f \notin \mathrm{E}$, ce qui est absurde.

Correction de l'exercice II.2. :

Posons $\mathrm{K} = \{x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} = 1\}$.

- → (1) ⇒ (2) L'application $f : x \longmapsto \|u(x)\|_{\mathrm{F}}$ est continue, donc $\{x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} = 1\} = f^{-1}(\{1\})$ est fermé car c'est l'image réciproque d'un fermé par une application continue.
- → (2) ⇒ (1) Procédons par contraposée. Supposons que u n'est pas continue, u n'est donc pas bornée sur $\mathrm{S}(0,1)$. On dispose alors d'une suite $(x_n) \in \mathrm{S}(0,1)^{\mathbb{N}}$ telle que $\|u(x_n)\| \xrightarrow[n\to+\infty]{} +\infty$. Soit n_0 assez grand pour qu'on ait pour tout $n \geq n_0$, $\|u(x_n)\| > 0$. On pose alors pour tout $n \geq n_0$, $y_n = \dfrac{x_n}{\|u(x_n)\|}$.
  Pour tout $n \geq n_0$, on a $y_n \in \{x \in \mathrm{E},\ \|u(x)\|_{\mathrm{F}} = 1\}$ et

  $$\|y_n\|_{\mathrm{E}} = \frac{\|x_n\|_{\mathrm{E}}}{\|u(x_n)\|_{\mathrm{F}}} = \frac{1}{\|u(x_n)\|_{\mathrm{F}}} \xrightarrow[n\to+\infty]{} 0,$$

  mais $0 \notin \mathrm{K}$. on a trouvé une suite à valeurs dans K qui converge vers un élément qui n'est pas dans cet ensemble. On en déduit que K n'est pas fermé.

Correction de l'exercice II.3. :

On a tout d'abord

$$\forall n \in \mathbb{Z},\ \forall x \in \mathrm{E},\ f(nx) = nf(x).$$

À partir de cette propriété, on peut facilement en déduire que

$$\forall r \in \mathbb{Q},\ \forall x \in \mathrm{E},\ f(rx) = rf(x).$$

Soit M un majorant de f sur $\mathrm{B}_f(0,1)$. Soit $\lambda \in \mathbb{R}$, et $(r_n) \in \mathbb{Q}^{\mathbb{N}}$ telle que $r_n \xrightarrow[n\to+\infty]{} \lambda$. Soit $x \in \mathrm{E}$. Pour tout $p \in \mathbb{N}^*$ il existe $\mathrm{N}_p \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $(r_n - \lambda)x \in \mathrm{B}\left(0, \frac{1}{p}\right)$. On a alors pour tout $p \in \mathbb{N}$, il existe $\mathrm{N}_p \in \mathbb{N}^*$ tel que pour tout $n \geq \mathrm{N}_p$,

$$\|f((r_n - \lambda)x)\| = \frac{1}{p} \|f(\underbrace{p(r_n - \lambda)x}_{\in \mathrm{B}_f(0,1)})\| \leq \frac{\mathrm{M}}{p}$$

i.e.

$$\|r_n f(x) - f(\lambda x)\| \leq \frac{\mathrm{M}}{p}.$$

On a alors

$$\|\lambda f(x) - f(\lambda x)\| \leq \frac{\mathrm{M}}{p} + \|(r_n - \lambda) f(x)\| \xrightarrow[n\to+\infty]{} \frac{\mathrm{M}}{p}.$$

On en déduit donc que pour tout $p \in \mathbb{N}^*$, $\|\lambda f(x) - f(\lambda x)\| \leq \dfrac{\mathrm{M}}{p}$ et que finalement $f(\lambda x) = \lambda f(x)$. On

(cc) BY-NC-SA

conclut que f est une application linéaire bornée sur $\mathrm{B}_f(0,1)$, donc continue.

Correction de l'exercice II.4. :

→ (1) ⇒ (2) On a $\operatorname{Ker} u = u^{-1}(\{0\})$. Il s'agit de l'image réciproque d'un fermé par une application continue, c'est donc un fermé.

→ (2) ⇒ (1) Si $u=0$, $\operatorname{Ker} u = \mathrm{E}$ est fermé. Supposons donc maintenant que $u \neq 0$. $\mathrm{H} = \operatorname{Ker} u$ est un hyperplan et u est surjective. Il existe donc $a \in \mathrm{E}$ tel que $u(a)=1$. On a alors

$$\forall x \in \mathrm{E},\ u(x) = 1 \Longleftrightarrow u(x-a) = 0 \Longleftrightarrow x \in a + \mathrm{H}.$$

De même, il est facile de voir que

$$\forall x \in \mathrm{E},\ u(x) = -1 \Longleftrightarrow u(x+a) = 0 \Longleftrightarrow x \in -a + \mathrm{H}.$$

Posons donc $\mathrm{F} = (a+\mathrm{H}) \cup (-a+\mathrm{H})$. $a+\mathrm{H}$ et $-a+\mathrm{H}$ sont fermés car ce sont les images directes du fermé H par respectivement $x \longmapsto a+x$ et $x \longmapsto -a+x$ qui sont deux applications linéaires bijectives d'inverse continu, donc F est fermé. Or $0 \notin \mathrm{F}$ et F^c est ouvert, donc il existe $r>0$ tel que $\mathrm{B}_f(0,r) \subset \mathrm{F}^c$.
Montrons que pour tout $x \in \mathrm{B}_f(0,r)$, $|u(x)| \leq 1$. Soit $x \in \mathrm{B}(0,r)$ non nul. On suppose par l'absurde que $|u(x)| > 1$. On a alors $y := \frac{x}{|u(x)|} \in \mathrm{B}(0,r)$ et $u(y) \in \{1,-1\}$, *i.e.* $y \in \mathrm{F}$ ce qui est absurde étant donné que $y \in \mathrm{B}(0,r) \subset \mathrm{F}^c$. On a donc trouvé une boule non réduite à un point où u est bornée, ce qui nous permet d'affirmer que u est continue.

Correction de l'exercice II.5. :

1. On a pour tout $f,g \in \mathrm{E}$,

$$f \geq g \Longrightarrow f-g \geq 0 \Longrightarrow u(f-g) \geq 0 \Longrightarrow u(f) \geq u(g).$$

Posons e la fonction constante égale à 1 sur X. On a pour tout $f \in \mathrm{E}$,

$$-\left\|f\right\|_\infty e \leq f \leq \left\|f\right\|_\infty e.$$

En posant $\mathrm{C} = u(e)$ et en appliquant u de chaque côté de l'inégalité, on obtient

$$-\mathrm{C}\left\|f\right\|_\infty \leq u(f) \leq \mathrm{C}\left\|f\right\|_\infty,$$

et enfin

$$\forall f \in \mathrm{E},\ |u(f)| \leq \mathrm{C}\left\|f\right\|_\infty.$$

Donc u est continue et $|||u||| \leq \mathrm{C}$, mais $u(e) = \mathrm{C}$ et $e \in \mathrm{B}_f(0,1)$, donc $|||u||| \geq \mathrm{C}$, *i.e.* $|||u||| = \mathrm{C}$.

2. Soit $f \in \mathrm{E}$ tel que $f \geq 0$, on pose $g = \sqrt{f}$. On a $g \in \mathrm{E}$ et

$$\chi(f) = \chi(g^2) = \chi(g)^2 \geq 0.$$

Il suffit alors d'appliquer la question 1.

(cc) BY-NC-SA

Chapitre 11.7

Espaces vectoriels normés de dimension finie

Dans tout ce chapitre, $\mathbb{K}$ est égal à $\mathbb{R}$ ou $\mathbb{C}$. On considère E un $\mathbb{K}-$espace vectoriel normé de dimension finie $n \geq 1$ et on pose $(e_1, \ldots, e_n)$ une base de cet espace.

I Équivalence des normes

Proposition I.1.

Dans $\mathbb{R}^n$, toutes les normes sont équivalentes.

Démonstration. Soit N une norme sur $\mathbb{R}^n$. En posant $(\varepsilon_1, \ldots, \varepsilon_n)$ la base canonique de $\mathbb{R}^n$, on a pour tout $x = x_1\varepsilon_1 + \cdots + x_n\varepsilon_n \in \mathbb{R}^n$,

$$\mathrm{N}(x) = \mathrm{N}(x_1\varepsilon_1 + \cdots + x_n\varepsilon_n) \leq \sum_{k=1}^{n} |x_k| \, \mathrm{N}(\varepsilon_k) \leq \left(\sum_{k=1}^{n} \mathrm{N}(\varepsilon_k)\right) \|x\|_\infty .$$

De plus, en posant $\mathrm{C} = \sum_{k=1}^{n} \mathrm{N}(\varepsilon_k)$ pour tout $x, y \in \mathrm{E}$

$$|\mathrm{N}(x) - \mathrm{N}(y)| \leq \mathrm{N}(x-y) \leq \mathrm{C} \, \|x-y\|_\infty .$$

N est donc lipschitzienne et alors continue pour $\|.\|_\infty$.
Posons $\mathrm{S} = \{x \in \mathbb{R}^n, \ \|x\|_\infty = 1\}$. S est fermé et borné dans un espace de dimension finie pour $\|.\|_\infty$, c'est donc un compact d'après la proposition III.2 du chapitre 11.5. N étant continue, elle est bornée sur S et y atteint sa borne inférieure qu'on note α. Il existe alors $y \in \mathrm{S}$ pour lequel pour tout $x \in \mathrm{E} \setminus \{0\}$,

$$\mathrm{N}\left(\frac{x}{\|x\|_\infty}\right) \geq \mathrm{N}(y) = \alpha > 0,$$

et alors

$$\alpha \, \|x\|_\infty \leq \mathrm{N}(x) \leq \mathrm{C} \, \|x\|_\infty .$$

On en déduit donc que toute norme N sur $\mathbb{R}^n$ est équivalent à $\|.\|_\infty$ et que donc par transitivité, toutes les normes sur $\mathbb{R}^n$ sont équivalentes.

Corollaire I.2.

E étant isomorphe à $\mathbb{R}^n$, toutes les normes sur E sont équivalentes.

Démonstration. Soit $\varphi : \mathbb{R}^n \longrightarrow \mathrm{E}$ un isomorphisme et soit N_1 et N_2 deux normes sur E. Sur $\mathbb{R}^n$, les deux normes $\mathrm{N}_1 \circ \varphi$ et $\mathrm{N}_2 \circ \varphi$ sont équivalentes. Il existe donc $\beta, \gamma > 0$ tels que pour tout x

$$\gamma \mathrm{N}_2(\varphi(x)) \leq \mathrm{N}_1(\varphi(x)) \leq \beta \mathrm{N}_2(\varphi(x)).$$

Or φ étant surjective, pour tout $y \in \mathrm{E}$, il existe $x \in \mathrm{E}$ tel que $\varphi(x) = y$, on a donc pour tout $y \in \mathrm{E}$,

$$\gamma \mathrm{N}_2(y) \leq \mathrm{N}_1(y) \leq \beta \mathrm{N}_2(y).$$

(cc) BY-NC-SA

N_1 et N_2 sont alors équivalentes.

II Suites et coordonnées

Proposition II.1.

Soit $(u_p) \in E^{\mathbb{N}}$, on pose pour tout $p \in \mathbb{N}$, $u_p = \sum_{k=1}^{n} u_{p,k} e_k$ et soit $\|.\|$ une norme sur E. Les propositions suivantes sont équivalentes.

1. (u_p) converge pour $\|.\|$.
2. $\forall k \in [\![1;n]\!]$, $\exists \ell_k \in \mathbb{K}$, $u_{n,k} \xrightarrow[n\to+\infty]{} \ell_k$.

Dans ce cas, on a $u_n \xrightarrow[n\to+\infty]{} \sum_{k=1}^{n} \ell_k e_k$.

Démonstration. On pose pour tout $x = \sum_{k=1}^{n} x_k e_k$, $N(x) = \max_{k\in[\![1;n]\!]} |x_k|$. N est une norme sur E équivalente à $\|.\|$.
On a alors

$$\begin{aligned}
(1) &\Longleftrightarrow (u_n) \text{ converge vers un certain } \ell = \sum_{k=1}^{n} \ell_k e_k \text{ pour } \|.\|. \\
&\Longleftrightarrow (u_n) \text{ converge vers un certain } \ell \text{ pour N.} \\
&\Longleftrightarrow \exists \ell = (\ell_1, \dots, \ell_n) \in \mathbb{K}^n,\ \max_{k\in[\![1;n]\!]} |u_{n,k} - \ell_k| \xrightarrow[n\to+\infty]{} 0 \\
&\Longleftrightarrow \exists \ell \in E,\ \forall k \in [\![1;n]\!],\ u_{n,k} \xrightarrow[n\to+\infty]{} \ell_k \Longleftrightarrow (2).
\end{aligned}$$

Application : Une suite de matrices de $\mathcal{M}_n(\mathbb{R})$ converge si et seulement les suites de ses coordonnées convergent.

Proposition II.2.

Soit X un espace métrique et A une partie de X. Soit $a \in \overline{A}$ et l'application

$$f : \begin{cases} A & \longrightarrow E \\ x & \longmapsto \sum_{k=1}^{n} f_k(x) e_k \end{cases}$$

telle que pour tout k, f_k une application de A dans $\mathbb{K}$.
Les propositions suivantes sont équivalentes.

1. f possède une limite finie en a selon A, notée ℓ.
2. Chaque f_k possède une limite finie en a, notée ℓ_k.

Démonstration. La démonstration est laissée comme exercice au lecteur.

(cc) BY-NC-SA

III Compacité et complétude

Théorème (Bolzano Weierstraß) III.1.

Soit (u_p) une suite à valeurs dans E. Si (u_p) est bornée pour $\|.\|$, alors elle admet au moins une valeur d'adhérence.

Démonstration. Posons pour tout $x = \sum_{k=1}^{n} x_k e_k \in \mathrm{E}$, $\mathrm{N}(x) = \max_{k\in[\![1;n]\!]} |x_k|$.
N étant équivalente à $\|.\|$, (u_p) est aussi bornée pour N. Toutes les suites $(u_{p,k})_{p\in\mathbb{N}}$ sont alors bornées. Montrons par récurrence sur r la propriété suivante.

Il existe une extractrice φ_r telle que pour tout $i \in [\![1;r]\!]$, $(u_{\varphi_r(p),i})$ est convergence vers $\ell_i \in \mathbb{R}$.

La propriété pour $r = 1$ est vraie car il s'agit du théorème de Bolzano Weierstraß pour les suites réelles et complexes. Supposons maintenant que la propriété est vraie pour $r \in [\![1;n-1]\!]$. La suite $(u_{\varphi_r(p),r+1})$ est bornée, il existe alors une extractrice φ telle que $(u_{\varphi_r\circ\varphi(p),r+1})$ converge vers un certain ℓ_{r+1}. On a alors pour tout $i \in [\![1;r+1]\!]$, $(u_{\varphi_r\circ\varphi(p),i})$ converge vers ℓ_i. La propriété est alors vraie pour $r = n$ ce qui implique que, pour $\|.\|_\infty$, $(u_{\varphi_n(p)})$ converge vers $\ell = \sum_{k=1}^{n} \ell_k e_k$ et donc par équivalence des normes N et $\|.\|_\infty$, $(u_{\varphi_n(p)})$ converge vers ℓ pour N, d'où le résultat voulu.

Corollaire III.2.

Une partie A de E est compacte si elle est fermée et bornée.

Démonstration. Soit $(u_p) \in \mathrm{A}^{\mathbb{N}}$. (u_p) est bornée, le théorème précédent nous permet d'affirmer qu'il existe une extractrice φ telle que $u_{\varphi(p)} \xrightarrow[p\to+\infty]{} a \in \mathrm{E}$. A étant fermé, on a $a \in \mathrm{A}$, d'où la compacité de A.

Proposition III.3.

L'espace vectoriel normé de dimension finie $(\mathrm{E}, \|.\|)$ est complet.

Démonstration. Soit $(u_p) \in \mathrm{E}^{\mathbb{N}}$ une suite de Cauchy et $\mathrm{N} \in \mathbb{N}$ tel que pour tout $m, n \geq \mathrm{N}$,

$$\|u_m - u_n\| \leq 1.$$

On a alors pour tout $n \in \mathbb{N}$,

$$\|u_n\| \leq \max\{\|u_0\|, \ldots, \|u_{\mathrm{N}-1}\|, \|u_{\mathrm{N}}\| + 1\}.$$

La suite (u_p) est bornée, donc d'après le théorème III.1, elle admet une suite extraite convergente. D'après la remarque à la fin de la page 4 du chapitre 11.5, toute suite de Cauchy admettant une valeur d'adhérence converge, d'où le résultat voulu.

Corollaire III.4.

Soit (u_p) une suite à valeurs dans E.

$$\sum \|u_p\| \text{ converge} \Longrightarrow \sum u_p \text{ converge.}$$

BY-NC-SA

Démonstration. Supposons que $\sum \|u_p\|$ converge et soit $\varepsilon > 0$. Il existe $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$,

$$\sum_{k=n}^{\infty} \|u_p\| \leq \varepsilon.$$

On a alors pour tout $m \geq n \geq \mathrm{N}$,

$$\left\| \sum_{k=n}^{m} u_k \right\| \leq \sum_{k=n}^{m} \|u_k\| \leq \varepsilon.$$

La suite $(\mathrm{U}_n) = \left(\sum_{k=0}^{n} u_k \right)_{n\in\mathbb{N}}$ est de Cauchy et E est complet, donc elle converge, d'où la convergence de la série $\sum u_p$.

Remarque. La propriété ci-dessus est en particulier vraie pour tout espace vectoriel normé complet.

Proposition III.5.

Soit A une partie fermée de E et $f : \mathrm{A} \longrightarrow \mathrm{A}$. Si f vérifie la propriété

$$\exists q \in]0,1[, \ \forall x, y \in \mathrm{A}, \ \|f(x) - f(y)\| \leq q \|x - y\|,$$

alors f possède un unique point fixe dans A.

Démonstration.

→ Unicité
Soit ℓ et ℓ' deux éléments de A tels que $f(\ell) = \ell$ et $f(\ell') = \ell'$. On alors

$$q \|\ell - \ell'\| \geq \|f(\ell) - f(\ell')\| = \|\ell - \ell'\|,$$

ce qui donne nécessairement $\ell = \ell'$.

→ Existence
Soit $a \in \mathrm{A}$. On considère la suite $(u_n)_{n\in\mathbb{N}}$ définie par

$$\begin{cases} u_0 = a \\ \forall n \in \mathbb{N}, \ u_{n+1} = f(u_n). \end{cases}$$

Pour tout $n > 1$,

$$\|u_{n+1} - u_n\| = \|f(u_n) - f(u_{n-1})\| \leq q \|u_n - u_{n-1}\| \leq \cdots \leq q^n \|u_1 - u_0\|.$$

On a $q \in]0,1[$, $\sum_{n\in\mathbb{N}} q^n$ est alors convergente, et par conséquent $\sum_{n\in\mathbb{N}} \|u_{n+1} - u_n\|$ est convergente. E étant complet, le corollaire III.4 nous permet d'affirmer que $\sum_{n\in\mathbb{N}} u_{n+1} - u_n$ est convergente, *i.e.* (u_n) est convergente vers une limite qu'on note ℓ. A est fermé donc $\ell \in \mathrm{A}$ et en passant à la limite dans la relation de récurrence $f(u_n) = u_{n+1}$, on obtient par continuité de f que $\ell = f(\ell)$, d'où le résultat voulu.

Remarque. Pour démontrer le résultat ci-dessus, on n'a pas eu besoin du fait que E est de dimension finie. Il suffit que A soit complet.

(cc) BY-NC-SA

Proposition III.6.

Si E est un sous espace vectoriel de dimension finie d'un $\mathbb{K}-$espace vectoriel $(\mathrm{F}, \|.\|)$, alors E est fermé.

Démonstration. Soit (u_p) une suite à valeurs dans E qui converge vers $a \in \mathrm{F}$. E est de dimension finie et (u_p) est convergente donc bornée dans E, il existe donc une extractrice φ telle que $u_{\varphi(p)} \xrightarrow[p \to +\infty]{} b \in \mathrm{E}$, or par unicité de la limite, $a = b \in \mathrm{E}$, donc E est bien fermé.

Exemple. $\mathbb{R}_n[\mathrm{X}]$ est fermé dans $\mathcal{C}([0,1], \mathbb{R})$ muni de la norme $\|.\|_\infty$.

Théorème (Théorème de Riesz) III.7.

Soit $(\mathrm{F}, \|.\|)$ un espace vectoriel normé. $\mathrm{B}_f(0,1)$ est compacte si et seulement si F est de dimension finie.

Démonstration. La démonstration sera présentée sous forme d'exercice dans la suite du chapitre.

IV Fonctions polynômes sur $\mathbb{K}^n$

On rappelle ici que E est un espace vectoriel normé de **dimension finie** muni de la norme $\|.\|_{\mathrm{E}}$ et $(e_1, \ldots, e_n)$ est une base de E.

Définition IV.1.

Soit f une application de E dans $\mathbb{K}$. Soit $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_n)$ une base de F. On dit que f est polynomiale dans la base $\mathcal{E}$, lorsqu'il existe une suite $(a_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{K}^{\mathbb{N}^n}$ avec un nombre fini de termes non nuls telle que pour tout $x = x_1\varepsilon_1 + \cdots + x_n\varepsilon_n \in \mathrm{E}$,

$$f(x) = \sum_{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n} a_{\alpha_1, \ldots, \alpha_n} x_1^{\alpha_1} \ldots x_n^{\alpha_n}.$$

Proposition IV.2.

Soit F un espace vectoriel normé quelconque.

1. Toute application u linéaire de E dans F est continue.
2. Toute application f polynomiale dans une base de E à valeurs dans $\mathbb{K}$ est continue.

Démonstration.

1. Soit N une norme sur E définie de la manière suivante

$$\forall x = x_1e_1 + \cdots + x_ne_n \in \mathrm{E},\ \mathrm{N}(x) = \sum_{k=1}^{n} |x_k|.$$

On alors pour tout $x = x_1e_1 + \cdots + x_ne_n \in \mathrm{E}$,

$$\|u(x)\|_{\mathrm{F}} = \|x_1u(e_1) + \cdots + x_nu(e_n)\|_{\mathrm{F}} \le \sum_{k=1}^{n} |x_i| \, \|u(e_i)\|_{\mathrm{F}} \le \mathrm{C} \cdot \mathrm{N}(x),$$

avec $\mathrm{C} = \max_{k \in [\![1;n]\!]} \|u(e_k)\|_{\mathrm{F}}$. E étant de dimension finie, donc N est équivalente à $\|.\|_{\mathrm{E}}$, u est continue.

(cc) BY-NC-SA

2. Posons pour tout $x = x_1e_1 + \cdots + x_ne_n \in \mathrm{E}$,

$$f(x) = \sum_{(\alpha_1,\dots,\alpha_n)\in\mathbb{N}^n} a_{\alpha_1,\dots,\alpha_n} x_1^{\alpha_1}\dots x_n^{\alpha_n}.$$

Pour tout $k \in [\![1;n]\!]$, l'application $x \longmapsto x_k$ est continue, et f est somme et produit d'applications de cette forme qui sont toutes continues, donc f est continue.

Corollaire IV.3.

Toute application f multilinéaire de E^p dans un espace vectoriel normé F est continue.

Démonstration. Considérons les p éléments de E suivants.

$$\begin{aligned} x_1 &= x_{1,1}e_1 + \cdots + x_{1,n}e_n \\ x_2 &= x_{2,1}e_1 + \cdots + x_{2,n}e_n \\ &\vdots \\ x_p &= x_{p,1}e_1 + \cdots + x_{p,n}e_n. \end{aligned}$$

On a alors

$$\begin{aligned} f(x_1,\dots,x_p) &= f\left(\sum_{i_1=1}^{n} x_{1,i_1}e_{i_1}, \sum_{i_2=1}^{n} x_{2,i_2}e_{i_2}, \dots, \sum_{i_p=1}^{n} x_{p,i_p}e_{i_p}\right) \\ &= \sum_{(i_1,\dots,i_p)\in[\![1;n]\!]^p} x_{i_1}\dots x_{i_p} f(e_{i_1},\dots,e_{i_p}). \end{aligned}$$

f est donc combinaison linéaire et produit d'applications de la forme $(x_1,\dots,x_p) \longmapsto x_{k,\ell}$ avec $(k,\ell) \in [\![1;p]\!] \times [\![1;n]\!]$. Toutes ces applications sont continues, donc f est continue.

V Exercices

Exercice V.1.

Soit $f : \mathbb{R} \longrightarrow \mathbb{R}^+$ de classe $\mathcal{C}^n$ telle que f et $f^{(n)}$ sont bornées. Montrer que pour tout $k \in [\![1;n-1]\!]$, $f^{(k)}$ est bornée.

Exercice V.2.

Soit $n \in \mathbb{N}^*$. On note $\Omega = \{\mathrm{P} \in \mathbb{R}_n[\mathrm{X}],\ \deg \mathrm{P} = n \text{ et P est scindé à racines simples}\}$. Montrer que Ω est ouvert dans $\mathbb{R}_n[\mathrm{X}]$.

Exercice V.3.

Soit $(\mathrm{P}_k)_{k\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{R}_n[\mathrm{X}]$. On suppose que la suite $(\tilde{\mathrm{P}}_k)$ de fonctions polynômes associée à (P_k) converge simplement vers $f \in \mathcal{C}([0,1],\mathbb{R})$ sur $[0,1]$. Montrer que la convergence est uniforme.

BY-NC-SA

Exercice V.4.

Soit F un fermé non vide de l'espace vectoriel normé de dimension finie $(E, \|.\|)$ et soit $a \in E$. Montrer qu'il existe $b \in F$ tel que $\|a - b\| = d(a, F)$.

Exercice V.5.

Soit F un sous-espace de dimension finie d'un espace vectoriel normé $(G, \|.\|)$ de dimension quelconque. Soit $a \in G$. Montrer qu'il existe $b \in F$ tel que $\|a - b\| = d(a, F)$.

Exercice V.6.

Soit K un compact non vide de l'espace euclidien $\mathbb{R}^n$. Montrer qu'il existe une boule fermée de rayon minimum contenant K et que cette boule est unique.

Exercice V.7.

Soit f une fonction continue de l'espace vectoriel normé de dimension finie $(E, \|.\|)$ dans $\mathbb{R}$. On suppose que $f(x) \xrightarrow[\|x\| \to +\infty]{} +\infty$. Montrer que f est minorée et qu'il existe $a \in E$ tel que $f(a) = \min\limits_{x \in E} f(x)$.

Exercice V.8.

Ceci est un exercice visant à prouver le théorème III.7 (Théorème de Riesz).
Soit $(F, \|.\|)$ un espace vectoriel normé de dimension infinie.

1. Soit G un sous-espace vectoriel de F de dimension finie. Montrer qu'il existe $a \in S(0,1)$ tel que $d(a, F) \geq 1$.
2. Construire une suite $(a_n) \in S(0,1)^{\mathbb{N}}$ telle que $\forall n \neq m,\ \|a_n - a_m\| \geq 1$.
3. Conclure.

(cc) BY-NC-SA

Correction de l'exercice V.1. :

f est de classe $\mathcal{C}^n$, on peut alors appliquer l'égalité de Taylor-Lagrange à l'ordre n pour f sur $[0,1]$,

$$\forall (x,h) \in \mathbb{R}^+ \times [0,1],\ \exists c_{x,h} \in [0,1],\ f(x+h) - f(x) - \frac{h^n}{n!} f^{(n)}(c_{x,h}) = \underbrace{\sum_{k=1}^{n-1} \frac{f^{(k)}(x)}{k!} h^k}_{\mathrm{P}_x(h)}.$$

D'après les hypothèses, $(x,h) \longmapsto f(x+h) - f(x) - \frac{h^n}{n!} f^{(n)}(c_{x,h})$ est bornée par un certain $\mathrm{M} > 0$. On a alors en posant $g_x : h \longmapsto f(x+h) - f(x) - \frac{h^n}{n!} f^{(n)}(c_{x,h})$, on a par construction, pour tout $x \in \mathbb{R}^+$, $\|g_x\|_\infty \leq \mathrm{M}$, et alors $\|\mathrm{P}_x\|_\infty \leq \mathrm{M}$. Posons pour tout $\mathrm{P(X)} = \sum\limits_{k=0}^{n} a_k \mathrm{X}^k \in \mathbb{R}_n[\mathrm{X}]$, $\mathrm{N(P)} = \max\limits_{k \in [\![0;n]\!]} |a_k|$. $\mathbb{R}_n[\mathrm{X}]$ est de dimension finie, donc toutes les normes y sont équivalentes. En particulier, N et $\|.\|_\infty$ sont équivalentes, ce qui nous permet d'affirmer qu'il existe $\mathrm{M}' > 0$ tel que pour tout $x \in \mathbb{R}^+$, $\mathrm{N}(\mathrm{P}_x) < \mathrm{M}'$.
On a donc pour tout

$$x \in \mathbb{R}^+,\ \forall k \in [\![1;n]\!],\ \left| \frac{f^{(k)}(x)}{k!} \right| < \mathrm{M}',$$

ce qui nous permet d'affirmer que pour tout $k \in [\![1;n]\!]$, $f^{(k)}$ est bornée.

Correction de l'exercice V.2. :

Soit $\mathrm{P} \in \Omega$. P admet n racines distinctes et est scindé à racines simples, il change donc $n+1$ fois de signe sur $\mathbb{R}$. Soit $y_1, \dots, y_{n+1} \in \mathbb{R}$ tels que sans perte de généralité, pour tout $k \in [\![1;n+1]\!]$, $\mathrm{signe}(\mathrm{P}(y_k)) = (-1)^k$. Posons pour tout $\mathrm{Q} \in \mathbb{R}_n[\mathrm{X}]$, $\mathrm{N(Q)} = \sum\limits_{k=1}^{n+1} |\mathrm{Q}(y_i)|$. N est bien une norme sur $\mathbb{R}_n[\mathrm{X}]$, car elle vérifie l'inégalité triangulaire, homogène, et pour tout polynôme Q de degré au plus n, $\mathrm{N(Q)} = 0 \Longrightarrow \mathrm{Q} = 0$.
Posons $\varepsilon = \min\limits_{k \in [\![1;n+1]\!]} |\mathrm{P}(y_k)|$. Pour tout $\mathrm{Q} \in \mathbb{R}_n[\mathrm{X}]$, si $\mathrm{N(P-Q)} < \frac{\varepsilon}{2}$, alors pour tout $k \in [\![1;n+1]\!]$,

$$|\mathrm{Q}(y_k) - \mathrm{P}(y_k)| < \frac{\varepsilon}{2}$$

i.e.

$$\mathrm{P}(y_k) - \frac{\varepsilon}{2} < \mathrm{Q}(y_k) < \mathrm{P}(y_k) + \frac{\varepsilon}{2}.$$

Or pour tout $k \in [\![1;n+1]\!]$, $|\mathrm{P}(y_k)| \geq \varepsilon$, donc

$$\mathrm{signe}\left(\mathrm{P}(y_k) + \frac{\varepsilon}{2}\right) = \mathrm{signe}\left(\mathrm{P}(y_k) - \frac{\varepsilon}{2}\right) = \mathrm{signe}(\mathrm{P}(y_k)).$$

Par conséquent, pour tout $k \in [\![1;n+1]\!]$, $\mathrm{Q}(y_k)$ a le même signe que $\mathrm{P}(y_k)$. En appliquant le théorème des valeurs intermédiaires sur $[y_k, y_{k+1}]$ pour tout $k \in [\![1;n]\!]$, on trouve que Q, étant de degré n, admet n racines distinctes et donc est scindé à racines simples. On en déduit donc que $\mathrm{B(P}, \varepsilon) \subset \Omega$ et finalement que Ω est ouvert.

Correction de l'exercice V.3. :

Soit $x_0, x_1, \dots, x_n$ des éléments deux à deux distincts de $[0,1]$. Posons pour tout $k \in \mathbb{N}$,

$$\mathrm{P}_k(\mathrm{X}) = \sum_{i=0}^{n} a_{k,i} \mathrm{X}^i.$$

On a alors pour tout $\ell \in [\![0;n]\!]$,

$$\sum_{i=1}^{n} a_{k,i} x_\ell^i \xrightarrow[k \to +\infty]{} f(x_\ell)$$

(cc) BY-NC-SA

i.e.

$$\begin{pmatrix}1 & x_0 & \dots & x_0^n\\ 1 & x_1 & \dots & x_1^n\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & \dots & x_n^n\end{pmatrix}\begin{pmatrix}a_{k,0}\\ a_{k,1}\\ \vdots\\ a_{k,n}\end{pmatrix} \xrightarrow[k\to+\infty]{} \begin{pmatrix}f(x_0)\\ f(x_1)\\ \vdots\\ f(x_n)\end{pmatrix}.$$

En posant

$$\mathrm{V} = \mathrm{V}(x_0, x_1, \dots, x_n) = \begin{pmatrix}1 & x_0 & \dots & x_0^n\\ 1 & x_1 & \dots & x_1^n\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & \dots & x_n^n\end{pmatrix},$$

on remarque que V est une matrice de Vandermonde, elle est donc inversible. $\mathrm{X} \longmapsto \mathrm{V}^{-1}\mathrm{X}$ est une application linéaire en dimension finie, elle est donc continue d'après la proposition IV.2. On a alors, en multipliant par V^{-1} et en passant à la limite

$$\begin{pmatrix}a_{k,0}\\ a_{k,1}\\ \vdots\\ a_{k,n}\end{pmatrix} \xrightarrow[k\to+\infty]{} \mathrm{V}^{-1}\begin{pmatrix}f(x_0)\\ f(x_1)\\ \vdots\\ f(x_n)\end{pmatrix}. \tag{$*$}$$

En posant pour tout $\mathrm{Q}(\mathrm{X}) = \sum_{i=0}^{n} b_i \mathrm{X}^i$, $\mathrm{N}(\mathrm{Q}) = \sum_{i=0}^{n} |b_i|$, on voit que N est une norme dans l'espace de dimension finie $\mathbb{R}_n[\mathrm{X}]$, elle est donc équivalente à $\|.\|_\infty : \mathrm{Q} \longmapsto \sup_{x\in[0,1]} \tilde{\mathrm{Q}}(x)$. D'après $(*)$, $(\mathrm{P}_k)_{k\in\mathbb{N}}$ converge pour N (chaque coefficient converge) et donc par équivalence des normes, (P_k) converge aussi pour $\|.\|_\infty$. $(\tilde{\mathrm{P}}_k)$ converge donc uniformément vers une certaine fonction continue dans $[0, 1]$ g. Il reste à montrer que $f = g$. La convergence uniforme implique la convergence simple, donc $(\tilde{\mathrm{P}}_k)$ converge simplement vers f et g, et donc par unicité de la limite, $f = g$.

Remarque. L'espace des fonctions polynômes de degré au plus n étant de dimension finie, il est fermé dans l'espace des fonctions continues sur $[0, 1]$. Cet argument nous permet d'affirmer que (P_k) converge uniformément vers un polynôme de degré au plus n.
Une démonstration plus directe de ce résultat peut être faite de la manière suivante. Soit $a_0, \dots, a_n$ les limites respectives des suites de coordonnées $(a_{k,0})_{k\in\mathbb{N}}, \dots, (a_{k,n})_{k\in\mathbb{N}}$. En posant $\mathrm{P}(\mathrm{X}) = \sum_{i=0}^{n} a_i \mathrm{X}^i$, on a

$$\left\|\tilde{\mathrm{P}}_k - \tilde{\mathrm{P}}\right\|_\infty \leq \sum_{i=0}^{n} |a_{k,i} - a_i| \xrightarrow[k\to+\infty]{} 0.$$

Correction de l'exercice V.4. :

Soit $\mathrm{R} > d(a, \mathrm{F})$. On a $\mathrm{B}_f(a, \mathrm{R}) \cap \mathrm{F} = \mathrm{K} \neq \varnothing$. K est une intersection de deux fermés incluse dans l'ensemble borné $\mathrm{B}_f(a, \mathrm{R})$ en dimension finie, c'est donc une partie compacte de E. On considère la fonction

$$g : \begin{cases}\mathrm{K} & \longrightarrow \mathbb{R}\\ x & \longmapsto \|a - x\|.\end{cases}$$

g est continue sur le compact K, il y atteint donc son minimum en un élément de K qu'on note b.
Soit $y \in \mathrm{F}$. Si $y \notin \mathrm{K}$, alors $\|y - a\| > \mathrm{R} \geq \|a - b\|$. Sinon, si $y \in \mathrm{K}$, alors $\|y - a\| \geq \|a - b\|$. On a donc bien $\|a - b\| = \min_{x\in\mathrm{F}} \|a - x\| = d(a, \mathrm{F})$.

BY-NC-SA

Correction de l'exercice V.5. :

Cet exercice se fait d'une manière identique au précédent. Soit $\mathrm{R} > d(a,\mathrm{F})$ et $\mathrm{K} = \mathrm{B}_f(a,\mathrm{R}) \cap \mathrm{F}$. K est fermé borné inclus dans un espace de dimension finie F, c'est donc un compact. en reprenant les notations de l'exercice précédent, g est continue et donc atteint son minimum sur le compact K en un certain $b \in \mathrm{F}$. Avec un raisonnement identique à l'exercice précédent, on peut affirmer que $\|a-b\| = d(a,\mathrm{F})$.

Correction de l'exercice V.6. :

Posons $\mathrm{S_K} = \{\gamma \geq 0,\ \exists a \in \mathbb{R}^n,\ \mathrm{K} \subset \mathrm{B}_f(a,\gamma)\}$. K est borné, il existe donc une boule fermée contenant K. L'ensemble $\mathrm{S_K}$ est alors non vide, minoré par 0. Il admet donc une borne inférieure $r = \inf \mathrm{S_K}$.
Soit $((a_n, r_n))_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{R}^n \times \mathbb{R}^+$ vérifiant les deux propriétés suivantes

$$\begin{cases} \forall n \in \mathbb{N},\ \mathrm{K} \subset \mathrm{B}(a_n, r_n) \\ r_n \xrightarrow[n\to+\infty]{} r. \end{cases}$$

D'après les hypothèses, il existe N assez grand tel que pour tout $n \geq \mathrm{N}$ on ait $r_n \leq r+1$. Soit $b \in \mathrm{K}$. On a pour tout $n \geq \mathrm{N}$,

$$\|a_n - b\| \leq r_n \leq r+1$$

et alors

$$\|a_n\| \leq r + 1 + \|b\|.$$

La suite (a_n) est bornée, ce qui nous permet en utilisant le théorème III.1 d'affirmer que (a_n) admet une suite extraite convergente. Quitte à extraire de la suite (a_n), on suppose qu'elle est convergente vers un certain $a \in \mathbb{R}^n$.

→ Pour montrer l'existence de la boule, montrons que $\mathrm{K} \subset \mathrm{B}_f(a,r)$.
On a pour tout $x \in \mathrm{K}$ et pour tout $n \in \mathbb{N}$, $\|a_n - x\| \leq r_n$. En passant à la limite, on obtient $\|a-x\| \leq r$, *i.e.* $x \in \mathrm{B}_f(a,r)$. On a donc bien que $\mathrm{K} \subset \mathrm{B}_f(a,r)$.

→ Montrons que cette boule est unique.
On sait que le rayon de la boule est unique car il est défini comme la borne inférieure de $\mathrm{S_K}$, il suffit donc de montrer l'unicité du centre.
Soit $a' \in \mathbb{R}^n$ différent de a tel que $\mathrm{K} \subset \mathrm{B}_f(a',r)$. Soit $x \in \mathrm{B}_f(a,r) \cap \mathrm{B}_f(a',r)$, on a

$$\begin{aligned} 2r^2 \geq \|\underbrace{x-a}_{u}\|^2 + \|\underbrace{x-a'}_{v}\|^2 &= \frac{1}{2}(\|u+v\|^2 + \|u-v\|^2) \\ &= 2\left(\left\|x - \frac{a+a'}{2}\right\|^2 + \underbrace{\left\|\frac{a-a'}{2}\right\|^2}_{\alpha}\right). \end{aligned}$$

En posant $b = \frac{a+a'}{2}$, on en déduit

$$\|x-b\| \leq \sqrt{r^2 - \underbrace{\alpha}_{>0}} < r,$$

donc $\mathrm{K} \subset \mathrm{B}_f(b, \sqrt{r^2-\alpha})$ ce qui est absurde par définition de r, d'où le résultat voulu.

Correction de l'exercice V.7. :

Le fait que $f(x) \xrightarrow[\|x\|\to+\infty]{} +\infty$ se réécrit

$$\forall \mathrm{M} > 0,\ \exists \mathrm{R} > 0,\ \forall x \in \mathrm{E},\ \|x\| > \mathrm{R} \Longrightarrow \|f(x)\| > \mathrm{M}.$$

Appliquons cette proposition à $\mathrm{M} = |f(0)| + 1$. f est continue sur $\mathrm{B}_f(0,\mathrm{R})$ qui est fermé borné en dimension finie donc compact, elle y atteint donc son minimum en un certain $a \in \mathrm{B}_f(0,\mathrm{R})$.

(cc) BY-NC-SA

Or pour tout $y \in \mathrm{E}$,

→ Si $\|y\| > \mathrm{R}$, $f(y) > |f(0)| + 1 > f(a)$.

→ Sinon, par construction, $f(y) \geq f(a)$.

Donc f atteint bien son minimum sur E en a.

Correction de l'exercice V.8. : *(Démonstration du théorème de Riesz)*

1. Soit $a_0 \in \mathrm{F} \setminus \mathrm{G}$. L'exercice V.4 nous permet d'affirmer qu'il existe $b \in \mathrm{G}$ tel que

$$\|a_0 - b\| = \inf_{x \in \mathrm{G}} \|x - a_0\| = d(a_0, \mathrm{F}).$$

Posons alors $a = \dfrac{b - a_0}{\|b - a_0\|} \in \mathrm{S}(0,1)$, on a alors pour tout $x \in \mathrm{G}$,

$$\|x - a\| = \left\| x - \frac{a_0 - b}{\|a_0 - b\|} \right\| = \frac{1}{d(a_0, \mathrm{G})} \|\underbrace{b + x\,\|a_0 - b\|}_{=y \in \mathrm{G}} - a_0\| = \frac{1}{d(a_0, \mathrm{G})} \|y - a_0\| \geq 1.$$

2. Construisons la suite a_n par récurrence. le premier terme a_0 peut être pris quelconque vérifiant $\|a_0\| = 1$.
Soit $n \in \mathbb{N}$. Supposons que les termes $a_0, \ldots, a_n$ sont bien définis. Posons $\mathrm{G}_n = \mathrm{Vect}(a_0, \ldots, a_n)$. D'après la question précédente, il existe $a_{n+1} \in \mathrm{S}(0,1)$ tel que $d(a_{n+1}, \mathrm{G}_n) \geq 1$. On a alors

$$\forall m \leq n,\ \|a_m - a_{n+1}\| \geq 1.$$

Il est clair que la suite (a_n) telle qu'on l'a définie vérifie

$$\forall m, n \in \mathbb{N},\ (n \neq m \Longrightarrow \|a_n - a_m\| \geq 1) \text{ et } \forall n \in \mathbb{N},\ \|a_n\| = 1.$$

3. Reprenons le théorème III.7. L'énoncé du théorème est le suivant : Si H un espace vectoriel normé, alors on a l'équivalence

$$\text{H est de dimension finie} \iff \text{La boule unité fermée de H est compacte.}$$

→ ($\Leftarrow$) nous allons faire cette implication par contraposée. Supposons que H soit de dimension infinie. La question précédente nous permet de considérer une suite $(a_n) \in \mathrm{S}(0,1)^{\mathbb{N}}$ telle que

$$\forall m, n \in \mathbb{N},\ n \neq m \Longrightarrow \|a_n - a_m\| \geq 1.$$

Si la boule unité était compacte, on disposerait d'une extractrice φ telle que $(a_{\varphi(n)})$ soit convergente, et alors

$$1 \leq \left\| a_{\varphi(n+1)} - a_{\varphi(n)} \right\| \xrightarrow[n \to +\infty]{} 0,$$

ce qui est absurde.

→ ($\Rightarrow$) Si H est de dimension finie, alors la boule unité fermée de H est fermé bornée en dimension finie, elle est alors compacte.

(cc) BY-NC-SA

Chapitre 11.8 — Connexité

I Généralités

Définition I.1.

Soit (X, d) un espace métrique. Les propositions suivantes sont équivalentes

1. Si O_1 et O_2 sont deux ouverts de X tels que $\mathrm{X} = \mathrm{O}_1 \cup \mathrm{O}_2$ et $\mathrm{O}_1 \cap \mathrm{O}_2 = \varnothing$, alors $\mathrm{O}_1 = \varnothing$ ou $\mathrm{O}_2 = \varnothing$.
2. Si F_1 et F_2 sont deux fermés de X tels que $\mathrm{X} = \mathrm{F}_1 \cup \mathrm{F}_2$ et $\mathrm{F}_1 \cap \mathrm{F}_2 = \varnothing$, alors $\mathrm{F}_1 = \varnothing$ ou $\mathrm{F}_2 = \varnothing$.
3. Si A est une partie à la fois fermée et ouverte de X, alors $\mathrm{A} = \mathrm{X}$ ou $\mathrm{A} = \varnothing$.
4. Si f est une fonction continue de X dans $\{0, 1\}$, alors f est constante.

Lorsque l'une de ces propriétés est vérifiée, on dit que X est connexe.

Démonstration.

- → (1) ⇒ (2) Soit F_1 et F_2 deux fermés vérifiant $\mathrm{F}_2 \cup \mathrm{F}_2 = \mathrm{X}$ et $\mathrm{F}_1 \cap \mathrm{F}_2 = \varnothing$, et posons $\mathrm{O}_1 = \mathrm{X} \setminus \mathrm{F}_1$ et $\mathrm{O}_2 = \mathrm{X} \setminus \mathrm{F}_2$. O_1 et O_2 sont ouverts et $\mathrm{O}_1 \cup \mathrm{O}_2 = \mathrm{X}$ et $\mathrm{O}_1 \cap \mathrm{O}_2 = \varnothing$, donc $\mathrm{O}_1 = \mathrm{X}$ et $\mathrm{O}_2 = \varnothing$ ou $\mathrm{O}_1 = \varnothing$ et $\mathrm{O}_2 = \mathrm{X}$, *i.e.* $\mathrm{F}_1 = \mathrm{X}$ et $\mathrm{F}_2 = \varnothing$ ou $\mathrm{F}_1 = \varnothing$ et $\mathrm{F}_2 = \mathrm{X}$.
- → (2) ⇒ (3) Soit A une partie fermée et ouverte de X. $\mathrm{F}_1 = \mathrm{A}$ et $\mathrm{F}_2 = \mathrm{A}^c$ vérifient les hypothèses de (2). On a $\mathrm{A} = \mathrm{X}$ et $\mathrm{A}^c = \varnothing$ ou $\mathrm{A} = \varnothing$ et $\mathrm{A}^c = \mathrm{X}$, d'où le résultat voulu.
- → (3) ⇒ (4) Posons $\mathrm{F}_1 = f^{-1}(\{0\})$ et $\mathrm{F}_2 = f^{-1}(\{1\})$. F_1 et F_2 sont fermés car il s'agit d'images réciproques de deux fermés par une fonction continue. De plus, on a $\mathrm{F}_1 \cup \mathrm{F}_2 = \mathrm{X}$ et $\mathrm{F}_1 \cap \mathrm{F}_2 = \varnothing$. On a alors $\mathrm{F}_1 = \mathrm{X}$ ou $\mathrm{F}_2 = \mathrm{X}$, *i.e.* $f^{-1}(\{0\}) = \mathrm{X}$ ou $f^{-1}(\{1\}) = \mathrm{X}$. f est alors toujours égale à 0 ou toujours égale à 1. On a donc bien le résultat voulu.
- → (4) ⇒ (1) Soit O_1 et O_2 deux ouverts vérifiant $\mathrm{O}_1 \cup \mathrm{O}_2 = \mathrm{X}$ et $\mathrm{O}_1 \cap \mathrm{O}_2 = \varnothing$. On considère l'application de X dans $\{0, 1\}$ suivante.

$$f : \begin{cases} \mathrm{X} & \longrightarrow \{0,1\} \\ x & \longmapsto \begin{cases} 0 \text{ si } x \in \mathrm{O}_1 \\ 1 \text{ si } x \in \mathrm{O}_2. \end{cases} \end{cases}$$

  f est continue car O_1 et O_2 sont ouverts, elle est donc constante. On a alors $\mathrm{O}_1 = f^{-1}(\{0\}) = \varnothing$ et $\mathrm{O}_2 = f^{-1}(\{1\}) = \mathrm{X}$ ou $\mathrm{O}_1 = f^{-1}(\{0\}) = \mathrm{X}$ et $\mathrm{O}_2 = f^{-1}(\{1\}) = \varnothing$, ce qui est bien ce qu'on cherchait à montrer.

Proposition I.2.

Soit I un ensemble et $(\Omega_i)_{i \in \mathrm{I}}$ une partition d'ouverts d'un espace métrique connexe X. Il existe alors au plus un indice i_0 tel que $\Omega_{i_0} \neq \varnothing$.

Démonstration. Soit i_0 tel que $\Omega_{i_0} \neq \varnothing$. Posons $\Omega = \bigcup_{i \in \mathrm{I} \setminus \{i_0\}} \Omega_i$. On a alors $\Omega_{i_0} \cup \Omega = \mathrm{X}$ et $\Omega_{i_0} \cap \Omega = \varnothing$. De plus Ω_{i_0} et Ω sont ouverts (car c'est une union d'ouverts), donc X étant connexe, on a nécessairement $\Omega = \varnothing$, *i.e.* pour tout $i \neq i_0$, $\Omega_i = \varnothing$.

(cc) BY-NC-SA

Proposition I.3.

Si A est une partie connexe de l'espace métrique X alors $\overline{A}$ aussi.

Démonstration. Soit $f : \overline{A} \longrightarrow \{0,1\}$ une application continue. f est continue sur A par restriction et A est connexe, donc f est constante sur A. $\overline{A}$ est dense dans A donc f étant continue, elle est aussi constante sur $\overline{A}$, donc $\overline{A}$ est bien connexe.

Remarque. Attention, si A est connexe, $\mathring{A}$ ne l'est pas forcément.

Proposition I.4.

Soit X et Y deux espaces métriques et $f \in \mathcal{C}(X, Y)$. Si A est une partie connexe de X, alors $f(A)$ est aussi une partie connexe de Y.

Démonstration. Soit $g : f(A) \longrightarrow \{0,1\}$ une application continue. $g \circ f$ est une fonction continue de A dans $\{0,1\}$ car il s'agit de la composition de deux applications continues, elle est alors constante. On en déduit donc directement que $\exists c \in \{0,1\}$, $g \circ f(A) = \{c\}$, *i.e.* $g(f(A)) = \{c\}$. g est alors constante sur $f(A)$ ce qui nous permet d'affirmer que $f(A)$ est bien une partie connexe de Y.

II Parties connexes par arcs

Dans cette partie, on considère (E, d) un espace métrique.

Proposition II.1.

$[0,1]$ est une partie connexe.

Démonstration. Soit $f \in \mathcal{C}([0,1], \{0,1\})$. Supposons sans perte de généralité que $f(0) = 0$. On considère l'ensemble

$$A = \{x \in [0,1],\ \forall y \in [0,x],\ f(y) = 0\}.$$

Posons $a = \sup A$. a est limite d'une suite à valeurs dans A, donc par continuité de f, on a $f(a) = 0$. Supposons que $a \neq 1$. On a pour tout $n \in \mathbb{N}$, il existe $x_n \in \left[a, a + \frac{1}{n}\right]$, $f(x_n) = 1$. La suite (x_n) converge vers a et donc

$$f(x_n) \xrightarrow[n \to +\infty]{} 1 \neq f(a),$$

ce qui est absurde car f est continue.

Définition II.2.

1. On appelle arc continu toute application continue d'un segment $[a,b]$ de $\mathbb{R}$ dans E.
2. Une partie A de E est dite connexe par arcs si tous éléments x et y de A peuvent être reliés par un arc continu sur A, *i.e.*

$$\forall (x,y) \in A^2,\ \exists \gamma \in \mathcal{C}([0,1], A),\ \gamma(0) = x \text{ et } \gamma(1) = y.$$

Remarque. Dans la définition de la connexité par arc, on aurait pu remplacer 0 par $a \in \mathbb{R}$ et 1 par $b \in \mathbb{R}$ tel que $a < b$ quelconques.

BY-NC-SA

Définition II.3.

Soit $x, y, z \in \mathrm{E}$, γ un arc continu reliant x et y et δ un arc continu reliant y et z. L'arc $\gamma \bullet \delta$ est défini par

$$\gamma \bullet \delta : \begin{cases} [0,1] & \longrightarrow \mathrm{E} \\ y & \longmapsto \begin{cases} \gamma(2t) & \text{si } t \in \left[0, \frac{1}{2}\right] \\ \delta(2t-1) & \text{si } t \in \left]\frac{1}{2}, 1\right] \end{cases} \end{cases}.$$

Par construction $\gamma \bullet \delta$ est un arc continu reliant x à z. On dit que $\gamma \bullet \delta$ est la reliure des deux arcs γ et δ. $\bullet$ n'étant pas associative, pour trois arcs $\gamma_1, \gamma_2, \gamma_3$, on définit $\gamma_1 \bullet \gamma_2 \bullet \gamma_3$ comme $\gamma_1 \bullet (\gamma_2 \bullet \gamma_3)$.

Définition II.4.

Supposons que pour tout $t \in [0,1]$, $(1-t)a + tb \in \mathrm{E}$. Pour tout $x, y \in \mathrm{E}$, on définit l'arc $\theta(x,y)$ comme

$$\theta(x,y) : \begin{cases} [0,1] & \longrightarrow \mathrm{E} \\ t & \longmapsto (1-t)x + ty. \end{cases}$$

Par construction, $\theta(x,y)$ est bien un arc continu reliant x à y.

Remarque. Attention, ces deux dernières notations sont propres à ce cours en particulier et ne sont pas communes en classes préparatoires. Si vous souhaitez les utiliser, il faut les définir avant.

Proposition II.5.

Soit F un espace métrique et f une fonction continue de E dans F. L'image d'une partie de E connexe par arcs par f est connexe par arcs.

Démonstration. Soit A une partie connexe par arcs de E. Soit $a, b \in f(\mathrm{A})$, et $x, y \in \mathrm{A}$ tels que $a = f(x)$ et $b = f(y)$. A est connexe par arcs, il existe donc un arc $\gamma \in \mathcal{C}([0,1], \mathrm{A})$ tel que $\gamma(0) = x$ et $\gamma(1) = y$. On a alors $f \circ \gamma$ est un arc continu reliant $a = f(x)$ et $b = f(y)$.

Proposition II.6.

Une partie A de E connexe par arcs est connexe.

Démonstration. Soit $g : \mathrm{A} \longrightarrow \{0,1\}$ une fonction continue. Soient x et y dans A. A est connexe par arcs, il existe donc un arc continu $\gamma \in \mathcal{C}([0,1], \mathrm{A})$ tel que $\gamma(0) = x$ et $\gamma(1) = y$.
L'application $g \circ \gamma : [0,1] \longrightarrow \{0,1\}$ est continue et $[0,1]$ est connexe, donc elle est constante et alors $g(x) = g(y)$. On en déduit donc que g est constante et que par conséquent A est connexe.

Exemple. Pour tout $n \in \mathbb{N}^*$, $\mathrm{GL}_n(\mathbb{R})$ n'est pas connexe par arcs. En effet, l'application $f = \det$ est continue, donc si $\mathrm{GL}_n(\mathbb{R})$ était connexe par arcs, $f(\mathrm{GL}_n(\mathbb{R}))$ le serait aussi, mais $f(\mathrm{GL}_n(\mathbb{R})) = \mathbb{R}^*$ qui n'est pas connexe par arcs.

(cc) BY-NC-SA

Proposition II.7.

Soit A une partie de E.

1. La relation sur A^2 définie par

$$\forall x, y \in \mathrm{A},\ x \sim_{\mathrm{A}} y \Longleftrightarrow \exists \gamma \in \mathcal{C}([0,1], \mathrm{A}),\ \gamma(0) = x \text{ et } \gamma(1) = y$$

   est une relation d'équivalence.
2. Les classes d'équivalence pour $\sim_{\mathrm{A}}$ sont connexes par arcs. On appelle ces classes composantes connexes par arcs.

Lorsqu'il n'y a pas d'ambiguïté sur l'ensemble A, on note $\sim$ au lieu de $\sim_{\mathrm{A}}$.

Démonstration.

1. → Réflexivité
   Soit Soit $x, y \in \mathrm{A}$. Si $x \sim y$, alors il existe un arc $\gamma \in \mathcal{C}([0,1], \mathrm{A})$ reliant x et y. Il suffit alors de considérer $\tilde{\gamma} : a \longmapsto \gamma(1-a)$. L'arc $\tilde{\gamma}$ relie x et y et $\tilde{\gamma}(0) = y$ et $\tilde{\gamma}(1) = x$ et alors $y \sim x$.

   → Symétrie
   Soit $x \in \mathrm{A}$. L'arc $\gamma : t \longmapsto x$ constant sur $[0,1]$ relie x à x, donc $x \sim x$.

   → Transitivité
   Soit $x, y, z \in \mathrm{A}$ tels que $x \sim y$ et $y \sim z$. Soit $\gamma, \delta \in \mathcal{C}([0,1], \mathrm{A})$ deux arcs reliant respectivement x à y et y à z. On considère l'arc suivant

$$\beta = \gamma \bullet \delta : \begin{cases} [0,1] & \longrightarrow \mathrm{A} \\ x & \longmapsto \begin{cases} \gamma(2t) & \text{si } x \in \left[0, \frac{1}{2}\right] \\ \delta(2t-1) & \text{si } x \in \left[\frac{1}{2}, 1\right] \end{cases} \end{cases}.$$

   β est continu et $\beta(0) = x$ et $\beta(1) = z$, donc $x \sim z$.
2. Par définition, il existe toujours un arc continu liant deux éléments a, b de la même classe d'équivalence pour $\sim_{\mathrm{A}}$ car on a $a \sim_{\mathrm{A}} b$.

Proposition II.8.

1. Une réunion de parties de E connexes par arcs ayant un point en commun est connexe par arcs.
2. Soit $n \in \mathbb{N}^*$ et $\mathrm{E}_1, \ldots, \mathrm{E}_n$ des espaces métriques. Si pour tout $k \in [\![1; n]\!]$, A_k est une partie connexe par arcs de E_k, alors $\mathrm{A}_1 \times \cdots \times \mathrm{A}_n$ est une partie connexe par arcs de l'espace métrique produit $\mathrm{E}_1 \times \cdots \times \mathrm{E}_n$.

Démonstration.

1. Soit $n \in \mathbb{N}^*$ et $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une famille de parties connexes par arcs tels que pour tout $k \in [\![1; n]\!]$, $a \in \mathrm{A}_k$. Posons $\mathrm{A} = \bigcup\limits_{n\in\mathbb{N}} \mathrm{A}_n$. soit x et y deux points de A. Soit $k, \ell \in \mathbb{N}$ tels que $x \in \mathrm{A}_k$ et $y \in \mathrm{A}_\ell$. a et x sont deux éléments de l'ensemble connexe par arcs A_k, donc $x \sim_{\mathrm{A}_k} a$ donc $x \sim_{\mathrm{A}} a$ et a et y sont deux éléments de l'ensemble connexe par arcs A_ℓ, donc $a \sim_{\mathrm{A}_\ell} y$ donc $a \sim_{\mathrm{A}} y$. On en déduit donc par transitivité que $x \sim y$ et que finalement A est connexe par arcs.
2. Soit $a = (a_1, \ldots, a_n)$ et $b = (b_1, \ldots, b_n)$ deux éléments de $\mathrm{E}_1 \times \cdots \times \mathrm{E}_n$. Pour tout $k \in [\![1; k]\!]$, il existe un arc γ_k reliant a_k et b_k. Considérons l'arc sur $\mathrm{E}_1 \times \cdots \times \mathrm{E}_n$, $\gamma : t \longmapsto (\gamma_1(t), \ldots, \gamma_n(t))$. L'arc γ est continu et relie a et b, donc $\mathrm{A}_1 \times \cdots \times \mathrm{A}_n$ est bien connexe par arcs.

(cc) BY-NC-SA

Exercice II.9.

Soit Ω une union d'ouverts disjoints de $\mathbb{R}$. On écrit

$$\Omega = \bigcup_{n\in\mathbb{N}}]a_n, b_n[$$

avec pour tout $n \in \mathbb{N}$, $a_n < b_n$. Montrer que Ω n'est pas connexe par arcs.

Exercice II.10.

Soit $p \in \mathbb{N}^*$ et $z_1, \ldots, z_p \in \mathbb{C}$. Montrer que $\mathbb{C} \setminus \{z_1, \ldots, z_p\}$ est connexe par arcs.

Exercice II.11.

Soit $n \in \mathbb{N}^*$ et H un hyperplan de $\mathbb{C}^n$. Montrer que $\mathbb{C}^n \setminus \mathrm{H}$ est connexe par arcs.

Exercice II.12.

Soit $n \in \mathbb{N}^*$ et H un hyperplan de $\mathbb{R}^n$. Montrer que $\mathbb{R}^n \setminus \mathrm{H}$ n'est pas connexe par arcs.

Exercice II.13.

Soit Ω un ouvert connexe de $\mathbb{R}^n$. Montrer que Ω est connexe par arcs polygonaux, *i.e.* pour tous points x, y de E, x et y peuvent être reliés par un arc affine par morceaux.

III Application aux fonctions à variables réelles

Proposition III.1.

Soit A une partie de $\mathbb{R}$. Les propositions suivantes sont équivalentes.

1. A est connexe par arcs.
2. A est un intervalle.

Démonstration.

→ (2) ⇒ (1) Pour tout $a, b \in \mathrm{A}$ tel que $a < b$, si on considère l'application continue $f_{a,b} : x \longmapsto (b-a)x + a$, alors $f_{a,b}$ est un arc continu liant a à b donc $a \sim b$. A est donc connexe par arcs.

→ (1) ⇒ (2) Soit $x, y \in \mathrm{A}$ et $z \in]x, y[$. Soit γ un arc continu reliant x et y. Posons

$$\mathrm{S}_z = \{t \in [0,1],\ \gamma(t) \leq z\} = \gamma^{-1}([x, z]).$$

S_z est fermé car il s'agit de l'image réciproque d'un fermé par la fonction continue γ et de plus, S_z est non vide car $1 \in \mathrm{S}_z$, donc $c = \sup \mathrm{S}_z \in \mathrm{S}_z$, *i.e.* $\gamma(c) \leq z$ et $c < 1$ car $1 \notin \mathrm{S}_z$. De plus, pour tout $t \in]0, 1-c]$, $\gamma(c+t) > z$. En faisant tendre t vers 0, on obtient par continuité de γ que $\gamma(c) \geq z$, et alors $\gamma(c) = z$. On en déduit donc que $z \in \mathrm{A}$, donc $[x, y] \subset \mathrm{A}$ et finalement que A est un intervalle.

Conséquence. Si I est un intervalle et f une fonction continue de I dans $\mathbb{R}$, alors I est continu par arc donc $f(\mathrm{I})$ aussi, *i.e.* $f(\mathrm{I})$ est un intervalle.

BY-NC-SA

Proposition III.2.

Soit I un intervalle de $\mathbb{R}$ et $f \in \mathcal{C}(\mathrm{I}, \mathbb{R})$. Les propositions suivantes sont équivalentes.

1. f est injective.
2. f est strictement monotone.

Démonstration.

→ (2) ⇒ (1) Supposons que f est injective et donc dans perte de généralité strictement croissante. Pour tout $x, y \in \mathrm{I}$, si $x \neq y$, on suppose sans perte de généralité que $x > y$. Par stricte monotonie de f, on a que $f(x) > f(y)$ et alors $f(x) \neq f(y)$. f est donc bien injective,.

→ (1) ⇒ (2) Posons $\Delta = \{(x,y) \in \mathrm{I}^2,\ x < y\}$ et considérons l'application

$$\varphi : \begin{cases} \Delta & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto f(y) - f(x). \end{cases}$$

Le fait que f soit injective est équivalent au fait que φ ne s'annule pas sur Δ, donc φ est à valeurs dans $\mathbb{R}^*$. Δ est convexe donc connexe par arcs. En effet, si a et b sont deux points de Δ, l'arc continu

$$\gamma : \begin{cases} [0,1] & \longrightarrow \Delta \\ t & \longmapsto ta + (1-t)b \end{cases}$$

lie a et b dans Δ. Or φ est continue, $\varphi(\Delta)$ est alors une partie connexe par arcs de $\mathbb{R}^*$ et donc un intervalle. Deux cas se présentent alors.

- $\varphi(\Delta) \subset]0, +\infty[$ donc f est strictement croissante.
- $\varphi(\Delta) \subset]-\infty, 0[$ donc f est strictement décroissante.

On a donc bien le résultat voulu.

Proposition III.3.

Soit I un intervalle de $\mathbb{R}$ et f une application de I dans $\mathbb{R}$.

1. Si f est monotone, alors les propositions suivantes sont équivalentes.
   (a) f est continue.
   (b) $f(\mathrm{I})$ est un intervalle.
2. Si f est continue et injective, alors $f^{-1} : f(\mathrm{I}) \longrightarrow \mathrm{I}$ est continue.

Démonstration.

1. Supposons que f est monotone.
   → $(a) \Rightarrow (b)$ I est connexe par arcs et f est continue, donc $f(\mathrm{I})$ est aussi connexe par arcs, *i.e.* c'est un intervalle.
   → $(b) \Rightarrow (a)$ Supposons sans perte de généralité que f est croissante. Nous allons traiter uniquement le cas où f admet un point de discontinuité sur l'intérieur de I. Le lecteur pourra essayer de faire le cas d'une discontinuité aux bords de l'intervalle. Supposons que f admette un point de discontinuité x_0 sur l'intérieur de I. On note respectivement $f(x_0^+)$ et $f(x_0^-)$ les limites à droite et à gauche de x_0 de f. On a alors $f(x_0^-) \leq f(x_0) \leq f(x_0^+)$ et $f(x_0^-) < f(x_0^+)$.
   Soit $y_0 \in]f(x_0^-), f(x_0^+)[\setminus \{f(x_0)\} := \mathrm{A}_0$. On a pour tout $t \in \mathrm{I}$,
   - Si $t < x_0$, alors $f(t) \leq f(x_0^-)$ et donc $f(t) \neq y_0$.
   - Si $t = x_0$, alors $f(t) = f(x_0) \neq y_0$.

(cc) BY-NC-SA

- Si $t > x_0$, alors $f(t) \geq f(x_0^+)$, donc $f(t) \neq y_0$.

On en déduit que $y_0 \notin f(\mathrm{I})$ et y_0 est compris entre deux éléments de $f(\mathrm{I})$. $f(\mathrm{I})$ n'est donc pas un intervalle.

2. Il suffit d'appliquer la question précédente et la proposition III.2. f est une fonction continue sur l'intervalle I et injective, elle est donc strictement monotone d'après la proposition III.2. f^{-1} est alors aussi strictement monotone. De plus, on a $f^{-1}(f(\mathrm{I})) = \mathrm{I}$ et I et $f(\mathrm{I})$ sont des intervalles, donc d'après (1), f^{-1} est continue.

Exercice III.4.

Soit f une fonction de $[0,1]$ dans $[0,1]$ telle qu'il existe $p \in \mathbb{N}^*$, $\underbrace{f \circ \cdots \circ f}_{p \text{ fois}} = \mathrm{Id}$. Supposons que $f(0) = 0$ Montrer que $f = \mathrm{Id}$.

Exercice III.5.

Soit $f \in \mathcal{C}^1([0,1], \mathbb{R})$ telle que $f \times (1 - f') = 0$. Déterminer f.

Exercice III.6.

Montrer que $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes.

Exercice III.7.

Posons $\mathrm{S}(0,1) = \{z \in \mathbb{C},\ |z| = 1\}$. Montrer que $\mathrm{S}(0,1)$ n'est homéomorphe à aucune partie $\mathbb{R}$.

Exercice III.8.

Soit $f : \mathrm{S}(0,1) \longrightarrow \mathrm{S}(0,1)$ une fonction injective et continue. Montrer que f est un homéomorphisme.

(cc) BY-NC-SA

Correction de l'exercice II.9. :

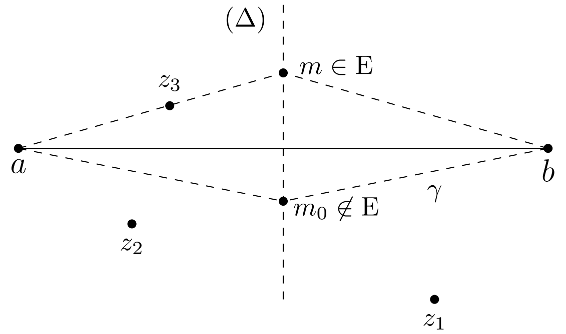
Méthode 1 : Soit $n, m \in \mathbb{N}$ deux entiers naturels différents, $x \in]a_n, b_n[$ et $y \in]a_m, b_m[$. Supposons par l'absurde que Ω soit connexe par arcs. Il existe un arc γ liant $x$et y. On suppose sans perte de généralité que $b_n \leq a_m$. On a $\gamma(0) = x < b_n < y = \gamma(1)$ et γ est continue. Par le théorème des valeurs intermédiaires, il existe $c \in]0, 1[$, $\gamma(c) = b_n$, mais $b_n \notin \Omega$ ce qui absurde car γ est à valeurs dans Ω par hypothèse.
Méthode 2 (plus simple) : En posant pour tout $k \in \mathbb{N}$, $\mathrm{A}_k =]a_k, b_k[$, on peut partitionner Ω en deux ouverts, par exemple $\mathrm{O}_1 = \mathrm{A}_0$ et $\mathrm{O}_2 = \bigcup_{k\geq 1} \mathrm{A}_k$. On a $\mathrm{O}_1 \cup \mathrm{O}_2 = \Omega$, $\mathrm{O}_1 \cap \mathrm{O}_2 = \varnothing$, $\mathrm{O}_1 \neq \varnothing$ et $\mathrm{O}_2 \neq \varnothing$. Ω n'est donc pas connexe et alors pas connexe par arcs.

Correction de l'exercice II.10. :

Soit $a, b \in \mathrm{S} = \mathbb{C} \setminus \{z_1, \ldots, z_p\}$. Considérons les deux ensembles suivants.

→ $\Delta = \{x \in \mathbb{C},\ |x - a| = |x - b|\}$ la médiatrice du segment $[a, b]$.

→ $\mathrm{E} = \{m \in \Delta,\ \exists k \in [\![1; p]\!],\ z_k \in [a, m] \cup [b, m]\}$.



$|\mathrm{E}| \leq p$, donc E est fini. On peut donc considérer $m_0 \in \Delta \setminus \mathrm{E}$ car cet ensemble est non vide. Il suffit alors de considérer l'arc

$$\gamma = \theta(a, m_0) \bullet \theta(m_0, b) : \begin{cases} [0,1] & \longrightarrow \mathrm{S} \\ t & \longmapsto \begin{cases} 2tm_0 + (1-2t)a & \text{si } t \in \left[0, \frac{1}{2}\right] \\ (2t-1)b + 2(1-t)m_0 & \text{si } t \in \left]\frac{1}{2}, 1\right] \end{cases} \end{cases}.$$

γ est un arc continu à valeurs dans S qui lie a à b en passant par m_0, donc S est bien connexe par arcs.

Correction de l'exercice II.11. :

Soit H un hyperplan de $\mathbb{C}^n$ et φ une forme linéaire non nulle telle que $\operatorname{Ker} \varphi = \mathrm{H}$. Posons $\mathrm{S} = \mathbb{C}^n \setminus \mathrm{H}$ et montrons que S est connexe par arcs. Soit $a, b \in \mathrm{S}$.

→ Si $\varphi(a) = \varphi(b)$, alors pour tout $\lambda \in [0, 1]$, $\varphi((1 - \lambda)a + \lambda b) = \varphi(a) \neq 0$. Il suffit alors de considérer l'arc $\gamma = \theta(a, b) : t \longmapsto (1 - t)a + tb$. Cet arc est bien à valeurs dans S et lie a et b.

→ Si $\varphi(a) \neq \varphi(b)$, alors on considère $\mathrm{K}_{a,b} = \{\lambda \in \mathbb{C},\ \varphi((1 - \lambda)a + \lambda b) = 0\}$.
$\varphi((1 - \lambda)a + \lambda b) = 0 \Longleftrightarrow \lambda = \dfrac{\varphi(a)}{\varphi(b) - \varphi(a)} := \lambda_{a,b}$, donc $\mathrm{K}_{a,b} = \{\lambda_{a,b}\}$.
D'après l'exercice précédent, on dispose d'un arc γ liant 0 et 1 dans $\mathbb{C} \setminus \{\lambda_{a,b}\}$. Considérons Γ l'arc défini par

$$\Gamma : \begin{cases} [0,1] & \longrightarrow \mathrm{S} \\ t & \longmapsto (1 - \gamma(t))a + \gamma(t)b. \end{cases}$$

Γ est bien à valeurs dans S car pour tout t, $\Gamma(t) \in \mathrm{H}$ impliquerait que $\gamma(t) = \lambda_{a,b}$ ce qui est impossible. Γ lie a à b dans S, S est donc bien connexe par arcs.

(cc) BY-NC-SA

Correction de l'exercice II.12. :

Soit H un hyperplan de $\mathbb{R}^n$ et φ une forme linéaire non nulle telle que $\operatorname{Ker}\varphi = \mathrm{H}$. Considérons les deux ensembles $\mathrm{A}_+ = \{x \in \mathbb{R}^n,\ \varphi(x) > 0\}$ et $\mathrm{A}_- = \{x \in \mathbb{R}^n,\ \varphi(x) < 0\}$. On a $\mathbb{R}^n \setminus \mathrm{H} = \mathrm{A}_+ \cup \mathrm{A}_-$. Soit $x \in \mathrm{A}_+$ et $y \in \mathrm{A}_-$. Supposons par l'absurde qu'il existe un arc γ liant x et y. L'application $\varphi \circ \gamma$ est continue et $\varphi \circ \gamma(0) = \varphi(x) > 0 > \varphi(y) = \varphi \circ \gamma(1)$, donc d'après le théorème des valeurs intermédiaires, il existe $c \in]0,1[$, $\varphi \circ \gamma(c) = 0$, *i.e.* $\gamma(c) \in \mathrm{H}$ ce qui est absurde. On en déduit donc que $\mathbb{R}^n \setminus \mathrm{H}$ n'est pas connexe par arcs.
Méthode alternative (plus simple) : A_+ et A_- sont deux ouverts disjoints non vides dont l'union fait $\mathbb{R}^n \setminus \mathrm{H}$, $\mathbb{R}^n \setminus \mathrm{H}$ n'est alors pas connexe donc pas connexe par arcs.

Correction de l'exercice II.13. :

Soit $a \in \Omega$. Considérons A la composante connexe par arcs polygonaux de a dans Ω. Montrer que Ω est connexe par arcs polygonaux revient à montrer que $\mathrm{A} = \Omega$. Ω étant connexe par arcs donc connexe, une idée serait de montrer que A est ouvert et fermé dans Ω.

- $\rightarrow$ Montrons que A est ouvert.
  Soit $x \in \mathrm{A}$. Il existe un arc polygonal γ reliant a et x. Ω étant ouvert, il existe $\varepsilon > 0$ tel que $\mathrm{B}(x,\varepsilon) \subset \Omega$. soit $y \in \mathrm{B}(x,\varepsilon)$. $\mathrm{B}(x,\varepsilon)$ est convexe, $\theta(x,y)$ est donc un arc continu polygonal (c'est un segment) reliant x à y dans $\mathrm{B}(x,\varepsilon)$ donc dans Ω. On en déduit donc que $\gamma \bullet \theta(x,y)$ est un arc polygonal qui relie a à y dans Ω, donc $y \in \mathrm{A}$ et alors $\mathrm{B}(x,\varepsilon) \subset \mathrm{A}$. A est donc ouvert.
- $\rightarrow$ Montrons que $\Omega \setminus \mathrm{A}$ est aussi ouvert.
  Soit $b \in \Omega \setminus \mathrm{A}$ et B la composante connexe par arcs de b. Par le raisonnement précédent, B est ouvert et contient b, c'est donc un voisinage de b inclus dans $\Omega \setminus \mathrm{A}$. $\Omega \setminus \mathrm{A}$ est donc ouvert. Ω est connexe et A est ouvert et fermé dans Ω, ce qui nous permet d'affirmer finalement que $\mathrm{A} = \Omega$.

Correction de l'exercice III.4. :

Soit $p \in \mathbb{N}^*$ tel que $\underbrace{f \circ \cdots \circ f}_{p \text{ fois}} = \mathrm{Id}$. Cette inégalité nous permet de dire que f est bijective sur un intervalle et alors par la proposition III.2., f est strictement monotone. $f(0) = 0$, donc f est strictement croissante. Supposons par l'absurde qu'il existe $x \in [0,1]$ tel que $f(x) < x$, on a alors

$$x > f(x) > f \circ f(x) > \cdots > \underbrace{f \circ \cdots \circ f}_{p \text{ fois}}(x) = x,$$

ce qui est absurde. De même, on obtient la même contradiction lorsqu'on suppose que $f(x) < x$. Donc a bien que $f = \mathrm{Id}$.

Correction de l'exercice III.5. :

Supposons que $f \neq 0$ et posons $\mathrm{A} = \{x \in [0,1],\ f'(x) = 1\}$. Montrons que A est fermé et ouvert dans $[0,1]$ et est donc égal à $[0,1]$ par connexité de $[0,1]$.

- $\rightarrow$ A est fermé car il s'agit de l'image réciproque du fermé $\{1\}$ par l'application continue f'.
- $\rightarrow$ Montrons que A est ouvert. Soit $a \in \mathrm{A}$.
  - Si $f(a) \neq 0$, alors par continuité de f, il existe un voisinage U de a où f est non nulle, et alors l'égalité $f \times (f' - 1) = 0$ entraîne que $f' = 1$ sur U, *i.e.* $\mathrm{U} \subset \mathrm{A}$.
  - Si $f(a) = 0$, alors on a

    $$f(a+h) = f'(a)h + o(h) = h\underbrace{(1 + o(1))}_{\varepsilon(h)}.$$

    $\varepsilon(h) \xrightarrow[h \to 0]{} 1$, donc il existe $\eta > 0$ tel que pour tout $h \in \mathrm{B}_f(0,\eta) \setminus \{0\}$, $|\varepsilon(h)| > \frac{1}{2}$ et alors

(cc) BY-NC-SA

pour tout $h \in \mathrm{B}_f(0,\eta) \setminus \{0\}$, $f(a+h) > \frac{h}{2} > 0$ et alors l'égalité $f \times (f'-1) = 0$ entraîne $f'(a+h) = 1$. On en déduit donc que $[a-h, a+h] \subset \mathrm{A}$, donc A est bien ouvert et finalement $\mathrm{A} = [0,1]$.

Finalement on déduit de ce qui précède que les fonctions $f \in \mathcal{C}^1([0,1],\mathbb{R})$ vérifiant $f \times (f'-1) = 0$ sont la fonction nulle et les fonctions de la forme $f : x \longmapsto x + \mathrm{C}$ avec $\mathrm{C} \in \mathbb{R}$.

Correction de l'exercice III.6. :

Supposons par l'absurde que $\mathbb{R}$ et $\mathbb{R}^2$ sont homéomorphes. Il existe alors une fonction f bijective continue de $\mathbb{R}$ dans $\mathbb{R}^2$ d'inverse continu. $\mathbb{R}^2 \setminus \{f(0)\}$ est connexe par arcs, f^{-1} est continue donc $f^{-1}(\mathbb{R}^2 \setminus \{f(0)\})$ est connexe par arcs d'après la proposition II.5 mais $f^{-1}(\mathbb{R}^2 \setminus \{f(0)\}) = \mathbb{R}^*$ et $\mathbb{R}^*$ n'est pas connexe par arcs, ce qui est absurde.

Correction de l'exercice III.7. :

Soit A une partie de $\mathbb{R}$. Supposons que A est homéomorphe à $\mathrm{S}(0,1)$, *i.e.* il existe un homéomorphisme $f : \mathrm{S}(0,1) \longrightarrow \mathrm{A}$. $\mathrm{S}(0,1)$ est connexe par arcs, donc $\mathrm{A} = f(\mathrm{S}(0,1))$ est compact (c'est l'image d'un compact par une application continue) connexe par arcs, *i.e.* c'est un intervalle fermé, on pose alors $\mathrm{A} = [a,b]$ avec $a < b$. Soit $c = \frac{a+b}{2}$ et $\mathrm{B} = \mathrm{S}(0,1) \setminus \{f^{-1}(c)\}$. B est connexe par arcs et f est continue, donc $f(\mathrm{B})$ est connexe par arcs. Or $f(\mathrm{B}) = \mathrm{A} \setminus \{c\}$ et $\mathrm{A} \setminus \{c\}$ n'est pas connexe par arcs ce qui est absurde.

Correction de l'exercice III.8. :

→ Si f est surjective, alors elle est bijective. Montrons que c'est bien un homéomorphisme. Pour tout fermé F de $\mathrm{S}(0,1)$, $(f^{-1})^{-1}(\mathrm{F}) = f(\mathrm{F})$ qui est fermé d'après le corollaire VI.2 du chapitre 11.5. On a montré que l'image réciproque de tout fermé de $\mathrm{S}(0,1)$ par f^{-1} est fermé, donc f^{-1} est continue et finalement f est un homéomorphisme.

→ Si f n'est pas surjective, alors posons $f(\mathrm{S}(0,1)) = \mathrm{S}' \subset \mathrm{S}(0,1)$. Quitte à composer par une rotation, on suppose que $-1 \notin f(\mathrm{S}(0,1))$. Considérons l'application

$$\mathrm{Arg} : \begin{cases} \mathrm{S}' & \longrightarrow]-\pi,\pi[\\ x+iy & \longmapsto 2\arctan\left(\dfrac{y}{x+1}\right) \end{cases}.$$

Soit $x+iy \in \mathrm{S}'$. Posons $\theta = \mathrm{Arg}(x+iy)$. On a

$$\cos(\theta) = \frac{1-\tan\left(\dfrac{\theta}{2}\right)^2}{1+\tan\left(\dfrac{\theta}{2}\right)^2} = \frac{1-\left(\dfrac{y}{x+1}\right)^2}{1+\left(\dfrac{y}{x+1}\right)^2} = \frac{x^2+2x+1-y^2}{x^2+2x+1+y^2} = \frac{2x^2+2x}{2x+2} = \frac{2x(x+1)}{2(x+1)} = x.$$

On peut montrer de la même manière que $\sin(\theta) = y$, donc pour tout $z = x+iy$ et $z' = x'+iy'$ dans $\mathrm{S}(0,1)$, $\mathrm{Arg}(z) = \mathrm{Arg}(z')$ donne en composant par cos et sin que $x = x'$ et $y = y'$, *i.e.* $z = z'$. Arg est donc bien injective.

Soit $g = \mathrm{Arg} \circ f$. g est une injection continue de $\mathrm{S}(0,1)$ dans $\mathbb{R}$ et $\mathrm{S}(0,1)$ est connexe par arcs, donc il existe $a, b \in \mathbb{R}$ tels que $a < b$ et $g(\mathrm{S}(0,1)) = [a,b]$ (c'est un intervalle fermé car il s'agit de l'image d'un compact par une application continue, *i.e.* un compact). g est donc une bijection continue de $\mathrm{S}(0,1)$ dans $[a,b]$ qui sont deux compacts, et alors par un raisonnement similaire au début de l'exercice, g est un homéomorphisme. On a montré que $\mathrm{S}(0,1)$ est homéomorphe à un segment de $\mathbb{R}$ ce qui est impossible d'après l'exercice précédent. On en déduit donc que f est bijective et que finalement que f est un homéomorphisme en se ramenant au premier cas.

BY-NC-SA

CHAPITRE 11.9

Convexité

Dans ce chapitre, on considère E un $\mathbb{R}-$espace vectoriel et A une partie de E.

I Enveloppe convexe

Définition I.1.

Posons $\mathrm{P_A} = \{\mathrm{C} \in \mathcal{P}(\mathrm{E}),\ \mathrm{A} \subset \mathrm{C} \text{ et } \mathrm{C} \text{ est convexe}\}$. L'enveloppe convexe de A est définie par

$$\mathrm{conv(A)} = \bigcap_{\mathrm{C}\in\mathrm{P_A}} \mathrm{C}.$$

Lemme I.2.

Soit C une partie convexe de E. On a

$$\forall p \in \mathbb{N}^*,\ \forall (z_1, \dots, z_p) \in \mathrm{C}^p,\ \forall (\lambda_1, \dots, \lambda_p) \in \mathbb{R}_+^p,\ \lambda_1 + \cdots + \lambda_p = 1 \Longrightarrow \sum_{k=1}^{p} \lambda_k z_k \in \mathrm{C}$$

Démonstration. Montrons la propriété par récurrence sur p.

→ Le cas $p = 1$ est évident et le cas $p = 2$ correspond à la définition de la convexité.

→ Soit $p \in \mathbb{N}^*$, supposons que la propriété soit vraie pour p. Soit $z_1, \dots, z_{p+1} \in \mathrm{C}$ et $\lambda_1, \dots, \lambda_{p+1} \in \mathbb{R}^+$ tels que $\lambda_1 + \cdots + \lambda_{p+1} = 1$. On a lorsque $\lambda_{p+1} \neq 1$,

$$\sum_{k=1}^{p+1} \lambda_k z_k = (1 - \lambda_{p+1}) \sum_{k=1}^{p} \underbrace{\frac{\lambda_k}{1 - \lambda_{p+1}}}_{\mu_k} z_k + \lambda_{p+1} z_{p+1}.$$

De plus,

$$\mu_1 + \cdots + \mu_p = \frac{\lambda_1 + \cdots + \lambda_p}{1 - \lambda_{p+1}} = \frac{1 - \lambda_{p+1}}{1 - \lambda_{p+1}} = 1.$$

Donc par hypothèse de récurrence, on a $\displaystyle\sum_{k=1}^{p} \mu_k z_k \in \mathrm{C}$, et alors en appliquant la définition de la convexité (*i.e.* la propriété pour $p = 2$), on a

$$\sum_{k=1}^{p+1} \lambda_k z_k = (1 - \lambda_{p+1}) \underbrace{\sum_{k=1}^{p} \frac{\lambda_k}{1 - \lambda_{p+1}} z_k}_{\in \mathrm{C}} + \lambda_{p+1} \underbrace{z_{p+1}}_{\in \mathrm{C}} \in \mathrm{C}.$$

On a donc bien le résultat voulu. Il reste le cas $\lambda_{p+1} = 1$ qui est évident.

(cc) BY-NC-SA

Proposition I.3.

1. L'intersection de deux ensembles convexes est convexe.
2. conv(A) est le plus petit convexe contenant A.
3. $\operatorname{conv}(A) = \left\{ x \in E,\ \exists p \in \mathbb{N}^*,\ \exists x_1, \dots, x_p \in A, \exists \lambda_1, \dots, \lambda_p \in \mathbb{R}^+,\ \middle| \begin{array}{c} \lambda_1 + \dots + \lambda_p = 1 \\ x = \displaystyle\sum_{k=1}^{p} \lambda_k x_k \end{array} \right\}$.

Démonstration.

1. Cette démonstration est laissée comme exercice au lecteur.
2. Par définition de conv(A), tout convexe contenant A contient conv(A), conv(A) est alors le plus petit contenant A.
3. Montrons l'égalité par double inclusion
   - $\to$ ($\supset$) $A \subset \operatorname{conv}(A)$ et conv(A) est convexe, donc pour tout $x_1, \dots, x_p \in A \subset \operatorname{conv}(A)$ et $\lambda_1, \dots, \lambda_p \in \mathbb{R}^+$ tels que $\lambda_1 + \dots + \lambda_p = 1$, $\displaystyle\sum_{k=1}^{p} \lambda_k x_k \in \operatorname{conv}(A)$.
   - $\to$ ($\subset$) Posons $B = \left\{ x \in E,\ \exists p \in \mathbb{N}^*,\ \exists x_1, \dots, x_p \in A, \exists \lambda_1, \dots, \lambda_p \in \mathbb{R}^+,\ \middle| \begin{array}{c} \lambda_1 + \dots + \lambda_p = 1 \\ x = \displaystyle\sum_{k=1}^{p} \lambda_k x_k \end{array} \right\}$.

     On a $A \subset B \subset \operatorname{conv}(A)$, donc il suffit de montrer que B est convexe pour montrer l'égalité, car conv(A) est le plus petit convexe au sens de l'inclusion contenant A.

     Soit $x = \displaystyle\sum_{k=1}^{p} \lambda_k x_k \in B$, $y = \displaystyle\sum_{k=1}^{p} \mu_k y_k \in B$ et $\lambda \in [0, 1]$. On a

     $$\lambda x + (1 - \lambda) y = \sum_{k=1}^{p} \lambda \lambda_k x_k + (1 - \lambda) \mu_k y_k.$$

     Posons

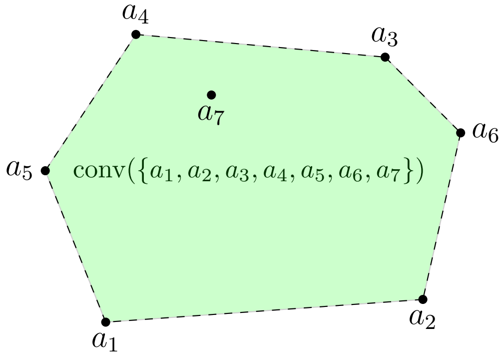
     $$\begin{cases} (z_1, \dots, z_{2p}) = (x_1, \dots, x_p, y_1, \dots, y_p) \\ (\delta_1, \dots, \delta_{2p}) = (\lambda\lambda_1, \dots, \lambda\lambda_p, (1-\lambda)\mu_1, \dots (1-\lambda)\mu_p). \end{cases}$$

     On a $(z_1, \dots, z_{2p}) \in A^{2p}$, $(\delta_1, \dots, \delta_{2p}) \in \mathbb{R}_+^{2p}$ et $\delta_1 + \dots + \delta_{2p} = 1$, donc

     $$\lambda x + (1 - \lambda) y = \sum_{k=1}^{2p} \delta_k z_k \in B.$$

     B est donc bien convexe, ce qui nous permet d'affirmer que conv(A) = B.

Exemple. On considère l'exemple suivant, où $\dim E = 2$. Le dessin suivant montre comment obtenir graphiquement l'enveloppe convexe des 7 points $a_1, a_2, a_3, a_4, a_5, a_6, a_7 \in E$.



(cc) BY-NC-SA

Exercice I.4.

Supposons que E est de dimension finie égale à d. Soit $p \geq d+1$ et $(a_0, \ldots, a_p) \in \mathrm{E}^{p+1}$. Montrer qu'il existe $(\alpha_0, \ldots, \alpha_p) \in \mathbb{R}^{p+1} \setminus \{(0, \ldots, 0)\}$ tel que

$$\sum_{k=0}^{p} \alpha_k a_k = 0 \text{ et } \alpha_0 + \cdots + \alpha_p = 0.$$

Le résultat de l'exercice ci-dessous est connu sous le nom de théorème de Carathéodory.

Exercice I.5.

Supposons que E soit de dimension finie égale à d. Soit $p \geq d+1$, $(a_0, \ldots, a_p) \in \mathrm{E}^{p+1}$, $(\lambda_0, \ldots, \lambda_p) \in \mathbb{R}_+^{p+1}$ tel que $\lambda_0 + \cdots + \lambda_p = 1$. Posons $b = \sum_{k=0}^{p} \lambda_k a_k$. Monter qu'il existe $j \in [\![0; p]\!]$ tel que

$$\exists (\mu_0, \ldots, \mu_{j-1}, \mu_{j+1}, \ldots, \mu_p) \in \mathbb{R}_+^p, \quad \sum_{k=0,\ k \neq j}^{p} \mu_k = 1 \text{ et } \sum_{k=0,\ k \neq j}^{p} \mu_k a_k = b.$$

Exercice I.6.

On suppose que E est de dimension finie. Soit K une partie compacte de E. Montrer que conv(K) est une partie compacte de E.

II Projection et séparation

Exercice II.1.

Soit C un convexe non vide fermé de l'espace euclidien E muni du produit scalaire $\langle ., . \rangle$.

1. Montrer que pour tout $a \in \mathrm{E}$ il existe un unique $b \in \mathrm{C}$ tel que
$$d(a, \mathrm{C}) = \|a - b\|.$$
2. Montrer que $\forall x \in \mathrm{C},\ \langle a - b,\ x - b \rangle \leq 0$.
3. On pose $b = p_{\mathrm{C}}(a)$. Montrer que p_{C} est $1-$lipschitzienne.

Exercice II.2.

Soit C un convexe non vide fermé de l'espace euclidien E muni du produit scalaire $\langle ., . \rangle$. Soit $a \notin \mathrm{C}$. Montrer qu'il existe un hyperplan affine de E qui sépare strictement C de a.

BY-NC-SA

Correction de l'exercice I.4. :

Posons pour tout $k \in [\![1;p]\!]$, $x_k = a_k - a_0$. La famille $(x_1, \ldots, x_p)$ est liée car le nombre de vecteurs de cette famille est supérieur à la dimension de E. Il existe donc $(\beta_1, \ldots, \beta_p) \in \mathbb{R}^p \setminus \{(0, \ldots, 0)\}$ tel que $\sum_{k=1}^{p} \beta_k x_k = 0$, *i.e.*

$$-\left(\sum_{k=1}^{p} \beta_k\right) a_0 + \beta_1 a_1 + \cdots + \beta_p a_p = 0.$$

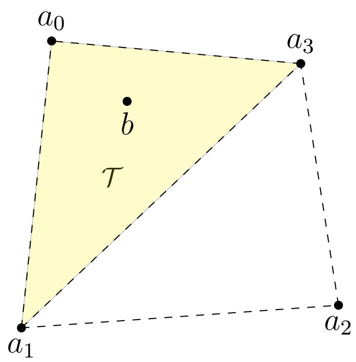
En posant $\alpha_0 = -\left(\sum_{k=1}^{p} \beta_k\right)$ et pour tout $k \in [\![1;p]\!]$, $\alpha_k = \beta_k$, on a

$$\alpha_0 + \cdots + \alpha_p = 0 \text{ et } \sum_{k=0}^{p} \alpha_k a_k = 0,$$

d'où le résultat voulu.

Correction de l'exercice I.5. : *(Théorème de Carathéodory)*

Remarquons qu'une autre manière de formuler ce théorème est la suivante.
Soit $p \geq d+1$, $b \in$ E et $a_0, \ldots, a_p \in$ E tel que $b \in \operatorname{conv}(\{a_0, \ldots, a_p\})$. Il existe alors $j \in [\![0;p]\!]$ tel que $b \in \operatorname{conv}(\{a_0, \ldots, a_{j-1}, a_{j+1}, \ldots, a_p\})$.
Pour avoir une intuition du résultat, observons le dessin suivant.



Dans le dessin ci-dessus, on regarde le cas $d = 2$ et $p = 3 \geq d+1$. b appartient au quadrilatère $a_1a_2a_3a_0$, *i.e.* $b \in \operatorname{conv}(\{a_0, a_1, a_2, a_3\})$. Il existe donc $\lambda_0, \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^+$ tels que

$$\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 = 1 \text{ et } b = \lambda_0 a_0 + \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3,$$

mais on voit aussi que b appartient au triangle jaune $\mathcal{T} = \operatorname{conv}(\{a_0, a_1, a_3\})$, il existe donc μ_0, μ_1, μ_3 tels que

$$\mu_0 + \mu_1 + \mu_3 = 1 \text{ et } b = \mu_0 a_0 + \mu_1 a_1 + \mu_3 a_3.$$

Remarquons que le point qu'on peut enlever à la combinaison convexe n'est pas unique : on aurait pu faire la même chose en enlevant le point a_1 et en considérant le triangle $\mathcal{T}' = \operatorname{conv}(\{a_0, a_3, a_2\})$.
Montrons à présent cette propriété dans le cas général.
On considère $\alpha_0, \ldots, \alpha_p$ $p+1$ réels vérifiant la propriété montrée dans l'exercice précédent, *i.e.*

$$\alpha_0 + \cdots + \alpha_p = 0 \text{ et } \sum_{k=0}^{p} \alpha_k a_k = 0.$$

Posons pour tout $k \in [\![0;p]\!]$ et $\tau \in \mathbb{R}$, $\nu_k(\tau) = \lambda_k - \tau\alpha_k$. On a alors pour tout $\tau \in \mathbb{R}$,

$$\sum_{k=0}^{p} \nu_k(\tau) a_k = \sum_{k=0}^{p} \lambda_k a_k - \tau \sum_{k=0}^{p} \alpha_k a_k = \sum_{k=0}^{p} \lambda_k a_k = b$$

(cc) BY-NC-SA

et

$$\sum_{k=0}^{p}\nu_k(\tau) = \sum_{k=0}^{p}\lambda_k - \tau\sum_{k=0}^{p}\alpha_k = 1.$$

On souhaite choisir τ de manière à ce que tous les $\nu_k(\tau)$ soient positifs et qu'au moins un seul parmi ces coefficients soit nul. Posons pour tout $k \in [\![0;p]\!]$, $\mathrm{A}_k = \{\tau \in \mathbb{R},\ \nu_k(\tau) \geq 0\}$. Les A_k sont fermés et non vides, car ils contiennent tous 0.
On a alors pour tout $k \in [\![0;p]\!]$,

→ Si $\alpha_k > 0$, alors $\mathrm{A}_k = \left]-\infty, -\frac{\lambda_k}{\alpha_k}\right]$.

→ Si $\alpha_k = 0$, alors $\mathrm{A}_k = \mathbb{R}$.

→ Sinon, $\mathrm{A}_k = \left[-\frac{\lambda_k}{\alpha_k}, +\infty\right[$.

Posons $\mathrm{S} = \bigcap_{k=0}^{p} \mathrm{A}_k$. On sait que $(\alpha_0, \ldots, \alpha_p) \neq (0, \ldots, 0)$ et $\sum_{k=0}^{p}\alpha_k = 0$. Il existe donc k et ℓ deux entiers dans $[\![0;p]\!]$ tels que $\alpha_k < 0$ et $\alpha_\ell > 0$. On peut donc affirmer que $\mathrm{S} \subset \left]-\infty, -\frac{\lambda_\ell}{\alpha_\ell}\right]$ et $\mathrm{S} \subset \left[-\frac{\lambda_k}{\alpha_k}, +\infty\right[$. S est donc une intersection d'intervalles majorée et minorée et contenant 0, il s'agit alors un segment borné non vide. S est une intersection finie de segments de la forme $\left[-\frac{\lambda_k}{\alpha_k}, +\infty\right[$ et $\left]-\infty, -\frac{\lambda_\ell}{\alpha_\ell}\right]$ avec $\alpha_\ell < 0$ et $\alpha_k > 0$, il y a donc au moins un $j \in [\![0;p]\!]$ tel que $-\frac{\lambda_j}{\alpha_j} \in \mathrm{S}$. Posons $\tau_0 = -\frac{\lambda_j}{\alpha_j}$.
On a alors pour tout $k \in [\![0;p]\!]$, $\nu_k(\tau_0) \geq 0$ et $\nu_j(\tau_0) = 0$. Finalement en posant pour tout $k \in [\![0;p]\!]$, $\mu_k = \nu_k(\tau_0)$,

$$\sum_{k=0,\ k\neq j}^{p} \mu_k a_k = b \text{ et } \sum_{k=0,\ k\neq j}^{p} \underbrace{\mu_k}_{\geq 0} = 1,$$

ce qui est bien le résultat recherché.
Une conséquence de ce théorème est la suivante : Soit $p \geq d+1$, $b \in \mathrm{E}$ et $a_0, \ldots, a_p \in \mathrm{E}$ tel que $b \in \operatorname{conv}(\{a_0, \ldots, a_p\})$. Il existe alors $\mathrm{S} \subset [\![0;p]\!]$ tel que $|\mathrm{S}| = d+1$ vérifiant $b \in \operatorname{conv}(\{a_k,\ k \in \mathrm{S}\})$. Ce résultat est facilement prouvable en enlevant de la combinaison convexe des points jusqu'à ce que la condition $p \geq d+1$ ne soit plus vérifiée. Ceci entraîne que si A est une partie de E, alors tout élément de conv(A) peut s'écrire comme combinaison convexe d'au plus $d+1$ éléments de A.

Correction de l'exercice I.6. :

Posons $d = \dim \mathrm{E}$ et considérons l'ensemble

$$\Lambda = \{(\lambda_0, \ldots, \lambda_d) \in \mathbb{R}_+^{d+1},\ \lambda_0 + \cdots + \lambda_d = 1\}$$

et considérons l'application

$$\varphi : \begin{cases} \mathrm{K}^{d+1} \times \Lambda & \longrightarrow \operatorname{conv}(\mathrm{K}) \\ ((a_k)_{k\in[\![0;d]\!]}, (\lambda_k)_{k\in[\![0;d]\!]}) & \longmapsto \displaystyle\sum_{k=0}^{p}\lambda_k a_k. \end{cases}$$

Le résultat énoncé à la fin de l'exercice précédent nous permet d'affirmer que tout élément de conv(K) s'écrit comme combinaison linéaire d'au plus $d+1$ éléments de K, *i.e.* s'écrit de la forme $\varphi((a_k)_{k\in[\![0;d]\!]}, (\lambda_k)_{k\in[\![0;d]\!]})$ avec $((a_k)_{k\in[\![0;d]\!]}, (\lambda_k)_{k\in[\![0;d]\!]}) \in \mathrm{K}^{d+1} \times \Lambda$. φ est donc surjective. De plus, Λ est un fermé borné en dimension finie, il s'agit donc un compact. $\mathrm{K}^{d+1} \times \Lambda$ est un produit de compacts, donc d'après la proposition III.1 du chapitre 11.5, il s'agit d'un compact. φ est continue, car il s'agit de produit et somme de fonctions continues, donc d'après la proposition VI.1 du chapitre 11.5, $\varphi(\mathrm{K}^{d+1} \times \Lambda)$ est compact, mais φ est surjective, donc $\varphi(\mathrm{K}^{d+1} \times \Lambda) = \operatorname{conv}(\mathrm{K})$. On en déduit donc finalement que conv(K) est compact.

(cc) BY-NC-SA

Correction de l'exercice II.1. : *(Projection sur un convexe fermé)*

1. Soit $a \in \mathrm{E}$.

   → Existence
   Posons $\delta = \inf_{x \in \mathrm{C}} \|x - a\| = d(a, \mathrm{C})$ et soit (x_n) une suite à valeurs dans C vérifiant $\|x_n - a\| \xrightarrow[n \to +\infty]{} \delta$. Pour n assez grand, $\|x_n - a\| \leq \delta + 1$, *i.e.* $x_n \in \mathrm{B}_f(a, \delta + 1)$. Quitte à extraire, on suppose que (x_n) est à valeurs dans $\mathrm{C} \cap \mathrm{B}_f(a, \delta + 1)$. $\mathrm{C} \cap \mathrm{B}_f(a, \delta + 1)$ est un fermé borné en dimension finie, c'est donc un compact. il existe donc une extractrice φ telle que $x_{\varphi(n)} \xrightarrow[n \to +\infty]{} b \in \mathrm{C}$. On a alors $\left\|x_{\varphi(n)} - a\right\| \xrightarrow[n \to +\infty]{} \|b - a\|$ et finalement $\|b - a\| = \delta = d(a, \mathrm{C})$.

   → Unicité
   Soit b et b' deux éléments de C tels que $\|a - b\| = \|a - b'\| = d(a, \mathrm{C})$. Supposons par l'absurde que $b \neq b'$. On a

$$\begin{aligned} 2d(a, \mathrm{C})^2 = \|\underbrace{a - b}_{u}\|^2 + \|\underbrace{a - b'}_{v}\|^2 &= 2\left(\left\|\frac{u + v}{2}\right\|^2 + \left\|\frac{u - v}{2}\right\|^2\right) \\ &= 2\left\|a - \frac{b + b'}{2}\right\|^2 + \frac{1}{2}\|b - b'\|^2 > 2d(a, \mathrm{C})^2, \end{aligned}$$

   ce qui est absurde, donc $b = b'$. On a donc bien l'unicité. La dernière inégalité est vraie car $\dfrac{b + b'}{2} \in \mathrm{C}$.

2. Posons pour tout $t \in]0, 1]$ et $x \in \mathrm{C}$, $x_t = tx + (1 - t)b$. On a alors pour tout $x \in \mathrm{C}$ et $t \in]0, 1]$, $x_t \in \mathrm{C}$, donc $\|a - x_t\|^2 \geq \|a - b\|^2$. De plus,

$$\begin{aligned} \|a - x_t\|^2 \geq \|a - b\|^2 &\iff \|t(a - x) + (1 - t)(a - b)\|^2 \geq \|a - b\|^2 \\ &\iff t^2\|a - x\|^2 + (t^2 - 2t + 1)\|a - b\|^2 + 2t(1 - t)\langle a - x,\, a - b\rangle \geq \|a - b\|^2 \\ &\iff t^2\|a - x\|^2 + (t^2 - 2t)\|a - b\|^2 + 2t(1 - t)\langle a - x,\, a - b\rangle \geq 0 \\ &\iff t\|a - x\|^2 + (t - 2)\|a - b\|^2 + 2(1 - t)\langle a - x,\, a - b\rangle \geq 0. \end{aligned}$$

   En faisant tendre t vers 0 on obtient que pour tout $x \in \mathrm{C}$,

$$\begin{aligned} \|a - x_t\|^2 \geq \|a - b\|^2 &\implies 2\langle a - x,\, a - b\rangle \geq 2\|a - b\|^2 \quad (*) \\ &\iff 2\langle a - b + b - x,\, a - b\rangle \geq 2\|a - b\|^2 \\ &\iff \langle b - x,\, a - b\rangle \geq 0 \\ &\iff \langle x - b,\, a - b\rangle \leq 0. \end{aligned}$$

   On a donc bien le résultat voulu.

3. Soient $a, a' \in \mathrm{E}$, $b = p_{\mathrm{C}}(a)$ et $b' = p_{\mathrm{C}}(a')$. D'après la question précédente on a

$$\langle a - b,\, b' - b\rangle \leq 0$$

   et

$$\langle a' - b',\, b - b'\rangle \leq 0 \ \textit{i.e.}\ \langle b' - a',\, b' - b\rangle \leq 0.$$

   En sommant ces deux inégalités, on obtient

$$\langle b' - a' + a - b,\, b' - b\rangle \leq 0 \ \textit{i.e.}\ \|b' - b\|^2 \leq \langle a' - a,\, b' - b\rangle.$$

(cc) BY-NC-SA

On a donc finalement

$$\|b'-b\|^2 \leq \langle a'-a,\ b'-b\rangle \underset{\substack{\uparrow \\ \text{Cauchy-Schwarz}}}{\leq} \|a'-a\| \times \|b'-b\|\,.$$

Si $b = b'$, alors

$$\|p_{\mathrm{C}}(a) - p_{\mathrm{C}}(a')\| = \|b-b'\| = 0 \leq \|a-a'\|\,.$$

Sinon, en simplifiant des deux côtés par $\|b'-b\|$ on trouve que

$$\|b'-b\| \leq \|a'-a\| \ \ i.e.\ \ \|p_{\mathrm{C}}(a) - p_{\mathrm{C}}(a')\| \leq \|a-a'\|\,,$$

ce qui est bien le résultat voulu.

Correction de l'exercice II.2. : *(Séparation par un hyperplan)*

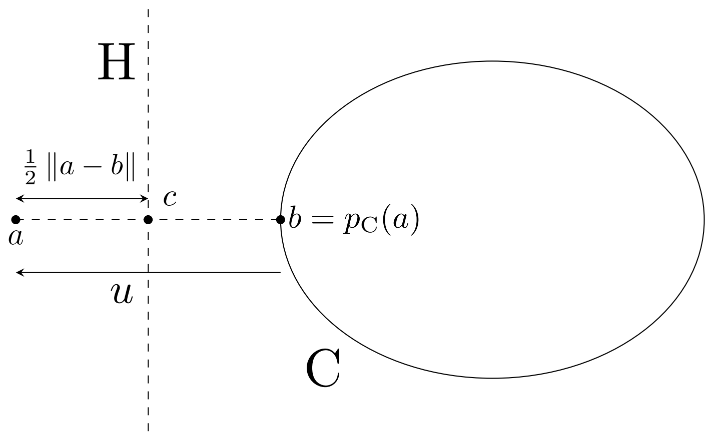
Tout hyperplan affine H de E s'écrit de la manière suivante.

$$\mathrm{H} = \{z \in \mathbb{E},\ \langle u,\ z\rangle = k\}$$

avec $k \in \mathbb{R}$ et $u \in \mathrm{E}$. L'hyperplan H sépare strictement a et C s'il vérifie la condition suivante.

$$\forall x \in \mathrm{C},\ \langle u,\ x\rangle < k \text{ et } \langle u,\ a\rangle > k.$$

Pour comprendre l'intuition derrière la solution de cet exercice, observons le dessin suivant en dimension 2.



Nous allons construire l'hyperplan H en nous aidant du dessin ci-dessus.
Posons $b = p_{\mathrm{C}}(a)$ la projection de a sur C, $u = a - b$ et $c = \dfrac{a+b}{2}$. Rappelons l'inégalité $(*)$ de l'exercice précédent :

$$\forall x \in \mathrm{C},\ \langle a-x,\ a-b\rangle \geq \|a-b\|^2\,.$$

On souhaite trouver un hyperplan orthogonal à $u = a - b$ séparant a et C. On a pour tout $x \in \mathrm{C}$,

$$\begin{aligned}
\langle a-x,\ a-b\rangle \geq \|a-b\|^2 &\iff \langle a-c+c-x,\ a-b\rangle \geq \|a-b\|^2 \\
&\iff \langle c-x,\ a-b\rangle \geq \frac{1}{2}\|a-b\|^2 \\
&\iff \langle x,\ u\rangle \leq \langle c,\ u\rangle - \frac{1}{2}\|a-b\|^2 < \langle c,\ u\rangle\,.
\end{aligned}$$

De plus, on a

$$\langle a-c,\ a-b\rangle = \left\langle \frac{a-b}{2},\ a-b\right\rangle = \frac{1}{2}\|a-b\|^2$$

(cc) BY-NC-SA

i.e.

$$\langle a,\ u\rangle = \langle c,\ u\rangle + \frac{1}{2}\left\|a-b\right\|^2 > \langle c,\ u\rangle\,.$$

En prenant donc $k = \langle c,\ u\rangle$ et $\mathrm{H} = \{z \in \mathbb{E},\ \langle u,\ z\rangle = k\}$. On a bien

$$\forall x \in \mathrm{C},\ \langle u,\ x\rangle < k \text{ et } \langle u,\ a\rangle > k.$$

H sépare donc bien a de C strictement.

(cc) BY-NC-SA

CHAPITRE 12

Suites de fonctions

Dans ce chapitre, on considère X un espace métrique muni d'une distance d, E un $\mathbb{R}-$espace vectoriel normé muni d'une norme $\|.\|$ et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de X dans E.

I Convergences simple et uniforme

1. Premières définitions

Définition I.1.

Soit f une fonction de X dans E. On dit que la suite (f_n) converge simplement vers f lorsque pour tout $x \in$ X, $f_n(x) \xrightarrow[n\to+\infty]{} f(x)$. Dans ce cas, on appelle f la limite simple de la suite (f_n).

Remarque. Bien évidement, une formulation équivalente de cette définition découlant de la définition de la limite est

$$\forall x \in \mathrm{X},\ \forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \|f_n(x) - f(x)\| < \varepsilon.$$

Exercice I.2.

Déterminer la limite simple de la suite de fonctions (f_n) de $\mathbb{R}$ (muni de la distance $(x, y) \longmapsto |x - y|$) dans $\mathbb{R}$ (muni de la norme $|.|$) définie par

$$\forall n \in \mathbb{N},\ \forall x \in \mathbb{R},\ f_n(x) = \frac{1 - x^{2n}}{1 + x^{2n}}.$$

Définition I.3.

Soit f une fonction de X dans E. Les propositions suivantes sont équivalentes.

1. $\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \forall x \in \mathrm{X},\ \|f_n(x) - f(x)\| < \varepsilon.$
2. $\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \sup\limits_{x\in\mathrm{X}} \|f_n(x) - f(x)\| < \varepsilon.$
3. $f_n - f$ est bornée à partir d'un certain rang et $\|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0.$

Lorsque ces propositions sont vérifiées, on dit que la suite (f_n) converge uniformément vers f sur X. Dans ce cas, on appelle f la limite uniforme de (f_n).

Remarques.

→ Bien entendu, si (f_n) converge uniformément vers f, alors (f_n) converge simplement vers f.

→ En pratique, pour trouver la limite uniforme d'une suite (f_n), on commence par trouver sa limite simple si elle existe en étudiant la suite $(f_n(x))_{n\in\mathbb{N}}$ pour tout $x \in$ X et voir ensuite s'il s'agit d'une limite uniforme.

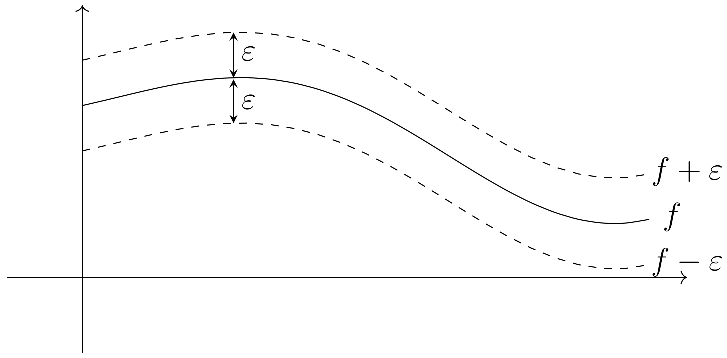
→ La négation de la convergence uniforme vers f s'écrit

$$\sup_{x\in\mathrm{X}} \|f_n(x) - f(x)\| \nrightarrow 0 \quad i.e. \quad \exists \delta > 0,\ \sup_{x\in\mathrm{X}} \|f_n(x) - f(x)\| \geq \delta \text{ pour une infinité de } n.$$

$$i.e. \quad \exists \delta > 0,\ \exists (x_n) \in \mathrm{X}^{\mathbb{N}},\ \|f_n(x_n) - f(x_n)\| \geq \delta \text{ pour une infinité de } n.$$

BY-NC-SA

→ L'interprétation de la convergence uniforme lorsque $X = \mathbb{R}$ et $E = \mathbb{R}$ est la suivante : si (f_n) converge uniformément vers f, alors pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que toutes les fonctions f_n d'indice $n \geq N$ sont comprises dans le tube de largeur 2ε autour de f.



Exercice I.4.

Soit $(f_n) \in \mathcal{C}([0,1],\mathbb{R})^{\mathbb{N}}$. Montrer que si (f_n) converge uniformément vers une fonction f, alors f et bornée et $\sup\limits_{x\in[0,1]} f_n(x) \xrightarrow[n\to+\infty]{} \sup\limits_{x\in[0,1]} f(x)$.

2. Opérations simples

Proposition I.5.

Soit f une fonction de X dans $(E, +, \ .\ , \times)$ une algèbre munie d'une norme d'algèbre $\|.\|$. Les propositions suivantes sont vraies.

1. Soit $p \in \mathbb{N}^*$ et $(A_k)_{k\in[\![1;p]\!]}$ p parties de X. Si la suite (f_n) converge uniformément vers f sur A_k pour tout $k \in [\![1;p]\!]$, alors (f_n) converge uniformément vers f sur $A_1 \cup \cdots \cup A_p$.
2. Si (f_n) converge uniformément vers une fonction f et (g_n) est une suite de fonctions qui converge uniformément vers une fonction g, alors pour tout $\alpha, \beta \in \mathbb{R}$, $(\alpha f_n + \beta g_n)_{n\in\mathbb{N}}$ converge uniformément vers $\alpha f + \beta g$.
3. Si g est une fonction bornée sur X et (f_n) converge uniformément vers une fonction f, alors $(g \times f_n)_{n\in\mathbb{N}}$ converge uniformément vers $g \times f$.
4. Si (f_n) est une suite de fonctions bornées qui converge uniformément vers f, alors f est bornée et la suite (f_n) est uniformément bornée.
5. Si de plus (g_n) est une suite de fonctions bornées qui converge uniformément vers une fonction g, alors la suite $(f_n \times g_n)$ converge uniformément vers $f \times g$.

Démonstration.

1. Soit $\varepsilon > 0$. Pour tout $k \in [\![1;p]\!]$, (f_n) converge uniformément vers f sur A_k. Il existe donc $N_k \in \mathbb{N}$ tel que pour tout $x \in A_k$ et $n \geq N_k$, $\|f_n(x) - f(x)\| < \varepsilon$. On en déduit donc qu'en posant $N = \max(N_1, \ldots, N_p)$, pour tout $x \in A_1 \cup \cdots \cup A_p$ et $n \geq N$, $\|f_n(x) - f(x)\| < \varepsilon$ et alors (f_n) converge uniformément vers f sur $A_1 \cup \cdots \cup A_p$.
2. Si $\alpha = 0$ ou $\beta = 0$, le résultat est évident. Supposons donc que α et β sont non nuls. Soit $\varepsilon > 0$. Il existe $N_1 \in \mathbb{N}$ tel que pour tout $n \geq N_1$ et $x \in X$, $\|g_n(x) - g(x)\| < \dfrac{\varepsilon}{2\,|\beta|}$. De même, il existe $N_2 \in \mathbb{N}$ tel que pour tout $n \geq N_2$ et $x \in X$, $\|f_n(x) - f(x)\| < \dfrac{\varepsilon}{2\,|\alpha|}$. On en déduit alors que pour

(cc) BY-NC-SA

tout $n \geq \max(\mathrm{N}_1, \mathrm{N}_2)$ et $x \in \mathrm{X}$,

$$\begin{aligned}\|\alpha f_n(x) + \beta g_n(x) - \alpha f(x) - \beta g(x)\| &\leq |\alpha| \, \|f_n(x) - f(x)\| + |\beta| \, \|g_n(x) - g(x)\| \\ &< |\alpha| \frac{\varepsilon}{2|\alpha|} + |\beta| \frac{\varepsilon}{2|\beta|} = \varepsilon.\end{aligned}$$

On conclut donc que la suite de fonctions $(\alpha f_n + \beta g_n)$ converge bien uniformément vers $(\alpha f + \beta g)$

3. Soit $\mathrm{M} > 0$ un majorant de $|g|$. Soit $\varepsilon > 0$ et $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$ et $x \in \mathrm{X}$, $\|f_n(x) - f(x)\| < \dfrac{\varepsilon}{\mathrm{M}}$. On a alors pour tout $n \geq \mathrm{N}$ et $x \in \mathrm{X}$,

$$\|g(x)f_n(x) - g(x)f(x)\| \leq \|g(x)\| \, \|f_n(x) - f(x)\| < \mathrm{M}\frac{\varepsilon}{\mathrm{M}} = \varepsilon.$$

et donc la suite de fonctions $(g \times f_n)$ converge bien uniformément vers $g \times f$.

4. (f_n) converge uniformément vers f, donc en utilisant la définition de l'uniforme convergence pour $\varepsilon = 1$, on obtient

$$\exists \mathrm{N} \in \mathbb{N}, \ \forall n \geq \mathrm{N} \ \forall x \in \mathrm{X}, \ \|f_n(x) - f(x)\| < 1.$$

On en déduit donc qu'en considérant $\mathrm{N} \in \mathbb{N}$ vérifiant la propriété ci-dessus, on a

$$\|f(x)\| \leq \|f(x) - f_{\mathrm{N}}(x)\| + \|f_{\mathrm{N}}(x)\| \leq 1 + \mathrm{A},$$

où A est un majorant de $\|f_{\mathrm{N}}\|$. f est donc bien bornée. Montrons maintenant que (f_n) est uniformément bornée. Soit $\mathrm{M}_1, \ldots, \mathrm{M}_{\mathrm{N}}$ des majorants respectifs de $\|f_1\|, \ldots, \|f_{\mathrm{N}}\|$. Pour tout $n \in \mathbb{N}$ et $x \in \mathrm{X}$

$$\|f_n(x)\| \leq \max(\|f_n(x) - f(x)\| + \|f(x)\|, \mathrm{M}_1, \ldots, \mathrm{M}_{\mathrm{N}}) \leq \max(2 + \mathrm{A}, \mathrm{M}_1, \ldots, \mathrm{M}_{\mathrm{N}}),$$

et alors la suite de fonctions (f_n) est bien uniformément bornée.

5. D'après le point précédent, les deux fonctions f et g sont bornées. Considérons alors M un majorant strictement positif de $\|f\|$. On sait également d'après le point précédent que la suite (g_n) est uniformément bornée, ce qui nous permet de considérer A un majorant strictement positif de $(\|g_n\|)$. Soit $\varepsilon > 0$ et soit $\mathrm{N}_1 \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}_1$ et $x \in \mathrm{X}$, $\|f_n(x) - f(x)\| < \dfrac{\varepsilon}{2\mathrm{A}}$. Considérons également $\mathrm{N}_2 \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}_2$ et $x \in \mathrm{X}$, $\|g_n(x) - g(x)\| < \dfrac{\varepsilon}{2\mathrm{M}}$. On a alors pour tout $n \geq \max(\mathrm{N}_1, \mathrm{N}_2)$ et $x \in \mathrm{X}$,

$$\begin{aligned}\|f_n(x)g_n(x) - f(x)g(x)\| &= \|f_n(x)g_n(x) - f(x)g_n(x) + f(x)g_n(x) - f(x)g(x)\| \\ &\leq \|g_n(x)\| \, \|f_n(x) - f(x)\| + \|f(x)\| \, \|g_n(x) - g(x)\| \\ &< \mathrm{A}\frac{\varepsilon}{2\mathrm{A}} + \mathrm{M}\frac{\varepsilon}{2\mathrm{M}} = \varepsilon.\end{aligned}$$

La suite $(f_n \times g_n)$ converge donc bien uniformément vers $f \times g$.

Exercice I.6.

Étudier la convergence simple puis la convergence uniforme des suites de fonctions réelles définies de la manière suivante.

1. $\forall n \in \mathbb{N}, \ \forall x \in [0,1], \ f_n(x) = n^a x^n (1-x)$ avec $a \in \mathbb{R}$.
2. $\forall n \in \mathbb{N}, \ \forall x \in \mathbb{R}^+, \ f_n(x) = n^a x e^{-nx}$ avec $a \in \mathbb{R}$.
3. $\forall n \in \mathbb{N}, \ \forall x \in \mathbb{R}, \ f_n(x) = \dfrac{x^n}{n!} e^{-x}$.

(cc) BY-NC-SA

II Transfert

1. Convergence simple

Proposition II.1.

Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$ telle que (f_n) converge simplement vers f. Les propositions suivantes sont vraies.

1. Si pour tout $n \in \mathbb{N}$, f_n est croissante, alors f est croissante.
2. Si pour tout $n \in \mathbb{N}$, f_n est convexe, alors f est convexe.

Démonstration.

1. Pour tout $n \in \mathbb{N}$ et $x, y \in \mathbb{R}$ tel que $x \leq y$, on a $f_n(x) \leq f_n(y)$. En passant à la limite lorsque n tend vers l'infini des deux côtés, on obtient $f(x) \leq f(y)$ ce qui signifie que f est croissante.
2. Pour tout $n \in \mathbb{N}$, $x, y \in \mathbb{R}$ et $\lambda \in [0,1]$, on a $f_n(\lambda x+(1-\lambda)y) \leq \lambda f_n(x)+(1-\lambda)f_n(y)$. En passant à la limite lorsque n tend vers l'infini des deux côtés de l'inégalité, on obtient que $f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)$. On en déduit donc que f est bien convexe.

Remarque. Bien entendu, les deux propositions ci-dessous restent vraies lorsqu'on remplace croissance par décroissance et convexité par concavité.

Théorème (Premier théorème de Dini) II.2.

Supposons que (f_n) est une suite de fonctions continues de X dans $\mathbb{R}$ et X est **compact**. Soit f une fonction continue de X dans $\mathbb{R}$. Supposons de plus que pour tout $x \in \mathrm{X}$, la suite $(f_n(x))$ est croissante convergente vers $f(x)$, alors (f_n) converge uniformément vers f.

Démonstration. Ce résultat a été prouvé au chapitre 11.5 (Compacité). Il s'agit du corollaire VII.4.

2. Convergence uniforme

a) Continuité

Proposition II.3.

Soit f une fonction de X dans E et $a \in \mathrm{X}$. Si les propriétés suivantes sont vérifiées,

→ Pour tout $n \in \mathbb{N}$, f_n est continue en a.

→ Il existe U un voisinage de a tel que $(f_n|_{\mathrm{U}})$ converge uniformément vers $f|_{\mathrm{U}}$ sur U.

alors f est continue en a.

Démonstration. Soit $\varepsilon > 0$. Soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$ et $x \in \mathrm{U}$, $\|f_n(x) - f(x)\| < \varepsilon$. f_{N} est continue en a, on peut donc considérer $\eta > 0$ tel que $\mathrm{B}(a,\eta) \subset \mathrm{U}$ et pour tout $x \in \mathrm{B}(a,\eta)$, $\|f_{\mathrm{N}}(x) - f_{\mathrm{N}}(a)\| < \varepsilon$. On a alors pour tout $x \in \mathrm{B}(a,\eta)$

$$\|f(x) - f(a)\| \leq \|f(x) - f_{\mathrm{N}}(x)\| + \|f_{\mathrm{N}}(x) - f_{\mathrm{N}}(a)\| + \|f_{\mathrm{N}}(a) - f(a)\| < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.$$

On en déduit donc que f est bien continue en a.

Remarque. Une conséquence directe de cette proposition est la suivante. Si pour tout $n \in \mathbb{N}$, f_n est continue et si (f_n) converge uniformément dans X vers une fonction f, alors f est continue.

(cc) BY-NC-SA

b) Interversion de limites

Rappel II.4.

Soit A une partie de X, $a \in \overline{\mathrm{A}}$ et $\lambda \in \mathrm{E}$. On dit que $f(x)$ tend vers λ lorsque x tend vers a selon A si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \mathrm{A},\ d(x,a) < \eta \Longrightarrow \|f(x) - \lambda\| < \varepsilon.$$

Dans ce cas, on écrit $f(x) \underset{\substack{x \to a \\ x \in \mathrm{A}}}{\longrightarrow} \lambda$.

Proposition II.5.

Soit A une partie de X, f une fonction de X dans E et $a \in \overline{\mathrm{A}}$. Supposons que E est complet et que pour tout $n \in \mathbb{N}$, il existe $\lambda_n \in \mathrm{E}$ tel que $f_n(x) \underset{\substack{x \to a \\ x \in \mathrm{A}}}{\longrightarrow} \lambda_n$. Supposons de plus que (f_n) converge uniformément dans A vers f. Il existe $\lambda \in \mathrm{E}$ tel que

$$\lambda_n \underset{n \to +\infty}{\longrightarrow} \lambda \text{ et } f(x) \underset{\substack{x \to a \\ x \in \mathrm{A}}}{\longrightarrow} \lambda.$$

Démonstration. Écrivons la définition de la convergence uniforme de (f_n) vers f.

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n \geq \mathrm{N},\ \forall x \in \mathrm{X},\ \|f_n(x) - f(x)\| < \varepsilon.$$

On en déduit donc que

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n, m \geq \mathrm{N},\ \forall x \in \mathrm{A},\ \|f_n(x) - f_m(x)\| \leq \|f_n(x) - f(x)\| + \|f_m(x) - f(x)\| < 2\varepsilon.$$

En faisant tendre x vers a, cette proposition devient

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall m, n \geq \mathrm{N},\ \|\lambda_n - \lambda_m\| \leq 2\varepsilon.$$

On en déduit alors que la suite (λ_n) est de Cauchy. E étant complet, on peut affirmer que (λ_n) converge vers un élément λ de E. Montrons alors que $f(x) \underset{\substack{x \to a \\ x \in \mathrm{A}}}{\longrightarrow} \lambda$. Soit $\varepsilon > 0$. Soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $\|\lambda_n - \lambda\| < \varepsilon$ et $\eta > 0$ tel que pour tout $x \in \mathrm{A} \cap \mathrm{B}(a, \eta)$, $\|f_\mathrm{N}(x) - \lambda_\mathrm{N}\| < \varepsilon$. On a alors pour tout $x \in \mathrm{B}(a, \eta) \cap \mathrm{A}$,

$$\|f(x) - \lambda\| \leq \|f(x) - f_\mathrm{N}(x)\| + \|f_\mathrm{N}(x) - \lambda_\mathrm{N}(x)\| + \|\lambda_\mathrm{N} - \lambda\| \leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.$$

On en déduit donc qu'on a bien $f(x) \underset{\substack{x \to a \\ a \in \mathrm{A}}}{\longrightarrow} \lambda$.

Remarques.

→ Dans le cas général, on utilisera cette propriété dans le cas $\mathrm{X} = \mathrm{A} = \mathbb{R}$ ou alors $\mathrm{X} = \mathbb{R}$ et A un voisinage de a. On pourra dans ce cas tout simplement écrire $f(x) \underset{x \to a}{\longrightarrow} \lambda$ au lieu de $f(x) \underset{\substack{x \to a \\ a \in \mathrm{A}}}{\longrightarrow} \lambda$.

→ Cette proposition montre qu'on peut faire l'inversion de limites suivante.

$$\lim_{n \longrightarrow +\infty} \lim_{\substack{x \longrightarrow a \\ x \in \mathrm{A}}} f_n(x) = \lim_{n \longrightarrow +\infty} \lambda_n = \lambda = \lim_{\substack{x \longrightarrow a \\ x \in \mathrm{A}}} f(x) = \lim_{\substack{x \longrightarrow a \\ x \in \mathrm{A}}} \lim_{n \longrightarrow +\infty} f_n(x).$$

→ On peut étendre de la même manière cette propriété au cas $a = +\infty$. Le lecteur est encouragé à essayer de le faire.

(cc) BY-NC-SA

c) Intégration

Proposition II.6.

Soit $a, b \in \mathbb{R}$ tels que $a \leq b$. Si pour tout $n \in \mathbb{N}$, $f_n \in \mathcal{C}([a,b],\mathbb{C})$ est continue par morceaux et (f_n) converge uniformément vers une fonction $f \in \mathcal{C}([a,b],\mathbb{C})$ de X dans E continue par morceaux, alors

$$\int_a^b f_n(x)dx \xrightarrow[n\to+\infty]{} \int_a^b f(x)dx \text{ et } \int_a^b |f_n(x)-f(x)|^2\, dx \xrightarrow[n\to+\infty]{} 0.$$

Démonstration. On a

$$\left|\int_a^b f_n(x)dx - \int_a^b f(x)dx\right| \leq \int_a^b |f_n(x)-f(x)|\, dx \leq (b-a)\, \|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0.$$

De la même manière, on a également

$$\int_a^b |f_n(x)-f(x)|^2\, dx \leq (b-a)\, \|f_n - f\|_\infty^2 \xrightarrow[n\to+\infty]{} 0.$$

Remarque.

→ On peut également montrer cette propriété dans le cas où on remplace $[a,b]$ par une union finie d'intervalles bornée.

→ Attention, cette propriété est en général fausse lorsque I n'est pas borné. En effet, considérons le cas où $\mathrm{I} = [0,+\infty[$ et pour tout $n \in \mathbb{N}^*$ et $x \in \mathbb{R}$, $f_n(x) = \dfrac{x^n}{n!}e^{-x}$. On a $\displaystyle\int_0^{+\infty} e^{-x}dx = 1$ et pour tout $n \in \mathbb{N}$,

$$\int_0^{+\infty} x^{n+1}e^{-x}dx = \underbrace{\left[-x^{n+1}e^{-x}\right]_0^{+\infty}}_{=0} + (n+1)\int_0^{+\infty} x^n e^{-x}dx = \cdots = (n+1)!\int_0^{+\infty} e^{-x} = (n+1)!.$$

On en déduit alors que pour tout $n \in \mathbb{N}$,

$$\int_0^{+\infty} f_n(x)dx = \int_0^{+\infty} \frac{x^n}{n!}e^{-x} = 1.$$

De plus, on sait d'après l'exercice I.6 que (f_n) converge uniformément vers 0. On en déduit donc qu'on a

$$\begin{cases} f_n \text{ converge uniformément vers } 0 \\ \displaystyle\int_0^{+\infty} f_n(x) \xrightarrow[n\to+\infty]{} 1, \end{cases}$$

ce qui montre que l'énoncé de la proposition II.6 pour $\mathrm{I} = [0,+\infty[$ est faux.

Notation. On note pour toute union finie d'intervalles I, $\mathcal{C}_{pm}(\mathrm{I},\mathbb{C})$ l'ensemble des fonctions continues par morceaux de I dans $\mathbb{C}$.

(cc) BY-NC-SA

Théorème (Théorème de convergence dominée) II.7.

Soit I un intervalle et $(f_n) \in \mathcal{C}_{pm}(\mathrm{I}, \mathbb{C})^{\mathbb{N}}$. Si (f_n) vérifie les propriétés suivantes,

→ (f_n) converge simplement vers une fonction f continue par morceaux sur I.

→ $\exists \varphi \in \mathrm{L}^1(\mathrm{I}, \mathbb{C}),\ \forall n \in \mathbb{N},\ \forall x \in \mathrm{I},\ |f_n(x)| \leq |\varphi(x)|$.

alors pour tout $n \in \mathbb{N}$, $f_n \in \mathrm{L}^1(\mathrm{I}, \mathbb{C})$, $f \in \mathrm{L}^1(\mathrm{I}, \mathbb{C})$ et $\displaystyle\int_{\mathrm{I}} f_n(x)dx \xrightarrow[n\to+\infty]{} \int_{\mathrm{I}} f(x)dx$.

Démonstration. La démonstration de ce résultat a été faite au chapitre 17 (Intégration II).

Remarque. Ici, l'intervalle I n'est pas forcément borné.

d) Dérivation

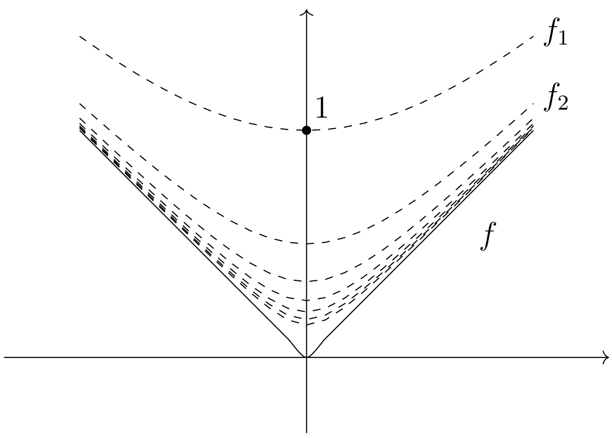
Avant d'énoncer les résultats de cette partie, attirons l'attention du lecteur sur le fait que la convergence uniforme d'une suite de fonctions dérivables n'implique pas forcément que la limite uniforme est dérivable. En effet, en considérant la suite de fonctions (f_n) définie par

$$\forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R},\ f_n(x) = \sqrt{x^2 + \frac{1}{n^2}},$$

on a clairement que (f_n) converge uniformément vers $f : x \longmapsto |x|$. En effet, pour tout $n \in \mathbb{N}^*$ et $x \in \mathbb{R}$,

$$0 \leq f_n(x) - \sqrt{x^2} = \sqrt{x^2 + \frac{1}{n^2}} - \sqrt{x^2} \underset{(*)}{\leq} \sqrt{x^2 + \frac{1}{n^2} - x^2} = \frac{1}{n}.$$

L'inégalité $(*)$ peut être facilement démontrée en utilisant le fait que pour tout $a, b \in \mathbb{R}^+$ tels que $a \geq b$, $(a-b)^2 \leq a^2 - b^2$. On en déduit donc que $\|f_n - f\|_\infty \leq \dfrac{1}{n} \xrightarrow[n\to+\infty]{} 0$ et donc (f_n) converge uniformément vers f et pour tout $n \in \mathbb{N}^*$, f_n est dérivable sur $\mathbb{R}$ mais f n'est pas dérivable en 0.



Dans la figure ci-dessus, les termes de la suite (f_n) sont en pointillés et la fonction valeur absolue est en traits pleins. On voit bien qu'à la limite, un point anguleux se forme en 0 alors que tous les termes de la suite (f_n) sont dérivables (lisses) en 0.

Pour pouvoir obtenir la dérivabilité à partir de la convergence uniforme, il faut ajouter des hypothèses supplémentaires.

Notation. Pour toute fonction f de A dans B (où A et B sont deux ensembles quelconques) et $\mathrm{C} \subset \mathrm{A}$, on note $\|f\|_{\infty,\mathrm{C}} = \sup\limits_{x \in \mathrm{C}} f(x)$.

(cc) BY-NC-SA

Proposition II.8.

Soit I un intervalle de $\mathbb{R}$ et g une fonction de I dans $\mathbb{R}$. Supposons que $(f_n) \in \mathcal{C}^1(\mathrm{I}, \mathbb{R})^{\mathbb{N}}$. Si les hypothèses suivantes sont vérifiées,

- $\rightarrow$ Il existe $a \in \mathrm{I}$ tel que la suite $(f_n(a))$ est convergente.
- $\rightarrow$ Pour tout segment $\mathrm{S} \subset \mathrm{I}$, $(f_n'|_\mathrm{S})$ converge uniformément vers $g|_\mathrm{S}$.

alors (f_n) converge simplement vers f sur I et uniformément sur tout segment inclus dans I où f est une fonction dans $\mathcal{C}^1(\mathrm{I}, \mathbb{R})$ telle que $f' = g$.

Démonstration. On a par hypothèse, pour tout $x \in \mathrm{I}$, (f_n') est une suite de fonctions continues qui converge uniformément vers g sur $[\min(a,x), \max(a,x)]$. On en déduit donc d'après la proposition II.3 que g est continue sur $[\min(a,x), \max(a,x)]$. On a de plus d'après la proposition II.6,

$$\int_a^x f_n'(t)dt \xrightarrow[n\to+\infty]{} \int_a^x g(t)dt \quad i.e. \quad f_n(x) - f_n(a) \xrightarrow[n\to+\infty]{} \int_a^x g(t)dt.$$

On pose alors pour tout $x \in \mathrm{I}$, $f(x) = \lim\limits_{n\longrightarrow+\infty} f_n(a) + \displaystyle\int_a^x g(t)dt$. g est continue, donc $f \in \mathcal{C}(\mathrm{I}, \mathbb{R})$ et (f_n) converge simplement vers f. On a de plus, pour tout $c, d \in \mathbb{R}$ tels que $c \leq d$ et $[c,d] \subset \mathrm{I}$, pour tout $x \in [c,d]$,

$$|f_n(x) - f(x)| = \left| f(c) - f_n(c) + \int_c^x (g(t) - f_n'(t))dt \right| \leq |f(c) - f_n(c)| + (d-c) \left\| g - f_n' \right\|_{\infty,[c,d]}.$$

On en déduit donc que

$$\|f_n - f\|_{\infty,[c,d]} \leq |f(c) - f_n(c)| + (d-c)\left\| g - f_n' \right\|_{\infty,[c,d]} \xrightarrow[n\to+\infty]{} 0.$$

La suite (f_n) converge donc bien simplement vers f et uniformément vers f sur tout segment inclus dans I. Enfin, par construction on a bien que $f' = g$.

Remarque. On peut aisément généraliser cette proposition de la manière suivante : soit $p \in \mathbb{N}^*$. Si les propriétés suivantes sont vérifiées,

- $\rightarrow$ $(f_n) \in \mathcal{C}^p(\mathrm{I}, \mathbb{R})^{\mathbb{N}}$.
- $\rightarrow$ Il existe $a \in \mathrm{I}$ tel que pour tout $k \in [\![1; p]\!]$, la suite $(f_n^{(k)}(a))$ converge.
- $\rightarrow$ Pour tout segment $\mathrm{S} \subset \mathrm{I}$, $(f_n^{(p)}|_\mathrm{S})$ converge uniformément vers $(g|_\mathrm{S})$.

alors (f_n) converge simplement vers f sur I et uniformément sur tout segment inclus dans I où f est une fonction dans $\mathcal{C}^p(\mathrm{I}, \mathbb{R})$ telle que $f^{(p)} = g$. Nous ne démontrerons pas cette généralisation malgré le fait que la démonstration n'est pas très difficile. Le lecteur est encouragé encore une fois à démontrer lui-même cette généralisation à partir de la proposition pour $p = 1$.

III Compléments

1. Convergence diagonale

Proposition III.1.

Supposons que (f_n) est une suite de fonctions continues sur X qui converge uniformément vers $f \in \mathcal{C}(\mathrm{X}, \mathrm{E})$. Si $a \in \mathrm{X}$ et $(x_n) \in \mathrm{X}^{\mathbb{N}}$ est une suite à valeurs dans X vérifiant $x_n \xrightarrow[n\to+\infty]{} a$, alors $f_n(x_n) \xrightarrow[n\to+\infty]{} f(a)$.

(cc) BY-NC-SA

Démonstration. Soit $\varepsilon > 0$. On considère les éléments suivants.

→ (f_n) converge uniformément vers f, on peut donc considérer $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$ et $x \in \mathrm{X}$, $\|f_n(x) - f(x)\| \leq \varepsilon$.

→ La suite de fonctions (f_n) est à valeur dans l'ensemble des fonctions continues, donc d'après la proposition II.3, f est continue et alors on peut considérer $\eta > 0$ tel que pour tout $x \in \mathrm{B}(a, \eta)$, $\|f(x) - f(a)\| < \eta$.

→ $x_n \xrightarrow[n \to +\infty]{} a$, on peut donc considérer $\mathrm{N}' \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $d(x_n, a) < \varepsilon$.

On a alors pour tout $n \geq \max(\mathrm{N}, \mathrm{N}')$,

$$\|f_n(x_n) - f(a)\| \leq \|f_n(x_n) - f(x_n)\| + \|f(x_n) - f(a)\| < \varepsilon + \varepsilon = 2\varepsilon,$$

ce qui nous permet d'affirmer que $f_n(x_n) \xrightarrow[n \to +\infty]{} f(a)$.

Exercice III.2.

Soit K un convexe compact de l'espace E vectoriel normé muni d'une norme $\|.\|$ et soit $f : \mathrm{K} \longrightarrow \mathrm{K}$ une fonction $1-$lipschitzienne. Montrer que f admet un point fixe dans K.
Indication : Utiliser le résultat de l'exercice VI.6 du chapitre 11.5 (Compacité).

Remarque. Lorsque E est de dimension finie, il suffit que f soit bornée pour qu'on ait l'existence d'un point fixe. Ce résultat est connu sous le nom de théorème du point fixe de Brouwer (hors programme).

2. Critère de Cauchy uniforme

Théorème (Critère de Cauchy uniforme) III.3.

Les propriétés suivantes sont vraies.

1. Si (f_n) converge uniformément, alors
$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall m, n \geq \mathrm{N},\ \forall x \in \mathrm{X},\ \|f_n(x) - f_m(x)\| < \varepsilon. \tag{γ}$$
2. Réciproquement, si E est un espace complet et la condition (γ) est vérifiée, alors la suite (f_n) converge uniformément.

Démonstration.

1. Supposons que (f_n) converge uniformément et soit f la limite uniforme de (f_n). Soit $\varepsilon > 0$. Soit $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $\forall x \in \mathrm{E}$, $\|f_n(x) - f(x)\| < \dfrac{\varepsilon}{2}$. On a alors pour tout $m, n \geq \mathrm{N}$ et $x \in \mathrm{X}$,
$$\|f_n(x) - f_m(x)\| \leq \|f_n(x) - f(x)\| + \|f_m(x) - f(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$
ce qui est bien le résultat voulu.
2. Supposons que E est complet et que la condition (γ) est vérifiée. Il est clair que pour tout $x \in \mathrm{E}$, la suite $(f_n(x))$ est une suite de Cauchy (voir le chapitre 4 des valeurs d'adhérence et 11.5 sur la compacité) donc elle converge. Posons donc pour tout $x \in \mathrm{X}$, $f(x)$ la limite de la suite $(f_n(x))$. f est la limite simple de la suite (f_n) et notre but est de montrer qu'il s'agit d'une limite uniforme.

   Soit $\varepsilon > 0$ et $\mathrm{N} \in \mathbb{N}$, tel que pour tout $m, n \geq \mathrm{N}$ et $x \in \mathrm{X}$, $\|f_n(x) - f_m(x)\| < \varepsilon$. En faisant tendre m vers l'infini dans l'inégalité, on obtient que pour tout $n \geq \mathrm{N}$ et $x \in \mathrm{X}$, $\|f_n(x) - f(x)\| \leq \varepsilon$. (f_n) converge donc bien uniformément (vers f).

(cc) BY-NC-SA

Notation. On note $\mathcal{C}_b(\mathrm{X},\mathrm{E})$ l'ensemble des fonctions continues bornées de X dans E

Remarques.

- $\rightarrow$ Lorsqu'une suite de fonctions vérifie la propriété (γ), on dit qu'elle vérifie le critère de Cauchy uniforme.

- $\rightarrow$ Une suite de fonctions bornées (f_n) de X dans E vérifie le critère de Cauchy uniforme si et seulement s'il s'agit d'une suite de Cauchy dans l'espace métrique $\mathcal{C}_b(\mathrm{X},\mathrm{E})$ muni de la distance induite par $\|.\|_\infty$ *i.e.*
$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall m, n \geq \mathrm{N},\ \|f_n - f_m\|_\infty < \varepsilon.$$

- $\rightarrow$ Une conséquence du théorème ci-dessus est que si E est complet, alors l'espace métrique $\mathcal{C}_b(\mathrm{X},\mathrm{E})$ muni de la distance induite par $\|.\|_\infty$ est complet.

- $\rightarrow$ La négation du critère de Cauchy uniforme s'écrit de la manière suivante : il existe φ et ψ deux extractrices et une suite (x_n) tels que pour tout $\left\|f_{\varphi(n)}(x_n) - f_{\psi(n)}(x_n)\right\| \nrightarrow 0$ ou d'une manière équivalente, quitte à extraire, qu'il existe $\delta > 0$ tel que pour tout $n \in \mathbb{N}$, $\left\|f_{\varphi(n)}(x_n) - f_{\psi(n)}(x_n)\right\| > \delta$.

Proposition III.4.

Soit A une partie de X. On suppose que $(f_n) \in \mathcal{C}_b(\mathrm{X},\mathrm{E})^{\mathbb{N}}$ et que E est complet. Soit f une fonction continue de X dans E. Si (f_n) converge uniformément vers f sur A, alors elle converge uniformément vers f sur $\overline{\mathrm{A}}$.

Démonstration. La suite (f_n) converge uniformément, elle vérifie donc le critère de Cauchy uniforme sur A, *i.e.*
$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbb{N},\ \forall n, m \geq \mathrm{N},\ \forall x \in \mathrm{A},\ \|f_n(x) - f_m(x)\| < \varepsilon.$$
Soit $\varepsilon > 0$ et $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n, m \geq \mathrm{N}$ et $x \in \mathrm{A}$, $\|f_n(x) - f_m(x)\| < \dfrac{\varepsilon}{2}$. Pour tout $n, m \geq \mathrm{N}$, la fonction $g : x \longmapsto f_n(x) - f_m(x)$ est continue, donc l'ensemble
$$\left\{x \in \mathrm{X},\ \|f_n(x) - f_m(x)\| \leq \frac{\varepsilon}{2}\right\} = g^{-1}\left(\mathrm{B}_f(0, \varepsilon/2)\right)$$
est fermé car il s'agit de l'image réciproque d'un fermé par une image continue (voie chapitre 11.3 des limites et continuité). On a de plus $\mathrm{A} \subset \{x \in \mathrm{X},\ \|f_n(x) - f_m(x)\| \leq \varepsilon\}$ et l'ensemble de droite est fermé, donc
$$\overline{\mathrm{A}} \subset \left\{x \in \mathrm{X},\ \|f_n(x) - f_m(x)\| \leq \frac{\varepsilon}{2}\right\} \subset \{x \in \mathrm{X},\ \|f_n(x) - f_m(x)\| < \varepsilon\}.$$
On en déduit que pour tout $n, m \geq \mathrm{N}$ et $x \in \overline{\mathrm{A}}$, $\|f_n(x) - f_m(x)\| < \varepsilon$. (f_n) vérifie le critère de Cauchy uniforme sur $\overline{\mathrm{A}}$ et E est complet donc converge uniformément sur $\overline{\mathrm{A}}$ vers une certaine fonction g continue qui coïncide avec f dans A. On a alors $(f-g)^{-1}(\{0\}) = \{x \in \mathrm{X}, f(x) - g(x) = 0\} \supset \mathrm{A}$. $(f-g)^{-1}(\{0\})$ est l'image réciproque d'un fermé par une fonction continue, il s'agit donc d'un fermé. On en déduit alors que
$$(f-g)^{-1}(\{0\}) = \overline{(f-g)^{-1}(\{0\})} \supset \overline{\mathrm{A}}.$$
On en déduit donc que f et g coïncident également sur $\overline{\mathrm{A}}$ et que finalement (f_n) converge uniformément vers f sur $\overline{\mathrm{A}}$.

(cc) BY-NC-SA

3. Passage de la convergence simple à la convergence uniforme

Théorème (Deuxième théorème de Dini) III.5.

Soit $a, b \in \mathbb{R}$ tels que $a \leq b$. Si (f_n) est une suite de fonctions croissantes de $[a, b]$ dans $\mathbb{R}$ qui converge simplement vers une fonction f continue, alors (f_n) converge uniformément vers f.

Démonstration. Soit $\varepsilon > 0$. f est une fonction continue sur le segment $[a, b]$, donc d'après le théorème de Heine (voir chapitre 11.3 sur les limites et continuité) f est uniformément continue sur $[a, b]$ pour tout $n \in \mathbb{N}$. Soit $\eta > 0$ tel que

$$\forall s, t \in [a, b],\ |s - t| < \eta \Longrightarrow |f(s) - f(t)| < \varepsilon.$$

Soit $p \in \mathbb{N}^*$ et $x_0, \ldots, x_p \in [a, b]$ tels que $x_0 = a < x_1 < \cdots < x_p = b$ et pour tout $i \in [\![0; p-1]\!]$, $|x_i - x_{i+1}| < \eta$. L'ensemble $\mathcal{X} = \{x_0, \ldots, x_p\}$ est fini donc (f_n) converge uniformément vers f sur $\mathcal{X}$. En effet, la convergence simple de (f_n) vers f sur $[a, b]$ nous permet d'affirmer pour tout $\varepsilon > 0$ l'existence de $\mathrm{N}_0, \ldots, \mathrm{N}_p \in \mathbb{N}$ tels que pour tout $i \in \mathbb{N}$ et pour tout $n \geq \mathrm{N}_i$, $|f_n(x_i) - f(x_i)| < \varepsilon$. On en déduit donc que pour tout $n \geq \max(\mathrm{N}_0, \ldots, \mathrm{N}_p)$ et tout $i \in [\![0; p]\!]$, $|f_n(x_i) - f(x_i)| < \varepsilon$ *i.e.* (f_n) converge uniformément sur $\mathcal{X}$. On en déduit donc que pour tout $n \geq \max(\mathrm{N}_0, \ldots, \mathrm{N}_p)$ et $x \in [a, b]$,

→ S'il existe $i \in [\![0; p]\!]$, $x = x_i$, alors on a

$$|f_n(x) - f(x)| = |f_n(x_i) - f(x_i)| < \varepsilon.$$

→ Sinon, il existe $i \in [\![0; p-1]\!]$ tel que $x \in]x_i, x_{i+1}[$ et donc

$$f(x) - 2\varepsilon \underset{(1)}{<} f(x_i) - \varepsilon \underset{(2)}{<} f_n(x_i) \underset{(3)}{\leq} f_n(x) \underset{(4)}{\leq} f_n(x_{i+1}) \underset{(5)}{<} f(x_{i+1}) + \varepsilon \underset{(6)}{<} f(x) + 2\varepsilon.$$

Justifions ces inégalités. Les inégalités (3) et (4) sont vraies par croissance de f_n. Les inégalités (2) et (5) sont vraies car (f_n) converge uniformément sur $\mathcal{X}$. Les inégalités (1) et (6) sont vraies car f est uniformément continue sur $[a, b]$ et $|x_i - x| < |x_{i+1} - x_i| < \eta$ et $|x_{i+1} - x| < |x_{i+1} - x_i| < \eta$.

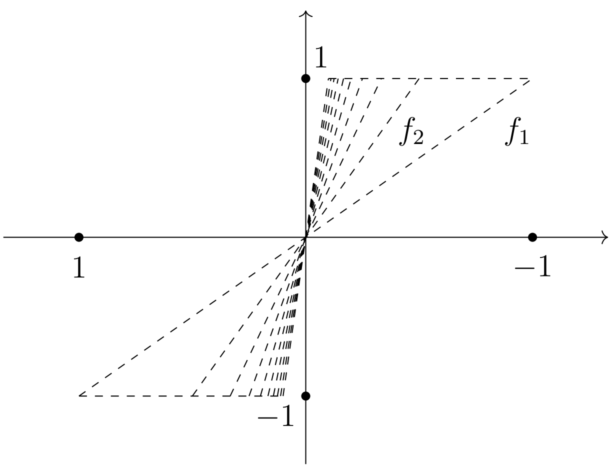
On en déduit donc que $n \geq \max(\mathrm{N}_0, \ldots, \mathrm{N}_p)$ et $x \in [a, b]$, $|f_n(x) - f(x)| < 2\varepsilon$ et alors (f_n) converge uniformément vers f.

Remarque.

→ Ce théorème ne nécessite pas que les termes de la suite (f_n) soient des fonctions continues.

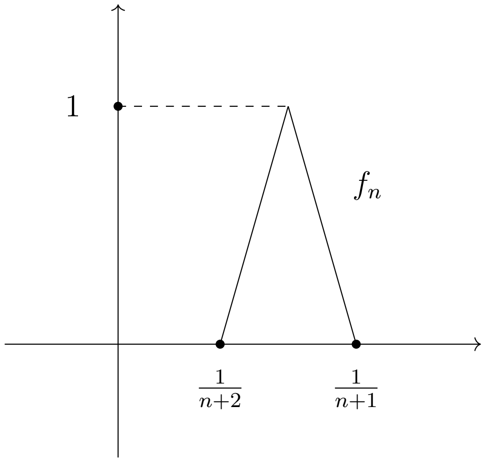
→ Ce résultat est faux en général si on enlève la croissance des termes de la suite (f_n) ou la continuité de f. En effet, si on suppose que (f_n) est définie par

$$\forall x \in [-1, 1],\ f_n(x) = \begin{cases} nx & \text{si } x \in \left[-\dfrac{1}{n}, \dfrac{1}{n}\right] \\ 1 & \text{si } x \in \left]\dfrac{1}{n}, 1\right] \\ -1 & \text{si } x \in \left[-1, -\dfrac{1}{n}\right[. \end{cases}$$



CC BY-NC-SA

On voit clairement sur le dessin que (f_n) converge simplement vers la fonction f qui vaut 1 sur $]0,1]$, 0 en 0 et -1 sur $[-1,0[$ et on peut facilement le montrer. Cependant, (f_n) est une suite de fonctions continues et f n'est pas continue, donc d'après la proposition II.3, (f_n) ne peut pas converger uniformément vers f. On peut même trouver un exemple où le résultat est faux lorsqu'on ajoute la continuité de f sans la croissance des f_n. En effet, si on considère la suite de fonctions (f_n) définie par le dessin ci-dessous, on peut montrer (et on le voit sur le dessin) que (f_n) converge simplement vers la fonction nulle mais ne converge clairement pas uniformément.



Exercice III.6.

Soit $(g_n) \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ une suite de fonctions convexes et g une fonction de $\mathbb{R}$ dans $\mathbb{R}$ telle que (g_n) converge simplement vers g et (g_n') converge simplement vers une fonction continue. Montrer que $g \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ et que (g_n') converge simplement vers g'.

Théorème III.7.

Soit $a, b \in \mathbb{R}$. tels que $a \leq b$. Supposons que $(f_n) \in \mathcal{C}([a,b],\mathbb{R})^{\mathbb{N}}$. Soit f une fonction de $[a,b]$ dans $\mathbb{R}$. Si les deux propriétés suivantes sont vérifiées,

1. Il existe $\mathrm{K} \geq 0$ tel que pour tout $n \in \mathbb{N}$, f_n est $\mathrm{K}-$lipschitzienne.
2. (f_n) converge simplement vers f.

Alors (f_n) converge uniformément vers f.

Démonstration. Montrons tout d'abord que f est $\mathrm{K}-$lipschitzienne. On a pour tout $n \in \mathbb{N}$ et $x, y \in [a,b]$, $|f_n(x) - f_n(y)| \leq \mathrm{K}\,|x-y|$. En faisant tendre n vers l'infini, on obtient que $|f(x) - f(y)| \leq \mathrm{K}\,|x-y|$. Soit $\varepsilon > 0$ et $\delta = \dfrac{\varepsilon}{\mathrm{K}+1}$. Soit $p \in \mathbb{N}^*$ et $x_0, \ldots, x_p \in [a,b]$ tels que $x_0 = a < x_1 < \cdots < x_p = b$ et pour tout $i \in [\![0; p-1]\!]$, $|x_i - x_{i+1}| < \delta$. Par le même raisonnement que dans la démonstration du théorème III.6, (f_n) converge uniformément vers f sur $\mathcal{X} = \{x_0, \ldots, x_p\}$. Considérons donc $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$ et tout $i \in [\![0; p-1]\!]$, $|f_n(x_i) - f(x_i)| < \varepsilon$. On a alors pour tout $n \geq \mathrm{N}$ et $x \in [a,b]$,

→ S'il existe $i \in [\![0; p]\!]$ tel que $x = x_i$, alors

$$|f_n(x) - f(x)| = |f_n(x_i) - f(x_i)| < \varepsilon.$$

→ Sinon, il existe $i \in [\![0; p-1]\!]$ tel que $x \in]x_i, x_{i+1}[$. On a alors

$$\begin{aligned}|f_n(x) - f(x)| &\leq |f_n(x) - f_n(x_i)| + |f_n(x_i) - f(x_i)| + |f(x_i) - f(x)| \\ &\leq \mathrm{K}\,|x - x_i| + \varepsilon + \mathrm{K}\,|x_i - x| < 2\mathrm{K}\delta + \varepsilon \leq 3\varepsilon.\end{aligned}$$

BY-NC-SA

On en déduit donc que pour tout $n \geq \mathrm{N}$ et $x \in [a,b]$, $|f_n(x) - f(x)| \leq 3\varepsilon$ et donc (f_n) converge uniformément vers f.

(cc) BY-NC-SA

Correction de l'exercice I.2. :

On commence par trouver la limite simple de (f_n). On a pour tout $x \in \mathbb{R}$

$$f_n(x) \xrightarrow[n\to+\infty]{} \begin{cases} 1 & \text{si } x \in]-1,1[\\ 0 & \text{si } x \in \{-1,1\} \\ -1 & \text{si } x \in]-\infty,-1[\cup]1,+\infty[. \end{cases}$$

On en déduit que la limite simple de la suite (f_n) est la fonction f définie par

$$\forall x \in \mathbb{R},\ f(x) = \begin{cases} 1 & \text{si } x \in]-1,1[\\ 0 & \text{si } x \in \{-1,1\} \\ -1 & \text{si } x \in]-\infty,-1[\cup]1,+\infty[. \end{cases}$$

Correction de l'exercice I.4. :

Bien entendu, pour tout $n \in \mathbb{N}$, f_n est bornée car il s'agit d'une fonction continue sur un segment. Pour montrer que $\|f_n\|_\infty = \sup_{x\in[0,1]} f_n(x)$ converge vers $\|f\|_\infty$, il faut montrer d'abord que $\|f\|_\infty$ est bien définie *i.e.* que f est bornée. En effet, il existe $\mathrm{N} \in \mathbb{N}$ tel que pour tout $n \geq \mathrm{N}$, $|f_n(x) - f(x)| < 1$ et alors $|f(x)| \leq 1 + |f_\mathrm{N}(x)| \leq 1 + \|f_\mathrm{N}\|_\infty$ donc f est bien bornée. On a alors pour tout $n \in \mathbb{N}$,

$$|\|f_n\|_\infty - \|f\|_\infty| \leq \|f_n - f\|_\infty \xrightarrow[n\to+\infty]{} 0.$$

Correction de l'exercice I.6. :

On procède à chaque fois en deux étapes.

1. → **Convergence simple :** Pour tout $x \in [0,1]$, $f_n(x) = n^a x^n(1-x) \xrightarrow[n\to+\infty]{} 0$, donc (f_n) converge simplement vers la fonction nulle.

   → **Convergence uniforme :** (f_n) converge uniformément vers 0 si et seulement si $\|f_n\|_\infty \xrightarrow[n\to+\infty]{} 0$. Il faut donc trouver la borne supérieure de f_n pour tout $n \in \mathbb{N}$. On a pour tout $n \geq 2$ et $x \in [0,1]$,

   $$f_n'(x) = 0 \iff n^a x^{n-1}(n-(n+1)x) = 0 \iff x = 0 \text{ ou } x = \frac{n}{n+1}.$$

   On peut alors tracer de tableau de variations de f_n.

x	0		$\frac{n}{n+1}$		1
$f_n'(x)$		$+$	0	$-$	
$f_n(x)$	0	$\nearrow$	$\frac{n^{a+n}}{(n+1)^{n+1}}$	$\searrow$	0

   On en déduit que pour $n \in \mathbb{N}$, $\|f_n\|_\infty = \dfrac{n^{a+n}}{(n+1)^{n+1}} = n^{a-1}\left(1+\dfrac{1}{n}\right)^{-(n+1)} \underset{+\infty}{\sim} \dfrac{n^{a-1}}{e}$. On en déduit donc que $\|f\|_\infty \xrightarrow[n\to+\infty]{} 0$ si et seulement si $a < 1$, *i.e.* que (f_n) converge uniformément vers la fonction nulle si et seulement si $a < 1$.

(cc) BY-NC-SA

2. → **Convergence simple :** Pour tout $x \in \mathbb{R}_+^*$, $f_n(x) = n^a x e^{-nx} \xrightarrow[n\to+\infty]{} 0$, donc (f_n) converge simplement vers la fonction nulle.

→ **Convergence uniforme :** On procède de la même manière que précédemment. On a pour tout $n \in \mathbb{N}^*$ et pour tout $x \in \mathbb{R}^+$,

$$f_n'(x) = 0 \Longleftrightarrow n^a(1-nx)e^{-nx} = 0 \Longleftrightarrow x = \frac{1}{n}.$$

On peut alors tracer le tableau de variations de f_n.

x	0		$\frac{1}{n}$		$+\infty$
$f_n'(x)$		$+$	0	$-$	
$f_n(x)$	0	$\nearrow$	$n^{a-1}e^{-1}$	$\searrow$	0

On en déduit donc pour tout $n \in \mathbb{N}^*$, $\|f_n\|_\infty = n^{a-1}e^{-1}$, donc $\|f_n\|_\infty \xrightarrow[n\to+\infty]{} 0$ si et seulement si $a < 1$ *i.e.* (f_n) converge uniformément vers la fonction nulle si et seulement si $a < 1$.

3. → **Convergence simple :** Soit $x \in \mathbb{R}$. Posons $(u_n) = (f_n(x))_{n\in\mathbb{N}}$. On a pour tout $n \in \mathbb{N}$, $\dfrac{u_{n+1}}{u_n} = \dfrac{x}{n+1} \xrightarrow[n\to+\infty]{} 0$ donc d'après le critère de d'Alembert, $u_n \xrightarrow[n\to+\infty]{} 0$. On en déduit donc que (f_n) converge simplement vers la fonction nulle.

→ **Convergence uniforme :** Encore une fois, on procède de la même manière. On a pour tout $x \in \mathbb{R}^+$ et $n \geq 2$

$$f_n'(x) = 0 \Longleftrightarrow \frac{(nx^{n-1} - x^n)}{n!}e^{-x} \Longleftrightarrow x = 0 \text{ ou } x = n.$$

On peut donc tracer le tableau de variations de f_n.

x	0		n		$+\infty$
$f_n'(x)$		$+$	0	$-$	
$f_n(x)$	0	$\nearrow$	$\frac{n^n}{n!}e^{-n}$	$\searrow$	0

On a donc pour tout $n \geq 2$, en utilisant la formule de Stirling,

$$\|f_n\|_\infty = f_n(n) = \frac{n^n}{n!}e^{-n} \underset{+\infty}{\sim} \frac{n^n}{\sqrt{2n\pi}(n/e)^n}e^{-n} = \frac{1}{\sqrt{2n\pi}} \xrightarrow[n\to+\infty]{} 0.$$

On en déduit alors que (f_n) converge uniformément vers la fonction nulle sur $\mathbb{R}^+$.

Correction de l'exercice III.2. :

Le résultat de l'exercice VI.6 du chapitre 11.5 est le suivant.

(cc) BY-NC-SA

Lemme III.8.

Soit K est un compact non vide de E, et $f : \mathrm{K} \longrightarrow \mathrm{K}$ est une fonction telle que pour tout $x, y \in \mathrm{K}$, si $x \neq y$, alors $\|f(x) - f(y)\| < \|x - y\|$. f admet un unique point fixe de K.

La différence entre les hypothèses de l'exercice du lemme et de l'exercice c'est que l'inégalité du lemme est stricte alors que cette de l'exercice résultat de la 1−lipschitzianité ne l'est pas. Pour pouvoir appliquer le résultat du lemme, on va approcher f par une suite de fonctions qui vérifient l'inégalité stricte.

Soit $a \in \mathrm{K}$. On considère la suite de fonctions (f_n) définie par

$$\forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{K},\ f_n(x) = \frac{1}{n} f(a) + \left(1 - \frac{1}{n}\right) f(x).$$

Pour tous $x, y \in \mathrm{K}$ différents et $n \in \mathbb{N}^*$,

$$\|f_n(x) - f_n(y)\| = \left(1 - \frac{1}{n}\right) \|f(x) - f(y)\| \leq \left(1 - \frac{1}{n}\right) d \|x - y\| < \|x - y\|.$$

De plus, pour tout $n \in \mathbb{N}^*$ et $x \in \mathbb{R}$,

$$\|f_n(x) - f(x)\| = \frac{1}{n} \|f(a) - f(x)\| \leq \frac{1}{n} \underbrace{\sup\{\|x - y\|,\ x, y \in \mathrm{K}\}}_{\text{fini car K borné}},$$

et donc

$$\|f_n - f\|_\infty \leq \frac{1}{n} \sup\{\|x - y\|,\ x, y \in \mathrm{K}\} \xrightarrow[n \to +\infty]{} 0.$$

(f_n) converge alors uniformément vers f et pour tout $n \in \mathbb{N}$, le lemme III.8 nous permet de dire que f_n admet un unique point fixe x_n. (x_n) est une suite à valeurs dans un compact, il existe donc une extractrice φ et $x \in \mathrm{K}$ tels que $x_{\varphi(n)} \xrightarrow[n \to +\infty]{} x$. (f_n) est une suite de fonctions continues qui converge uniformément, donc on a d'après la proposition III.1, $f_{\varphi(n)}(x_{\varphi(n)}) \xrightarrow[n \to +\infty]{} f(x)$ et finalement $x = f(x)$. f admet donc bien un point fixe.

Correction de l'exercice III.6. :

La suite de fonctions $(f_n) = (g'_n)$ est une suite de fonctions croissantes car (g_n) est une suite de fonctions convexes. De plus, (f_n) converge simplement vers une fonction continue donc converge uniformément sur tout segment de $\mathbb{R}$ d'après le deuxième théorème de Dini. Posons alors f la limite simple de (f_n). On a alors

- → (g_n) converge simplement vers g, donc il existe bien $a \in \mathbb{R}$ tel que la suite $(g_n(a))$ converge.
- → Pour tout segment $\mathrm{S} \subset \mathbb{R}$, $(g'_n|_\mathrm{S})$ converge uniformément vers $f|_\mathrm{S}$.

donc d'après la proposition II.8 (g_n) converge uniformément vers g sur tout segment S dans $\mathbb{R}$ et $g' = f$. Soit $\mathrm{M} > 0$, remarquons maintenant que $g'|_\mathrm{U}$ est limite uniforme de la suite de fonctions continues $(g'_n|_\mathrm{U})$ sur l'ouvert $\mathrm{U} =]-\mathrm{M}, \mathrm{M}[$, ce qui nous permet de dire d'après la proposition II.3 que g' est continue sur $]-\mathrm{M}, \mathrm{M}[$. Or $\mathbb{R} = \bigcup_{\mathrm{M} \in \mathbb{N}^*}]-\mathrm{M}, \mathrm{M}[$, on déduit que g' est continue sur $\mathbb{R}$ et que donc $g \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$.

(cc) BY-NC-SA

Chapitre 19 — Groupes

I Généralités

Définition I.1.

Soit G un ensemble non vide muni d'une application $* : \mathrm{G} \times \mathrm{G} \to \mathrm{G}$, appelée loi de composition interne. On dit que $(\mathrm{G}, *)$ est un groupe si

- $\to$ $*$ est associative, *i.e.* $\forall x, y, z \in \mathrm{G},\ x * (y * z) = (x * y) * z$.
- $\to$ $*$ admet un élément neutre, *i.e.* il existe $e \in \mathrm{G}$ tel que pour tout$x \in \mathrm{G},\ x * e = x = e * x$.
- $\to$ Tout élément admet un inverse, *i.e.* $\forall x \in \mathrm{G},\ \exists x^{-1} \in \mathrm{G},\ x * x^{-1} = e = x^{-1} * x$

Si $*$ est commutative, *i.e.* $\forall x, y \in \mathrm{G}\ x * y = y * x$, on dit alors que $(\mathrm{G}, *)$ est un groupe commutatif ou groupe abélien.

Notations.

- $\to$ Lorsque $(\mathrm{G}, *)$ est un groupe commutatif, on notera plutôt $(\mathrm{G}, +)$. Dans tout le reste du chapitre, on considère G un groupe d'élément neutre e.
- $\to$ Lorsque A et B sont deux ensembles disjoints, on note $\mathrm{A} \sqcup \mathrm{B}$ au lieu de $\mathrm{A} \cup \mathrm{B}$.

Exercice I.2.

Soit $(\mathrm{H}, *)$ un monoïde (*i.e.* $*$ est associative et admet un élément neutre) tel que tout élément admet un inverse à gauche. Montrer que H est un groupe.

Définition I.3.

On dit que $\mathrm{H} \subset \mathrm{G}$ est un sous-groupe du groupe $(\mathrm{G}, *)$ si H est stable par $*$ et $(\mathrm{H}, *|_{\mathrm{H}\times\mathrm{H}}^{\mathrm{H}}) = (\mathrm{H}, *)$ forme un groupe. On notera $\mathrm{H} \leq \mathrm{G}$.

Proposition I.4.

$(\mathrm{H}, *)$ est un sous-groupe de $(\mathrm{G}, *)$ si et seulement si $\mathrm{H} \neq \varnothing$ et

$$\forall x, y \in \mathrm{H},\ x * y^{-1} \in \mathrm{H}.$$

Proposition I.5.

Soit $(\mathrm{H}, *)$ un sous-groupe de $(\mathrm{G}, *)$. Les propositions suivantes sont vraies.

1. G et H ont le même élément neutre.
2. Pour tout $x \in \mathrm{H}$, l'inverse de x dans $(\mathrm{H}, *)$ est égal à l'inverse de x dans $(\mathrm{G}, *)$.
3. $\forall g, h \in \mathrm{G},\ (g * h)^{-1} = h^{-1} * g^{-1}$.

Démonstration.

(cc) BY-NC-SA

1. Posons e_{G} et e_{H} les éléments neutres respectifs de $(\mathrm{G}, *)$ et $(\mathrm{H}, *)$. On a $e_{\mathrm{H}} * e_{\mathrm{H}} = e_{\mathrm{H}}$, donc en multipliant des deux côtés par l'inverse de e_{H} dans $(\mathrm{G}, *)$, on obtient $e_{\mathrm{H}} = e_{\mathrm{G}}$.

2. Soit $a \in \mathrm{H}$ et $a_{\mathrm{H}}^{-1}, a_{\mathrm{G}}^{-1}$ ses inverses respectifs dans $(\mathrm{H}, *)$ et $(\mathrm{G}, *)$. On a $a_{\mathrm{H}}^{-1} * a = e_{\mathrm{H}} = e_{\mathrm{G}}$, donc en multipliant des deux côtés à droite par a_{G}^{-1}, on obtient $a_{\mathrm{G}}^{-1} = a_{\mathrm{H}}^{-1}$.

3. Pour tout $g, h \in \mathrm{G}$, on a $(g * h) * (h^{-1} * g^{-1}) = e$ et donc en multipliant par $(g * h)^{-1}$ des deux côtés à gauche, on obtient $(g * h)^{-1} = h^{-1} * g^{-1}$.

Proposition I.6.

Soit $(\mathrm{G}, *)$ un groupe et $\mathrm{S} \subset \mathrm{G}$. Il existe un (unique) sous-groupe minimum (pour l'inclusion) de G contenant S qu'on nomme sous-groupe engendré par S et note $\langle \mathrm{S} \rangle$. On a alors

$$\langle \mathrm{S} \rangle = \bigcap_{\mathrm{H} \leq \mathrm{G},\, \mathrm{S} \subset \mathrm{H}} \mathrm{H}.$$

Démonstration. Il suffit de vérifier qu'une intersection quelconque de sous-groupes est toujours un sous-groupe. Ceci étant clair, on a alors que $\bigcap\limits_{\mathrm{H} \leq \mathrm{G},\, \mathrm{S} \subset \mathrm{H}} \mathrm{H}$ est bien un sous-groupe minimum (pour l'inclusion) contenant S et donc on a bien l'existence. L'unicité provient du fait que, dans un ensemble partiellement ordonné, le minimum est toujours unique s'il existe.

Exercice I.7.

Soit H une partie non vide finie de G. Montrer que

$$\mathrm{H} \leq \mathrm{G} \Longleftrightarrow \forall (x, y) \in \mathrm{H}^2,\ x * y \in \mathrm{H}.$$

Notations.

→ (Translations dans un groupe G) Soit $a \in \mathrm{G}$. On note

$$\gamma_a : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{G} \\ x & \longmapsto a * x \end{cases} \quad \text{et} \quad \delta_a : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{G} \\ x & \longmapsto x * a. \end{cases}$$

Il s'agit de permutations de G qui sont des morphismes seulement si $a = e$.

→ (Puissances dans un groupe G) Soit $n \in \mathbb{Z}^*$ et $x \in \mathrm{G}$. On définit

$$x^n := \underbrace{x^\delta * \cdots * x^\delta}_{|n| \text{ fois}} \quad \text{où } \delta = \begin{cases} 1 & \text{si } n \geq 1 \\ -1 & \text{si } n \leq -1. \end{cases}$$

Pour $n = 0$ on définit $x^0 = e$. Compte tenu de cette définition, il est aisé de vérifier que

$$\forall (n, m) \in \mathbb{Z}^2,\ x^{n+m} = x^n * x^m.$$

→ On note $\mathrm{Z}(\mathrm{G}) = \{x \in \mathrm{G},\ \forall y \in \mathrm{G}\ x * y = y * x\}$.

Remarque. Lorsqu'on note $+$ au lieu de $*$, on note souvent nx au lieu de x^n.

(cc) BY-NC-SA

Définition I.8.

Soit f une application de G vers H. On dit que f est un morphisme (ou homomorphisme) de groupe de $(\mathrm{G}, *)$ vers $(\mathrm{H}, \Diamond)$ si

$$\forall x, y \in \mathrm{G},\ f(x * y) = f(x) \Diamond f(y).$$

On notera $\mathrm{Hom}(\mathrm{G}, \mathrm{H})$ l'ensemble des morphismes de $(\mathrm{G}, *)$ vers $(\mathrm{H}, \Diamond)$.

Exemples.

1. Soit $a \in \mathrm{G}$. L'application $j_a : \begin{cases} (\mathbb{Z}, +) & \longrightarrow (\mathrm{G}, *) \\ n & \longmapsto a^n \end{cases}$ est un morphisme de groupe.

2. Soit $a \in \mathrm{G}$. L'application $\sigma_a : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{G} \\ x & \longmapsto a * x * a^{-1} \end{cases}$ est un morphisme de groupe bijectif, d'inverse $\sigma_{a^{-1}}$. Ce morphisme est nommé automorphisme intérieur associé à a.
   En effet, en considérant $a, b, g \in \mathrm{G}$, on peut affirmer que

   $$\sigma_a \circ \sigma_b(g) = \sigma_a(b * g * b^{-1}) = a * b * g * b^{-1} * a^{-1} = (a * b) * g * (a * b)^{-1} = \sigma_{a*b}(g),$$

   d'où $\sigma_{a*b} = \sigma_a \circ \sigma_b$ et donc, en particulier, $\sigma_a \circ \sigma_{a^{-1}} = \sigma_e = \mathrm{Id} = \sigma_e = \sigma_{a^{-1}} \circ \sigma_a$ ce qui nous permet bien de dire que σ_a est un morphisme bijectif.

Vocabulaire. On dit que $x, y \in \mathrm{G}$ sont conjugués s'il existe $a \in \mathrm{G}$ tel que $y = a * x * a^{-1} = \sigma_a(x)$. La relation de conjugaison est en fait d'une relation d'équivalence sur G, qui donc le partitionne en classes d'équivalences de conjugaison.

Notation. Pour tous groupes G, G' et $f \in \mathrm{Hom}(\mathrm{G}, \mathrm{G}')$ on note

$$\mathrm{Ker}\, f = \{g \in \mathrm{G},\ f(g) = e_{\mathrm{G}'}\} = f^{-1}(\{e_{\mathrm{G}'}\}) \text{ et } \mathrm{Im}\, f = \{g' \in \mathrm{G}',\ \exists g \in \mathrm{G},\ f(g) = g'\} = f(\mathrm{G}),$$

où $e_{\mathrm{G}'}$ est l'élément neutre de G'.

Proposition I.9.

Soit $(\mathrm{G}', \Diamond)$ un groupe, e, e' les éléments neutres respectifs de G et G' et $f \in \mathrm{Hom}(\mathrm{G}, \mathrm{G}')$.

1. $f(e) = e$.
2. Si $\mathrm{H} \leq \mathrm{G}$ alors $f(\mathrm{H}) \leq \mathrm{G}'$. En particulier, $\mathrm{Im}\, f = f(\mathrm{G}) \leq \mathrm{G}'$.
3. Si $\mathrm{K} \leq \mathrm{G}'$ alors $f^{-1}(\mathrm{K}) \leq \mathrm{G}$. En particulier, $\mathrm{Ker}\, f = f^{-1}(\{e'\}) \leq \mathrm{G}$.
4. f est injective $\Longleftrightarrow \mathrm{Ker}\, f = \{e\} \Longleftrightarrow \mathrm{Ker}\, f \subset \{e\}$.

Démonstration. Les démonstrations de ces résultats sont laissées comme exercice au lecteur.

(cc) BY-NC-SA

Proposition I.10.

Soit $(\mathrm{H}, \lozenge)$ et $(\mathrm{H}', \lozenge')$ deux groupes.

1. Les homomorphismes se composent *i.e.* on peut composer $f \in \mathrm{Hom}(\mathrm{H}, \mathrm{H}')$ avec $g \in \mathrm{Hom}(\mathrm{G}, \mathrm{H})$ pour obtenir $f \circ g \in \mathrm{Hom}(\mathrm{G}, \mathrm{H}')$.
2. Les isomorphismes (morphismes bijectifs) se composent *i.e.* on peut composer $f \in \mathrm{Hom}(\mathrm{H}, \mathrm{H}')$ isomorphisme avec $g \in \mathrm{Hom}(\mathrm{G}, \mathrm{H})$ isomorphisme pour obtenir $f \circ g \in \mathrm{Hom}(\mathrm{G}, \mathrm{H}')$ isomorphisme.
3. L'ensemble des isomorphismes de G vers G, nommés automorphismes de G et notés $\mathrm{Aut}(\mathrm{G})$, forment un groupe pour la loi $\circ$.

Démonstration. La démonstration de cette proposition ne présente pas de difficulté et est laissée comme exercice au lecteur.

Notation. S'il existe un isomorphisme entre deux groupes G et H, on notera $\mathrm{G} \simeq \mathrm{H}$.

Exemple. L'application $\varphi : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{Aut}(\mathrm{G}) \\ a & \longmapsto \sigma_a \end{cases}$ est un morphisme de groupes de noyau

$$\mathrm{Ker}\,\varphi = \mathrm{Z}(\mathrm{G}) = \{x \in \mathrm{G},\ \forall y \in \mathrm{G},\ x * y = y * x\}.$$

En effet, pour tout $a \in \mathrm{G}$,

$$a \in \mathrm{Ker}\, f \iff \sigma_a = \mathrm{Id} \iff \forall x \in \mathrm{G},\ a * x * a^{-1} = x \iff \forall x \in \mathrm{G},\ a * x = x * a.$$

Exercice I.11.

Montrer que l'application

$$\varphi : \begin{cases} (\mathrm{G}, *) & \longrightarrow (\mathrm{Bij}(\mathrm{G}), \circ) \\ a & \longmapsto \gamma_a \end{cases}$$

où $\mathrm{Bij}(\mathrm{G})$ désigne l'ensemble des applications bijectives de l'ensemble G dans lui-même, est un morphisme injectif de groupes.

Remarque. Cet exercice montre que tout groupe peut être vu comme un sous-groupe du groupe symétrique (groupe de permutations, éventuellement infini) et que donc si $|\mathrm{G}| = n < \infty$, alors G peut être identifié à un sous-groupe de $\mathcal{S}_n$.

Notation. À partir de maintenant, lorsqu'il n'y a pas ambigüité, pour tout $a, b \in \mathrm{G}$, nous noterons ab ou $a \cdot b$ au lieu de $a * b$.

Définition I.12.

Soit H un sous-groupe de G. On dit que H est distingué (ou normal) lorsque

$$\forall a \in \mathrm{G},\ \sigma_a(\mathrm{H}) = a\mathrm{H}a^{-1} \subset \mathrm{H}.$$

On note dans ce cas $\mathrm{H} \trianglelefteq \mathrm{G}$. Lorsque cette propriété est vérifiée, on a

$$\forall a \in \mathrm{G},\ a\mathrm{H}a^{-1} = \mathrm{H} \text{ et } a\mathrm{H} = \mathrm{H}a.$$

Remarque. Si G est abélien, alors tout sous-groupe $\mathrm{H} \leq \mathrm{G}$ est distingué.

(cc) BY-NC-SA

Exemples.

→ $\{e\}$ et G sont des sous-groupes distingués de G.

→ Si G′ est un groupe, alors pour tout $f \in \mathrm{Hom}(\mathrm{G}, \mathrm{G}')$, $\mathrm{Ker}\, f$ est un sous-groupe distingué de G.

→ Plus généralement, si G′ est un groupe, $\mathrm{H}' \trianglelefteq \mathrm{G}'$ et $f \in \mathrm{Hom}(\mathrm{G}, \mathrm{G}')$ alors $f^{-1}(\mathrm{H}') \trianglelefteq \mathrm{G}$.

→ Si G′ est un groupe, $\mathrm{H} \trianglelefteq \mathrm{G}$ et $f \in \mathrm{Hom}(\mathrm{G}, \mathrm{G}')$ alors $f(\mathrm{H}) \trianglelefteq \mathrm{Im}\, f = f(\mathrm{G})$.

Exercice I.13.

Soit H et K deux sous-groupes de G. Montrer les propositions suivantes.

1. $\mathrm{HK} = \mathrm{KH} \iff \mathrm{HK}$ est un sous-groupe de G.
2. Si H est un sous groupe distingué de G alors HK est un sous groupe de G.
3. $\mathrm{H} \cap \mathrm{K} = \{e\} \iff f : \begin{cases} \mathrm{H} \times \mathrm{K} & \longrightarrow \mathrm{HK} \\ (h,k) & \longmapsto hk \end{cases}$ est bijective.
4. Montrer que si H et K sont distingués et $\mathrm{H} \cap \mathrm{K} = \{e\}$ alors $\forall (h,k) \in \mathrm{H} \times \mathrm{K},\ hk = kh$.
5. Montrer que si $\mathrm{H} \cap \mathrm{K} = \{e\}$ et $\forall (h,k) \in \mathrm{H} \times \mathrm{K},\ hk = kh$ alors f est un isomorphisme.

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

Soit $n \geq 2$. On note $\sim$ la relation d'équivalence sur $\mathbb{Z}$ définie par $a \sim b \iff n | a - b \iff a - b \in n\mathbb{Z}$. On note finalement $\mathbb{Z}/n\mathbb{Z}$ l'ensemble des classes d'équivalence pour $\sim$. Une telle classe s'écrit $\overline{x} = x + n\mathbb{Z}$. Il est aisé de vérifier que $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \ldots, \overline{n-1}\} = \{\overline{m}, \ldots, \overline{m+n-1}\}$ pour tout m et que ces n éléments sont distincts.

Proposition II.1.

Si X, Y sont deux classes dans $\mathbb{Z}/n\mathbb{Z}$ et $x \in \mathrm{X}$, $y \in \mathrm{Y}$, alors $\overline{x+y}$ ne dépend que de X et Y.

Remarque. La proposition ci-dessus permet donc de définir une loi + sur $\mathbb{Z}/n\mathbb{Z}$ qui en en fait un groupe abélien.

Proposition II.2.

1. L'application $\varphi : \begin{cases} \mathbb{Z}/n\mathbb{Z} & \longrightarrow \mathbb{U}_n \\ \overline{k} & \longmapsto e^{\frac{i2\pi k}{n}} \end{cases}$ est bien définie et est un isomorphisme.
2. L'application $\pi : \begin{cases} \mathbb{Z} & \longrightarrow \mathbb{Z}/n\mathbb{Z} \\ x & \longmapsto \overline{x} \end{cases}$ est un morphisme surjectif de noyau $\mathrm{Ker}\, \pi = n\mathbb{Z}$.

III Ordre d'un élément

Soit $a \in \mathrm{G}$. Remarquons que $\langle a \rangle = \{a^k,\ k \in \mathbb{Z}\}$. Introduisons tout d'abord le morphisme suivant

$$j_a : \begin{cases} \mathbb{Z} & \longrightarrow \langle a \rangle \\ m & \longmapsto a^m. \end{cases}$$

Pour tout $a \in \mathrm{G}$, j_a est un morphisme de groupe surjectif. D'après la proposition I.9, $\mathrm{Ker}\, j_a$ est un sous groupe de $(\mathbb{Z}, +)$. Il est bien connu que les sous groupes de $\mathbb{Z}$ sont les groupes de la forme $n\mathbb{Z}$ avec $n \in \mathbb{N}$.

BY-NC-SA

Le lecteur ayant un doute sur ce résultat est encouragé à essayer de le prouver lui même. Deux cas se présentent donc.

→ Si $n \neq 0$, alors $\operatorname{Ker} j_a = n\mathbb{Z}$. On note alors $\omega(a) = n$ et on dit que l'ordre de a est égal à n.

→ Sinon, $\operatorname{Ker} j_a = \{0\}$ et alors j_a est un isomorphisme. Dans ce cas, on note $\omega(a) = \infty$ et on dit que l'ordre de a est infini.

Remarque. Pour tout $a \in \mathrm{G}$, on peut également définir $\omega(a)$ comme

$$\omega(a) = \min\{k \in \mathbb{N}^*,\ a^k = e\}.$$

Proposition III.1.

Supposons que $\operatorname{Ker} j_a \neq \{0\}$ et posons $n = \omega(a)$. Les propositions suivantes sont vraies.

1. $\langle a\rangle = \{e, a, \dots, a^{n-1}\}$ et ces éléments sont distincts. De plus, $\forall k \in \mathbb{Z},\ a^k = e \Longleftrightarrow n|k$.
2. $n = \min\{m \in \mathbb{N}^*,\ a^m = e\}$.
3. $\forall m \in \mathbb{Z},\ \exists! k \in [\![0; n-1]\!],\ a^m = a^k$.
4. $\forall k \in \mathbb{Z},\ \omega(a^k) = \dfrac{n}{n \wedge k}$.
5. Si $q|n$ alors $\omega(a^q) = \dfrac{n}{q}$.

Démonstration.

1. Montrons d'abord la deuxième partie de ce point. Pour tout $k \in \mathbb{Z}$,

$$a^k = e \Longleftrightarrow k \in \operatorname{Ker} j \Longleftrightarrow k \in n\mathbb{Z} \Longleftrightarrow n|k,$$

et en particulier $a^n = e$. Montrons ensuite que $\langle a\rangle = \{e, a, \dots, a^{n-1}\}$.

→ ($\supset$) Cette inclusion est évidente.

→ ($\subset$) Soit $k \in \mathbb{Z}$. Faisons la division euclidienne de k par n : $k = qn + b$ avec $b \in [\![0; n-1]\!]$. On a alors $a^k = (a^n)^k a^b = e^k a^b = a^b$. Ces éléments sont de plus distincts car si $i, j \in [\![0; n-1]\!]$ vérifient $a^i = a^j$ alors $a^{i-j} = e$ et donc $n|i-j$ d'où $i = j$.

2. D'après le point 1, $\forall k \in [\![1; n-1]\!]$ $a^k \neq e$ et $a^n = e$ d'où le résultat.

3. Ce point est une conséquence directe du point (1).

4. Soit $k, \ell \in \mathbb{Z}$. On a

$$(a^k)^\ell = e \Longleftrightarrow a^{k\ell} = e \Longleftrightarrow n|k\ell \Longleftrightarrow \frac{n}{n \wedge k} \,\Big|\, \frac{k}{n \wedge k}\ell \underset{\text{Gauß}}{\Longleftrightarrow} \frac{n}{n \wedge k} \Big| \ell.$$

L'avant-dernière équivalence provient du fait que $\dfrac{k}{n \wedge k} \wedge \dfrac{n}{n \wedge k} = \dfrac{n \wedge k}{n \wedge k} = 1$.
a^k étant clairement d'ordre fini, en utilisant le point 2, on a $\omega(a^k) = \min\{\ell \in \mathbb{N}^*\ (a^k)^\ell = e\} = \dfrac{n}{n \wedge k}$.

5. Ce point est une conséquence directe du point précédent.

(cc) BY-NC-SA

IV Actions de morphismes

Proposition IV.1.

Soit G' un groupe d'élément neutre e', $a \in G$, et $f \in \mathrm{Hom}(G, G')$. Les propositions suivantes sont vraies.

1. Si $\omega(a) < \infty$, alors $\omega(f(a)) | \omega(a)$.
2. Si f est injective, alors $\omega(a) = \omega(f(a))$.
3. Si a et b sont deux éléments de G conjugués, *i.e.* il existe $c \in G$ tel que $a = cbc^{-1}$, alors $\omega(a) = \omega(b)$. En particulier, $\forall x, y \in G,\ \omega(xy) = \omega(yx)$.

Démonstration.

1. Supposons que $\omega(a) < \infty$. On a $f(a)^{\omega(a)} = f(a^{\omega(a)}) = f(e) = e'$ et donc $\omega(f(a)) < \infty$ et $\omega(f(a)) | \omega(a)$.
2. Supposons que f est injective. Deux cas se présentent.
   - → Si $\omega(a) = \infty$, alors il est facile de vérifier que $\langle f(a) \rangle = f(\langle a \rangle)$. f est injective et $\langle a \rangle$ est infini, donc $\langle f(a) \rangle$ aussi et donc en utilisant le point 1 de la proposition III.1, on voit qu'on ne peut pas avoir $\omega(a) < \infty$.
   - → Si $\omega(a) = n \in \mathbb{N}^*$, alors
   $$f(a)^k = e' \Longleftrightarrow f(a^k) = f(e) \underset{f \text{ injective}}{\Longleftrightarrow} a^k = e \Longleftrightarrow \omega(a) | k.$$
   On en déduit donc que $\omega(a) | \omega(f(a))$. En combinant ce résultat avec le point (1), on obtient bien que $\omega(a) = \omega(f(a))$.
3. Le fait que a et b soient conjugués se traduit par le fait qu'il existe $z \in G$ tel que $b = \sigma_z(a)$. σ_z est un morphisme injectif, donc en utilisant le point (2), on voit que $\omega(b) = \omega(\sigma_z(a)) = \omega(a)$. Enfin, pour montrer la deuxième partie de ce point, il suffit de voir que $xy = \sigma_x(yx)$.

Exercice IV.2.

Soit $m, n \geq 1$ tel que $m \wedge n = 1$. Déterminer $\mathrm{Hom}(\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}/n\mathbb{Z})$.

Exercice IV.3.

Soit $(a, b) \in G^2$ tels que $\omega(a) = m \leq \omega(b) = n < \infty$.

1. A-t-on $\omega(ab) < \infty$? *Indice : Penser aux symétries dans* $\mathbb{R}^2$.
2. Supposons que $ab = ba$ et $\omega(a) \wedge \omega(b) = 1$. Montrer que $\omega(ab) = mn$.
3. On suppose toujours que $ab = ba$. Montrer qu'il existe $c \in G$ tel que $\omega(c) = n \vee m$.
4. On suppose maintenant que G est abélien et fini. Montrer qu'il existe $z \in G$ tel que $\forall x \in G,\ \omega(x) | \omega(z)$. En particulier, on montrera l'existence de $z \in G$ tel que $\omega(z) = \bigvee_{x \in G} \omega(x)$.

Proposition (Théorème faible de Lagrange) IV.4.

Supposons que G est commutatif et soit $a \in G$. On a $\omega(a) | \, |G|$.

BY-NC-SA

Démonstration. On sait que γ_a est bijective, donc

$$\prod_{g\in G} g = \prod_{g\in G} \gamma_a(g) = \prod_{g\in G} (ag) = a^{|G|} \prod_{g\in G} g.$$

En multipliant par l'inverse de $\prod\limits_{g\in G} g$ on obtient que $a^{|G|} = e$, ce qui implique d'après la proposition III.1 que $\omega(a) \mid |G|$.

V Groupes cycliques

Définition V.1.

On dit que G est monogène s'il est engendré par un seul élément, *i.e.* il existe $a \in G$ tel que $G = \langle a \rangle$. Si de plus G est fini, alors on dit qu'il est cyclique.

Remarque V.2.

Supposons que G est monogène, *i.e.* qu'il existe $a \in G$ tel que $G = \langle a \rangle$. Deux cas se présentent.

1. Si $\omega(a) = \infty$, alors $j : \begin{cases} \mathbb{Z} & \longrightarrow \langle a \rangle \\ m & \longmapsto a^m \end{cases}$ est un isomorphisme.
2. Si $\omega(a) < \infty$, alors $G = \langle a \rangle = \{e, a, \dots, a^{n-1}\}$ avec ces éléments distincts. En particulier, on a $|G| = \omega(a)$.

Vocabulaire. Pour tout $a \in G$, lorsque $G = \langle a \rangle$, on dit que a est un élément générateur de G.

Proposition V.3.

Soit $n \in \mathbb{N}^*$. G est cyclique de cardinal n si et seulement si G est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. De plus, si $\psi : (G, *) \longrightarrow (\mathbb{Z}/n\mathbb{Z}, +)$ et $b \in G$ un générateur de G, *i.e.* $G = \langle b \rangle$, alors $\psi(b) = \overline{1}$ implique que ψ est un isomorphisme.

Démonstration. Montrons tout d'abord l'équivalence.

→ ($\Leftarrow$) Cette implication est facile à montrer.

→ ($\Rightarrow$) Supposons que G est cyclique de cardinal fini égal à n. Il existe donc $a \in G$ tel que $G = \{e, a, \dots, a^{n-1}\}$. Considérons le morphisme de groupe suivant

$$\psi : \begin{cases} (\mathbb{Z}/n\mathbb{Z}, +) & \longrightarrow (G, *) \\ \overline{k} & \longmapsto a^k \end{cases}.$$

On peut facilement montrer que ψ est bien défini et bijectif et alors $G \simeq \mathbb{Z}/n\mathbb{Z}$.
Le dernier point de la proposition découle directement du raisonnement effectué ci-dessus.

Exercice V.4.

Supposons que G est cyclique d'ordre (de cardinal) n et soit a un générateur de G. Soit H un sous groupe de G. Posons $d = |H|$. Montrer que $H = \left\langle a^{\frac{n}{d}} \right\rangle$.

BY-NC-SA

Remarque. Cette remarque peut être très utile dans quelques exercices difficiles. Soit $n \geq 1$. L'ensemble $\mathrm{D}_n = \{k \geq 1\ k|n\}$ vérifie $|\mathrm{D}_n| \leq 2\sqrt{n}+1$. En effet, l'application

$$\varphi : \begin{cases} \mathrm{D}_n \cap [\![1; \lfloor\sqrt{n}\rfloor]\!] & \longrightarrow \mathrm{D}_n \cap [\![\lfloor\sqrt{n}\rfloor; n]\!] \\ d & \longmapsto \dfrac{n}{d} \end{cases}$$

est une bijection et donc

$$\begin{aligned} |\mathrm{D}_n| &= \left|\mathrm{D}_n \cap [\![1; \lfloor\sqrt{n}\rfloor]\!]\right| + \left|\mathrm{D}_n \cap [\![1+\lfloor\sqrt{n}\rfloor; n]\!]\right| \\ &\leq \left|\mathrm{D}_n \cap [\![1; \lfloor\sqrt{n}\rfloor]\!]\right| + \left|\mathrm{D}_n \cap [\![\lfloor\sqrt{n}\rfloor; n]\!]\right| + 1 \\ &= 2\left|\mathrm{D}_n \cap [\![1; \lfloor\sqrt{n}\rfloor]\!]\right| + 1 \leq 2\sqrt{n}+1. \end{aligned}$$

Exercice V.5.

Supposons que G est cyclique et posons $|\mathrm{G}| = n \geq 2$ et soit a un générateur de G.

1. Montrer que pour tout $k \in \mathbb{Z}$, a^k génère G si et seulement si $k \wedge n = 1$.
2. Posons $\varphi(n) = |\{k \in [\![1;n]\!]\ k \wedge n = 1\}|$. Cette application est nommée l'indicatrice d'Euler. En utilisant la question précédente, montrer que $\displaystyle\sum_{d|n\ d\geq 1} \varphi(d) = n$.

Exercice V.6.

1. Soit $(\mathrm{G}_1, *)$ et $(\mathrm{G}_2, \lozenge)$ deux groupes cycliques. On considère la loi de composition interne $\otimes : \mathrm{G}_1 \times \mathrm{G}_2 \longrightarrow \mathrm{G}_1 \times \mathrm{G}_2$ définie par
$$\forall (g_1, g_2, g_1', g_2') \in \mathrm{G}_1 \times \mathrm{G}_2 \times \mathrm{G}_1 \times \mathrm{G}_2,\ (g_1, g_2) \otimes (g_1', g_2') = (g_1 * g_1',\ g_2 \lozenge g_2')$$
Trouver une condition nécessaire et suffisante pour que $(\mathrm{G}_1 \times \mathrm{G}_2, \otimes)$ soit cyclique.
2. (Théorème des restes chinois) Soit $a, b \geq 2$. Montrer que $\varphi : \begin{cases} \mathbb{Z}/ab\mathbb{Z} & \longrightarrow \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \\ \overline{x} & \longmapsto (\overline{x}, \overline{x}) \end{cases}$ est bien définie puis que φ est un isomorphisme si et seulement si $a \wedge b = 1$.

Exercice V.7.

Soit A un anneau commutatif unitaire.

1. Soit $\mathrm{P}, \mathrm{Q} \in \mathrm{A}[\mathrm{X}]$ avec Q de coefficient dominant égal à 1 (*i.e.* $\mathrm{Q} = 1$ ou Q est unitaire) (ici, on désigne par 1 l'élément neutre pour la multiplication de l'anneau A).
Montrer que l'on peut effectuer "la" division euclidienne de P par Q.
2. On suppose que A est intègre et $\deg \mathrm{P} \geq 0$. Montrer que P possède au plus $\deg \mathrm{P}$ racines distinctes.
3. Soit $(\mathbb{K}, +, \times)$ un corps et $(\mathrm{G}, \times)$ un sous-groupe fini de $(\mathbb{K}^*, \times)$, où $\mathbb{K}^*$ est l'ensemble des éléments de $\mathbb{K}$ inversible pour la loi $\times$. Montrer que G est cyclique.

Remarque. La question 1 de l'exercice ci-dessus demande de montrer que lorsqu'un polynôme $\mathrm{Q} \in \mathrm{A}[\mathrm{X}]$ est de coefficient dominant égal à 1, on peut effectuer la division euclidienne de tout polynôme $\mathrm{P} \in \mathrm{A}[\mathrm{X}]$ par Q. Ce résultat est aussi vrai lorsque le coefficient dominant de Q est inversible, mais devient faux lorsque ce n'est pas le cas. Lorsque A est un corps, tous les éléments de A sauf 0 sont inversibles, on peut

(cc) BY-NC-SA

donc toujours effectuer la division euclidienne par un polynôme non nul.

VI Groupe engendré par une partie

Proposition VI.1.

Pour tout $\mathrm{A} \subset \mathrm{G}$, il existe un plus petit sous-groupe $\langle \mathrm{A} \rangle$ de G contenant A appelé sous-groupe engendré par A. De plus, on a

$$\langle \mathrm{A} \rangle = \bigcap_{\mathrm{H} \leq \mathrm{G}\ \mathrm{A} \subset \mathrm{H}} \mathrm{H} = \{a_1^{\alpha_1} \dots a_n^{\alpha_n},\ n \in \mathbb{N},\ a_1, \dots, a_n \in \mathrm{A},\ \alpha_1, \dots, \alpha_n \in \mathbb{Z}\}.$$

Démonstration. L'existence et la première égalité ont déjà étés établis à la proposition I.6. Montrons la seconde égalité, *i.e.*

$$\langle \mathrm{A} \rangle = \underbrace{\{a_1^{\alpha_1} \dots a_n^{\alpha_n},\ n \in \mathbb{N},\ a_1, \dots, a_n \in \mathrm{A},\ \alpha_1, \dots, \alpha_n \in \mathbb{Z}\}}_{\mathrm{B}}.$$

Il est facile de montrer que B est bien un sous-groupe de G contenant A. On en déduit donc que $\langle \mathrm{A} \rangle \subset \mathrm{B}$. Montrons l'inclusion réciproque. Pour tout $n \in \mathbb{N}$ $a_1, \dots, a_n \in \mathrm{A} \subset \langle \mathrm{A} \rangle$ et $\alpha_1, \dots, \alpha_n \in \mathbb{Z}$, par stabilité, on a $a_1^{\alpha_1} \dots a_n^{\alpha_n} \in \langle \mathrm{A} \rangle$, d'où l'égalité voulue.

Vocabulaire. On dit que $\mathrm{S} \neq \varnothing$ est générateur (de G) si $\mathrm{G} = \langle \mathrm{S} \rangle$.

Exemple. $(\mathcal{S}_n, \circ)$ est généré par les transpositions de la forme $(i\ i+1)$, *i.e.*

$$\mathcal{S}_n = \langle \{(i\ i+1),\ i \in [\![1; n-1]\!]\} \rangle .$$

Exercice VI.2.

Supposons que $(\mathrm{G}, *)$ soit un groupe fini abélien et soit p un nombre premier tel que pour tout $g \in \mathrm{G},\ \omega(g) | p$ (ou d'une manière équivalente, $\forall g \in \mathrm{G},\ g^p = e$). Montrer qu'il existe $n \in \mathbb{N}$ tel que $\mathrm{G} \simeq (\mathbb{Z}/p\mathbb{Z})^n$.

Exercice VI.3.

Soit $\mathrm{H} \neq \{e\}$ un groupe où e est son élément neutre. Notons $\mathbb{P}$ l'ensemble des nombres premiers. On pose $\mathrm{Sg(H)}$ l'ensemble des sous-groupes de H. Montrer que les propositions suivantes sont équivalentes.

1. $\mathrm{Sg(H)} = \{\{e\}, \mathrm{H}\}$
2. $\exists p \in \mathbb{P},\ \mathrm{H} \simeq \mathbb{Z}/p\mathbb{Z}$
3. $\forall g \in \mathrm{H} \setminus \{e\},\ \langle g \rangle = \mathrm{H}$
4. $\exists p \in \mathbb{P},\ |\mathrm{H}| = p$

(cc) BY-NC-SA

VII Compléments

1. Classes latérales et groupe quotient

Proposition VII.1.

Soit H un sous groupe de G. La relation $\sim_g$ sur G^2 définie par

$$\forall a, b \in G,\ a \sim_g b \Longleftrightarrow b \in aH$$

est une relation d'équivalence. La classe d'équivalence de a pour $\sim_g$ est aH et est appelée classe à gauche selon H de a. On note cette classe $\overline{a}$.

Remarque.

→ En fait, on peut également définir la même notion d'équivalence à droite $\sim_d$ par $a \sim_d b \Longleftrightarrow b \in Ha$. Ces deux relations sont les mêmes si et seulement si H est distingué. Dans ce cas, on la notera simplement $\sim$.

→ Une manière équivalente de définir cette relation est

$$a \sim_g b \Longleftrightarrow b^{-1}a \in H \text{ et } a \sim_d b \Longleftrightarrow ab^{-1} \in H.$$

Définition et Proposition VII.2.

Soit H un sous-groupe distingué de G. Les propriétés suivantes sont vraies.

1. Le produit de deux classes à gauche selon H est une classe à gauche selon H.
2. Pour ce produit, l'ensemble des classes à gauche selon H forme un groupe, noté G/H.
3. Pour tout $a, b \in G$, $\gamma_{ba^{-1}}$ est une bijection entre aH et bH et par conséquent lorsque H est fini, $\forall a \in G,\ |aH| = |H|$, *i.e.* toute classe à gauche selon H est en bijection avec H.
4. L'élément neutre de G/H est H (ou d'une manière équivalente $\overline{e}$).

Démonstration.

1. Pour tout $a, b \in G$, on a $aH * bH = a * b * H * H = abH$.
2. Il suffit d'appliquer la définition d'un groupe pour démontrer ce point.
3. Remarquons que
   → $\gamma_{ba^{-1}}(aH) \subset bH$ et $\gamma_{ab^{-1}}(bH) \subset aH$.
   → L'application $\gamma_{ab^{-1}} : bH \longrightarrow aH$ est un inverse (et donc l'unique inverse) de $\gamma_{ba^{-1}}$ et alors $\gamma_{ba^{-1}}$ est bijective.

   Par conséquent, pour tout $a, b \in G$, $|aH| = |bH|$ et en particulier $|aH| = |eH| = |H|$.
4. $H = \overline{e}$ donc pour tout $a \in G$, $\overline{ea} = \overline{ae} = \overline{a}$ donc H est bien l'élément neutre de G/H.

Remarque.

→ Pour bien comprendre cette notion de groupe quotient, il faut voir que pour tout $a \in G$, la classe $\overline{a}$ comme une sorte d'ensemble d'éléments de même "reste" que a dans G pour une certaine "division". Lorsque $*$ est additive, on peut faire l'analogie avec $\mathbb{Z}/n\mathbb{Z}$. En effet, $(n\mathbb{Z}, +) \trianglelefteq (\mathbb{Z}, +)$ et le groupe quotient de $\mathbb{Z}$ par $n\mathbb{Z}$ est tout simplement égal au groupe bien connu $\mathbb{Z}/n\mathbb{Z}$. Dans ce cas, lorsque $+$ est additive, en pensant à cet exemple, on peut voir les classes $\overline{a} = a + H$ comme l'ensemble des éléments de même "restes" que a "modulo H". Dans l'exemple de $\mathbb{Z}/n\mathbb{Z}$, on a bien entendu $H = n\mathbb{Z}$.

→ L'égalité $|H| = |aH| = |Ha|$ reste vraie même si H n'est pas distingué dans G.

(cc) BY-NC-SA

Notation. Même lorsque H est un sous-groupe non distingué de G, on notera G/H l'ensemble des classes à gauche de G selon H, *i.e.* l'ensemble $\{\overline{g},\ g \in \mathrm{G}\} = \{g\mathrm{H},\ g \in \mathrm{G}\}$. Dans ce cas, G/H n'est pas un groupe, car lorsque G est non abélien, le produit de deux classes à gauche n'est pas forcément une classe à gauche. De la même manière, on notera par H\G l'ensemble des classes à droite de G selon H *i.e.* l'ensemble $\{\mathrm{H}g,\ g \in \mathrm{G}\}$.

Proposition VII.3.

Soit H un sous-groupe distingué de G. Les propositions suivantes sont vraies.

1. L'application $\pi : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{G/H} \\ a & \mapsto a\mathrm{H} \end{cases}$ est un morphisme surjectif de noyau $\operatorname{Ker}\pi = \mathrm{H}$. On l'appellera la surjection canonique.
2. Soit G′ un groupe et $f \in \operatorname{Hom}(\mathrm{G}, \mathrm{G}')$. Il existe $\tilde{f} \in \operatorname{Hom}(\mathrm{G}/\operatorname{Ker} f, \mathrm{G}')$ injective d'image $\operatorname{Im} f$ tel que $f = \tilde{f} \circ \pi$. En particulier, $\mathrm{G}/\operatorname{Ker} f$ est isomorphe à $\operatorname{Im} f$.

Démonstration.

1. C'est clairement un morphisme par ce qui précède. Soit $a \in \operatorname{Ker}\pi$. On a $a\mathrm{H} = \mathrm{H}$, il existe donc $g \in \mathrm{H}$ tels que $ae = g$, *i.e.* $a = g \in \mathrm{H}$. On en déduit donc que $\operatorname{Ker}\pi \subset \mathrm{H}$. L'implication réciproque est évidente.
2. Posons $\mathrm{H} = \operatorname{Ker} f$ et considérons l'application

$$\tilde{f} : \begin{cases} \mathrm{G/H} & \longrightarrow \mathrm{G}' \\ a\mathrm{H} & \longmapsto f(a). \end{cases}$$

Cette application est bien définie. En effet, elle est indépendante du représentant de la classe aH, *i.e.* si $a, b \in \mathrm{G}$ et $a\mathrm{H} = b\mathrm{H}$, alors $\tilde{f}(a\mathrm{H}) = \tilde{f}(b\mathrm{H})$. Le lecteur ayant un doute peut le vérifier lui-même. De plus, cette application est clairement un morphisme et est injective. En effet, pour tout $\mathrm{U} \in \operatorname{Ker}\tilde{f}$, il existe $a \in \mathrm{G}$ tels que $\mathrm{U} = a\mathrm{H}$. On a alors

$$\tilde{f}(\mathrm{U}) = e' \iff \tilde{f}(a\mathrm{H}) = e' \iff f(a) = e' \iff a \in \operatorname{Ker} f \iff a\mathrm{H} = \mathrm{H},$$

et donc $\operatorname{Ker}\tilde{f} = \{\mathrm{H}\}$, et H est l'élément neutre de G/H. On en déduit donc que $\tilde{f}$ induit un isomorphisme entre G/H et $\operatorname{Im} f$ et qu'en particulier $\mathrm{G}/\operatorname{Ker} f$ et $\operatorname{Im} f$ sont isomorphes.

Théorème (Théorème de Lagrange) VII.4.

Supposons que G est fini. Soit H un sous groupe de G. On a $|\mathrm{H}| \mid |\mathrm{G}|$.

Démonstration. Les classes à gauche selon H, $\{a\mathrm{H},\ a \in \mathrm{G}\}$, sont disjointes et en nombre fini. Il existe donc $p \in \mathbb{N}^*$ et $a_1, \dots, a_p \in \mathrm{G}$ tels que $\mathrm{G} = \bigsqcup_{k=1}^{p} a_k\mathrm{H}$. Cette union étant disjointe et d'après la proposition VII.1 tous ces ensembles sont de même cardinal égal à $|\mathrm{H}|$, on peut écrire

$$|\mathrm{G}| = \left|\bigsqcup_{k=1}^{p} a_k\mathrm{H}\right| = \sum_{k=1}^{p} |a_k\mathrm{H}| = p\,|\mathrm{H}|\,,$$

et alors finalement $|\mathrm{H}| \mid |\mathrm{G}|$.

Remarques.

→ Avec ce théorème, on aurait pu facilement montrer la proposition IV.4. En effet, pour tout $a \in \mathrm{G}$, $\langle a\rangle$ est un sous-groupe de G, donc d'après le théorème de Lagrange, $|\langle a\rangle| \mid |\mathrm{G}|$. D'après la proposition III.1, on a $|\langle a\rangle| = \omega(a)$, ce qui permet de conclure.

BY-NC-SA

→ En rejetant un coup d'œil à la démonstration du théorème, il est facile de voir que lorsque H est un sous-groupe de G avec G fini, alors $|G| = |G/H| \times |H|$ et par conséquent $|G/H| = |G| \, / \, |H|$.

On notera $[G : H] := |G/H|$ qu'on nommera l'indice de H dans G. Ce dernier peut être défini même si G est infini et que, si G est fini, on obtient $|G : H] = |G|/|H|$.

Exercice VII.5.

Soit p, q deux nombres premiers distincts et $(H, *)$ un groupe abélien tel que $|H| = pq$. Montrer que $H \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$

Exercice VII.6.

Supposons que G est un groupe fini et H un sous-groupe de G d'indice $[G : H] = 2$ *i.e.* $|G| = 2\,|H|$. Montrer que H est un sous groupe distingué de G et que donc H contient tous les carrés.

2. Actions de groupes

Dans cette partie, on considère X un ensemble non vide.

Définition VII.7.

Une action de groupe (à gauche) est une application $\bullet : G \times X \to X$, telle que

→ $\forall x \in X, \; e \bullet x = x$

→ $\forall g, h \in G \; \forall x \in X \; (gh) \bullet x = g \bullet (h \bullet x)$.

Remarque. On peut également définir les actions de groupe d'une autre manière. En effet, en considérant $\phi : G \longrightarrow \mathrm{Bij}(X)$ un morphisme de groupe de $(G, *)$ dans $(\mathrm{Bij}(X), \circ)$, pour tout $g \in G$ et $x \in X$, on peut remplacer $g \bullet x$ par $\phi(g)(x)$.

Exemples.

→ Lorsque X = G, $\phi_1 : (g, h) \longmapsto \gamma_g(h)$ (*i.e.* $g \bullet h = \gamma_g(h)$) et $\phi_2 : (g, h) \longmapsto \sigma_g(h)$ (*i.e.* $g \bullet h = \sigma_g(h)$) sont des actions de groupe. ϕ_1 est appelée action de translation à gauche et ϕ_2 est appelée action de conjugaison.

→ $\phi_3 : (g, h) \longmapsto \delta_g(h)$ une action de groupe seulement si G est abélien.

Définition VII.8.

Soit $x \in X$.

→ L'ensemble $O(x) = \{g \bullet x, \; g \in G\}$ est appelé orbite de x.

→ L'ensemble $\mathrm{Stab}(x) = \{g \in G, \; g \bullet x = x\}$ est appelé stabilisateur de x.

BY-NC-SA

Proposition VII.9.

Soit $\bullet : \mathrm{G} \times \mathrm{X} \longmapsto \mathrm{X}$ une action de groupe. Les propositions suivantes sont vraies.

1. Pour tous $x, y \in \mathrm{X}$, on a $\mathrm{O}(x) \cap \mathrm{O}(y) = \varnothing$ ou $\mathrm{O}(x) = \mathrm{O}(y)$. De plus, on a $\mathrm{X} = \bigcup_{x \in \mathrm{X}} \mathrm{O}(x)$.
2. Pour tout $x \in \mathrm{X}$, $\mathrm{Stab}(x)$ est un sous-groupe de G. De plus, on a l'implication suivante pour tout $y \in \mathrm{X}$
$$y = g \bullet x \Longrightarrow \mathrm{Stab}(y) = g\mathrm{Stab}(x)g^{-1} = \sigma_g(\mathrm{Stab}(x)).$$
3. Soit $x \in \mathrm{X}$. L'application $\phi_x : \begin{cases} \mathrm{G}/\mathrm{Stab}(x) & \longrightarrow \mathrm{O}(x) \\ \overline{g} & \longmapsto g \bullet x \end{cases}$ est bien définie et est une bijection. En particulier, si G est fini, $|\mathrm{G}/\mathrm{Stab}(x)| = |\mathrm{G}|/|\mathrm{Stab}(x)| = |\mathrm{O}(x)|$.
4. (Fomule des classes) Supposons G et X sont finis et soit $k \in \mathbb{N}^*$ et $(x_i)_{i \in [\![1;k]\!]}$ un système de représentants des orbites de G *i.e.* $\mathrm{X} = \bigsqcup_{i \in [\![1;k]\!]} \mathrm{O}(x_i)$ et pour tout $i, j \in [\![1;k]\!]$ différents, $\mathrm{O}(x_i) \neq \mathrm{O}(x_j)$. On a $|\mathrm{X}| = \sum_{i \in [\![1;k]\!]} |\mathrm{O}(x_i)| = \sum_{i \in [\![1;k]\!]} \frac{|\mathrm{G}|}{|\mathrm{Stab}(x_i)|}$.

Démonstration.

1. Soit $\sim$ la relation sur X^2 définie par $\forall x, y \in \mathrm{X},\ x \sim y \iff y \in \mathrm{O}(x)$. Montrons que la relation $\sim$ est une relation d'équivalence.
   - $\to$ Réflexivité : Pour tout $x \in \mathrm{X}$, $e \bullet x = x$ donc $x \in \mathrm{O}(x)$ et alors $x \sim x$.
   - $\to$ Symétrie : Soit $x, y \in \mathrm{X}$. Supposons que $x \sim y$. On dispose donc de $g \in \mathrm{G}$ tel que $g \bullet x = y$. On a alors $g^{-1} \bullet y = g^{-1} \bullet g \bullet x = x$ et alors $x \in \mathrm{O}(y)$, *i.e.* $y \sim x$.
   - $\to$ Transitivité : Soit $x, y, z \in \mathrm{X}$ tels que $x \sim y$ et $y \sim z$. Il existe donc $g, h \in \mathrm{G}$ tels que $g \bullet x = y$ et $h \bullet y = z$ et alors $g \bullet h \bullet x = z$. On en déduit donc que $gh \bullet x = z$ et donc $z \in \mathrm{O}(x)$, *i.e.* $x \sim z$.

   On peut également facilement voir que pour tout $x \in \mathrm{X}$, $\mathrm{O}(x)$ est la classe d'équivalence de x pour $\sim$. On peut donc partitionner X en classes d'équivalences pour $\sim$, *i.e.* pour tout $x, y \in \mathrm{X}$, on a $\mathrm{O}(x) \cap \mathrm{O}(y) = \varnothing$ ou $\mathrm{O}(x) = \mathrm{O}(y)$ et $\mathrm{X} = \bigcup_{x \in \mathrm{X}} \mathrm{O}(x)$.
2. Soit $x \in \mathrm{X}$. La démonstration du fait que $\mathrm{Stab}(x)$ est un sous-groupe de G ne présente pas de difficulté et est donc laissé comme exercice au lecteur. Soit $y \in \mathrm{G}$ et $h \in \mathrm{G}$, on a alors
$$\begin{aligned} h \in \mathrm{Stab}(y) &\iff h \in \mathrm{Stab}(g \bullet x) \iff h \bullet (g \bullet x) = g \bullet x \\ &\iff (g^{-1}hg) \bullet x = x \iff g^{-1}hg \in \mathrm{Stab}(x) \\ &\iff h \in g\mathrm{Stab}(x)g^{-1}, \end{aligned}$$
   et donc $\mathrm{Stab}(g \bullet x) = g\mathrm{Stab}(x)g^{-1}$.
3. Le fait que ϕ_x est définie ne présente pas de difficulté majeure et est donc laissé comme exercice au lecteur : il s'agit simplement de montrer que l'image $\phi_x(\overline{g})$ d'une classe $\overline{g}$ est indépendante du représentant). Montrons que ϕ_x est injective. Soit $g, h \in \mathrm{G}$. On a
$$\begin{aligned} \phi_x(\overline{g}) = \phi_x(\overline{h}) &\Longrightarrow g \bullet x = h \bullet x \Longrightarrow (g^{-1}h) \bullet x = x \\ &\Longrightarrow g^{-1}h \in \mathrm{Stab}(x) \Longrightarrow h \in g\mathrm{Stab}(x) \\ &\Longrightarrow h \in \overline{g} \Longrightarrow \overline{h} = \overline{g}, \end{aligned}$$

BY-NC-SA

donc ϕ_x est bien injective. Elle est de plus clairement surjective par définition de $\mathrm{O}(x)$, ce qui nous permet de conclure qu'elle est bien bijective.

4. On sait d'après le point (1) que les ensembles $(\mathrm{O}(x_i))_{i\in[\![1;k]\!]}$ sont disjoints. De plus, d'après le point 3, pour tout $x \in \mathrm{X}$, $\mathrm{G}/\mathrm{Stab}(x)$ est en bijection avec $\mathrm{O}(x)$, *i.e.* $|\mathrm{G}/\mathrm{Stab}(x)| = |\mathrm{O}(x)|$. On en déduit donc que
$$|\mathrm{X}| = \left|\bigsqcup_{i\in[\![1;k]\!]} \mathrm{O}(x_i)\right| = \sum_{i=1}^{k} |\mathrm{O}(x_i)| = \sum_{i=1}^{k} |\mathrm{G}/\mathrm{Stab}(x_i)| = \sum_{i=1}^{k} \frac{|\mathrm{G}|}{|\mathrm{Stab}(x_i)|}.$$

Exercice VII.10.

Soit G un groupe fini et $\mathrm{H} \leq \mathrm{G}$ tel que $[\mathrm{G} : \mathrm{H}] = p$ où p est le plus petit premier qui divise l'ordre de G. Montrer que $\mathrm{H} \trianglelefteq \mathrm{G}$ et que donc H contient toutes les puissances $p-$ièmes.

3. Application : action de conjugaison

Cette section traite le cas particulier de l'action par conjugaison (ou automorphisme intérieur) sur G fini
$$\bullet : \begin{cases} \mathrm{G}\times\mathrm{G} & \longrightarrow \mathrm{G} \\ (g,a). & \longmapsto gag^{-1} \end{cases}$$

Définition VII.11.

Soit $a \in \mathrm{G}$.

→ L'ensemble $\mathrm{O}(a) = \{xax^{-1},\ x \in \mathrm{G}\}$ est appelé orbite de a.

→ L'ensemble $\mathrm{C}(a) = \{x \in \mathrm{G},\ xa = ax\} = \{x \in \mathrm{G},\ xax^{-1} = a\} = \mathrm{Stab}(a)$ est appelé commutant de a. Il s'agit de l'ensemble des éléments de G qui commutent avec a.

Remarque. Pour tout $a \in \mathrm{G}$, $a \in \mathrm{Z}(\mathrm{G}) \iff \mathrm{C}(a) = \mathrm{G} \iff \mathrm{O}(a) = \{a\}$.

Proposition VII.12.

Soit $a \in \mathrm{G}$. Les propositions suivantes sont vraies.

1. $\mathrm{C}(a)$ est un sous-groupe de G et $|\mathrm{G}/\mathrm{C}(a)| = |\mathrm{O}(a)|$.
2. $|\mathrm{G}| = |\mathrm{C}(a)| \times |\mathrm{O}(a)|$.
3. (Formule des classes) Soit R un ensemble de représentants des classes de conjugaison (ou des orbites) non réduites à un singleton et R′ les représentants des orbites réduits à un singleton (*i.e.* les éléments qui commutent avec tous les éléments de G), alors
$$|\mathrm{G}| = \sum_{x\in\mathrm{R}'} |\mathrm{O}(x)| + \sum_{x\in\mathrm{R}} |\mathrm{O}(x)| = |\mathrm{Z}(\mathrm{G})| + \sum_{x\in\mathrm{R}} \frac{|\mathrm{G}|}{|\mathrm{C}(x)|}.$$

Démonstration. Les démonstrations de ces résultats sont des applications directes de la proposition VII.9.

L'exercice suivant est une application de la proposition ci-dessus.

BY-NC-SA

Exercice VII.13.

Soit $n \in \mathbb{N}^*$ et p un nombre premier. On suppose que $|\mathrm{G}| = p^n$. Montrer que le centre de G, Z(G), n'est pas réduit à un singleton. En particulier, si $n \leq 2$, alors G est abélien

Le théorème suivant est également une application de la proposition ci-dessus.

Théorème (Théorème de Cauchy) VII.14.

Supposons que G est fini et soit p un nombre premier tel que $p||\mathrm{G}|$. Il existe $x \in \mathrm{G}$, $\omega(x) = p$.

Démonstration. Posons $n = |\mathrm{G}|$ et considérons l'ensemble $\mathrm{E} = \{(x_1, \ldots, x_p) \in \mathrm{G}^p \ x_1 \ldots x_p = e\}$. Il est aisé de voir que

$$\begin{aligned} |\mathrm{E}| &= |\{(x_1, \ldots, x_p) \in \mathrm{G}^p \ x_1 \ldots x_p = e\}| \\ &= \left|\{(x_1, \ldots, x_{p-1}, x_{p-1}^{-1} \ldots x_1^{-1}),\ (x_1, \ldots, x_{p-1}) \in \mathrm{G}^{p-1}\}\right| = \left|\mathrm{G}^{p-1}\right|, \end{aligned}$$

et alors $|\mathrm{E}| = n^{p-1}$. Considérons l'action de groupe de permutation circulaire sur E

$$\bullet : \begin{cases} \mathbb{Z}/p\mathbb{Z} \times \mathrm{E} & \longrightarrow \mathrm{E} \\ (\overline{k}, (x_1, \ldots, x_p)) & \longmapsto (x_{\sigma^k(1)}, \ldots x_{\sigma^k(p)}), \end{cases}$$

où $\sigma = (1\ 2\ \ldots\ p)$. On notera par abus de notation $k \bullet x$ au lieu de $\overline{k} \bullet x$ (remarquer au passage que ce n'est pas vraiment un abus de notation vu que c'est plutôt l'action de $\mathbb{Z}$ qui est court-circuitée par $\mathbb{Z}/p\mathbb{Z}$). Remarquons que pour tout $k, \ell \in \mathbb{Z}$, $k \bullet (\ell \bullet x) = (k + \ell) \bullet x$. S'il existe $k \in \mathbb{Z}$ tel que $\overline{k} \neq \overline{0}$ et $k \bullet x = x$, alors d'après Bézout, étant donné que k est premier avec p, il existe $u, v \in \mathbb{Z}$ tels que $uk + vp = 1$ et alors, en supposant sans perte de généralité que $u \geq 0$ et $v \leq 0$

$$1 \bullet x = (uk + vp) \bullet x = vp \bullet (uk \bullet x) = \underbrace{-p \bullet \cdots \bullet -p}_{-v \text{ fois}} \bullet \underbrace{k \bullet \cdots \bullet k}_{u \text{ fois}} \bullet x = x.$$

Le fait que $1 \bullet x = x$ est équivalent à $(x_1, x_2, \ldots, x_p) = (x_2, x_3, \ldots, x_p, x_1)$ *i.e.*

$$x_1 = x_2,\ x_2 = x_3,\ \ldots,\ x_{p-1} = x_p,$$

et alors pour tout $k \in \mathbb{Z}$, $k \bullet x = x$ *i.e.* $\mathrm{O}(x)$ ne contient qu'un seul élément. On en déduit que pour tout $x \in \mathrm{E}$, deux cas sont possibles : $\forall k, \ell \in \mathbb{Z},\ k \bullet x \neq \ell \bullet x$ et alors $\mathrm{O}(x)$ contient p éléments, ou alors $\forall k \in \mathbb{Z},\ k \bullet x = x$ et donc $\mathrm{O}(x)$ contient un seul élément. En posant S_1 l'ensemble des orbites à un seul élément et S_2 l'ensemble des orbites à p éléments, le fait que E est union disjointe des orbites de l'action de groupe $\bullet$ nous permet de dire que

$$|\mathrm{G}|^{p-1} = |\mathrm{E}| = \left|\bigcup_{\mathrm{X} \in \mathrm{S}_1} \mathrm{X}\right| + \left|\bigcup_{\mathrm{X} \in \mathrm{S}_2} \mathrm{X}\right| = |\mathrm{S}_1| + p\,|\mathrm{S}_2|\,.$$

p divise $|\mathrm{G}|$, donc p divise aussi $|\mathrm{S}_1| = |\mathrm{K}|$. $\mathrm{O}((e, \ldots, e)) \in \mathrm{S}_1$ et $p \geq 2$ donc $|\mathrm{K}| = |\mathrm{S}_1| \geq 2$. Il existe donc $x \in \mathrm{E} \setminus \{e\}$ tel que $x^p = e$.

(cc) BY-NC-SA

Lorsque G est abélien, on dispose d'une démonstration plus rapide.

Exercice (Inspiré du TD d'Alain Troesch) VII.15.

Supposons que G est un groupe abélien fini et soit p un nombre premier.

1. Soit K un sous groupe distingué de G. Montrer que s'il existe $x \in \mathrm{G}/\mathrm{K}$ d'ordre p alors il en existe un aussi d'ordre p dans G.
2. On suppose que $p \mid |\mathrm{G}|$, montrer qu'il existe $x \in \mathrm{G}$ tel que $\omega(x) = p$.
3. Soit $a \geq 1$. On suppose que $p^a \mid |\mathrm{G}|$. Monter qu'il exsite $\mathrm{H} \leq \mathrm{G}$ d'ordre p^a.
4. Soit H et L deux sous groupes de G tels que $|\mathrm{H}| \wedge |\mathrm{L}| = 1$. G est commutatif, donc HL est un sous-groupe de G. Considérons le morphisme de groupes
$$\psi : \begin{cases} \mathrm{H} \times \mathrm{L} & \longrightarrow \mathrm{HL} \\ (h, \ell) & \longmapsto h\ell. \end{cases}$$
Montrer que ψ est bijectif et en déduire que $|\mathrm{HL}| = |\mathrm{H}| \times |\mathrm{L}|$.
5. En déduire que pour tout $n \in \mathbb{N}^*$, si $n \mid |\mathrm{G}|$, alors il existe H un sous groupe de G tel que $|\mathrm{H}| = n$.

(cc) BY-NC-SA

Correction de l'exercice I.2. :

Soit e' l'élément neutre de H. Pour tout $a \in \mathrm{H}$, on note a_g^{-1} un inverse à gauche de a, *i.e.* un élément de H tel que $a_g^{-1}a = e'$. Soit $a \in \mathrm{H}$. On a

$$a_g^{-1}aa_g^{-1} = e'a_g^{-1} = a_g^{-1},$$

et alors

$$a \cdot a_g^{-1} = e' \cdot a \cdot a_g^{-1} = (a_g^{-1})_g^{-1}a_g^{-1} \cdot a \cdot a_g^{-1} = (a_g^{-1})_g^{-1}a_g^{-1} = e'.$$

a admet donc aussi un inverse à droite qui est forcément le même qu'à gauche et donc H est un groupe.

Correction de l'exercice I.7. :

→ ($\Rightarrow$) Cette implication est évidente.

→ ($\Leftarrow$) $*$ est associative sur H, il faut donc simplement vérifier l'existence de l'inverse dans H et l'élément neutre. Soit $a \in \mathrm{H}$. H étant fini, on dispose de $i, j \in \mathbb{N}$ tels que $1 \leq i < j$ et $a^i = a^j$. Si $j = i+1$ alors $a = e$ et donc $a^{-1} = e \in \mathrm{H}$. Sinon, $j \geq i+2$ et donc $b = a^{j-i-1} \in \mathrm{H}$. Montrons que $b = a^{-1}$. On a

$$ba = a^{j-i-1}a = a^{j-i} = e = aa^{j-i-1} = ab,$$

donc tout élément de H admet un inverse dans H pour $*$. L'existence de l'élément neutre vient simplement du fait que si $a \in \mathrm{H}$, alors $a^{-1} \in \mathrm{H}$ et alors $e = a * a^{-1} \in \mathrm{H}$.

Correction de l'exercice I.11. :

Le fait que ce soit un morphisme est clair. En effet, pour tout $a, b \in \mathrm{G}$ et $x \in \mathrm{G}$, on a

$$\varphi(a*b)(x) = \gamma_{a*b}(x) = a*b*x = \gamma_a \circ \gamma_b(x) = \varphi(a) \circ \varphi(b)(x),$$

et alors $\varphi(a*b) = \varphi(a) \circ \varphi(b)$.

Montrons à présent que φ est injective. Pour cela, on va montrer que $\operatorname{Ker}\varphi = \{\mathrm{Id}\}$. Soit $a \in \operatorname{Ker}\varphi$. On a $\varphi(a)(e) = e$ et donc $a \cdot e = e$ d'où $a = e$, ce qui donne $\operatorname{Ker}\varphi = \{e\}$.

Remarque. Le seul élément a de G tel que γ_a admet un point fixe est e.

Correction de l'exercice I.13. :

1. Montrons cette proposition par double implication.

   → ($\Rightarrow$) Supposons que HK est un sous-groupe de G. On a alors

   $$\mathrm{HK} \underset{\mathrm{HK} \leq \mathrm{G}}{=} (\mathrm{HK})^{-1} = \mathrm{K}^{-1}\mathrm{H}^{-1} \underset{\mathrm{H\ et\ K} \leq \mathrm{G}}{=} \mathrm{KH}.$$

   La première égalité est due au fait qu'étant donné que HK est un sous-groupe de G, alors l'application $i : \begin{cases} \mathrm{HK} & \longrightarrow \mathrm{HK} \\ x & \longmapsto x^{-1} \end{cases}$ est bijective.

   → ($\Rightarrow$) Supposons que HK = KH. Utilisons la proposition I.4 pour montrer que HK est un sous-groupe de G. HK $\neq \varnothing$ il suffit donc de vérifier que

   $$\forall (h, k, h', k') \in \mathrm{H} \times \mathrm{K} \times \mathrm{H} \times \mathrm{K},\ (hk)\,(h'k')^{-1} \in \mathrm{HK}.$$

   Méthode 1 (Rapide) : On a pour tout $(h, k, h', k') \in \mathrm{H} \times \mathrm{K} \times \mathrm{H} \times \mathrm{K}$,

   $$(hk)(h'k')^{-1} = hkk'^{-1}h'^{-1} \underset{\mathrm{K} \leq \mathrm{G}}{\in} h\mathrm{K}h'^{-1} \underset{\mathrm{H} \leq \mathrm{G}}{\subset} \mathrm{HKH} \underset{\mathrm{HK}=\mathrm{KH}}{=} \mathrm{HHK} \underset{\mathrm{H} \leq \mathrm{G}}{=} \mathrm{HK}.$$

(cc) BY-NC-SA

Méthode 2 (Même chose mais en plus détaillé) : Soit $(h,k,h',k') \in \mathrm{H}\times\mathrm{K}\times\mathrm{H}\times\mathrm{K}$. On a $(hk)(h'k')^{-1} = \underbrace{hkk'^{-1}}_{\in \mathrm{HK}=\mathrm{KH}} h'^{-1}$. Il existe donc $(h'',k'') \in \mathrm{H}\times\mathrm{K}$ tel que $hkk'^{-1} = k''h''$, et alors on peut écrire $(hk)(h'k')^{-1} = k''h''h' \in \mathrm{KH} = \mathrm{HK}$.

2. Supposons que H est un sous-groupe distingué de G. Deux méthodes sont possibles.
   → **Méthode 1 :** On a pour tout $(h,k,h',k') \in \mathrm{H}\times\mathrm{K}\times\mathrm{H}\times\mathrm{K}$, $(hk)(h'k') = hkh'k' = hkh'k^{-1}kk'$. H est un sous-groupe distingué de G, donc $kh'k^{-1} \in \mathrm{H}$. Il existe donc $h'' \in \mathrm{H}$ tel que $kh'k^{-1}$. On a donc $(hk)(h'k') = hh'''k' \in \mathrm{HK}$. De plus, HK est non vide, donc d'après l'exercice I.7, HK est bien un sous-groupe de G.
   → **Méthode 2 :** En utilisant la proposition VII.1, on peut écrire
   $$\mathrm{HK} = \bigcup_{k\in\mathrm{K}} \mathrm{H}k \underset{\mathrm{H}\trianglelefteq\mathrm{G}}{=} \bigcup_{k\in\mathrm{K}} k\mathrm{H} = \mathrm{KH}.$$
3. Montrons le résultat par double implication.
   → ($\Leftarrow$) Soit $x \in \mathrm{H}\cap\mathrm{K}$. On a $f(x,e) = x = f(e,x)$ et donc, par injectivité, $x = e$.
   → ($\Rightarrow$) Supposons que $\mathrm{H}\cap\mathrm{K} = \{e\}$. f est surjective par définition. Soit $(h,k,h',k') \in \mathrm{H}\times\mathrm{K}\times\mathrm{H}\times\mathrm{K}$. On a
   $$f(h,k) = f(h',k') \iff hk = h'k' \iff (h')^{-1}h = k'k^{-1} \underset{\mathrm{H}\cap\mathrm{K}=\{e\}}{=} e \iff h = h' \text{ et } k = k',$$
   donc f est bijective.
   Remarque. Si les éléments de H et K ne commutaient pas entre eux, f ne serait pas forcément un morphisme.
4. Supposons que H et K sont des sous-groupes distingués de G et que $\mathrm{H}\cap\mathrm{K} = \{e\}$. Pour tout $(h,k,h',k') \in \mathrm{H}\times\mathrm{K}\times\mathrm{H}\times\mathrm{K}$, on a
   $$hkh^{-1}k^{-1} \in h\mathrm{K}h^{-1}k^{-1} \cap hk\mathrm{H}k^{-1} \underset{\mathrm{H}\text{ et }\mathrm{K}\trianglelefteq\mathrm{G}}{=} \mathrm{K}k^{-1} \cap h\mathrm{H} \underset{\mathrm{H}\text{ et }\mathrm{K}\leq\mathrm{G}}{=} \mathrm{K}\cap\mathrm{H} = \{e\},$$
   et donc $hk = kh$
5. Supposons que les éléments de H et K commutent entre eux et que $\mathrm{H}\cap\mathrm{K} = \{e\}$. D'après la question 3, f est bijective. De plus, f est un morphisme (facile à vérifier) et donc un isomorphisme.

Correction de l'exercice IV.2. :

Notons $\overline{1}_{\mathbb{Z}/m\mathbb{Z}}$ la classe de 1 dans $\mathbb{Z}/m\mathbb{Z}$. Soit $f \in \mathrm{Hom}(\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}/n\mathbb{Z})$. On sait que $f(\overline{1}_{\mathbb{Z}/m\mathbb{Z}}) \in \mathbb{Z}/n\mathbb{Z}$ et donc $nf(\overline{1}_{\mathbb{Z}/m\mathbb{Z}}) = 0$. En particulier, $\omega(f(\overline{1}_{\mathbb{Z}/m\mathbb{Z}}))|n$. De même, $\omega(f(\overline{1}_{\mathbb{Z}/m\mathbb{Z}}))|\omega(\overline{1}_{\mathbb{Z}/m\mathbb{Z}}) = m$ et donc $\omega(f(\overline{1}_{\mathbb{Z}/m\mathbb{Z}}))|m \wedge n = 1$ *i.e.* $f(1) = 0$ et alors $f = 0$. On en déduit donc que $\mathrm{Hom}(\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}/n\mathbb{Z}) = \{0\}$.

Attention : La loi considérée ici est additive, donc pour tout $a \in \mathbb{Z}/m\mathbb{Z}$ (ou $\mathbb{Z}/n\mathbb{Z}$), on note $\omega(a) = \min\{k \geq 1, ka = 0\}$ au lieu de $\omega(a) = \min\{k \geq 1, a^k = 1\}$.

Correction de l'exercice IV.3. :

1. Considérons l'exemple où $(\mathrm{G},*) = (\mathrm{GL}(\mathbb{R}^2), \circ)$. Soit $\theta, \theta' \in \mathbb{R}$ et $\mathrm{S}_\theta, \mathrm{S}_{\theta'}$ les symétries par rapport aux axes faisant respectivement un angle θ et θ' par rapport à l'axe des abscisses dans le sens trigonométrique. On rappelle que
   $$\mathrm{S}_\theta \circ \mathrm{S}_{\theta'} = \mathrm{R}_{\theta-\theta'}.$$
   où pour tout $\alpha \in \mathbb{R}$, R_α désigne la rotation d'angle θ dans le sens trigonométrique dans $\mathbb{R}^2$.
   En particulier, si on prend θ et θ' vérifiant $\dfrac{\theta-\theta'}{2\pi} \in \mathbb{R}\setminus\mathbb{Q}$ alors
   $$\mathrm{S}_\theta^2 = \mathrm{Id},\ \mathrm{S}_{\theta'}^2 = \mathrm{Id} \text{ et } \forall n \in \mathbb{N},\ \mathrm{R}_{\theta-\theta'}^n = \mathrm{R}_{n(\theta-\theta')}$$

(cc) BY-NC-SA

Pour tout $n \in \mathbb{N}$, $n(\theta-\theta') \notin 2\pi\mathbb{Z}$ car $\dfrac{\theta-\theta'}{2\pi} \notin \mathbb{Q}$ et alors $\mathrm{R}_{n(\theta-\theta')} \neq \mathrm{Id}$. On en déduit donc que S_θ et $\mathrm{S}_{\theta'}$ sont d'ordre 2 mais leur produit est d'ordre infini.

2. On a $(ab)^{mn} = (a^m)^n(b^n)^m = e$, donc ab est d'ordre fini et $\omega(ab)|mn$. Posons $\ell = \omega(ab)$. Remarquons maintenant que $(ab)^\ell = e$ et donc $(ab)^{\ell m} = e = (a^m)^\ell b^{\ell m} = b^{\ell m}$ d'où $n|m\ell$ et donc par Gauß, étant donné que $m \wedge n = 1$, on a $n|\ell$. On peut montrer de la même manière que $m|\ell$ et que donc, vu que $m \wedge n = 1$, on a $mn|\ell$ et finalement $\ell = mn$.

3. Décomposons m et n en facteurs premiers. Soit $r \in \mathbb{N}$, $p_1, \dots, p_r$ des nombres premiers distincts et $\alpha_1, \dots, \alpha_r \in \mathbb{N}$, $\beta_1, \dots, \beta_r \in \mathbb{N}$ tels que

$$m = p_1^{\alpha_1} \dots p_r^{\alpha_r} \text{ et } n = p_1^{\beta_1} \dots p_r^{\beta_r}.$$

Quitte à réordonner les p_i, supposons sans perte de généralité qu'il existe $k \in [\![1;r]\!]$ tel que

$$\forall i \in [\![1;k]\!],\ \max(\alpha_i, \beta_i) = \alpha_i \text{ et } \forall i \in [\![k+1;r]\!],\ \max(\alpha_i, \beta_i) = \beta_i.$$

Posons ensuite

$$m' = p_{k+1}^{\alpha_{k+1}} \dots p_r^{\alpha_r},\ n' = p_1^{\beta_1} \dots p_k^{\beta_k},\ a' = a^{m'} \text{ et } b' = b^{n'}.$$

On a alors $\omega(a') = \dfrac{m}{m'}$ et $\omega(b') = \dfrac{n}{n'}$ ce qui donne

$$\omega(a') \wedge \omega(b') = \frac{n \wedge m}{n'm'} = \frac{n \wedge m}{p_1^{\min(\alpha_1,\beta_1)} \dots p_r^{\min(\alpha_r,\beta_r)}} = \frac{n \wedge m}{n \wedge m} = 1,$$

et donc d'après la question précédente

$$\omega(a'b') = \omega(a')\omega(b') = \frac{mn}{m'n'} = \frac{mn}{m \wedge n} = m \vee n.$$

4. Posons $\mathrm{S} = \{\omega(x),\ x \in \mathrm{G}\}$. Étant donné que S est fini non vide, on peut, en utilisant la question précédente, trouver un élément $z \in \mathrm{G}$ tel que $\omega(z) = \bigvee\limits_{m \in \mathrm{S}} m$ qui vérifie bien la propriété voulue.

<u>Correction de l'exercice V.4. :</u>

On envisage deux méthodes.

→ **Méthode 1 (plus intuitive) :** H est un sous-groupe de G qui est généré par a, donc tout élément de H s'écrit sous forme de a^ℓ avec $\ell \in \mathbb{N}$. Soit $\ell = \min\{k \in \mathbb{N}^*,\ a^k \in \mathrm{H}\}$. Montrons que $\left\langle a^\ell \right\rangle = \mathrm{H}$. Supposons le contraire, *i.e.* qu'il existe $m \in \mathbb{N}$ tel que $a^m \notin \left\langle a^\ell \right\rangle$ et $a^m \in \mathrm{H}$. On peut donc écrire $m = q\ell + r$ avec $q \in \mathbb{Z}$ et $r \in [\![0; \ell-1]\!]$ (car k ne divise pas m). On a alors

$$a^r = a^{-q\ell}a^{q\ell+r} = \left(a^\ell\right)^{-q} a^m \in \mathrm{H},$$

ce qui est absurde par définition de ℓ. On en déduit donc que $\mathrm{H} = \left\langle a^\ell \right\rangle$. D'après la proposition IV.4, $\omega(a^\ell)|\mathrm{H}$, on a alors $\left(a^\ell\right)^d = e$ et donc $a^{d\ell} = e$. D'après le point 1 de la proposition 1, on a $n|d\ell$. De plus, d'après la proposition IV.4, H étant un sous-groupe cyclique de G, on a $d|n$ et donc $\dfrac{n}{d}\Big|\ \ell$. Posons donc $\ell = \alpha\dfrac{n}{d}$ avec $\alpha \in \mathbb{N}$. On a alors

$$\mathrm{H} = \left\langle a^\ell \right\rangle = \left\langle a^{\alpha \frac{n}{d}} \right\rangle \subset \left\langle a^{\frac{n}{d}} \right\rangle.$$

On a de plus d'après le point 5 de la proposition III.1, $\left|\left\langle a^{\frac{n}{d}} \right\rangle\right| = \dfrac{n}{n/d} = d = |\mathrm{H}|$ et donc $|\mathrm{H}| = \left\langle a^{\frac{n}{d}} \right\rangle$.

(cc) BY-NC-SA

$\rightarrow$ **Méthode 2 (plus rapide) :** a est un générateur de G donc l'application $\varphi : \begin{cases} \mathbb{Z} & \longrightarrow \mathrm{G} \\ k & \longmapsto a^k \end{cases}$ est surjective et $\operatorname{Ker}\varphi = n\mathbb{Z}$. On a donc $\mathrm{H} = \varphi(\varphi^{-1}(\mathrm{H}))$ et d'après la proposition I.9, $\varphi^{-1}(\mathrm{H})$ est un sous-groupe de $\mathbb{Z}$, *i.e.* de la forme $k\mathbb{Z}$ avec $k \in \mathbb{N}$. De plus, $\{e\} \subset \mathrm{H}$ et donc

$$n\mathbb{Z} = \varphi^{-1}(\{e\}) \subset \varphi^{-1}(\mathrm{H}) = k\mathbb{Z}.$$

On en déduit alors que $k|n$ et que

$$\mathrm{H} = \varphi(k\mathbb{Z}) = \{a^\ell,\ \ell \in k\mathbb{Z}\} = \left\langle a^k \right\rangle.$$

De plus, $d = |\mathrm{H}| = \left|\left\langle a^k \right\rangle\right| = \dfrac{n}{k}$ et donc $k = \dfrac{n}{d}$ et finalement $\mathrm{H} = \left\langle a^{\frac{n}{d}} \right\rangle$.

Correction de l'exercice V.5. :

1. Soit $k \in \mathbb{Z}$. D'après le point (4) de la proposition III.1, on a $\omega(a^k) = \dfrac{\omega(a)}{\omega(a) \wedge k} = \dfrac{n}{n \wedge k}$. On en déduit donc que
$$\mathrm{G} = \left\langle a^k \right\rangle \Longleftrightarrow \omega(a^k) = n \Longleftrightarrow \frac{n}{n \wedge k} = n \Longleftrightarrow n \wedge k = 1.$$

2. On envisage deux méthodes.

$\rightarrow$ **Méthode 1 :** Notons pour tout $d \in \mathbb{N}^*$ diviseur de n, $\mathrm{G}_d = \{b \in \mathrm{G},\ \omega(b) = d\}$. On sait que d'après l'exercice V.4, pour tout d diviseur positif de n, tout sous-groupe de cardinal d est égal à $\mathrm{H}_d = \left\langle a^{\frac{n}{d}} \right\rangle$. On a donc

$$b \in \mathrm{G}_d \Longleftrightarrow |\langle b \rangle| = d \underset{\langle b \rangle \leq \mathrm{G}}{\Longleftrightarrow} \langle b \rangle = \mathrm{H}_d.$$

On a donc pour tout d diviseur positif de n,

$$\begin{aligned} |\mathrm{G}_d| &= |\{b \in \mathrm{G},\ \omega(b) = d\}| \\ &= \left|\left\{ \left(a^{\frac{n}{d}}\right)^k,\ k \in \mathbb{N},\ \left\langle a^{\frac{kd}{n}} \right\rangle = \mathrm{H}_d \right\}\right| \\ &= |\{k \in [\![1; d]\!],\ k \wedge d = 1\}| = \varphi(d). \end{aligned}$$

L'avant-dernière égalité est due au fait que $a^{\frac{n}{d}}$ est un générateur de H_d et donc d'après la question précédente $\left(a^{\frac{n}{d}}\right)^k$ génère H_d si et seulement si $k \wedge d = 1$. On a alors

$$n = |\mathrm{G}| = \left| \bigsqcup_{d|n,\ d\geq 1} \mathrm{G}_d \right| = \sum_{d|n,\ d\geq 1} |\mathrm{G}_d| = \sum_{d|n,\ d\geq 1} \varphi(d).$$

(cc) BY-NC-SA

→ **Méthode 2 :** On a

$$\begin{aligned}
n = |[\![1;n]\!]| &= \left|\bigsqcup_{d|n,\ d\geq 1} \{k \in [\![1;n]\!]\ k\wedge n = d\}\right| \\
&= \sum_{d|n,\ d\geq 1} |\{k \in [\![1;n]\!]\ k\wedge n = d\}| \\
&= \sum_{d|n,\ d\geq 1} \left|\left\{k \in d\times [\![1;\frac{n}{d}]\!],\ \frac{k}{d}\wedge \frac{n}{d} = 1\right\}\right| \\
&= \sum_{d|n,\ d\geq 1} \left|\left\{k \in [\![1;\frac{n}{d}]\!],\ k\wedge \frac{n}{d} = 1\right\}\right| \\
&= \sum_{d|n,\ d\geq 1} \varphi\left(\frac{n}{d}\right) = \sum_{d|n,\ d\geq 1} \varphi(d).
\end{aligned}$$

La dernière égalité est vraie car

$$h : \begin{cases} \{d \geq 1,\ d|n\} & \longrightarrow \{d\geq 1,\ d|n\} \\ d & \longmapsto \dfrac{n}{d}. \end{cases}$$

est une bijection.

Correction de l'exercice V.6. :

1. Notons $n = |\mathrm{G}_1|$, $m = |\mathrm{G}_2|$ et a, b des générateurs de G_1 et G_2 respectivement. On note également e_1 et e_2 les éléments neutres respectifs de $(\mathrm{G}_1, *)$ et $(\mathrm{G}_2, *)$. On envisage deux cas possibles.
   → **Cas** 1 **:** $n \wedge m = 1$.
   Soit $z = (a, b) = (e_1, b) \otimes (a, e_2)$. On a clairement $\omega((a, e_2)) = n$ et $\omega((e_1, b)) = m$ donc d'après l'exercice IV.3. $\omega(z) = mn$ et donc étant donné que $|\mathrm{G}_1 \times \mathrm{G}_2| = mn$, alors $\mathrm{G}_1 \times \mathrm{G}_2 = \langle z\rangle$, *i.e.* $(\mathrm{G}_1 \times \mathrm{G}_2, \otimes)$ est cyclique.
   → **Cas** 2 **:** $n \wedge m \neq 1$.
   Posons alors $\ell = n \vee m$. Il est clair que $\forall z \in \mathrm{G}_1 \times \mathrm{G}_2,\ z^{\ell} = (e_1, e_2)$ et donc $\mathrm{G}_1 \times \mathrm{G}_2$ ne peut pas être cyclique car pour tout $z \in \mathrm{G}$, $|\langle z\rangle| \leq \ell < mn = |\mathrm{G}_1 \times \mathrm{G}_2|$.

   On en déduit donc que $\mathrm{G}_1 \times \mathrm{G}_2$ est cyclique si et seulement si $|\mathrm{G}_1| \wedge |\mathrm{G}_2| = 1$.
2. Pour vérifier que f est bien définie, il faut vérifier que tout élément de $\mathbb{Z}/ab\mathbb{Z}$ a une image unique par f. On doit donc vérifier que pour tout $x, y \in \mathbb{Z}$, si dans $\mathbb{Z}/ab\mathbb{Z}$ on a $\overline{x} = \overline{y}$, alors $\varphi(\overline{x}) = \varphi(\overline{y})$. Notons pour tout $z \in \mathbb{Z}$ et $r \geq 2$, $\overline{z}_{(r)}$ la classe de z dans $\mathbb{Z}/r\mathbb{Z}$.

   Considérons donc $x, y \in \mathbb{Z}$ tels que dans $\mathbb{Z}/ab\mathbb{Z}$, on ait $\overline{x} = \overline{y}$. Il existe donc $k \in \mathbb{Z}$ tel que $x = y + kab$. On a alors

   $$\varphi\left(\overline{x}_{(ab)}\right) = \left(\overline{x}_{(a)}, \overline{x}_{(b)}\right) = \left(\overline{y}_{(a)} + \overline{kab}_{(a)}, \overline{y}_{(b)} + \overline{kab}_{(b)}\right) = \left(\overline{y}_{(a)}, \overline{y}_{(b)}\right) = \varphi\left(\overline{y}_{(ab)}\right).$$

   φ est donc bien définie. Il est également facile de vérifier qu'il s'agit d'un morphisme. Montrons maintenant qu'il s'agit d'un isomorphisme si et seulement si $a \wedge b = 1$.
   → $(\Leftarrow)$ Supposons que $a \wedge b = 1$. On a $\varphi\left(\overline{1}_{(ab)}\right) = \left(\overline{1}_{(a)}, \overline{1}_{(b)}\right)$. D'après la question 1, étant donné que $\overline{1}_{(a)}$ et $\overline{1}_{(b)}$ sont d'ordre respectivement a et b et $a \wedge b = 1$, alors $\omega\left(\left(\overline{1}_{(a)}, \overline{1}_{(b)}\right)\right) = ab$. On a de plus

   $$\left\langle\left(\overline{1}_{(a)}, \overline{1}_{(b)}\right)\right\rangle = \left\langle\varphi\left(\overline{1}_{(ab)}\right)\right\rangle \subset \operatorname{Im}\varphi,$$

   et donc

   $$|\operatorname{Im}\varphi| \geq \left|\left\langle\left(\overline{1}_{(a)}, \overline{1}_{(b)}\right)\right\rangle\right| = ab.$$

(cc) BY-NC-SA

Or $|\operatorname{Im}\varphi| \leq |\mathbb{Z}/a\mathrm{Z} \times \mathbb{Z}/b\mathbb{Z}| = ab$, donc $|\operatorname{Im}\varphi| = ab$ et alors $\operatorname{Im}\varphi = \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$, *i.e.* φ est surjective. De plus, $|\mathbb{Z}/ab\mathbb{Z}| = ab = |\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}|$ donc la surjectivité de φ nous donne directement que φ est bijective, c'est alors un isomorphisme.

→ $(\Rightarrow)$ Supposons que $a \wedge b = d > 1$. On a alors $a \vee b \in]0, ab[$ et donc $\overline{a \vee b}_{(ab)} \neq \overline{0}_{(ab)}$, mais

$$\varphi\left(\overline{a \vee b}_{(ab)}\right) = \left(\overline{a \vee b}_{(a)}, \overline{a \vee b}_{(b)}\right) = \left(\overline{0}_{(a)}, \overline{0}_{(b)}\right).$$

On en déduit que $\operatorname{Ker}\varphi \neq \left\{\overline{0}_{(ab)}\right\}$ et que donc φ n'est pas injective. φ ne peut alors pas être un isomorphisme.

Correction de l'exercice V.7. :

1. Soit $\mathrm{P} \in \mathrm{A}[\mathrm{X}]$ et $\mathrm{Q} \in \mathrm{A}[\mathrm{X}]$ de coefficient dominant égal à 1. Il s'agit de montrer la proposition suivante

$$\exists! \mathrm{B}, \mathrm{R} \in \mathrm{A}[\mathrm{X}],\ \mathrm{P}(\mathrm{X}) = \mathrm{B}(\mathrm{X})\mathrm{Q}(\mathrm{X}) + \mathrm{R}(\mathrm{X}),\ \deg \mathrm{R} < \deg \mathrm{Q}.$$

→ **Existence :** Procédons par récurrence forte sur le degré de P. Posons $n = \deg \mathrm{P}$.

- Si $\deg \mathrm{P} = 0$, alors si $\deg \mathrm{Q} > 0$, $\mathrm{B} = 0$ et $\mathrm{R} = \mathrm{P}$ conviennent. si $\deg \mathrm{Q} = 0$ *i.e.* $\mathrm{Q} = 1$, alors $\mathrm{R} = 0$ et $\mathrm{B} = \mathrm{P}$ conviennent.
- Soit $n \in \mathbb{N}$. Supposons que la propriété est vraie pour tout n, *i.e.* pour tout polynôme P de degré inférieur ou égal à n,

$$\exists \mathrm{B}, \mathrm{R} \in \mathrm{A}[\mathrm{X}],\ \mathrm{P}(\mathrm{X}) = \mathrm{B}(\mathrm{X})\mathrm{Q}(\mathrm{X}) + \mathrm{R}(\mathrm{X}),\ \deg \mathrm{R} < \deg \mathrm{Q}.$$

Montrons que la propriété est vraie pour $n+1$. On suppose que P est de degré $n+1$ et on pose

$$\mathrm{P}(\mathrm{X}) = \sum_{k=0}^{n+1} a_k \mathrm{X}^k \text{ et } \mathrm{Q}(\mathrm{X}) = \mathrm{X}^r + \mathrm{H}(\mathrm{X}).$$

avec $r = \deg \mathrm{Q}$ et $\mathrm{H} \in \mathrm{A}[\mathrm{X}]$ tel que $\deg \mathrm{H} < r$. Si $r > n+1$, alors $\mathrm{B} = 0$ et $\mathrm{R} = \mathrm{Q}$ conviennent. Si $r \leq n+1$ on a alors

$$\mathrm{P}(\mathrm{X}) - a_{n+1}\mathrm{X}^{n+1-r}\mathrm{Q}(\mathrm{X}) = \sum_{k=0}^{n} a_k \mathrm{X}^k - a_{n+1}\mathrm{X}^{n+1-r}\mathrm{H}(\mathrm{X}).$$

On a donc

$$\deg(\mathrm{P}(\mathrm{X}) - a_{n+1}\mathrm{X}^{n+1-r}\mathrm{Q}(\mathrm{X})) = \deg\left(\sum_{k=0}^{n} a_k \mathrm{X}^k - a_{n+1}\mathrm{X}^{n+1-r}\mathrm{H}(\mathrm{X})\right) \leq n.$$

Par hypothèse de récurrence, il existe $\tilde{\mathrm{B}}, \tilde{\mathrm{R}} \in \mathrm{A}[\mathrm{X}]$ tels que $\deg \tilde{\mathrm{R}} < \deg \mathrm{Q}$ tels que

$$\mathrm{P}(\mathrm{X}) - a_{n+1}\mathrm{X}^{n+1-r}\mathrm{Q}(\mathrm{X}) = \tilde{\mathrm{B}}(\mathrm{X})\mathrm{Q}(\mathrm{X}) + \tilde{\mathrm{R}}(\mathrm{X}),$$

et alors

$$\mathrm{P}(\mathrm{X}) = (\tilde{\mathrm{B}}(\mathrm{X}) + a_{n+1}\mathrm{X}^{n+1-r})\mathrm{Q}(\mathrm{X}) + \mathrm{R}(\mathrm{X}).$$

On en déduit que $\mathrm{B}(\mathrm{X}) = \tilde{\mathrm{B}}(\mathrm{X}) + a_{n+1}\mathrm{X}^{n+1-r}$ et $\tilde{\mathrm{R}}(\mathrm{X}) = \mathrm{R}(\mathrm{X})$ conviennent, d'où l'existence.

→ **Unicité :** Soit $\mathrm{B}_1, \mathrm{B}_2, \mathrm{R}_1, \mathrm{R}_2 \in \mathrm{A}[\mathrm{X}]$ tels que $\deg \mathrm{R}_1 \leq \deg \mathrm{Q}$, $\deg \mathrm{R}_2 \leq \deg \mathrm{Q}$ et

$$\mathrm{P}(\mathrm{X}) = \mathrm{B}_1(\mathrm{X})\mathrm{Q}(\mathrm{X}) + \mathrm{R}_1(\mathrm{X}) = \mathrm{B}_2(\mathrm{X})\mathrm{Q}(\mathrm{X}) + \mathrm{R}_2(\mathrm{X}).$$

On a alors

$$(\mathrm{B}_1(\mathrm{X}) - \mathrm{B}_2(\mathrm{X}))\mathrm{Q}(\mathrm{X}) = \mathrm{R}_1(\mathrm{X}) - \mathrm{R}_2(\mathrm{X}).$$

(cc) BY-NC-SA

Si $B_1 - B_2 \neq 0$, alors étant donné que Q est non nul, on a

$$\begin{aligned} \deg(R_1 - R_2) &= \deg((B_1 - B_2)Q) &(11)\\ &= \deg\left((B_1 - B_2)X^r + (B_1 - B_2)H)\right) &(12)\\ &= \deg(B_1 - B_2) + \deg Q \geq \deg Q. &(13) \end{aligned}$$

ce qui est absurde. On en déduit que $B_1 = B_2$ et que $R_1 = R_2$, d'où l'unicité.

Attention : Lorsque $U, V \in A[X]$ et que A n'est pas intègre, on n'a pas forcément l'égalité $\deg(UV) = \deg U + \deg V$, mais ici, on peut passer de la ligne (11) à (13) car le coefficient dominant de Q est égal à 1. Ce passage serait également vrai si le coefficient dominant de Q n'était pas un diviseur de zéro.

2. Montrons ce résultat par récurrence sur le degré de P encore une fois.
   - $\rightarrow$ Lorsque $\deg P = 0$, P n'a clairement pas de racines (le cas $P = 0$ correspond à $\deg P = -\infty$).
   - $\rightarrow$ Soit $n \in \mathbb{N}$. Supposons que la propriété est vraie lorsque $\deg P \in [\![0; n]\!]$ et supposons maintenant que $\deg P = n+1$. Si P n'admet pas de racines, la propriété est vraie. Supposons que P admet une racine $a \in A$. La question précédente nous permet d'effectuer la division euclidienne de P par $X - a$ (ce polynôme est de coefficient dominant égal à 1). Il existe donc $B, R \in A[X]$ tel que
   $$\deg R < \deg(X - a) = 1 \text{ et } P(X) = (X - a)B(X) + R(X).$$
   On a de plus, $0 = P(a) = R(a)$ et R est constant donc $R = 0$. B est de degré n, donc par hypothèse de récurrence, B admet au plus n racines et donc $P(X) = (X - a)B(X)$ admet au plus $n + 1 = \deg P$ racines.
3. Si G est cyclique, alors nécessairement on a $|G| = \bigvee_{x\in G} \omega(x)$. Notre intuition est donc de considérer un élément de G d'ordre le plus grand qu'on peut trouver (ici égal à n) et d'essayer de montrer qu'il engendre G. Posons $n = \bigvee_{x\in G} \omega(x)$. D'après la question 4 de l'exercice IV.3, il existe $z \in G$ tel que $\omega(z) = \bigvee_{x\in G} \omega(x) = n$. n est un multiple de tous les ordres des éléments de G, donc pour tout $x \in G$, $x^n = 1$. En posant Z l'ensemble des racines de $X^n - 1$ dans $\mathbb{K}$, on voit que $G \subset Z$. D'après la question précédente, le polynôme $X^n - 1$ admet au plus n racines, donc on a $|G| \leq |Z| \leq n$. De plus, on a $n = |\langle z\rangle| \leq |G|$ et donc $|G| = n = |\langle z\rangle|$ et $\langle z\rangle \subset G$, et finalement $G = \langle z\rangle$, *i.e.* G est cyclique.

Correction de l'exercice VI.2. :

On peut voir G comme un $\mathbb{Z}/p\mathbb{Z}$-espace vectoriel. En effet, en considérant les lois (bien définies)

$$+ : \begin{cases} G \times G & \longrightarrow G \\ (x, y) & \longmapsto x * y \end{cases} \quad \text{et} \cdot : \begin{cases} \mathbb{Z}/p\mathbb{Z} \times G & \longrightarrow G \\ (\overline{k}, x) & \longmapsto x^k, \end{cases}$$

on peut voir que $(G, +, \cdot)$ est un $\mathbb{Z}/p\mathbb{Z}-$espace vectoriel. G est fini et donc de dimension finie. En posant $n = \dim_{\mathbb{Z}/p\mathbb{Z}} G$, on peut affirmer l'existence d'une base de G, $(x_1, \ldots, x_n)$. On peut donc affirmer que tout $x \in G$ s'écrit d'une manière unique sous forme de

$$x = \overline{k_1} \cdot x_1 + \cdots + \overline{k_n} \cdot x_n,\ \overline{k_1}, \ldots, \overline{k_n} \in \mathbb{Z}/p\mathbb{Z}.$$

En considérant donc l'isomorphisme (il est facile de montrer qu'il est bien défini et qu'il s'agit d'un isomorphisme)

$$\varphi : \begin{cases} G & \longrightarrow (\mathbb{Z}/p\mathbb{Z})^n \\ \overline{k_1}x_1 + \cdots + \overline{k_n}x_n & \longmapsto (\overline{k_1}, \ldots, \overline{k_n}). \end{cases}$$

(cc) BY-NC-SA

On voit donc que $\mathrm{G} \simeq (\mathbb{Z}/p\mathbb{Z})^n$. En particulier, on remarquera que $|\mathrm{G}| = |\mathbb{Z}/p\mathbb{Z}|^n = p^n$.

Remarque. En utilisant le théorème de Cauchy (VII.14), on peut très facilement montrer qu'il existe $n \in \mathbb{N}^*$ tel que $|\mathrm{G}| = p^n$. En effet, pour tout $g \in \mathrm{G} \setminus \{e\}$, $g^p = e$ et donc $\omega(g)|p$ *i.e.* $\omega(g) = p$. Pour tout nombre premier q, si $q|\,|\mathrm{G}|$, alors par le théorème de Cauchy il existe $g \in \mathrm{G}$ tel que $\omega(g) = q$ et alors $q = p$. p est donc le seul nombre premier qui divise $|\mathrm{G}|$ ce qui signifie qu'il existe $n \in \mathbb{N}^*$ tel que $|\mathrm{G}| = p^n$.

Correction de l'exercice VI.3. :

Montrons le résultat par implications successives.

→ (1) ⇒ (2) Montrer que $\mathrm{H} \simeq \mathbb{Z}/p\mathbb{Z}$ est équivalent à montrer que H est engendré par un élément d'ordre p premier. En effet, s'il existe $x \in \mathrm{H}$ tel que $\langle x\rangle = \mathrm{G}$ et $\omega(x) = p \in \mathbb{P}$, alors il est facile de montrer que l'application $\psi : \begin{cases} \mathrm{H} & \longrightarrow \mathbb{Z}/p\mathbb{Z} \\ x^k & \longmapsto \overline{k} \end{cases}$ est bien définie et est un isomorphisme et que donc en particulier $\mathrm{H} \simeq \mathbb{Z}/p\mathbb{Z}$.

Montrons maintenant que G est engendré par un élément d'ordre premier. Soit $x \in \mathrm{G} \setminus \{e\}$. $\langle x\rangle \in \mathrm{Sg(H)} \setminus \{\{e\}\}$ et donc $\langle x\rangle = \mathrm{H}$. Montrons que $\omega(x)$ est premier.

Si $\omega(x) = \infty$, alors $\langle x^2\rangle$ est un sous-groupe de G. Il est facile de vérifier que $\langle x^2\rangle \neq \langle x\rangle$ et alors on a $\langle x^2\rangle = \{e\}$, *i.e.* $\omega(x) \leq 2$ ce qui est en contradiction avec le fait que $\omega(x) = \infty$. On en déduit donc que $\omega(x)$ est fini.

Soit d un diviseur de $\omega(x)$ supérieur ou égal à 2. D'après la proposition III.1, on a $\omega(x^d) = \dfrac{\omega(x)}{d}$, ce qui donne nécessairement $\left\langle x^d\right\rangle \in \mathrm{Sg(H)} \setminus \{\mathrm{H}\}$ et donc $\left\langle x^d\right\rangle = \{e\}$, ce qui signifie que $x^d = e$, et alors encore d'après la proposition III.1 $\omega(x)|d$ et finalement $d = \omega(x)$. On en déduit donc que les seuls diviseurs positifs de $\omega(x)$ sont 1 et $\omega(x)$, *i.e.* que $\omega(x)$ est premier (ou égal à 1, mais ce cas est impossible car $x \neq e$). En posant $p = \omega(x)$, on en déduit d'après ce qui précède que $\mathrm{H} \simeq \mathbb{Z}/p\mathbb{Z}$ avec p premier.

→ (2) ⇒ (3) Si $\mathrm{H} \simeq \mathbb{Z}/p\mathbb{Z}$, alors et $|\mathrm{G}| = |\mathbb{Z}/p\mathbb{Z}| = p$, ce qui implique d'après le théorème de Lagrange faible (IV.4) que tout élément a de $\mathrm{H} \setminus \{e\}$ divise p et donc $\omega(a) \in \{1, p\}$. On en déduit alors que pour tout $a \in \mathrm{H} \setminus \{e\}$, $|\langle a\rangle| = \omega(a) = p = |\mathrm{H}|$ *i.e.* $\langle a\rangle = \mathrm{H}$, d'où le résultat.

→ (3) ⇒ (4) G est monogène, on peut donc considérer $g \in \mathrm{G}$ tel que $\langle g\rangle = \mathrm{G}$. Deux cas se présentent.

- Si $\omega(g) = \infty$, on a par hypothèse $\langle g^2\rangle = \mathrm{G}$ ce qui est impossible. En effet, $g \notin \langle g^2\rangle$ car sinon il existerait $k \in \mathbb{Z}$ tel que $g^{2k} = g$, *i.e.* $g^{2k-1} = e$ ce qui est absurde.
- Si $\omega(g) = n \in \mathbb{N}^*$, alors le morphisme

$$\psi' : \begin{cases} \mathbb{Z}/n\mathbb{Z} & \longrightarrow \mathrm{H} \\ \overline{k} & \longmapsto g^k \end{cases}$$

est bien défini et est bijectif. On en déduit alors que $\mathrm{H} \simeq \mathrm{Z}/n\mathbb{Z}$. De plus, si d est un diviseur positif de n différent de n, alors d'après le point 5 de la proposition III.1, $\left|\left\langle g^d\right\rangle\right| = \omega(g^d) = \dfrac{n}{d}$. On a donc

$$\frac{n}{d} = \omega(g^d) = \left|\left\langle g^d\right\rangle\right| = |\mathrm{H}| = n,$$

et donc $d = 1$. On en déduit alors que les seuls diviseurs de n sont 1 et n, *i.e.* que n est premier et alors $|\mathrm{H}| = |\mathbb{Z}/n\mathbb{Z}| = n$ est premier.

→ (4) ⇒ (1) Soit L un sous-groupe de H. H est cyclique donc d'après le résultat de l'exercice V.4 L est également cyclique. On a donc d'après le théorème faible de Lagrange (IV.4) $|\mathrm{L}|\,|\,|\mathrm{H}|$, mais H est premier, donc $|\mathrm{L}| \in \{1, |\mathrm{H}|\}$ et finalement $\mathrm{Sg(H)} = \{\mathrm{H}, \{e\}\}$.

BY-NC-SA

Correction de l'exercice VII.5. :

D'après le théorème faible de Lagrange (IV.4), l'ordre de tout élément de H divise pq et donc pour tout $a \in \mathrm{H} \setminus \{e\}$, $\omega(a) \in \{p, q, pq\}$.

- → Si tout élément de $\mathrm{H} \setminus \{e\}$ est d'ordre p, alors d'après l'exercice VI.2, il existe $n \in \mathbb{N}$ tel que $|\mathrm{H}| = p^n$ ce qui est absurde.
- → De même, si tout élément de $\mathrm{H} \setminus \{e\}$ est d'ordre q, alors il existe $n \in \mathbb{N}$ tel que $|\mathrm{H}| = q^n$ ce qui est absurde.
- → S'il existe un élément $a \in \mathrm{H}$ d'ordre pq, alors $\langle a \rangle = \mathrm{H}$ et donc d'après la proposition V.3, $\mathrm{H} \simeq \mathbb{Z}/pq\mathbb{Z}$.
- → S'il existe un élément $a \in \mathrm{H}$ d'ordre p et un élément $b \in \mathrm{H}$ d'ordre q, alors d'après la question 2 de l'exercice IV.3, $\omega(ab) = pq$ et donc $\mathrm{H} = \langle ab \rangle$ et alors pour les mêmes raisons qu'au point précédent, on a $\mathrm{H} \simeq \mathbb{Z}/pq\mathbb{Z}$.

D'après le théorème chinois (exercice V.6), on a $\mathbb{Z}/pq\mathbb{Z} \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$ et finalement $\mathrm{H} \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$.

Correction de l'exercice VII.6. :

On a par hypothèse $|\mathrm{G}/\mathrm{H}| = 2$ donc il y a deux classes à gauche. Soit $a \in \mathrm{G} \setminus \mathrm{H}$. On a $a\mathrm{H} \neq \mathrm{H}$, donc $\mathrm{G}/\mathrm{H} = \{\mathrm{H}, a\mathrm{H}\}$. De la même manière, G a deux classes à droite et donc étant donné que $\mathrm{H}a \neq \mathrm{H}$, ces deux classes sont H et $\mathrm{H}a$. On a donc $\mathrm{G} = a\mathrm{H} \sqcup \mathrm{H} = \mathrm{H}a \sqcup \mathrm{H}$ et alors $a\mathrm{H} = \mathrm{H}a = \mathrm{G} \setminus \mathrm{H}$. De même, lorsque $a \in \mathrm{H}$, on a $a\mathrm{H} = \mathrm{H} = \mathrm{H}a$ et alors en déduit donc que H est bien un sous-groupe distingué de G. Montrons à présent que H contient tous les carrés. $|\mathrm{G}/\mathrm{H}| = 2$ donc l'ordre de tout élément de G/H divise 2, *i.e.* pour tout $x \in \mathrm{H}$, $(x\mathrm{H})^2$ est égal à H, l'élément neutre du groupe G/H. On en déduit que pour tout $x \in \mathrm{G}$, $\mathrm{H} = x\mathrm{H}x\mathrm{H} = x^2\mathrm{H}$ et alors $x^2 \in \mathrm{H}$. H contient donc tous les carrés.

Correction de l'exercice VII.10. :

Intuition : Pour montrer que H est distingué, il faut et il suffit de montrer que pour tout $h \in \mathrm{H}$ et $a \in \mathrm{G}$, $ha\mathrm{H} = a\mathrm{H}$. En effet, lorsque cela est vérifié, alors on a pour tout $a \in \mathrm{G}$ et $h \in \mathrm{H}$, $a^{-1}ha\mathrm{H} = a^{-1}a\mathrm{H} = \mathrm{H}$ et donc $a^{-1}ha \in \mathrm{H}$. Ceci donne directement que $a^{-1}\mathrm{H}a = \mathrm{H}$ ce qui signifie que H est un sous-groupe distingué de G.

Pour montrer que pour tout $h \in \mathrm{H}$ et $a \in \mathrm{G}$, $ha\mathrm{H} = \mathrm{H}$, on va considérer une action de groupe $\bullet$ telle que pour tout $h \in \mathrm{H}$ et $a \in \mathrm{G}$, $h \bullet a\mathrm{H} = ha\mathrm{H}$.

Considérons donc l'action de groupe $\bullet$ de H sur G/H définie par

$$\bullet : \begin{cases} \mathrm{H} \times \mathrm{G}/\mathrm{H} & \longrightarrow \mathrm{G}/\mathrm{H} \\ (h, a\mathrm{H}) & \longmapsto ha\mathrm{H}. \end{cases}$$

Pour montrer que H est distingué, il suffit de montrer que pour tout $h \in \mathrm{H}$ et $a \in \mathrm{G}$, $ha\mathrm{H} = a\mathrm{H}$, *i.e.* que $\mathrm{O}(a\mathrm{H}) = \{a\mathrm{H}\}$. Soit $a \in \mathrm{G}$. Supposons que $|\mathrm{O}(a\mathrm{H})| \neq 1$. D'après le point 3 de la proposition VII.9, on a $\mathrm{O}(a\mathrm{H})| \, |\mathrm{H}|$. De plus, par le théorème de Lagrange (VII.4), H étant un sous-groupe de G, on a $|\mathrm{H}| \, | \, |\mathrm{G}|$ et donc $|\mathrm{O}(a\mathrm{H})| \, | \, |\mathrm{G}|$. p étant le plus petit diviseur de $|\mathrm{G}|$ strictement supérieur à 1, on en déduit que $|\mathrm{O}(a\mathrm{H})| \geq p$. Remarquons de plus que pour tout $h \in \mathrm{H}$, $h \bullet \mathrm{H} = \mathrm{H}$ et donc $\mathrm{O}(\mathrm{H}) = \{\mathrm{H}\}$.

On a alors d'après le point 4 de la proposition VII.9, si R est l'ensemble des représentants des orbites de l'action $\bullet$ de H sur G/H, alors

$$p = |\mathrm{G}/\mathrm{H}| = \sum_{x \in \mathrm{R}} |\mathrm{O}(x)| \geq |\mathrm{O}(\mathrm{H})| + |\mathrm{O}(a\mathrm{H})| \geq p + 1,$$

ce qui est absurde, donc $|\mathrm{O}(a\mathrm{H})| = 1$ *i.e.* $\mathrm{O}(a\mathrm{H}) = \{a\mathrm{H}\}$ et alors H est bien un sous-groupe distingué de G.

BY-NC-SA

Correction de l'exercice VII.13. :

Soit R l'ensemble des représentants de classes de conjugaison de G non réduites à un singleton. En appliquant la formule des classes à G (point 3 de la proposition VII.12), on obtient

$$|\mathrm{Z}(\mathrm{G})| = |\mathrm{G}| - \sum_{x\in\mathrm{R}} |\mathrm{O}(x)| = |\mathrm{G}| - \sum_{x\in\mathrm{R}} \frac{|\mathrm{G}|}{\mathrm{C}(x)}.$$

De plus, pour tout $x \in \mathrm{R}$, $\left.\dfrac{|\mathrm{G}|}{|\mathrm{C}(x)|}\right| |\mathrm{G}|$ et $|\mathrm{G}| = p^n$, donc tous les diviseurs strictement positifs de $|\mathrm{G}|$ sont de la forme p^k avec $k \in [\![0; n]\!]$. Aussi, pour tout $x \in \mathrm{R}$, $\dfrac{|\mathrm{G}|}{|\mathrm{C}(x)|} = |\mathrm{O}(x)| \neq 1$ et alors

$$\forall x \in \mathrm{R},\ \exists k \in [\![1; n]\!],\ \frac{|\mathrm{G}|}{|\mathrm{C}(x)|} = p^k,$$

et donc

$$\mathrm{Z}(\mathrm{G}) \equiv p^n - \sum_{x\in\mathrm{R}} \frac{|\mathrm{G}|}{|\mathrm{C}(x)|} \equiv 0[p],$$

De plus, $e \in \mathrm{Z}(\mathrm{G})$ donc $\mathrm{Z}(\mathrm{G}) \geq p$. $\mathrm{Z}(\mathrm{G})$ n'est alors pas réduit à un singleton.

Supposons à présent que $n \leq 2$ et montrons que G est abélien.

→ Si $n = 1$, alors on a $p = |\mathrm{G}| \geq |\mathrm{Z}(\mathrm{G})| \geq p$ donc $|\mathrm{Z}(\mathrm{G})| = p$ et alors $\mathrm{Z}(\mathrm{G}) = \mathrm{G}$ *i.e.* G est abélien.

→ Si $n = 2$, alors on sait que $\mathrm{Z}(\mathrm{G})$ est un sous-groupe de G, donc par le théorème de Lagrange (VII.4), $|\mathrm{Z}(\mathrm{G})| \,|\, |\mathrm{G}|$ et donc $|\mathrm{Z}(\mathrm{G})| = p$ ou $|\mathrm{Z}(\mathrm{G})| = p^2$. Si $|\mathrm{Z}(\mathrm{G})| = p^2$, alors on peut dire comme avant que $\mathrm{Z}(\mathrm{G}) = \mathrm{G}$ et alors que G est abélien.

Soit $x \in \mathrm{G} \setminus \mathrm{Z}(\mathrm{G})$. On sait encore une fois, d'après le théorème de Lagrange (VII.4), que $|\langle x\rangle| \,|\, |\mathrm{G}|$, et $x \neq e$, donc $|\langle x\rangle| \in \{p, p^2\}$. Si $|\langle x\rangle| = p^2$, alors $\mathrm{G} = \langle x\rangle$ et on en déduit immédiatement que G est commutatif. Sinon, on a encore une fois d'après Lagrange $|\langle\{x\} \cup \mathrm{Z}(\mathrm{G})\rangle| \in \{p, p^2\}$ et $|\langle\{x\} \cup \mathrm{Z}(\mathrm{G})\rangle| > |\langle x\rangle| = p$ et donc $|\langle\{x\} \cup \mathrm{Z}(\mathrm{G})\rangle| = p^2$ et alors $\langle\{x\} \cup \mathrm{Z}(\mathrm{G})\rangle = \mathrm{G}$. D'après la proposition VI.1, on peut écrire

$$\mathrm{G} = \langle\{x\} \cup \mathrm{Z}(\mathrm{G})\rangle = \{a_1^{\alpha_1} \dots a_n^{\alpha_n},\ n \in \mathbb{N},\ a_1, \dots, a_n \in \{x\} \cup \mathrm{Z}(\mathrm{G}),\ \alpha_1, \dots, \alpha_n \in \mathbb{Z}\}.$$

Il est facile de montrer que ce groupe est abélien, et donc on en déduit que G est abélien.

Correction de l'exercice VII.15. :

1. Soit $x \in \mathrm{G}$ tel que $\omega(\overline{x}) = p$. En considérant le morphisme introduit à la proposition VII.3

$$\pi : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{G}/\mathrm{K} \\ x & \longmapsto \overline{x} = x\mathrm{K}, \end{cases}$$

on peut affirmer d'après le point 1 de la proposition IV.1 que $\omega(\pi(x))|\omega(x)$ *i.e.* $p|\omega(x)$. On a alors d'après le point 5 de la proposition III.1, $\omega\left(x^{\frac{\omega(x)}{p}}\right) = \dfrac{\omega(x)}{\omega(x)/p} = p$.

2. Posons $\mathrm{G} = kp$ avec $k \in \mathbb{N}^*$ et procédons par récurrence forte sur k.

→ Si $k = 1$, alors $|\mathrm{G}| = p$, alors la propriété est vraie d'après l'exercice VI.3.

→ Soit $k \in \mathbb{N}$, supposons que pour tout $\ell \in [\![1; k]\!]$, si $|\mathrm{G}| = \ell p$, alors G vérifie la propriété voulue. Supposons maintenant que $|\mathrm{G}| = (k+1)p$ et montrons qu'il existe un élément de G d'ordre p. Soit $x \in \mathrm{G} \setminus \{e\}$.

(cc) BY-NC-SA

- Si $p|\omega(x)$, alors encore une fois, d'après le point 5 de la proposition III.1,

$$\omega\left(x^{\frac{\omega(x)}{p}}\right) = \frac{\omega(x)}{\omega(x)/p} = p.$$

- Si $p \nmid \omega(x)$, alors on a $|\mathrm{G}/\langle x\rangle| = \dfrac{|\mathrm{G}|}{\omega(x)} = \dfrac{k+1}{\omega(x)}p$ et $\dfrac{k+1}{\omega(x)} \in [\![1;k]\!]$ (bien entendu, $\dfrac{k+1}{\omega(x)} \in \mathbb{N}$ car étant donné que $\omega(x) \wedge p = 1$, d'après Gauss $\omega(x)|k+1$). On en déduit alors par hypothèse de récurrence qu'il existe $y \in \mathrm{G}/\langle x\rangle$ tel que $\omega(y) = p$ et alors d'après la question 1, il existe $z \in \mathrm{G}$ tel que $\omega(z) = p$ ce qui bien le résultat voulu.

3. Procédons par récurrence forte sur a.
   - $\rightarrow$ Lorsque $a = 1$, la propriété est vraie d'après la question précédente.
   - $\rightarrow$ Soit m le plus grand entier tel que $p^m|\,|\mathrm{G}|$ et soit $k \in [\![1;m-1]\!]$. Supposons que le résultat est vrai pour tout $a \in [\![1;k]\!]$ et montrons qu'il est également vrai pour $a = k+1$. Par hypothèse de récurrence, il existe H sous-groupe de G tel que $|\mathrm{H}| = p^k$. On a alors $|\mathrm{G}/\mathrm{H}| = |\mathrm{G}|/p^k$ et $k \leq m-1$ donc $p|\,|\mathrm{G}/\mathrm{H}|$. On sait donc d'après la question précédente (qui donne le résultat pour $a = 1$) qu'il existe M un sous-groupe de G/H tel que $|\mathrm{M}| = p$. Considérons l'endomorphisme (identique à celui considéré à la question 1)

$$\pi : \begin{cases} \mathrm{G} & \longrightarrow \mathrm{G}/\mathrm{H} \\ x & \longmapsto x\mathrm{H}. \end{cases}$$

Posons $\mathrm{L} = \pi^{-1}(\mathrm{M})$ (c'est un groupe) et considérons le morphisme de groupe

$$\tilde{\pi} : \begin{cases} \mathrm{L} & \longrightarrow \mathrm{M} \cap \operatorname{Im} \pi \\ x & \longmapsto \pi(x). \end{cases}$$

π est surjective par définition, donc $\mathrm{M} \cap \operatorname{Im}\pi = \mathrm{M}$. On a de plus d'après la proposition VII.3,

$$\mathrm{L}/\operatorname{Ker}\tilde{\pi} \simeq \operatorname{Im}\tilde{\pi}. \tag{14}$$

Aussi, M est un sous-groupe de G/H et contient donc son élément neutre H, et alors

$$\mathrm{H} = \operatorname{Ker}\pi = \pi^{-1}(\{\mathrm{H}\}) \subset \pi^{-1}(\mathrm{M}) = \mathrm{L},$$

ce qui donne

$$\operatorname{Ker}\tilde{\pi} = \mathrm{L} \cap \operatorname{Ker}\pi = \operatorname{Ker}\pi = \mathrm{H},$$

et donc $\operatorname{Ker}\tilde{\pi} = \operatorname{Ker}\pi$. De plus, on a également

$$\operatorname{Im}\tilde{\pi} = \pi(\mathrm{L}) = \pi(\pi^{-1}(\mathrm{M})) = \mathrm{M} \cap \operatorname{Im}\pi \underset{\pi \text{ surjective}}{=} \mathrm{M} \cap \mathrm{G}/\mathrm{H} = \mathrm{M}$$

et donc on peut réécrire l'égalité (1) comme $\mathrm{L}/\mathrm{H} = \mathrm{M}$ ce qui implique

$$|\mathrm{L}| = |\mathrm{H}| \times |\mathrm{M}| = p^k \times p = p^{k+1}.$$

L est donc un sous-groupe de G de cardinal p^{k+1}, ce qui est bien le résultat voulu.

4. D'abord, par commutation, HL = LH et donc ce dernier est bien un sous-groupe de G (voir question 1 exercice I.13). De même, par commutation, ψ est bien un morphisme de groupes. Montrons maintenant que ψ est injective. On envisage deux méthodes.
   - $\rightarrow$ **Méthode 1 (à la main) :** Soit $(h,\ell) \in \operatorname{Ker}\psi$. On a alors $h\ell = e$. Posons $p = \omega(h)$ et $q = \omega(\ell)$. D'après le théorème de Lagrange, on a $p|\,|\mathrm{H}|$ et $q|\,|\mathrm{L}|$. De plus, on a $|\mathrm{H}| \wedge \mathrm{L} = 1$ ce qui donne

(cc) BY-NC-SA

$p \wedge q = 1$. On a alors d'après Bézout, il existe $u, v \in \mathbb{Z}$ tels que $up + vq = 1$. On a alors

$$e = (h\ell)^{vq} = h^{1-up}\ell^{vq} = h \text{ et } e = (h\ell)^{up} = h^{up}\ell^{1-vq} = \ell,$$

et donc $(h, \ell) = (e, e)$. On en déduit donc que $\operatorname{Ker}\psi = \{(e, e)\}$ et que donc ψ est injective.

→ **Méthode 2 :** Soit $(h, \ell) \in \operatorname{Ker}\psi$. On a alors $h\ell = e$ et donc $h = \ell^{-1} \in \mathrm{H} \cap \mathrm{L}$.
Or, $\mathrm{H} \cap \mathrm{L} \leq \mathrm{H}$ et $\mathrm{H} \cap \mathrm{L} \leq \mathrm{L}$ et donc, par Lagrange, $|\mathrm{H} \cap \mathrm{L}| \,|\, |\mathrm{H}| \wedge |\mathrm{L}| = 1$ *i.e.* $\mathrm{H} \cap \mathrm{L} = \{e\}$.
Ainsi, $h = \ell = e$ et donc ψ est injective.

ψ est clairement surjective donc bijective et donc $|\mathrm{HL}| = |\mathrm{H} \times \mathrm{L}| = |\mathrm{H}| \times |\mathrm{L}|$.

5. On peut montrer facilement par récurrence en utilisant la question précédente le résultat suivant. Pour tout $r \in \mathbb{N}^*$, si $\mathrm{H}_1, \ldots, \mathrm{H}_r$ sont r sous-groupes de G tels que pour tout $i, j \in [\![1; r]\!]$ différents, $|\mathrm{H}_i| \wedge |\mathrm{H}_j| = 1$, alors $\mathrm{H}_1 \ldots \mathrm{H}_r$ est un sous-groupe de G et

$$|\mathrm{H}_1 \ldots \mathrm{H}_r| = |\mathrm{H}_1| \times \cdots \times |\mathrm{H}_r|\,.$$

Écrivons maintenant la décomposition en produits de nombres premiers de n (le cas $n = 1$ est trivial). Soit $r \in \mathbb{N}^*$, $\alpha_1, \ldots, \alpha_r \in \mathbb{N}^*$ et $p_1, \ldots, p_r$ r nombres premiers distincts tels que $n = p_1^{\alpha_1} \ldots p_r^{\alpha_r}$. D'après la question 3, pour tout $k \in [\![1; r]\!]$, il existe H_k un sous groupe de G tel que $|\mathrm{H}_k| = p_k^{\alpha_k}$. De plus, on a clairement pour tout $i, j \in [\![1; r]\!]$ différents $|\mathrm{H}_i| \wedge |\mathrm{H}_j| = 1$ et alors on en déduit que $\mathrm{H}_1 \ldots \mathrm{H}_r$ est un sous-groupe de G et que $|\mathrm{H}_1 \ldots \mathrm{H}_r| = p_1^{\alpha_1} \ldots p_r^{\alpha_r} = n$.

(cc) BY-NC-SA

Chapitre 23

Polynômes

Notations

Soit $\mathbb{K}$ un corps et $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$.

→ On note $\mathrm{Z}_{\mathbb{K}}(\mathrm{P})$ l'ensemble des racines de P dans $\mathbb{K}$. Lorsqu'il n'y a pas ambiguïté sur $\mathbb{K}$, on note simplement $\mathrm{Z}(\mathrm{P})$.

→ On note val P la valuation de P, c'est-à-dire le coefficient du terme de plus petit degré de P.

I Préambule

Dans tout le chapitre, $\mathbb{K}$ désigne un corps.

Proposition I.1.

Soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$.

1. Si $\deg \mathrm{P} = 1$, alors P est irréductible.
2. Si $\deg(\mathrm{P}) \in \{2, 3\}$, alors P est irréductible si et seulement s'il n'a pas de racine.

Démonstration.

1. Ce point est clair, car si $\mathrm{P} = \mathrm{QR}$ et $\deg \mathrm{P} = 1$ alors nécessairement Q ou R est constant.
2. Montrons les deux implications.
   → ($\Rightarrow$) Si P est irréductible de degré $n \geq 2$, il est sans racine (s'il admet une racine a, $\mathrm{P} = (\mathrm{X}-a)\mathrm{Q}$ avec Q non constant, absurde).
   → ($\Leftarrow$) Procédons par contraposée. si $\deg \mathrm{P} \in \{2, 3\}$ se factorise de manière non triviale, $\mathrm{P} = \mathrm{QR}$ avec $1 \leq \deg \mathrm{Q} \leq \deg \mathrm{R}$ et $\deg \mathrm{P} = \deg \mathrm{Q} + \deg \mathrm{R}$, donc $1 \leq \deg \mathrm{Q} \leq \dfrac{\deg(\mathrm{P})}{2} \leq 1.5$, donc $\deg \mathrm{Q} = 1$ et Q admet une racine dans $\mathbb{K}$ et finalement P aussi.

Remarque. Le second point de la proposition ci-dessus est faux en général lorsque $\deg \mathrm{P} \geq 4$. En effet, $\mathrm{X}^4 + 1$ dans $\mathbb{R}[\mathrm{X}]$ est réductible car sa factorisation dans $\mathbb{R}[\mathrm{X}]$ est $(\mathrm{X}^2 + \sqrt{2}\mathrm{X} + 1)(\mathrm{X}^2 - \sqrt{2}\mathrm{X} + 1)$ mais n'admet pas de racine dans $\mathbb{R}$.

Proposition I.2.

Si $\mathbb{K}, \mathbb{L}$ sont deux corps tels que $\mathbb{K} \subset \mathbb{L}$ et $(\mathrm{A}, \mathrm{B}) \in (\mathbb{K}[\mathrm{X}] - \{0\})^2$, alors $\mathrm{A} \wedge \mathrm{B}$ et $\mathrm{A} \vee \mathrm{B}$ sont les mêmes dans $\mathbb{L}[\mathrm{X}]$ et dans $\mathbb{K}[\mathrm{X}]$.

Démonstration.

→ Pour $\mathrm{A} \wedge \mathrm{B}$: le calcul de cette quantité se fait par l'algorithme d'Euclide qui manipule des éléments de $\mathbb{K}[\mathrm{X}]$. Tous ces éléments restent dans $\mathbb{K}[\mathrm{X}]$ au fil des étapes de l'algorithme, donc le résultat de l'algorithme, *i.e.* $\mathrm{A} \wedge \mathrm{B}$, reste dans $\mathbb{K}[\mathrm{X}]$.

→ Pour $\mathrm{A} \vee \mathrm{B}$: $\mathrm{A} \vee \mathrm{B} = \dfrac{\mathrm{AB}}{\mathrm{A} \wedge \mathrm{B}}$, il s'agit donc du rapport de deux éléments de $\mathbb{K}[\mathrm{X}]$ et est de plus le même par invariance de $\mathrm{A} \wedge \mathrm{B}$.

(cc) BY-NC-SA

II Polynômes complexes

Théorème (Théorème de D'Alembert-Gauss) II.1.

Les polynômes irréductibles de $\mathbb{C}[X]$ sont de degré 1. D'une manière équivalente, pour tout $P \in \mathbb{C}[X]$ non constant, il existe $\lambda_1, \ldots, \lambda_n, a \in \mathbb{C}$ tels que

$$P(X) = a \prod_{i=1}^{n} (X - \lambda_i).$$

Démonstration. La démonstration de ce théorème dépasse le cadre de ce cours et ne sera donc pas faite.

Proposition II.2.

Soit $P, Q \in \mathbb{K}[X]$, $\lambda, \mu \in \mathbb{K}^*$ et $z_1, \ldots, z_r \in \mathbb{K}$ distincts tels que

$$P(X) = \lambda \prod_{i=1}^{r} (X - z_i)^{\alpha_i} \text{ et } Q(X) = \mu \prod_{i=1}^{r} (X - z_i)^{\beta_i}.$$

Les deux égalités suivantes sont toujours vérifiées.

$$P \wedge Q = \prod_{i=1}^{r} (X - z_i)^{\min(\alpha_i, \beta_i)} \text{ et } P \vee Q = \prod_{i=1}^{r} (X - z_i)^{\max(\alpha_i, \beta_i)}.$$

Exemple. Pour tout $m, n \in \mathbb{N}^*$, on a, dans $\mathbb{C}[X]$, $X^m - 1 \wedge X^n - 1 = X^{m \wedge n} - 1$. En effet, on a

$$X^m - 1 \wedge X^n - 1 = \prod_{\omega \in \mathbb{U}_m} (X - \omega) \wedge \prod_{\omega \in \mathbb{U}_n} (X - \omega) = \prod_{\omega \in \mathbb{U}_m \cap \mathbb{U}_n} (X - \omega) = \prod_{\omega \in \mathbb{U}_{m \wedge n}} (X - \omega) = X^{m \wedge n} - 1.$$

Justifions le fait que $\mathbb{U}_n \cap \mathbb{U}_m = \mathbb{U}_{n \wedge m}$.

→ ($\subset$) Pour tout $z \in \mathbb{U}_{n \wedge m}$, on a $z^n = 1$ et $z^m = 1$ car $n \wedge m$ divise n et m.

→ ($\supset$) D'après Bézout, il existe $a, b \in \mathbb{Z}$ tels que $am + bn = a \wedge b$, donc pour tout $z \in \mathbb{U}_n \cap \mathbb{U}_m$, on a $z^{n \wedge m} = (z^n)^a \cdot (z^m)^b = 1$, *i.e.* $z \in \mathbb{U}_{n \wedge m}$.

Proposition (Localisation des racines) II.3.

Soit $P \in \mathbb{C}_n[X]$ unitaire : $P = X^n + a_{n-1} X^{n-1} + \cdots + a_0$. Pour tout $z \in \mathbb{C}$, si $P(z) = 0$, alors

$$|z| \leq 1 + \max_{k \in [\![0;n]\!]} |a_k|.$$

Démonstration. La démonstration de ce résultat a déjà été faite au chapitre 1.

Proposition II.4.

Pour toute suite $(P_k)_k \in \mathbb{N}$ à valeurs dans $\mathbb{C}_n[X]$, la suite de fonctions (P_k) converge simplement vers $P \in \mathbb{C}_n[X]$ si et seulement si les coefficients de (P_k) convergent vers ceux de P.

Démonstration.

→ ($\Leftarrow$) Supposons que les coefficients de (P_k) convergent vers ceux de P. Soit $a \in \mathbb{C}$. Il est facile de voir que $P_k(a) \xrightarrow[k \to +\infty]{} P(a)$ en passant à la limite.

(cc) BY-NC-SA

→ ($\Rightarrow$) Supposons que (P_k) converge simplement vers P. Soit $b_0, \dots, b_n \in \mathbb{C}$ deux à deux différents. Considérons les deux normes suivantes dans $\mathbb{C}_n[\mathrm{X}]$

$$\|.\|_{\infty,\text{coeff}} : \begin{cases} \mathbb{C}_n[\mathrm{X}] & \longrightarrow \mathbb{R}^+ \\ \mathrm{P} & \longmapsto \max\limits_{i\in[\![0;n]\!]} |\alpha_i| \end{cases} \qquad \|.\| : \begin{cases} \mathbb{C}_n[\mathrm{X}] & \longrightarrow \mathbb{R}^+ \\ \mathrm{P} & \longmapsto \sum\limits_{i=0}^{n} |\mathrm{P}(b_i)|, \end{cases}$$

avec $\mathrm{P}(\mathrm{X}) = \alpha_n \mathrm{X}^n + \cdots + \alpha_1 \mathrm{X} + \alpha_0$.
$\mathbb{C}_n[\mathrm{X}]$ est de dimension finie, donc les normes $\|.\|_{\infty,\text{coeff}}$ et $\|.\|$ sont équivalentes. Le fait que (P_k) converge simplement vers P implique que (P_k) converge vers P pour $\|.\|$ et donc par équivalence des normes, P_k converge vers P pour $\|.\|_{\infty,\text{coeff}}$, *i.e.* les coefficients des termes de (P_k) convergent vers ceux de P.

Exercice II.5.

Soit $n \geq 1$ et $(\mathrm{P}_k)_{k\in\mathbb{N}}$ une suite de polynômes unitaires de $\mathbb{C}_n[\mathrm{X}]$. On suppose que (P_k) converge simplement vers $\mathrm{P} \in \mathbb{C}_n[\mathrm{X}]$ unitaire de degré n. Soit U un ouvert non vide de $\mathbb{C}$ tel que U contient $\ell \in \mathbb{N}$ racines de P comptées avec multiplicité et qu'aucune racine de P n'est dans la frontière de U qu'on note $\partial \mathrm{U} = \overline{\mathrm{U}} \backslash \mathrm{U}$. Montrer qu'il existe $\mathrm{K} \in \mathbb{N}$ tel que pour tout $k \geq \mathrm{K}$, U contient exactement ℓ racines de P_k comptées avec multiplicité.

III Polynômes réels

Proposition III.1.

Les polynômes irréductibles (unitaires) de $\mathbb{R}[\mathrm{X}]$ sont ceux de la forme

→ $\mathrm{X} - a$ avec $a \in \mathbb{R}$.

→ $\mathrm{X}^2 - a\mathrm{X} + b$ avec $a, b, c \in \mathbb{R}$ et $\Delta = a^2 - 4b < 0$.

Démonstration. : Soit $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$ unitaire non constant. Si $\deg \mathrm{P} \geq 3$, alors deux cas se présentent.

→ P admet une racine $z \in \mathbb{R}$, et alors P est réductible car il est divisible par $\mathrm{X} - z$.

→ P admet une racine $z \in \mathbb{C} \setminus \mathbb{R}$ et P est à coefficients réels, donc $\overline{z}$ est aussi une racine de P. P est alors divisible par $(\mathrm{X} - z)(\mathrm{X} - \overline{z}) = \mathrm{X}^2 - 2\mathrm{Re}(z) + |z|^2 \in \mathbb{R}[\mathrm{X}]$ et est donc réductible.

Si $\deg \mathrm{P} = 1$, alors d'une manière évidente P s'écrit $\mathrm{X} - a$ avec $a \in \mathbb{R}$ et est irréductible et si $\deg \mathrm{P} = 2$, alors P est irréductible s'il n'admet pas de racine réelle, *i.e.* son discriminant est négatif.

Remarque. Les polynômes irréductibles de $\mathbb{R}[\mathrm{X}]$ de degré 2 s'écrivent aussi de la forme $(\mathrm{X} - a)^2 + b^2$ avec $a \in \mathbb{R}$ et $b > 0$.

Exercice III.2.

Montrer que tout polynôme $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$ positif sur $\mathbb{R}$ s'écrit de la forme $\mathrm{P} = \mathrm{A}^2 + \mathrm{B}^2$ avec $\mathrm{A}, \mathrm{B} \in \mathbb{R}[\mathrm{X}]$.

BY-NC-SA

Proposition III.3.

Soit $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$ non constant unitaire. On pose $\mathrm{P}(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \cdots + a_0$.

1. Si P est scindé alors P' et $\mathrm{P} + \alpha\mathrm{P}', \alpha \in \mathbb{R}$ aussi. Si P est dissocié (*i.e.* scindé à racines simples), P' aussi.
2. Si P est dissocié alors $\forall k \in [\![0, n-1]\!], a_k^2 + a_{k+1}^2 > 0$.
3. Si P est scindé, alors $\forall k \in [\![1, n-1]\!], a_k a_{k+2} \leq a_{k+1}^2$.
4. Si une suite $(\mathrm{P}_k) \in \mathbb{R}_n[\mathrm{X}]^{\mathbb{N}}$ de polynômes non constants unitaires scindés (pas forcément tous de même degré) converge simplement vers un polynôme $\mathrm{Q} \in \mathbb{R}[\mathrm{X}]$, alors Q est non constant et est scindé.

Démonstration. :

1. Montrons la propriété successivement pour P' et $\mathrm{P} + \alpha\mathrm{P}'$.

   → Si $a_1 < \cdots < a_r$ sont les racines de P de multiplicités respectives $\alpha_1, \ldots \alpha_r$, alors pour tout i, a_i est racine de P' de multiplicité $\alpha_i - 1$. On trouve ainsi $\alpha_1 + \cdots + \alpha_r - r = n - r$ racines de P' avec multiplicité. De plus, en appliquant le théorème de Rolle sur les segments $[a_i, a_{i+1}]$ on peut affirmer que P' s'annule en $b_1, \ldots, b_{r-1}$ tels que $a_1 < b_1 < a_2 < \cdots < a_{r-1} < b_{r-1} < a_r$, soit $r-1$ racines (simples) supplémentaires. On a donc un total de $n-r+r-1 = n-1$ racines de P' comptées avec multiplicité et P' est de degré n-1, donc P' est scindé. Cette démonstration montre aussi que P' est dissocié si P l'est.

   → Soit $\alpha \in \mathbb{R}$. Si $\alpha = 0$ alors il s'agit du cas précédent. Sinon considérons $f : t \longmapsto e^{\beta t}\mathrm{P}(t)$ où $\beta = \dfrac{1}{\alpha}$. On a alors pour tout $t \in \mathbb{R}$,

   $$f'(t) = 0 \Longleftrightarrow e^{\beta t}(\beta\mathrm{P}(t) + \mathrm{P}'(t)) = 0 \Longleftrightarrow \mathrm{P}(t) + \alpha\mathrm{P}'(t) = 0.$$

   Le reste est similaire à la discussion précédente : si λ est racine de P de multiplicité $\delta \geq 1$ alors elle l'est dans $\mathrm{P} + \alpha\mathrm{P}'$ de multiplicité exactement $\delta - 1$. Pour retrouver les autres racines manquantes, il suffit d'appliquer le théorème de Rolle à f entre chaque deux racines consécutives.

   Remarquons au passage que cette propriété n'est pas vraie pour tout corps $\mathbb{K}$. En effet, lorsque $\mathbb{K} = \mathbb{Q}[\sqrt{6}] = \{\mathrm{R}(\sqrt{6}),\ \mathrm{R} \in \mathbb{Q}[\mathrm{X}]\}$ et $\mathrm{P}(\mathrm{X}) = \mathrm{X}^3 - 6\mathrm{X}$, on voit que P est scindé dans $\mathbb{K}$ mais pas P'.

2. Supposons que P et dissocié. Procédons par l'absurde. Supposons qu'il existe $k \in [\![0; n-1]\!]$ tel que $a_k = a_{k+1} = 0$. On a alors

   $$\mathrm{P}^{(k)}(\mathrm{X}) = k!a_k + (k+1)!a_{k+1}\mathrm{X} + \frac{(k+2)!}{2}a_{k+2}\mathrm{X}^2 + \cdots + \frac{n!}{(n-k)!}a_n\mathrm{X}^{n-k}.$$

   0 est alors racine de multiplicité 2. De plus, P est dissocié et donc d'après la propriété précédente, $\mathrm{P}^{(k)}$ aussi, ce qui est absurde. On en déduit donc que pour tout $k \in [\![0; n-1]\!]$, $a_k^2 + a_{k+1}^2 > 0$.

3. On sait que

   $$\frac{\mathrm{P}'}{\mathrm{P}} = \sum_{i=1}^{n} \frac{1}{\mathrm{X} - \lambda_i},$$

   où les $\lambda_1, \ldots, \lambda_n$ sont les racines de P. Ceci nous permet de dire que

   $$\left(\frac{\mathrm{P}'}{\mathrm{P}}\right)' = -\sum_{i=1}^{m} \frac{1}{(\mathrm{X} - \lambda_i)^2} \text{ et alors } \frac{\mathrm{P}''\mathrm{P} - \mathrm{P}'^2}{\mathrm{P}^2} \leq 0.$$

(cc) BY-NC-SA

sur $\mathbb{R} \setminus \mathrm{Z(P)}$. On a alors en particulier,

$$\mathrm{P}''(0)\mathrm{P}(0) - \mathrm{P}'(0)^2 \leq 0.$$

Notons que dans le cas où 0 est racine de P, il suffit de multiplier par P^2 des deux côtés de l'inégalité et passer à la limite en 0 pour obtenir le résultat ci-dessus. On remplace maintenant P par $\mathrm{P}^{(k)}$, qui est également scindé, et on trouve

$$\mathrm{P}^{(k+2)}(0)\mathrm{P}^{(k)}(0) \leq \mathrm{P}^{(k+1)}(0)^2,$$

soit

$$(k+2)!a_{k+2}k!a_k \leq (k+1)!^2 a_{k+1}^2 \quad \text{puis} \quad a_{k+2}a_k \leq \frac{k+1}{k+2}a_{k+1}^2 \leq a_{k+1}^2,$$

d'où le résultat.

4. Soit $(\mathrm{P}_k) \in \mathbb{R}_n[\mathrm{X}]^{\mathbb{N}}$ une suite de polynômes non constants unitaires scindés convergeant simplement vers $\mathrm{Q} \in \mathbb{R}[\mathrm{X}]$. Posons pour tout $k \in \mathbb{N}$ et $z \in \mathbb{C}$,

$$\mathrm{P}_k(z) = \prod_{j=1}^{\deg \mathrm{P}_k} (z - a_{k,j}).$$

On a alors pour tout $z \in \mathbb{C} \setminus \mathbb{R}$,

$$|\mathrm{P}_k(z)| = \prod_{j=1}^{\deg \mathrm{P}_k} |z - a_{k,j}| \geq |\mathrm{Im}\ z|^{\deg \mathrm{P}_k} \geq \min(|\mathrm{Im}\ z|, |\mathrm{Im}\ z|^n).$$

En faisant tendre k vers l'infini, on obtient

$$|\mathrm{Q}(z)| \geq \min(|\mathrm{Im}\ z|^n, |\mathrm{Im}\ z|) > 0.$$

Q est donc soit constant, soit scindé dans $\mathbb{R}$ (il n'a pas de racine complexe non réelle) car pour tout $z \in \mathbb{C} \setminus \mathbb{R}$, $\mathrm{Q}(z) \neq 0$. Le premier cas est impossible car si $\mathrm{Q(X)} = \mathrm{C}$ est constant alors en prenant $z = i(1 + |\mathrm{C}|)$, on obtient une contradiction avec l'inégalité.

Exercice III.4.

Trouver les polynômes $\mathrm{P} \in \mathbb{R}[\mathrm{X}]$ unitaires à coefficients dans $\{-1, 0, 1\}$ qui sont scindés sur $\mathbb{R}$.

Lemme (Formules de Vieta) III.5.

Soit $n \in \mathbb{N}^*$, $\mathrm{P(X)} = a_0 + a_1\mathrm{X} + \cdots + a_n\mathrm{X}^n$ et $\lambda_1, \ldots, \lambda_n$ les racines (non nécessairement différentes) de P. Supposons que $a_n \neq 0$. On a pour tout $\ell \in [\![0; n]\!]$

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_\ell \leq n} \left(\prod_{j=1}^{\ell} \lambda_{i_j} \right) = (-1)^\ell \frac{a_{n-\ell}}{a_n}.$$

Démonstration. Nous ne ferons par la démonstration détaillée de ce résultat mais nous en donnerons uniquement l'idée principale. Pour voir d'où vient cette formule, il suffit de développer $a_n(\mathrm{X} - \lambda_1) \ldots (\mathrm{X} - \lambda_n)$ et d'identifier les coefficients de ce polynôme avec ceux de P. En particulier, il est facile de voir cette propriété pour $\ell = n$ et $\ell = 1$ qui s'écrivent

$$(-1)^n \frac{a_0}{a_n} = \lambda_1 \ldots \lambda_n \quad \text{et} \quad -\frac{a_{n-1}}{a_n} = \lambda_1 + \cdots + \lambda_n.$$

(cc) BY-NC-SA

IV Polynômes à coefficients rationnels

Exercice IV.1.

Soit $n \geq 1$, $a_0, \dots, a_n \in \mathbb{Z}$ tels que $a_0 \neq 0$ et $a_n \neq 0$. On pose $\mathrm{P}(\mathrm{X}) = a_n \mathrm{X}^n + \cdots + a_1 \mathrm{X} + a_0$. Soit $a \in \mathbb{Q}$. Trouver une condition nécessaire sur a pour qu'il soit racine de P.

Exercice IV.2.

Soit $\mathrm{P} \in \mathbb{Q}[\mathrm{X}]$ irréductible et a une racine complexe de P. Montrer que a est une racine simple de P.

Exercice IV.3.

Soit $\mathrm{P} \in \mathbb{Q}[\mathrm{X}]$ de degré égal à 5. Montrer que si P admet une racine double dans $\mathbb{C}$, alors il possède une racine dans $\mathbb{Q}$.

V Irréductibilité de $\mathbb{Z}[\mathrm{X}]$

Exercice V.1.

Soient $n \geq 1$ et $a_1, \dots a_n \in \mathbb{Z}$ deux à deux distincts. Soit $\mathrm{P} = 1 + \prod_{i=1}^{n} (\mathrm{X} - a_i)^2$. Montrer que P est irréductible dans $\mathbb{Z}[\mathrm{X}]$.

Exercice (Critère d'Eisenstein) V.2.

Soit $\mathrm{P}(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \cdots + a_0 \in \mathbb{Z}[\mathrm{X}]$. On suppose qu'il existe un nombre premier p tel que $p^2 \nmid a_0$ et pour tout $i \in [\![0; n-1]\!]$, $p | a_i$. Montrer que P est irréductible dans $\mathbb{Z}[\mathrm{X}]$.

Exercice V.3.

Montrer que $\mathrm{X}^4 + 1$ est irréductible dans $\mathbb{Q}[\mathrm{X}]$ et que pour tout p premier, $\mathrm{X}^4 + 1$ est réductible dans $\mathbb{Z}/p\mathbb{Z}$.

VI Complément : lemme de Gauß

Dans cette partie, on dit que $\mathrm{P} \in \mathbb{Z}[\mathrm{X}] \setminus \{0\}$ est irréductible dans $\mathbb{Z}[\mathrm{X}]$ si on ne peut pas l'écrire sous forme de QR avec $\mathrm{Q}, \mathrm{R} \in \mathbb{Z}[\mathrm{X}]$ et $\deg \mathrm{Q} \geq 1$ et $\deg \mathrm{R} \geq 1$. On mentionne ceci car pour certains auteurs, 2X par exemple est réductible car 2 n'est pas inversible dans $\mathbb{Z}$. En d'autres termes, dans notre définition, les termes constants ne comptent pas comme un facteur, même s'ils sont non inversibles dans $\mathbb{Z}$.

BY-NC-SA

Exercice VI.1.

Le but de cet exercice est de montrer qu'un polynôme $\mathrm{P} \in \mathbb{Z}[\mathrm{X}]$ non constant est irréductible dans $\mathbb{Z}[\mathrm{X}]$ si et seulement s'il l'est sur $\mathbb{Q}[\mathrm{X}]$. Le sens réciproque étant clair, on se propose alors de montrer le sens direct.

1. Posons pour tout polynôme $\mathrm{P} = \sum_{k=0}^{n} a_k \mathrm{X}^k$ non nul $c(\mathrm{P}) = \bigwedge_{i=0}^{n} a_i$ son contenu. Montrer que

$$\forall \mathrm{P}, \mathrm{Q} \in \mathbb{Z}[\mathrm{X}] - \{0\},\ c(\mathrm{PQ}) = c(\mathrm{P})c(\mathrm{Q}).$$

2. En déduire le résultat.

(cc) BY-NC-SA

Correction de l'exercice II.5 :

Notons pour tout polynôme Q, F(Q) le nombre de racines de Q dans U. Supposons par l'absurde que la propriété qu'on veut démontrer est fausse, *i.e.*

$$\forall \mathrm{K} \geq 0,\ \exists k \geq \mathrm{K},\ \mathrm{F}(\mathrm{P}_k) \neq \ell.$$

Ceci est équivalent à dire que l'ensemble $\mathrm{A} = \{k \in \mathbb{N},\ \mathrm{F}(\mathrm{P}_k) \neq \ell\}$ est infini. A étant infini, il existe une extractrice ψ telle que pour tout $k \in \mathbb{N}$, $\psi(k) \in \mathrm{A}$. Pour alléger les notations, on va noter (P_k) au lieu de $(\mathrm{P}_{\psi(k)})$ car $(\mathrm{P}_{\psi(k)})$ converge aussi vers P.

On sait d'après la proposition II.4 que les coefficients des termes de (P_k) convergent vers ceux de P. En particulier, P étant unitaire, le coefficient devant X^n des termes de (P_k) converge vers 1 et est donc non nul à partir d'un certain rang et alors il existe $\mathrm{A} > 0$ tel que pour tout $k \geq \mathrm{A}$, $\deg \mathrm{P}_k = n$. Posons alors pour tout $k \geq \mathrm{A}$, $\lambda_{k,1}, \ldots, \lambda_{k,n}$ les racines de P_k. Les coefficients de P_k convergent et sont donc tous bornés. En utilisant la proposition II.3, on peut en déduire que les suites $(\lambda_{k,1})_{k\geq \mathrm{A}}, \ldots, (\lambda_{k,n})_{k\geq \mathrm{A}}$ sont toutes bornées. Par conséquent, d'après Bolzano-Weierstraß, il existe une extractrice φ (la même pour les n suites, quitte à faire des extractions successives) et $\lambda_1, \ldots, \lambda_n$ tels que

$$\forall i \in [\![1;n]\!],\ \lambda_{\varphi(k),i} \xrightarrow[k\to+\infty]{} \lambda_i.$$

En utilisant ce qu'on vient de voir, il est facile de voir (soit en utilisant les relations coefficient racines, ou alors en passant par la convergence simple) que

$$\mathrm{P}_{\varphi(k)}(\mathrm{X}) = \prod_{i=1}^{n}(\mathrm{X} - \lambda_{\varphi(k),i}) \xrightarrow[k\to+\infty]{} \prod_{i=1}^{n}(\mathrm{X} - \lambda_i),$$

et donc que $\mathrm{P}(\mathrm{X}) = \prod_{i=1}^{n}(\mathrm{X} - \lambda_i)$. On a alors pour tout $i \in [\![1;n]\!]$,

- → Soit $\lambda_i \in \mathrm{U}$, et alors U étant ouvert, il existe $\mathrm{K} > 0$ tel que pour tout $k > \mathrm{K}$, $\lambda_{i,\varphi(k)} \in \mathrm{U}$.
- → Soit $\lambda_i \notin \mathrm{U}$ et alors par hypothèse (pas de racines dans la frontière) $\lambda_i \in \mathbb{C} \setminus \overline{\mathrm{U}}$. Cet ensemble est ouvert, il existe donc $\mathrm{K} > 0$ tel que pour tout $k \geq \mathrm{K}$, $\lambda_{i,\varphi(k)} \in \mathbb{C} \setminus \overline{\mathrm{U}}$ et alors $\lambda_{i,\varphi(k)} \notin \mathrm{U}$.

Ceci permet de dire que pour k assez grand, $\lambda_{i,\varphi(k)}$ et λ_i sont soit tous les deux dans U soit tout les deux à l'extérieur de U et donc en particulier que $\mathrm{F}(\mathrm{P}_{\varphi(k)}) = \mathrm{F}(\mathrm{P}) = \ell$, ce qui est absurde par hypothèse.

Correction de l'exercice III.2. :

Posons

$$\mathcal{S} = \{\mathrm{P} \in \mathbb{R}[\mathrm{X}],\ \exists \mathrm{A}, \mathrm{B} \in \mathbb{R}[\mathrm{X}],\ \mathrm{P} = \mathrm{A}^2 + \mathrm{B}^2\}.$$

Le but de l'exercice est de montrer que l'ensemble des polynômes positifs sur $\mathbb{R}$ est inclus dans $\mathcal{S}$ (au passage, ces deux ensembles sont égaux car l'inclusion réciproque est évidente).
Montrons d'abord que $\mathcal{S}$ est stable par multiplication. On a pour tout $\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D} \in \mathbb{R}[\mathrm{X}]$,

$$(\mathrm{A}^2 + \mathrm{B}^2)(\mathrm{C}^2 + \mathrm{D}^2) = (\mathrm{AC} + \mathrm{BD})^2 + (\mathrm{AD} - \mathrm{BC})^2,$$

ce qui nous donne bien la stabilité par multiplication.

Soit P un polynôme positif sur $\mathbb{R}$. Si P est constant alors le résultat est évident. Supposons désormais $\deg \mathrm{P} \geq 1$. On peut alors écrire

$$\mathrm{P}(\mathrm{X}) = \lambda \prod_{i=1}^{r}(\mathrm{X} - a_i)^{\alpha_i} \prod_{i=1}^{s} \underbrace{(\mathrm{X}^2 + b_i\mathrm{X} + c_i)}_{\text{irréductible}},$$

avec pour tout i, $a_i, b_i, c_i \in \mathbb{R}$, $b_i^2 - 4c_i < 0$ et les a_i distincts. On a alors

(cc) BY-NC-SA

→ $\lambda > 0$ car $\mathrm{P}(x) \xrightarrow[x\to+\infty]{} +\infty$, donc $\lambda \in \mathcal{S}$.

→ Pour tout $i \in [\![1;s]\!]$, $\mathrm{X}^2 + b_i\mathrm{X} + c_i = \left(\mathrm{X} + \dfrac{b_i}{2}\right)^2 + \sqrt{\dfrac{-b_i^2+4c_i}{4}}^{\,2} \in \mathcal{S}$.

→ Pour tout $i \in [\![1;r]\!]$, on peut écrire $\mathrm{P}(\mathrm{X}) = \mathrm{Q}(\mathrm{X})(\mathrm{X}-a_i)^{\alpha_i}$ avec $\mathrm{Q} \in \mathbb{R}[\mathrm{X}]$. a_i n'est pas racine de Q, donc $\mathrm{Q}(a_i) \neq 0$ et alors $\mathrm{Q}(a_i) > 0$. Il existe donc $\varepsilon > 0$ tel que pour tout $x \in]a_i - \varepsilon, a_i]$, $\mathrm{Q}(x) > 0$ et alors pour tout $x \in]a_i - \varepsilon, a_i]$, $(x-a_i)^{\alpha_i} \geq 0$, ce qui nous donne nécessairement que α_i est pair. On en déduit que $(\mathrm{X}-a_i)^{\alpha_i} = \left((\mathrm{X}-a_i)^{\frac{\alpha_i}{2}}\right)^2 \in \mathcal{S}$.

P est donc produit d'éléments de $\mathcal{S}$ et donc par stabilité par multiplication, il est bien dans $\mathcal{S}$.

Correction de l'exercice III.4. :

On suppose $\mathrm{P}(\mathrm{X}) = a_n\mathrm{X}^n + \cdots + a_1\mathrm{X} + a_0$ scindé avec $a_n = 1$. On se ramène à $\mathrm{P}(0) \neq 0$ en factorisant éventuellement par X^k pour k la valuation de P (0 sera alors racine réelle et le caractère scindé du polynôme en sera inchangé).

Soient $x_1, \ldots, x_n$ les racines de P. On a

$$\prod_{k=1}^{n} x_n = (-1)^n\mathrm{P}(0) \quad \text{donc} \quad \prod_{k=1}^{n} x_k^2 = \mathrm{P}(0)^2 = 1.$$

Pour obtenir la première égalité, il suffit de remplacer X par 0 dans la décomposition de P.

L'inégalité arithmético-géométrique donne alors

$$\sum_{k=1}^{n} x_k^2 \geq n\sqrt[n]{x_1^2 \ldots x_n^2} = n, \tag{15}$$

et en utilisant les formules de Vieta,

$$n \leq \sum_{k=1}^{n} x_k^2 = \left(\sum_{k=1}^{n} x_k\right)^2 - 2\sum_{k<\ell} x_k x_\ell = (-a_{n-1})^2 - 2a_{n-2} \leq a_{n-1}^2 + 2\,|a_{n-1}| \leq 3. \tag{16}$$

et donc $n \leq 3$. Le cas $n = 3$ impose le cas d'égalité dans les inégalités ci-dessus

→ Égalité dans (15), *i.e.* le cas d'égalité de l'inégalité arithmético-géométrique qui donne nécessairement $x_1^2 = \cdots = x_n^2$. En utilisant l'égalité $\displaystyle\prod_{k=1}^{n} x_k^2 = 1$, on obtient que $x_1^2 = \cdots = x_n^2 = 1$ *i.e.* $\mathrm{Z}(\mathrm{P}) \subset \{-1, 1\}$.

→ Égalité dans l'inégalité de droite de (16), *i.e.* $a_{n-1}^2 - 2a_{n-2} = 3 \iff a_{n-1} = \pm 1$ et $a_{n-2} = -1$.

En particulier, la condition $n = 3$ impose que

$$\mathrm{P} \in \{\underbrace{(\mathrm{X}-1)^3}_{\mathrm{P}_1},\ \underbrace{(\mathrm{X}-1)^2(\mathrm{X}+1)}_{\mathrm{P}_2},\ \underbrace{(\mathrm{X}-1)(\mathrm{X}+1)^2}_{\mathrm{P}_3},\ \underbrace{(\mathrm{X}+1)^3}_{\mathrm{P}_4}\}.$$

Il est clair que les coefficients de P_1 et P_4 ne sont pas tous dans $\{-1, 0, 1\}$ et donc ne vérifient pas les conditions voulues. Cependant P_2 et P_3 conviennent.
Pour le cas $n = 2$, on peut écrire $\mathrm{P} = \mathrm{X}^2 + a\mathrm{X} + b$ où $b = \pm 1$ et $a \in \{-1, 0, 1\}$. La condition est alors $\Delta = a^2 - 4b \geq 0$ et donc les seuls polynômes convenables sont $\mathrm{X}^2 - \mathrm{X} - 1, \mathrm{X}^2 - 1$ et $\mathrm{X}^2 + \mathrm{X} - 1$.
Pour $n = 1$ la propriété voulue est toujours vérifiée. En posant donc

$$\mathrm{A} = \{\mathrm{X}-1,\ \mathrm{X}+1,\ \mathrm{X}^2-\mathrm{X}-1,\ \mathrm{X}^2-1,\ \mathrm{X}^2+\mathrm{X}-1,\ (\mathrm{X}-1)^2(\mathrm{X}+1),\ (\mathrm{X}+1)^2(\mathrm{X}-1)\},$$

(cc) BY-NC-SA

on déduit que l'ensemble des polynômes scindés à coefficients dans $\{-1, 0, 1\}$ est égal à

$$\{\mathrm{X}^k\mathrm{Q},\ k \in \mathbb{N},\ \mathrm{Q} \in \mathrm{A}\}.$$

Correction de l'exercice IV.1. :

Soit a une racine de P. On a $a_0 \neq 0$ donc $a \neq 0$. Posons $a = \dfrac{p}{q}$ avec $p \wedge q = 1$. On a alors $\mathrm{P}\left(\frac{p}{q}\right) = 0$, *i.e.*

$$a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_1 pq^{n-1} + a_0 q^n = 0.$$

On en déduit que $p|a_0q^n$ et $q|a_np^n$ et donc par Gauß, $p|a_0$ et $q|a_n$.

Correction de l'exercice IV.2. :

Les seuls diviseurs unitaires de P sont P et 1 car P est irréductible, donc $\mathrm{P} \wedge \mathrm{P}' \in \{1, \mathrm{P}\}$. Mais $\deg \mathrm{P}' < \deg \mathrm{P}$ donc P ne peut pas diviser P' et alors $\mathrm{P} \wedge \mathrm{P}' = 1$. On a alors d'après Bezout,

$$\exists \mathrm{U}, \mathrm{V} \in \mathbb{Q}[\mathrm{X}],\ \mathrm{UP} + \mathrm{VP}' = 1.$$

En évaluant en a, on obtient $\underbrace{\mathrm{U}(a)\mathrm{P}(a)}_{=0} + \mathrm{V}(a)\mathrm{P}'(a) = 1$ et alors nécessairement $\mathrm{P}'(a) \neq 0$, *i.e.* a est une racine simple de P.
Remarquons aussi que dire $\mathrm{P} \wedge \mathrm{P}' = 1$ est équivalent à dire que P et P' n'ont aucune racine complexe commune et que donc, en particulier, a n'est pas racine de P', ce qui fournit le résultat.

Correction de l'exercice IV.3. :

Supposons sans perte de généralité que P est unitaire. P admet une racine double dans $\mathbb{C}$, donc d'après l'exercice précédent, P est réductible. On écrit $\mathrm{P} = \mathrm{QR}$ avec Q, R non constants, unitaires et irréductibles (s'il y a plus de deux facteurs irréductibles non constants, au moins un sera de degré 1 et donc P admet une racine rationnelle). Le cas $\mathrm{Q} = \mathrm{R}$ étant impossible vu que $\deg(\mathrm{P}) = 5$ est impair, on a alors que Q et R sont premiers entre eux. En effet, Q et R étant irréductibles unitaires, on a $\mathrm{Q} \wedge \mathrm{R} \in \{1, \mathrm{Q}\} \cap \{1, \mathrm{R}\} = \{1\}$.

→ Si $\deg \mathrm{Q} = 1$, alors $\mathrm{Q} = \mathrm{X} - r$, avec $r \in \mathbb{Q}$ et donc r est racine rationnelle de P.

→ Si $\deg \mathrm{Q} = 2$, alors a est racine double de P mais d'après l'exercice précédent n'est racine double ni de Q ni de R et donc $\mathrm{Q}(a) = \mathrm{R}(a) = 0$ ce qui est absurde car $\mathrm{Q} \wedge \mathrm{R} = 1$.

On a donc bien le résultat voulu.

Correction de l'exercice V.1. :

Supposons par l'absurde que P est réductible. Il existe donc $\mathrm{Q}, \mathrm{R} \in \mathbb{Z}[\mathrm{X}]$ tels que $\deg \mathrm{Q} \geq 1$, $\deg \mathrm{R} \geq 1$ et $\mathrm{P} = \mathrm{QR}$. On a alors

$$\forall i \in [\![1; n]\!],\ \underbrace{\mathrm{Q}(a_i)}_{\in \mathbb{Z}}\underbrace{\mathrm{R}(a_i)}_{\in \mathbb{Z}} = \mathrm{P}(a_i) = 1,$$

et donc

$$\forall i \in [\![1; n]\!],\ \mathrm{Q}(a_i) = \mathrm{R}(a_i) = \pm 1.$$

De plus, P est positif et ne s'annule pas sur R, donc Q et R aussi et sont alors de signe constant et ont le même signe. On suppose que $\mathrm{Q} > 0$ et $\mathrm{R} > 0$ (le cas $\mathrm{Q} < 0$ et $\mathrm{R} < 0$ se traite de la même manière), chose qui implique en particulier que Q et R sont unitaires (P est unitaire). On a alors

$$\forall i \in [\![1; n]\!],\ \mathrm{Q}(a_i) - 1 = 0 \text{ et } \mathrm{R}(a_i) - 1 = 0.$$

$\mathrm{Q} - 1$ et $\mathrm{R} - 1$ admettent n racines distinctes, sont donc de degré au moins n et la somme de leurs degrés

(cc) BY-NC-SA

est $2n$. On en déduit que $\deg \mathrm{Q} = \deg \mathrm{R} = n$ et

$$\mathrm{Q}(\mathrm{X}) - 1 = \mathrm{R}(\mathrm{X}) - 1 = \prod_{i=1}^{n} (\mathrm{X} - a_i),$$

et que finalement

$$\prod_{i=1}^{n} (\mathrm{X} - a_i)^2 + 1 = \mathrm{P}(\mathrm{X}) = \left(\prod_{i=1}^{n} (\mathrm{X} - a_i) + 1 \right)^2,$$

ce qui est absurde. P est donc bien irréductible.

Correction de l'exercice V.2. :

Supposons par l'absurde que P est réductible dans $\mathbb{Z}[\mathrm{X}]$, *i.e.* il existe $\mathrm{Q}, \mathrm{R} \in \mathbb{Z}[\mathrm{X}]$ de degré supérieur ou égal à 1 tels que $\mathrm{P} = \mathrm{QR}$, qu'on peut supposer unitaires (P est unitaire). Il existe donc $r, s \in \mathbb{N}^*$, $(b_i)_{i \in [\![0;r-1]\!]} \in \mathbb{Z}^r$ et $(c_i)_{i \in [\![0;s-1]\!]} \in \mathbb{Z}^r$ tels que

$$\mathrm{Q}(x) = \mathrm{X}^r + b_{r-1}\mathrm{X}^{r-1} + \cdots + b_1\mathrm{X} + b_0 \quad \text{et} \quad \mathrm{R}(\mathrm{X}) = \mathrm{X}^s + c_{s-1}\mathrm{X}^{s-1} \cdots + c_1\mathrm{X} + c_0.$$

On considère les équivalents de ces polynômes dans $\mathbb{Z}/p\mathbb{Z}[\mathrm{X}]$: comme $\forall i \in [\![0; n-1]\!], a_i \equiv 0[p]$, on peut écrire dans $\mathbb{Z}/p\mathbb{Z}[\mathrm{X}]$

$$\mathrm{P}(\mathrm{X}) = \mathrm{X}^n + \underbrace{\overline{a_{n-1}}\mathrm{X}^{n-1} + \cdots + \overline{a_1}\mathrm{X} + \overline{a_0}}_{=\overline{0}} = \mathrm{X}^n.$$

De plus, on a également dans $\mathbb{Z}/p\mathbb{Z}[\mathrm{X}]$

$$\mathrm{X}^n = \mathrm{P}(\mathrm{X}) = \mathrm{Q}(\mathrm{X})\mathrm{R}(\mathrm{X}) = (\mathrm{X}^r + \overline{b_{r-1}}\mathrm{X}^{r-1} + \cdots + \overline{b_1}\mathrm{X} + \overline{b_0})(\mathrm{X}^s + \overline{c_{s-1}}\mathrm{X}^{s-1} \cdots + \overline{c_1}\mathrm{X} + \overline{c_0}).$$

Or par unicité de la décomposition de X^n en facteurs irréductibles dans $\mathbb{Z}/p\mathbb{Z}[\mathrm{X}]$, Q(X) et R(X) doivent être de la forme X^{s_1} et X^{s_2}. Pour qu'on ait la bonne puissance, il est nécessaire que $s_1 = r$ et $s_2 = s$, *i.e.*, dans $\mathbb{Z}/p\mathbb{Z}[\mathrm{X}]$, $\mathrm{Q}(\mathrm{X}) = \mathrm{X}^r$ et $\mathrm{R}(\mathrm{X}) = \mathrm{X}^s$, et alors

$$\overline{b_{r-1}} = \cdots = \overline{b_1} = \overline{b_0} = \overline{c_{s-1}} = \cdots = \overline{c_1} = \overline{c_0} = \overline{0}.$$

Or $a_0 = c_0 b_0$ et d'après l'égalité ci-dessus, $p | b_0$ et $p | c_0$ et donc $p^2 | a_0$, ce qui est absurde.

Correction de l'exercice V.3. :

→ Irréductibilité dans $\mathbb{R}[\mathrm{X}]$.
On décompose $\mathrm{X}^4 + 1$

$$\mathrm{X}^4 + 1 = (\mathrm{X}^2 + 1)^2 - 2\mathrm{X}^2 = \underbrace{(\mathrm{X}^2 + \sqrt{2}\mathrm{X} + 1)}_{\mathrm{A}(\mathrm{X})}\underbrace{(\mathrm{X}^2 - \sqrt{2}\mathrm{X} + 1)}_{\mathrm{B}(\mathrm{X})}.$$

Si $\mathrm{X}^4 + 1$ était réductible dans $\mathbb{Q}[\mathrm{X}]$, on pourrait écrire $\mathrm{X}^4 + 1 = \mathrm{QR}$ tel que $\mathrm{Q}, \mathrm{R} \in \mathbb{Q}[\mathrm{X}]$, $\deg \mathrm{R} \geq \deg \mathrm{Q} \geq 1$ et Q et R sont unitaires. Le cas $\deg \mathrm{Q} = 1$ est impossible car $\mathrm{X}^4 + 1$ n'admet pas de racine rationnelle. Le cas $\deg \mathrm{Q} = 2$ est impossible car par unicité de la décomposition en facteurs irréductibles dans $\mathbb{R}[\mathrm{X}]$, on aurait $\mathrm{Q}, \mathrm{R} \in \{\mathrm{A}, \mathrm{B}\}$ ce qui est absurde car $\mathrm{A}, \mathrm{B} \notin \mathbb{Q}[\mathrm{X}]$. On en déduit donc que $\mathrm{X}^4 + 1$ est irréductible dans $\mathbb{Q}[\mathrm{X}]$.

→ Irréductibilité dans $\mathbb{Z}/p\mathbb{Z}$.
Si $p = 2$, alors on peut écrire dans $\mathbb{Z}/p\mathbb{Z}$

$$\mathrm{X}^4 + \overline{1} = \mathrm{X}^4 + \overline{2}\mathrm{X}^2 + \overline{1} = (\mathrm{X}^2 + \overline{1})^2,$$

(cc) BY-NC-SA

donc X^4+1 est réductible. Supposons maintenant $p \geq 3$.

- Si $\overline{2}$ est un carré modulo p, *i.e.* il existe $a \in \mathbb{Z}/p\mathbb{Z}$ tel que $a^2 = \overline{2}$, on peut écrire
$$X^4+1 = (X^2+\overline{1})^2 - \overline{2}X^2 = (X^2+\overline{1})^2 - a^2X^2 = (X^2+\overline{1}+a)(X^2+\overline{1}-a).$$
- Si $\overline{-2}$ est un carré modulo p, *i.e.* il existe $b \in \mathbb{Z}/p\mathbb{Z}$ tel que $b^2 = \overline{-2}$, on peut écrire
$$X^4+1 = (X^2-\overline{1})^2 - \overline{-2}X^2 = (X^2-\overline{1})^2 - b^2X^2 = (X^2+\overline{1}+b)(X^2+\overline{1}-b).$$
- Supposons maintenant que ni $\overline{2}$ ni $\overline{-2}$ ne sont des carrés modulo p. On a alors d'après le chapitre 21 sur anneaux, $\overline{2}^{\frac{p-1}{2}} = \overline{-1}$ et $\overline{-2}^{\frac{p-1}{2}} = \overline{-1}$ et donc en faisant le produit des deux, $\overline{-4}^{\frac{p-1}{2}} = \overline{1}$. On peut donc écrire $\overline{1} = \overline{-4}^{\frac{p-1}{2}} = \overline{-1}^{\frac{p-1}{2}} \times \overline{4}^{\frac{p-1}{2}} = \overline{-1}^{\frac{p-1}{2}}$ ce qui nous permet de dire que -1 est un carré modulo p. En posant $c^2 = \overline{-1}$ avec $c \in \mathbb{Z}/p\mathbb{Z}$, on peut finalement écrire
$$X^4+\overline{1} = X^4 - \overline{-1} = X^4 - c^2 = (X^2-c)(X^2+c).$$

On en déduit donc que X^4+1 est bien réductible dans $\mathbb{Z}/p\mathbb{Z}[X]$.

Correction de l'exercice VI.1. :

1. Il est facile de voir que pour tout $P(X) = \sum_{i=0}^{n} a_k X^k \in \mathbb{Z}[X] \setminus \{0\}$ et $a \in \mathbb{N}^*$ tel que $a | c(P)$,
$$\frac{1}{a}c(P) = \bigwedge_{i=0}^{n} \frac{a_i}{a} = c\left(\frac{P}{a}\right),$$
donc montrer le résultat voulu est équivalent à montrer que
$$c\left(\frac{P}{c(P)} \times \frac{Q}{c(Q)}\right) = 1,$$
et donc quitte à remplacer $\frac{P}{c(P)}$ par P et $\frac{Q}{c(Q)}$ par Q, il suffit de montrer que
$$\forall P, Q \in \mathbb{Z}[X] \setminus \{0\},\ c(P) = c(Q) = 1 \Longrightarrow c(PQ) = 1.$$
Soient $P, Q \in \mathbb{Z}[X] \setminus \{0\}$ et $p \in \mathbb{Z}$ premier. Si $p | c(PQ)$, alors dans $\mathbb{Z}/p\mathbb{Z}[X]$, $PQ = 0$ (car p divise tous les coefficients de PQ) et alors, $\mathbb{Z}/p\mathbb{Z}$ étant intègre, $P = 0$ ou $Q = 0$ dans $\mathbb{Z}/p\mathbb{Z}[X]$ et alors p divise tous les coefficients de P ou tous les coefficients de Q, *i.e.* $p | c(P)$ ou $p | c(Q)$ ce qui est absurde car $c(P) = c(Q) = 1$. On en déduit donc qu'on a bien $c(PQ) = 1$.
2. Soit $P \in \mathbb{Z}[X] \setminus \{0\}$ réductible dans $\mathbb{Q}[X]$. Étant donné que diviser P par un entier (qui divise tous les coefficients pour rester dans $\mathbb{Z}[X]$) ne change pas fait que P soit irréductible ou non dans $\mathbb{Z}[X]$, quitte à diviser par $c(P)$, on suppose que $c(P) = 1$.

   Soit $\lambda \in \mathbb{Z}$ le coefficient dominant de P. P étant réductible dans $\mathbb{Q}[X]$, il existe $Q, R \in \mathbb{Q}[X]$ unitaires tels que $\deg Q \geq 1$, $\deg R \geq 1$ et $P = \lambda QR$. On va montrer que λ s'écrit $\lambda = cd$ avec $c, d \in \mathbb{Z}$ et $cQ, dR \in \mathbb{Z}[X]$.

   Posons $a = \min \underbrace{\{k \geq 1,\ kQ \in \mathbb{Z}[X]\}}_{\neq \varnothing}$. On veut montrer que $c(aQ) = 1$. Supposons le contraire. On a alors $\dfrac{aQ}{c(aQ)} \in \mathbb{Z}[X]$, et $c(aQ)$ divise le coefficient dominant de aQ qui est égal à a, donc $\dfrac{a}{c(aQ)} \in \mathbb{N}^*$ ce qui contredit la minimalité de a. En posant donc $b = \min\{k \geq 1,\ kR \in \mathbb{Z}[X]\}$, on a de même

(cc) BY-NC-SA

$c(b\mathrm{R}) = 1$. On a alors

$$ab = ab \cdot c(\mathrm{P}) = c(ab\mathrm{P}) = c(\lambda \times a\mathrm{Q} \times b\mathrm{R}) = |\lambda| c(a\mathrm{Q}) c(b\mathrm{R}) = |\lambda|.$$

Quitte à changer le signe de a, on peut donc écrire $\mathrm{P} = \underbrace{a\mathrm{Q}}_{\in \mathbb{Z}[\mathrm{X}]} \times \underbrace{b\mathrm{R}}_{\in \mathbb{Z}[\mathrm{X}]}$ et alors P est bien réductible dans $\mathbb{Z}[\mathrm{X}]$.

On peut montrer d'une manière plus générale, par la même démonstration ci-haut, que si $\mathrm{P} = \lambda \prod_{i=1}^{r} \mathrm{Q}_i \in \mathbb{Z}[\mathrm{X}] \setminus \{0\}$ est une factorisation de P dans $\mathbb{Q}[\mathrm{X}]$ avec Q_i unitaires et λ le coefficient dominant de P, alors on peut écrire $\lambda = c(\mathrm{P})\lambda_1 \ldots \lambda_r$ avec pour tout $i \in [\![1; r]\!]$, $\lambda_i \in \mathbb{Z}$ et $\lambda_i \mathrm{Q}_i \in \mathbb{Z}[\mathrm{X}]$. En particulier,

→ Si $\mathrm{P} = \lambda \prod_{i=1}^{r} \mathrm{Q}_i$ est la factorisation en irréductibles dans $\mathbb{Q}[\mathrm{X}]$, alors $\mathrm{P} = c(\mathrm{P}) \prod_{i=1}^{r} (\lambda_i \mathrm{Q}_i)$ est une factorisation en irréductibles dans $\mathbb{Z}[\mathrm{X}]$.

→ Pour tout $\mathrm{Q} \in \mathbb{Z}[\mathrm{X}] \setminus \{0\}$,

$$\mathrm{Q}|\mathrm{P} \text{ dans } \mathbb{Z}[\mathrm{X}] \Longleftrightarrow \mathrm{Q}|\mathrm{P} \text{ dans } \mathbb{Q}[\mathrm{X}] \text{ et } c(\mathrm{Q})|c(\mathrm{P}).$$

Le cas particulier où $\pm\mathrm{Q}$ est unitaire est assez connu et est une conséquence du fait qu'on peut effectuer la division euclidienne dans $\mathbb{Z}[\mathrm{X}]$ par un polynôme unitaire (ou de coefficient dominant égal à -1) de $\mathbb{Z}[\mathrm{X}]$ (en effet, lorsque $\mathbb{K}$ n'est pas un corps, on ne peut pas toujours faire la division euclidienne dans $\mathbb{K}[\mathrm{X}]$) et du fait que la division euclidienne dans $\mathbb{Z}[\mathrm{X}]$ (lorsqu'elle est faisable dans $\mathbb{Z}[\mathrm{X}]$) et $\mathbb{Q}[\mathrm{X}]$ donne toujours le même résultat.

→ Si P est unitaire et $\mathrm{P} = \prod_{i=1}^{r} \mathrm{Q}_i$ est une factorisation dans $\mathbb{Q}[\mathrm{X}]$ avec les Q_i unitaires, alors pour tout $i \in [\![1; r]\!]$, $\mathrm{Q}_i \in \mathbb{Z}[\mathrm{X}]$. En effet, en reprenant les notations vu précédemment, $1 = c(\mathrm{P})\lambda_1 \ldots \lambda_r$ implique que pour tout $i \in [\![1; r]\!]$, $\lambda_i \in \{-1, 1\}$. Cette propriété est particulièrement intéressante dans le cas où l'égalité $\mathrm{P} = \prod_{i=1}^{r} \mathrm{Q}_i$ est une factorisation en irréductibles dans $\mathbb{Q}[\mathrm{X}]$, vu qu'elle donne que chaque facteur est en fait (unitaire et) dans $\mathbb{Z}[\mathrm{X}]$.

(cc) BY-NC-SA

Chapitre 26

Systèmes linéaires

Soit $\mathbb{K}$ un corps et $m, n \in \mathbb{N}^*$.

I Généralités

D'une manière générale, un système linéaire s'écrit de la manière suivante.

$$(\mathrm{S}) : \begin{cases} a_{1,1}x_1 & + \cdots + & a_{1,n}x_n & = & b_1 \\ & \vdots & & & \\ a_{m,1}x_1 & + \cdots + & a_{m,n}x_n & = & b_m \end{cases} \iff \mathrm{AX} = \mathrm{B}$$

où $\mathrm{X} = (x_1, \ldots, x_n)^{\mathrm{T}}$ sont les inconnus, $\mathrm{A} = (a_{i,j})_{\substack{i \in [\![1;m]\!] \\ j \in [\![1;n]\!]}}$ une matrice à coefficients dans $\mathbb{K}$ et $\mathrm{B} = (b_1, \ldots, b_m)^{\mathrm{T}}$ un vecteur de $\mathbb{K}^m$.

Vocabulaire. On appelle rang du système (S) le rang de la matrice A. De plus, on dit que (S) est

→ Compatible lorsque l'ensemble des solutions, noté $\mathcal{S}$, est non vide.

→ Homogène lorsque $\mathrm{B} = 0$.

On notera $\mathcal{S}_0$ l'ensemble des solutions de l'équation $\mathrm{AX} = 0$.

Proposition I.1.

1. $\mathcal{S}_0$ est un sous-espace vectoriel de $\mathbb{K}^n$ de dimension $n - \operatorname{rg} \mathrm{A}$.
2. $\mathcal{S}$ est soit vide, soit un sous-espace affine de $\mathbb{K}^n$ parallèle à $\mathcal{S}_0$. En d'autres termes, $\mathcal{S}$ est soit vide, soit $\mathcal{S} = \mathrm{X}_0 + \mathcal{S}_0$ pour n'importe quel $\mathrm{X}_0 \in \mathcal{S}$.
3. $\mathcal{S} \neq \varnothing \iff \mathrm{B} \in \operatorname{Im} \mathrm{A}$.

Démonstration.

1. $\mathcal{S}_0 = \operatorname{Ker} \mathrm{A}$ et donc par la formule du rang,

$$\dim \mathcal{S}_0 = \dim \operatorname{Ker} \mathrm{A} = n - \dim \operatorname{Im} \mathrm{A} = n - \operatorname{rg} \mathrm{A}.$$

2. Supposons $\mathcal{S} \neq \varnothing$ et soit $\mathrm{X}_0 \in \mathcal{S}$. On a

$$\mathrm{X} \in \mathcal{S} \iff \mathrm{AX} = \mathrm{B} \iff \mathrm{AX} = \mathrm{AX}_0 \iff \mathrm{A}(\mathrm{X} - \mathrm{X}_0) = 0 \iff \mathrm{X} - \mathrm{X}_0 \in \mathcal{S}_0.$$

3. $\mathcal{S} \neq \varnothing \iff \exists \mathrm{X} \in \mathbb{K}^n \ \mathrm{AX} = \mathrm{B} \iff \mathrm{B} \in \operatorname{Im} \mathrm{A}$.

On aura besoin du lemme assez connu suivant.

Exercice I.2.

Soit $\mathrm{A} \in \mathcal{M}_{n,m}(\mathbb{K})$, alors $\operatorname{rg} \mathrm{A} = \operatorname{rg} \mathrm{A}^{\mathrm{T}}$.

II Système de Cramer

Dans cette partie, on suppose que $m = n$ et donc $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$.

BY-NC-SA

Théorème II.1.

Les propositions suivantes sont équivalentes.

1. $\mathrm{A} \in \mathrm{GL}_n(\mathbb{K})$.
2. $\forall \mathrm{B} \in \mathbb{K}^n,\ \mathrm{AX} = \mathrm{B}$ possède exactement une solution.
3. $\forall \mathrm{B} \in \mathbb{K}^n,\ \mathrm{AX} = \mathrm{B}$ possède au moins une solution.
4. La seule solution de $\mathcal{S}_0$ est 0.
5. $\det \mathrm{A} \neq 0$.

Si elles sont satisfaites, on dit que (S) est un système de Cramer.

Démonstration.

→ (1) ⇒ (2) Si $\mathrm{A} \in \mathrm{GL}_n(\mathbb{K})$, alors pour tout $\mathrm{B} \in \mathbb{K}^n$ l'unique solution de (S) est $\mathrm{X} = \mathrm{A}^{-1}\mathrm{B}$.

→ (2) ⇒ (3) Cette implication est facile à montrer.

→ (3) ⇒ (4) L'application $f_{\mathrm{A}} : \begin{cases} \mathbb{K}^n & \longrightarrow \mathbb{K}^n \\ \mathrm{X} & \longmapsto \mathrm{AX} \end{cases}$ est surjective et donc, par égalité des dimensions finies, est aussi injective et donc son noyau $\mathcal{S}_0 = \{0\}$.

→ (4) ⇒ (5) Si la seule solution de $\mathcal{S}_0$ est 0, alors $\operatorname{Ker} \mathrm{A} = \{0\}$ et donc A est inversible, *i.e.* $\det \mathrm{A} \neq 0$.

→ (5) ⇒ (1) Cette implication est facile à montrer.

Proposition (Formule de Cramer) II.2.

Supposons que (S) est de Cramer et posons $\mathrm{C}_1, \ldots, \mathrm{C}_n$ les colonnes de A et $\mathrm{X} = (x_1, \ldots, x_n)^{\mathrm{T}}$ l'unique solution de (S). On a nécessairement

$$\forall k \in [\![1; n]\!],\ x_k = \frac{\det(\mathrm{C}_1, \ldots, \mathrm{C}_{k-1}, \mathrm{B}, \mathrm{C}_{k+1}, \ldots, \mathrm{C}_n)}{\det \mathrm{A}}.$$

Démonstration. On a $\mathrm{B} = \mathrm{AX} = x_1\mathrm{C}_1 + \cdots + x_n\mathrm{C}_n$ et alors pour tout $k \in [\![1; n]\!]$,

$$\begin{aligned}\frac{\det(\mathrm{C}_1, \ldots, \mathrm{C}_{k-1}, \mathrm{B}, \mathrm{C}_{k+1}, \ldots, \mathrm{C}_n)}{\det \mathrm{A}} &= \sum_{i=1}^{n} x_i \frac{\det(\mathrm{C}_1, \ldots, \mathrm{C}_{k-1}, \mathrm{C}_i, \mathrm{C}_{k+1}, \ldots, \mathrm{C}_n)}{\det \mathrm{A}} \\ &= x_k \frac{\det(\mathrm{C}_1, \ldots, \mathrm{C}_{k-1}, \mathrm{C}_k, \mathrm{C}_{k+1}, \ldots, \mathrm{C}_n)}{\det \mathrm{A}} = x_k.\end{aligned}$$

Exemple. Lorsque $n = 2$, on peut écrire (S) sous la forme suivante

$$(\mathrm{S}) : \begin{cases} ax + by = c \\ a'x + b'y = c'. \end{cases}$$

(S) est de cramer si et seulement si $\begin{vmatrix} a & b \\ a' & b' \end{vmatrix} = ab' - a'b \neq 0$ et dans ce cas, la seule solution de (S) est

$$x = \frac{\begin{vmatrix} c & b \\ c' & b' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}} = \frac{b'c - bc'}{ab' - a'b} \text{ et } y = \frac{\begin{vmatrix} a & c \\ a' & c' \end{vmatrix}}{\begin{vmatrix} a & b \\ a' & b' \end{vmatrix}} = \frac{ac' - a'c}{ab' - a'b}.$$

La propriété exposée dans cet exercice est essentielle et est utilisée dans le chapitre de réduction d'endomorphismes.

(cc) BY-NC-SA

Exercice (Lemme d'Hadamard) II.3.

Soit $\mathrm{A} = (a_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{C})$ tel que pour tout $i \in [\![1;n]\!]$, $|a_{i,i}| > \sum_{j=1,j\neq i}^{n} |a_{i,j}|$. Montrer que $\mathrm{A} \in \mathrm{GL}_n(\mathbb{C})$.

Remarque. En transposant, on obtient un énoncé similaire, mais en sommant sur la même colonne.

III Valeurs propres et systèmes homogènes

Vocabulaire. On dit que $\lambda \in \mathbb{K}$ est une valeur propre de A s'il existe $\mathrm{X} \in \mathbb{K}^n \setminus \{0\}$ tel que $\mathrm{AX} = \lambda\mathrm{X}$.

Proposition III.1.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, $\lambda \in \mathbb{K}$ et $f_{\mathrm{A}} : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^n \\ \mathrm{X} & \mapsto \mathrm{AX} \end{cases}$. Les propositions suivantes sont équivalentes.

1. λ valeur propre de A.
2. $\mathrm{Ker}(f_{\mathrm{A}} - \lambda\,\mathrm{Id}) \neq \{0\}$.
3. $\det(\mathrm{A} - \lambda\mathrm{I}_n) = 0$.
4. Le système $(\mathrm{A} - \lambda\mathrm{I}_n)\mathrm{X} = 0$ a une solution $\mathrm{X} \neq 0$.

Démonstration. Il s'agit d'une conséquence directe de ce qui précède.

Exercice III.2.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$. Montrer que les valeurs propres complexes de A appartiennent à

$$\bigcup_{i=1}^{n} \mathrm{B}_f\left(a_{i,i}, \sum_{j=1,j\neq i}^{n} |a_{i,j}|\right).$$

Remarque. On peut améliorer cette borne en considérant aussi la transposée de A.

IV Matrices de permutation

Définition IV.1.

Pour tout $\sigma \in \mathcal{S}_n$, on pose $\mathrm{P}_\sigma = (\delta_{i,\sigma(j)})_{i,j\in[\![1;n]\!]}$. On appelle cette matrice la matrice de permutation associée à σ.

On présente ces propriétés qu'on utilisera dans la suite comme exercice.

(cc) BY-NC-SA

Exercice IV.2.

Soit $\sigma, \sigma' \in \mathcal{S}_n$.

1. Calculer $\det \mathrm{P}_\sigma$.
2. Montrer que $\mathrm{P}_\sigma \mathrm{P}_{\sigma'} = \mathrm{P}_{\sigma\circ\sigma'}$ et en déduire que $\mathrm{P}_\sigma^{-1} = \mathrm{P}_{\sigma^{-1}}$.
3. En déduire que $\mathrm{G} = \{\mathrm{P}_\sigma,\ \sigma \in \mathcal{S}_n\}$ muni du produit matriciel est isomorphe à $\mathcal{S}_n$.
4. Soit $\mathrm{A} = (a_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{K})$ et $\mathrm{X} = (x_i)_{i\in[\![1;n]\!]} \in \mathbb{K}^n$. Calculer $\mathrm{P}_\sigma\mathrm{A}$, $\mathrm{A}\mathrm{P}_\sigma$ et $\mathrm{P}_\sigma\mathrm{X}$.

V Comment déterminer $\operatorname{Ker} \mathrm{A}$ en pratique ?

Dans cette partie, nous présentons un algorithme pour résoudre les équations de la forme $\mathrm{AX} = \mathrm{B}$. Nous recommandons au lecteur de connaître uniquement les grandes lignes de cet algorithme et de savoir l'appliquer dans des cas simples, car il arrive des fois qu'on demande de faire des calculs de ce type à l'oral. Nous rappelons que ce que nous présentons ci-dessous n'est pas le seul moyen de résoudre ce type d'équations et que le lecteur se sentant à l'aise avec une autre méthode peut parfaitement s'en contenter, tout en comprenant l'idée derrière cet algorithme.

1. Explication de l'algorithme

Le but de cette partie est de répondre à la question suivante : Étant donnée $\mathrm{A} = (a_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{K})$, comment déterminer $\operatorname{Ker} \mathrm{A}$? Cette question est équivalente à déterminer l'ensemble des solutions du système $\mathrm{AX} = 0$. Posons alors $r = \operatorname{rg} \mathrm{A}$. Bien entendu, la réponse est claire lorsque $r = n$ car il s'agit d'un système de Cramer, on suppose donc que $r < n$.

$r \in [\![1; n-1]\!]$ donc il existe r lignes (qui forment une famille libre) de A telles que les $n-r$ lignes restantes sont combinaison linéaire de ces r lignes. Quitte à permuter les lignes de A, supposons que ces lignes sont les r premières. Écrivons alors le système $\mathrm{AX} = 0$.

$$\begin{cases} a_{1,1}x_1 & +\cdots+ & a_{1,n}x_n & = & 0 \\ & \vdots & & & \\ a_{r,1}x_1 & +\cdots+ & a_{r,n}x_n & = & 0 \\ a_{r+1,1}x_1 & +\cdots+ & a_{r+1,n}x_n & = & 0 \\ & \vdots & & & \\ a_{m,1}x_1 & +\cdots+ & a_{m,n}x_n & = & 0. \end{cases}$$

Il est facile de voir que si les r premières équations sont satisfaites, alors les $n-r$ suivantes aussi, donc il faut et il suffit de résoudre le système suivant.

$$(\mathrm{S}) : \begin{cases} a_{1,1}x_1 & +\cdots+ & a_{1,n}x_n & = & 0 \\ & \vdots & & & \\ a_{r,1}x_1 & +\cdots+ & a_{r,n}x_n & = & 0. \end{cases}$$

C'est-à-dire en posant $\mathrm{B} = (a_{i,j})_{i\in[\![1;r]\!],j\in[\![1;n]\!]}$, le système devient équivalent à $\mathrm{BX} = 0$. Les lignes de B forment une famille libre, donc $\operatorname{rg} \mathrm{B} = r$ et alors quitte à permuter les colonnes de B, on peut affirmer que les r premières colonnes de B forment une famille libre, et donc que la matrice $\mathrm{B}' = (a_{i,j})_{i,j\in[\![1;r]\!]}$ est inversible. Posons alors $\mathrm{B}'' = (a_{i,j})_{i\in[\![1;r]\!],j\in[\![n-r;n]\!]}$. En posant pour tout $\mathrm{X} \in \mathbb{K}^n$, $\mathrm{X} = \begin{pmatrix} \mathrm{X}' \\ \mathrm{X}'' \end{pmatrix}$ avec $\mathrm{X}' \in \mathbb{K}^r$

(cc) BY-NC-SA

et $X'' \in \mathbb{K}^{n-r}$, on peut écrire

$$(S) \Longleftrightarrow BX = 0 \Longleftrightarrow \begin{pmatrix} B' & B'' \end{pmatrix} \begin{pmatrix} X' \\ X'' \end{pmatrix} = 0 \Longleftrightarrow B'X' + B''X'' = 0 \Longleftrightarrow X' = -(B')^{-1}B''X''.$$

Remarquons qu'étant donné X'', le système $B'X' = -B''X''$ est de Cramer et admet donc une unique solution $-B'^{-1}B''X''$. On en déduit alors que pour tout $X = \begin{pmatrix} X' \\ X'' \end{pmatrix} \in \mathbb{K}^n$,

$$X \in \operatorname{Ker} A \Longleftrightarrow X' = -B'^{-1}B''X'' \text{ et alors } \operatorname{Ker} A = \left\{ \begin{pmatrix} -B'^{-1}B''X'' \\ X'' \end{pmatrix},\ X'' \in \mathbb{K}^{n-r} \right\}.$$

Ensuite, en notant pour tout $j \in [\![1; n-r]\!]$, $E_j = (\delta_{i,j})_{i \in [\![1;n-r]\!]}$ le $j-$ème vecteur de la base canonique de $\mathbb{K}^{n-r}$ on peut voir facilement que la famille

$$\mathcal{B} = \left\{ \begin{pmatrix} -B'^{-1}B''E_1 \\ E_1 \end{pmatrix}, \dots, \begin{pmatrix} -B'^{-1}B''E_{n-r} \\ E_{n-r} \end{pmatrix} \right\}$$

est une base de $\operatorname{Ker} A$ (nous conseillons au lecteur de le montrer à titre d'exercice). En pratique, les vecteurs de cette base peuvent être retrouvés en résolvant les systèmes

$$B'X' = -B''E_1,\ \dots\ , B'X' = -B''E_{n-r},$$

qui permettent de retrouver respectivement les vecteurs $-B'^{-1}B''E_1, \dots, -B'^{-1}B''E_{n-r}$ en utilisant le pivot de Gauss. On en déduit donc que finalement

$$\operatorname{Ker} A = \operatorname{Vect}\left(\left\{ \begin{pmatrix} -B'^{-1}B''E_1 \\ E_1 \end{pmatrix}, \dots, \begin{pmatrix} -B'^{-1}B''E_{n-r} \\ E_{n-r} \end{pmatrix} \right\} \right).$$

Remarques.

→ Ici, on a supposé que les r premières lignes de A forment une famille libre. Cette propriété est en général fausse. En pratique, après avoir trouvé des indices $i_1, \dots, i_r$ tels que les lignes de A d'indices $i_1, \dots, i_r$ forment une famille libre, la matrice B qu'on considérera sera égale à $\begin{pmatrix} L_{i_1} \\ \vdots \\ L_{i_r} \end{pmatrix}$ où $L_1, \dots, L_n$ sont les lignes de A.

→ On a également supposé que les r premières colonnes de B forment une famille libre, quitte à permuter les colonnes. Pour être un peu plus rigoureux, on doit introduire quelques éléments. Si on pose $C_1, \dots, C_n$ les colonnes de B, $\sigma : [\![1; n]\!] \longrightarrow [\![1; n]\!]$ une permutation telle que $(C_{\sigma(1)}, \dots, C_{\sigma(r)})$ forment une famille libre et $P_\sigma = (\delta_{i,\sigma(j)})_{i,j \in [\![1;n]\!]}$. D'après l'exercice IV.2, on a

$$BP_\sigma = \begin{pmatrix} C_{\sigma(1)} & \dots & C_{\sigma(n)} \end{pmatrix},$$

et alors, on voit que le raisonnement ci-dessus est valable si on remplace B par $\tilde{B} = BP_\sigma$, *i.e.*

$$\tilde{B}X = 0 \Longleftrightarrow X \in \operatorname{Vect}\left(\left\{ \begin{pmatrix} -\tilde{B}'^{-1}\tilde{B}''E_1 \\ E_1 \end{pmatrix}, \dots, \begin{pmatrix} -\tilde{B}'^{-1}\tilde{B}''E_{n-r} \\ E_{n-r} \end{pmatrix} \right\} \right),$$

(cc) BY-NC-SA

et donc on peut écrire

$$\begin{aligned}
\mathrm{X} \in \operatorname{Ker} \mathrm{A} &\Longleftrightarrow \mathrm{BX} = 0 \Longleftrightarrow \tilde{\mathrm{B}}\mathrm{P}_\sigma^{-1}\mathrm{X} = 0 \\
&\Longleftrightarrow \mathrm{P}_\sigma^{-1}\mathrm{X} \in \operatorname{Vect}\left(\left\{\begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_1 \\ \mathrm{E}_1 \end{pmatrix}, \dots, \begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_{n-r} \\ \mathrm{E}_{n-r} \end{pmatrix}\right\}\right) \\
&\Longleftrightarrow \mathrm{X} \in \operatorname{Vect}\left(\left\{\mathrm{P}_\sigma\begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_1 \\ \mathrm{E}_1 \end{pmatrix}, \dots, \mathrm{P}_\sigma\begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_{n-r} \\ \mathrm{E}_{n-r} \end{pmatrix}\right\}\right),
\end{aligned}$$

et donc en conclusion, on en déduit que

$$\operatorname{Ker} \mathrm{A} = \operatorname{Vect}\left(\left\{\mathrm{P}_\sigma\begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_1 \\ \mathrm{E}_1 \end{pmatrix}, \dots, \mathrm{P}_\sigma\begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_{n-r} \\ \mathrm{E}_{n-r} \end{pmatrix}\right\}\right).$$

→ Dans le cas où A n'est pas carrée, on peut éliminer des lignes comme à l'étape 2 et appliquer exactement le même algorithme.

→ Le raisonnement ci-dessus montre en particulier que $\dim \operatorname{Ker} \mathrm{A} = n - r$, c'est-à-dire la formule du rang $\operatorname{rg} \mathrm{A} + \dim \operatorname{Ker} \mathrm{A} = n$.

2. Algorithme en pratique

Résumons maintenant les étapes de l'algorithme qui permet de trouver Ker A.

→ Trouver le rang r de A (en échelonnant par exemple) et ensuite $\mathrm{L}_{i_1}, \dots, \mathrm{L}_{i_r}$ r lignes de A formant une famille libre.

→ Poser $\mathrm{B} = \begin{pmatrix} \mathrm{L}_{i_1} \\ \vdots \\ \mathrm{L}_{i_r} \end{pmatrix}$. Trouver des indices $j_1, \dots, j_n$ tels que si $\mathrm{C}_1, \dots, \mathrm{C}_n$ sont les colonnes de B, alors les colonnes $\mathrm{C}_{j_1}, \dots, \mathrm{C}_{j_r}$ forment une famille libre et poser

$$\tilde{\mathrm{B}} = \begin{pmatrix} \mathrm{C}_{j_1} & \dots & \mathrm{C}_{j_n} \end{pmatrix},\ \tilde{\mathrm{B}}' = \begin{pmatrix} \mathrm{C}_{j_1} & \dots & \mathrm{C}_{j_r} \end{pmatrix},\ \tilde{\mathrm{B}}'' = \begin{pmatrix} \mathrm{C}_{j_{r+1}} & \dots & \mathrm{C}_{j_n} \end{pmatrix}.$$

→ Résoudre les systèmes

$$\tilde{\mathrm{B}}'\mathrm{X}' = -\tilde{\mathrm{B}}''\mathrm{E}_1,\ \dots\ , \tilde{\mathrm{B}}'\mathrm{X}' = -\tilde{\mathrm{B}}''\mathrm{E}_{n-r}$$

et poser $\mathrm{X}_1, \dots, \mathrm{X}_{n-r} \in \mathbb{K}^{n-r}$ les solutions de ces systèmes.

→ Poser pour tout $i \in [\![1; n-r]\!]$, $\mathrm{Y}_i = \begin{pmatrix} \mathrm{X}_i \\ \mathrm{E}_i \end{pmatrix} = \begin{pmatrix} y_{i,j_1} \\ \vdots \\ y_{i,j_n} \end{pmatrix}$ où $(\mathrm{E}_1, \dots, \mathrm{E}_{n-r})$ est la base canonique de $\mathbb{K}^{n-r}$.

→ En posant pour tout $i \in [\![1; n-r]\!]$, $\tilde{\mathrm{Y}}_i = \begin{pmatrix} y_{i,1} \\ \vdots \\ y_{i,n} \end{pmatrix}$ (on réordonne les coordonnées de Y_i en effectuant les opérations inverses appliquées à B), la famille $\mathcal{B} = (\tilde{\mathrm{Y}}_1, \dots, \tilde{\mathrm{Y}}_{n-r})$ est une base de Ker A.

Remarque. Si on reprend les notations de la remarque précédente, on a pour tout $k \in [\![1; n]\!]$, $j_k = \sigma(k)$ et donc $\tilde{\mathrm{Y}}_i = \mathrm{P}_\sigma \mathrm{Y}_i$ et $\tilde{\mathrm{B}} = \mathrm{BP}_\sigma$.

Appliquons cet algorithme via l'exemple suivant.

Application. On considère la matrice $\mathrm{A} = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 2 \\ 2 & 4 & 5 & 5 \\ -1 & -2 & -1 & -4 \end{pmatrix}$. Trouver Ker A.

(cc) BY-NC-SA

→ **Étape 1 :** Trouver le rang r de A (en échelonnant par exemple) et ensuite $\mathrm{L}_{i_1}, \ldots, \mathrm{L}_{i_r}$ r lignes de A formant une famille libre.

On échelonne la matrice A par lignes en effectuant les opérations suivantes.

$$\begin{aligned}
\text{ligne } 2 &\longleftarrow \text{ligne } 2 - \text{ligne } 1\\
\text{ligne } 3 &\longleftarrow \text{ligne } 3 - 2 \times \text{ligne } 1\\
\text{ligne } 4 &\longleftarrow \text{ligne } 4 + \text{ligne } 1\\
\text{ligne } 3 &\longleftarrow \text{ligne } 3 - \text{ligne } 2\\
\text{ligne } 4 &\longleftarrow \text{ligne } 4 - \text{ligne } 2.
\end{aligned}$$

On obtient la matrice $\begin{pmatrix} 1 & 2 & 2 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ ce qui nous permet d'affirmer que A est rang 2. De plus, on voit que la première et la deuxième ligne de A forment une famille libre.

→ **Étape 2 :** Poser $\mathrm{B} = \begin{pmatrix} \mathrm{L}_{i_1} \\ \vdots \\ \mathrm{L}_{i_r} \end{pmatrix}$. Trouver des indices $j_1, \ldots, j_n$ tels que si $\mathrm{C}_1, \ldots, \mathrm{C}_n$ sont les colonnes de B, alors les colonnes $\mathrm{C}_{i_1}, \ldots, \mathrm{C}_{i_r}$ forment une famille libre et poser

$$\tilde{\mathrm{B}} = \begin{pmatrix} \mathrm{C}_{j_1} & \ldots & \mathrm{C}_{j_n} \end{pmatrix}, \ \tilde{\mathrm{B}}' = \begin{pmatrix} \mathrm{C}_{j_1} & \ldots & \mathrm{C}_{j_r} \end{pmatrix}, \ \tilde{\mathrm{B}}'' = \begin{pmatrix} \mathrm{C}_{j_{r+1}} & \ldots & \mathrm{C}_{j_n} \end{pmatrix}.$$

On pose donc

$$\mathrm{B} = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 2 \end{pmatrix}.$$

Les deux premières colonnes de B ne forment pas une famille libre mais les deux dernières si, on permute alors les colonnes pour que les deux premières soient libres de la manière suivante

$$\begin{aligned}
\text{colonne } 1 \text{ de } \tilde{\mathrm{B}} &\longleftarrow \text{colonne } 3 \text{ de } \mathrm{B}\\
\text{colonne } 3 \text{ de } \tilde{\mathrm{B}} &\longleftarrow \text{colonne } 1 \text{ de } \mathrm{B}\\
\text{colonne } 2 \text{ de } \tilde{\mathrm{B}} &\longleftarrow \text{colonne } 4 \text{ de } \mathrm{B}\\
\text{colonne } 4 \text{ de } \tilde{\mathrm{B}} &\longleftarrow \text{colonne } 2 \text{ de } \mathrm{B}.
\end{aligned}$$

C'est-à-dire $j_1 = 3$, $j_2 = 4$, $j_3 = 1$ et $j_4 = 2$. On pose alors

$$\tilde{\mathrm{B}} = \begin{pmatrix} 2 & 3 & 1 & 2 \\ 3 & 2 & 1 & 2 \end{pmatrix}, \ \tilde{\mathrm{B}}' = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \text{ et } \tilde{\mathrm{B}}'' = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

→ **Étape 3 :** Résoudre les systèmes

$$\tilde{\mathrm{B}}'\mathrm{X}' = -\tilde{\mathrm{B}}''\mathrm{E}_1, \ \ldots \ , \tilde{\mathrm{B}}'\mathrm{X}' = -\tilde{\mathrm{B}}''\mathrm{E}_{n-r},$$

et poser $\mathrm{X}_1, \ldots, \mathrm{X}_{n-r} \in \mathbb{K}^{n-r}$ les solutions de ces systèmes.

Ici, on a deux systèmes

$$\tilde{\mathrm{B}}'\mathrm{X}' = -\tilde{\mathrm{B}}''\mathrm{E}_1 \text{ et } \tilde{\mathrm{B}}'\mathrm{X}' = -\tilde{\mathrm{B}}''\mathrm{E}_2,$$

BY-NC-SA

i.e.

$$\begin{pmatrix}2&3\\3&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=-\begin{pmatrix}1&2\\1&2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} \text{ et } \begin{pmatrix}2&3\\3&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=-\begin{pmatrix}1&2\\1&2\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix},$$

c'est-à-dire

$$\begin{cases}2x+3y=-1\\3x+2y=-1\end{cases} \text{ et } \begin{cases}2x+3y=-2\\3x+2y=-2.\end{cases}$$

En résolvant ces deux systèmes, on trouve que leurs solutions sont respectivement

$$\mathrm{X}_1=\begin{pmatrix}-1/5\\-1/5\end{pmatrix} \text{ et } \mathrm{X}_2=\begin{pmatrix}-2/5\\-2/5\end{pmatrix}.$$

→ **Étape 4 :** Poser pour tout $i \in [\![1; n-r]\!]$, $\mathrm{Y}_i=\begin{pmatrix}\mathrm{X}_i\\\mathrm{E}_i\end{pmatrix}=\begin{pmatrix}y_{i,j_1}\\\vdots\\y_{i,j_n}\end{pmatrix}$ où $(\mathrm{E}_1,\ldots,\mathrm{E}_{n-r})$ est la base canonique de $\mathbb{K}^{n-r}$.

On pose donc

$$\mathrm{Y}_1=\begin{pmatrix}-1/5\\-1/5\\1\\0\end{pmatrix} \text{ et } \mathrm{Y}_2=\begin{pmatrix}-2/5\\-2/5\\0\\1\end{pmatrix}.$$

→ **Étape 5 :** En posant pour tout $i \in [\![1; n-r]\!]$, $\tilde{\mathrm{Y}}_i=\begin{pmatrix}y_{i,1}\\\vdots\\y_{i,n}\end{pmatrix}$ (on réordonne les coordonnées de Y_i), la famille $\mathcal{B}=(\tilde{\mathrm{Y}}_1,\ldots,\tilde{\mathrm{Y}}_{n-r})$ est une base de $\operatorname{Ker}\mathrm{A}$.

On va effectuer la permutation inverse de celle de l'étape 2 aux lignes des Y_i, c'est-à-dire

$$\begin{aligned}&\text{ligne 3 de } \tilde{\mathrm{Y}}_i \longleftarrow \text{ligne 1 de } \mathrm{Y}_i\\&\text{ligne 1 de } \tilde{\mathrm{Y}}_i \longleftarrow \text{ligne 3 de } \mathrm{Y}_i\\&\text{ligne 4 de } \tilde{\mathrm{Y}}_i \longleftarrow \text{ligne 2 de } \mathrm{Y}_i\\&\text{ligne 2 de } \tilde{\mathrm{Y}}_i \longleftarrow \text{ligne 4 de } \mathrm{Y}_i,\end{aligned}$$

et donc on a dans ce cas

$$\tilde{\mathrm{Y}}_1=\begin{pmatrix}1\\0\\-1/5\\-1/5\end{pmatrix} \text{ et } \tilde{\mathrm{Y}}_2=\begin{pmatrix}0\\1\\-2/5\\-2/5\end{pmatrix},$$

et alors, on en déduit que

$$\operatorname{Ker}\mathrm{A}=\operatorname{Vect}\left(\left\{\begin{pmatrix}1\\0\\-1/5\\-1/5\end{pmatrix},\begin{pmatrix}0\\1\\-2/5\\-2/5\end{pmatrix}\right\}\right).$$

On peut facilement vérifier que ce résultat est bien correct. En effet, les deux vecteurs ci-dessus forment

(cc) BY-NC-SA

une famille libre et

$$\mathrm{A}\tilde{\mathrm{Y}}_1 = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 2 \\ 2 & 4 & 5 & 5 \\ -1 & -2 & -1 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1/5 \\ -1/5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ et } \mathrm{A}\tilde{\mathrm{Y}}_2 = \begin{pmatrix} 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 2 \\ 2 & 4 & 5 & 5 \\ -1 & -2 & -1 & -4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -2/5 \\ -2/5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Ces deux vecteurs sont donc bien dans $\operatorname{Ker} \mathrm{A}$. On sait de plus que $\dim \operatorname{Ker} \mathrm{A} = 2$ il est donc clair que $(\tilde{\mathrm{Y}}_1, \tilde{\mathrm{Y}}_2)$ est une base de $\operatorname{Ker} \mathrm{A}$ et alors $\operatorname{Ker} \mathrm{A} = \operatorname{Vect}(\tilde{\mathrm{Y}}_1, \tilde{\mathrm{Y}}_2)$.

Remarque culturelle. On a montré que

$$\operatorname{Ker} \mathrm{A} = \operatorname{Vect}\left(\left\{\mathrm{P}_\sigma \begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_1 \\ \mathrm{E}_1 \end{pmatrix}, \dots, \mathrm{P}_\sigma \begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_{n-r} \\ \mathrm{E}_{n-r} \end{pmatrix}\right\}\right).$$

En fait, grâce à ce résultat, en posant pour tout $j \in [\![1;n]\!]$, $e_j = (\delta_{i,j})_{i\in[\![1;n]\!]}$ le $j-$ème vecteur de la base canonique de $\mathbb{K}^n$, on remarque que le projecteur

$$\pi : \begin{cases} \operatorname{Ker} \mathrm{A} & \longrightarrow \operatorname{Vect}(e_{\sigma^{-1}(r+1)}, \dots, e_{\sigma^{-1}(n)}) \\ x_1 e_1 + \cdots + x_n e_n & \longmapsto x_{\sigma^{-1}(r+1)} e_{\sigma^{-1}(r+1)} + \cdots + x_{\sigma^{-1}(n)} e_{\sigma^{-1}(n)} \end{cases}$$

est un isomorphisme. En effet, si on pose pour tout $i \in [\![1; n-r]\!]$, $\mathrm{U}_i = \mathrm{P}_\sigma \begin{pmatrix} -\tilde{\mathrm{B}}'^{-1}\tilde{\mathrm{B}}''\mathrm{E}_i \\ \mathrm{E}_i \end{pmatrix}$, on peut facilement voir que pour tout $i \in [\![1; n-r]\!]$, $\pi(\mathrm{U}_i) = e_{\sigma^{-1}(r+i)}$ et que les deux familles $(\mathrm{U}_1, \dots, \mathrm{U}_{n-r})$ et $(e_{\sigma^{-1}(r+1)}, \dots, e_{\sigma^{-1}(n)})$ sont toutes les deux des bases de respectivement $\operatorname{Ker} \mathrm{A}$ et $\operatorname{Vect}(e_{\sigma^{-1}(r+1)}, \dots, e_{\sigma^{-1}(n)})$ donc π est bien un isomorphisme.

Ceci signifie que tout élément de $\operatorname{Ker} \mathrm{A}$ peut être identifié d'une manière unique via ses coordonnées d'indices $\{\sigma^{-1}(r+1), \dots, \sigma^{-1}(n)\}$.

VI Comment résoudre $\mathrm{AX} = \mathrm{B}$ en pratique ?

On suppose toujours que $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ et $\mathrm{B} \in \mathbb{K}^n$ et on pose $r = \operatorname{rg} \mathrm{A}$. Lorsque $r = n$, il s'agit d'un système de Cramer qu'on peut résoudre avec le pivot de Gauss par exemple. On sait d'après la proposition I.1 que l'ensemble des solutions $\mathcal{S}$ de l'équation $\mathrm{AX} = \mathrm{B}$ s'écrit sous forme de $\mathcal{S} = \mathrm{X}_0 + \mathcal{S}_0$ avec X_0 une solution de $\mathrm{AX} = \mathrm{B}$ et $\mathcal{S}_0$ l'ensemble des solutions de l'équation $\mathrm{AX} = 0$, *i.e.* $\mathcal{S}_0 = \operatorname{Ker} \mathrm{A}$. Pour déterminer $\operatorname{Ker} \mathrm{A}$, il suffit d'appliquer ce qu'on vient de voir dans la section précédente. Il suffit donc de trouver une solution particulière X_0 du système. Nous ne détaillerons pas comment la trouver, mais il suffit de vérifier que $\mathrm{B} \in \operatorname{Im} \mathrm{A}$ (si ce n'est pas le cas le système n'a pas de solution), d'éliminer $n-r$ lignes, fixer $n-r$ variables et résoudre un système de Cramer de taille r.

(cc) BY-NC-SA

Correction de l'exercice VII.15. :

Posons $r = \operatorname{rg} \mathrm{A}$. On sait qu'il existe $(\mathrm{P}, \mathrm{Q}) \in \mathrm{GL}_n(\mathbb{K}) \times \mathrm{GL}_m(\mathbb{K})$ tels que

$$\mathrm{A} = \mathrm{P} \underbrace{\begin{pmatrix} \mathrm{I}_r & 0_{r,m-r} \\ 0_{n-r,r} & 0_{n-r,m-r} \end{pmatrix}}_{\mathrm{J}_{r,n,m}} \mathrm{Q},$$

où J_r a r coefficients égaux à 1 sur la diagonale. On a alors $\mathrm{A}^{\mathrm{T}} = \mathrm{Q}^{\mathrm{T}} \mathrm{J}_{r,n,m}^{\mathrm{T}} \mathrm{P}^{\mathrm{T}} = \mathrm{Q}^{\mathrm{T}} \mathrm{J}_{r,m,n} \mathrm{P}^{\mathrm{T}}$ et alors $\operatorname{rg} \mathrm{A}^{\mathrm{T}} = \operatorname{rg}(\mathrm{Q}^{\mathrm{T}} \mathrm{J}_{r,m,n} \mathrm{P}^{\mathrm{T}}) = \operatorname{rg} \mathrm{J}_{r,n,m} = r$.

Correction de l'exercice II.3. :

Supposons par l'absurde que $\mathrm{A} \notin \mathrm{GL}_n(\mathbb{C})$. Il existe alors $\mathrm{X} = (x_1, \ldots, x_n)^{\mathrm{T}} \in \mathbb{C}^n$ tel que $\mathrm{AX} = 0$. Considérons alors $i_0 \in [\![1; n]\!]$ tel que $0 < |x_{i_0}| = \max\limits_{1 \le j \le n} |x_j|$. Le coefficient de position i_0 du vecteur AX est égal à 0, c'est à dire

$$\sum_{j=1}^{n} a_{i_0,j} x_j = 0 \quad i.e. \quad a_{i_0,i_0} x_{i_0} = - \sum_{j=1,\ j \ne i_0}^{n} a_{i_0,j} x_j,$$

et alors

$$|a_{i_0,i_0}|\,|x_{i_0}| = \left| \sum_{j=1,\ j \ne i_0}^{n} a_{i_0,j} x_j \right| \le |x_{i_0}| \sum_{j=1,\ j \ne i_0}^{n} a_{i_0,j}.$$

$|x_{i_0}|$ est strictement positif, on peut donc simplifier des deux côtés et obtenir

$$|a_{i_0,i_0}| \le \sum_{j=1,\ j \ne i_0}^{n} a_{i_0,j},$$

ce qui est en contradiction avec les hypothèses.

Correction de l'exercice II.3. :

C'est une application directe du lemme d'Hadamard. En effet, si λ est une valeur propre de A, alors $\mathrm{A} - \lambda \mathrm{I}_n$ n'est pas inversible et donc il existe $i \in [\![1; n]\!]$ tel que

$$|a_{i,i} - \lambda| \le \sum_{\substack{j=1 \\ j \ne i}}^{n} |a_{i,j}| \quad i.e. \quad \lambda \in \bigcup_{i=1}^{n} \mathrm{B}_f \left(a_{i,i}, \sum_{j=1, j \ne i}^{n} |a_{i,j}| \right).$$

Correction de l'exercice IV.2. :

1. On a
$$\det \mathrm{P}_\sigma = \sum_{\gamma \in \mathcal{S}_n} \varepsilon(\gamma) \prod_{i=1}^{n} [\mathrm{P}_\sigma]_{\gamma(i),i} = \sum_{\gamma \in \mathcal{S}_n} \varepsilon(\gamma) \underbrace{\prod_{i=1}^{n} \delta_{\gamma(i),\sigma(i)}}_{=0 \text{ lorsque } \gamma \ne \sigma} = \varepsilon(\sigma).$$

2. On a pour tout $i, j \in [\![1; n]\!]$,
$$[\mathrm{P}_\sigma \mathrm{P}_{\sigma'}]_{i,j} = \sum_{k=1}^{n} [\mathrm{P}_\sigma]_{i,k} \times [\mathrm{P}_{\sigma'}]_{k,j} = \sum_{k=1}^{n} \delta_{i,\sigma(k)} \delta_{k,\sigma'(j)}$$
$$\underset{(1)}{=} \delta_{i,\sigma(\sigma^{-1}(i))} \delta_{\sigma^{-1}(i),\sigma'(j)} = \delta_{\sigma^{-1}(i),\sigma'(j)} \underset{(2)}{=} \delta_{i,\sigma \circ \sigma'(j)} = [\mathrm{P}_{\sigma \circ \sigma'}]_{i,j}.$$

(cc) BY-NC-SA

L'égalité (1) est vraie car le seul terme à priori non nul dans la somme $\sum_{k=1}^{n}\delta_{\sigma(i),k}\delta_{k,\sigma'(j)}$ est celui d'indice $k=\sigma^{-1}(i)$. L'égalité (2) est vraie car $\sigma^{-1}(i)=\sigma'(j) \Longleftrightarrow i=\sigma\circ\sigma'(j)$.
On en déduit donc qu'on a bien $\mathrm{P}_\sigma\mathrm{P}_{\sigma'}=\mathrm{P}_{\sigma\circ\sigma'}$. En particulier, on a

$$\mathrm{P}_\sigma\mathrm{P}_{\sigma^{-1}}=\mathrm{P}_{\sigma\circ\sigma^{-1}}=\mathrm{P}_{\mathrm{Id}}=(\delta_{i,\mathrm{Id}(j)})_{i,j\in[\![1;n]\!]}=\mathrm{I},$$

c'est dire que $\mathrm{P}_\sigma^{-1}=\mathrm{P}_{\sigma^{-1}}$.

3. Il suffit de considérer le morphisme $\psi:\begin{cases}\mathcal{S}_n & \longrightarrow \mathrm{G}\\ \sigma & \longmapsto \mathrm{P}_\sigma\end{cases}$ et de montrer qu'il est bijectif.

4. On a pour tout $i,j\in[\![1;n]\!]$

$$[\mathrm{AP}_\sigma]_{i,j}=\sum_{k=1}^{n}[\mathrm{A}]_{i,k}\times[\mathrm{P}_\sigma]_{k,j}=\sum_{k=1}^{n}a_{i,k}\delta_{k,\sigma(j)}=a_{i,\sigma(j)},$$

et donc en posant $\mathrm{A}=\begin{pmatrix}\mathrm{C}_1 & \dots & \mathrm{C}_n\end{pmatrix}$ où $\mathrm{C}_1,\dots,\mathrm{C}_n$ sont les colonnes de A, on peut voir facilement que

$$\mathrm{AP}_\sigma=\begin{pmatrix}\mathrm{C}_{\sigma(1)} & \dots & \mathrm{C}_{\sigma(n)}\end{pmatrix}.$$

On en déduit donc que multiplier A à droite par P_σ revient à permuter les colonnes de A en appliquant σ aux indices.

On a encore une fois, pour tout $i,j\in[\![1;n]\!]$,

$$[\mathrm{P}_\sigma\mathrm{A}]_{i,j}=\sum_{k=1}^{n}[\mathrm{P}_\sigma]_{i,k}\times[\mathrm{A}]_{k,j}=\sum_{k=1}^{n}\delta_{i,\sigma(k)}a_{k,j}\underset{(1)}{=}a_{\sigma^{-1}(i),j}.$$

L'inégalité (1) ci-dessus est vraie car $\delta_{i,\sigma(k)}$ est non nul si et seulement si $i=\sigma(k)$, *i.e.* $k=\sigma^{-1}(i)$. On en déduit donc que multiplier A à gauche par P_σ revient à permuter les lignes de A en appliquant σ^{-1} aux indices.

On a enfin pour tout $i\in[\![1;n]\!]$,

$$[\mathrm{P}_\sigma\mathrm{X}]_i=\sum_{k=1}^{n}[\mathrm{P}_\sigma]_{i,k}\times x_k=\sum_{k=1}^{n}\delta_{i,\sigma(k)}x_k=x_{\sigma^{-1}(i)}.$$

On en déduit donc que multiplier X à gauche par P_σ revient à permuter les coordonnées de X en appliquant σ^{-1} aux indices.

(cc) BY-NC-SA

Chapitre 28

Réduction d'endomorphisme

Notations

Introduisons tout d'abord quelques notations. Soit $\mathbb{K}$ un corps commutatif et E un $\mathbb{K}-$espace vectoriel de dimension $n \in \mathbb{N}^*$. Soit $u \in \mathcal{L}(\mathrm{E})$ et $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$.

→ Si β est une base de E alors $[u]_\beta$ est la matrice de $\mathcal{M}_n(\mathbb{K})$ dont les colonnes sont les coordonnées de l'image de chaque élément de β par u dans la base β.

→ Pour toute base β de $\mathbb{K}^n$, on note $[\mathrm{A}]_\beta = \mathrm{PAP}^{-1}$ avec P la matrice dont les colonnes sont la représentation des éléments de β dans la base canonique.

→ Pour toutes matrices $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{C})$, on écrit $\mathrm{A} \simeq \mathrm{B}$ s'il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$ tel que $\mathrm{A} = \mathrm{PBP}^{-1}$.

→ Si F et G sont deux $\mathbb{K}-$espaces vectoriels, on écrit $\mathrm{F} \approx \mathrm{G}$ s'ils sont isomorphes, *i.e.* il existe $\phi \in \mathcal{L}(\mathrm{F}, \mathrm{G})$ un endomorphisme inversible tel que $\phi(\mathrm{F}) = \mathrm{G}$.

→ On note $\mathrm{Com}(u) = \{v \in \mathcal{L}(\mathrm{E}),\ v \circ u = u \circ v\}$.

→ On note $\mathrm{Com}(\mathrm{A}) = \{\mathrm{B} \in \mathcal{M}_n(\mathbb{K}),\ \mathrm{AB} = \mathrm{BA}\}$.

→ On note $\mathrm{VP}(u) = \{\lambda \in \mathbb{K},\ \mathrm{Ker}(u - \lambda\,\mathrm{Id}) \neq \{0\}\}$.

→ On note $\mathrm{VP}(\mathrm{A}) = \{\lambda \in \mathbb{K},\ \mathrm{Ker}(\mathrm{A} - \lambda \mathrm{I}) \neq \{0\}\}$.

→ Pour tout $a \in \mathbb{N}^*$, on note $\mathcal{T}_a^+(\mathbb{K})$ l'ensemble des matrices triangulaires supérieures de taille a.

→ Pour toute famille $(x_1, \ldots, x_r)$ de E (resp. $\mathbb{K}^n$) et toute base β de E (resp. $\mathbb{K}^n$), on note $[(x_1, \ldots, x_r)]_\beta$ la matrice de $\mathcal{M}_{n,r(\mathbb{K})}$ dont les colonnes sont les représentations des vecteurs x_i dans la base β.

→ Pour tout $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, on note $\mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$.

→ Pour tout $i, j \in [\![1; n]\!]$, on note $\mathrm{E}_{i,j} = (\delta_{i,k}\delta_{j,\ell})_{k,\ell \in [\![1;n]\!]} \in \mathcal{M}_n(\mathbb{K})$.

→ On note A^T la transposée de A.

→ On note $\langle \mathrm{A} \rangle = \{\mathrm{PAP}^{-1},\ \mathrm{P} \in \mathrm{GL}_n(\mathbb{K})\}$.

→ Pour tout polynôme $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$, on note $\langle \mathrm{P} \rangle = \{\mathrm{QP},\ \mathrm{Q} \in \mathbb{K}[\mathrm{X}]\}$.

→ On note $\mathbb{K}[u] = \{\mathrm{P}(u),\ \mathrm{P} \in \mathbb{K}[\mathrm{X}]\}$.

→ On note $\mathbb{K}[\mathrm{A}] = \{\mathrm{P}(\mathrm{A}),\ \mathrm{P} \in \mathbb{K}[\mathrm{X}]\}$.

→ Pour tout $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$, on note $\mathrm{Z}(\mathrm{P})$ l'ensemble des racines de P.

Préambule : sommes directes et bases adaptées

Soit $\mathbb{K}$ un corps commutatif et E un $\mathbb{K}-$espace vectoriel de dimension $n \in \mathbb{N}^*$. Soit $r \in \mathbb{N}^*$ et $\mathrm{F}_1, \ldots, \mathrm{F}_r$ des sous espaces vectoriel de E.
On considère l'application

$$j : \begin{cases} \mathrm{F}_1 \times \cdots \times \mathrm{F}_r & \longrightarrow \mathrm{E} \\ (x_1, \ldots, x_r) & \longmapsto x_1 + \cdots + x_r. \end{cases}$$

Par définition, $\mathrm{Im}\, j = \mathrm{F}_1 + \cdots + \mathrm{F}_r$. Cette somme est dite directe lorsque j est injective.

BY-NC-SA

Proposition .1.

Supposons que E est de dimension finie. On considère pour tout $k \in [\![1;r]\!]$, $(e_{k,1},\dots,e_{k,n_k})$ une base de F_k.

1. $\mathrm{F}_1 \times \cdots \times \mathrm{F}_r$ est un $\mathbb{K}-$espace vectoriel et $\dim \mathrm{F}_1 \times \cdots \times \mathrm{F}_r = \sum_{k=1}^{r} \dim \mathrm{F}_k$.
2. Si les F_j sont en somme directe, alors $\dim \bigoplus_{k=1}^{r} \mathrm{F}_k = \sum_{k=1}^{r} \dim \mathrm{F}_k$.
3. Si les F_j sont en somme directe, alors $(e_{1,1},\dots,e_{1,n_1},\dots,e_{r,1},\dots,e_{r,n_r})$ est une base de $\bigoplus_{k=1}^{r} \mathrm{F}_k$.

Démonstration.

1. On considère la famille $\mathcal{B} = \{(0,\dots,0,e_{k,\ell},0,\dots,0),\ k \in [\![1;r]\!],\ \ell \in [\![1;n_k]\!]\}$. $\mathcal{B}$ est une base de $\mathrm{F}_1 \times \cdots \times \mathrm{F}_r$ et $|\mathcal{B}| = \sum_{k=1}^{r} n_k = \sum_{k=1}^{r} \dim \mathrm{F}_k$, d'où le résultat.
2. j est un morphisme injectif, donc
$$\sum_{k=1}^{r} \dim \mathrm{F}_r = \dim \mathrm{F}_1 \times \cdots \times \mathrm{F}_r = \dim j(\mathrm{F}_1 \times \cdots \times \mathrm{F}_r) = \dim \bigoplus_{k=1}^{r} \mathrm{F}_k.$$
3. Posons $\mathcal{B}' = (e_{1,1},\dots,e_{1,n_1},\dots,e_{r,1},\dots,e_{r,n_r})$. On a
$$\begin{aligned}\mathrm{Vect}(\mathcal{B}') &= \mathrm{Vect}(e_{1,1},\dots,e_{1,n_1}) + \cdots + \mathrm{Vect}(e_{r,n_r},\dots,e_{r,n_r})\\ &= \mathrm{F}_1 \oplus \cdots \oplus \mathrm{F}_r,\end{aligned}$$
donc $\mathcal{B}'$ est une famille génératrice (de $\bigoplus_{k=1}^{r} \mathrm{F}_k$) de taille $n_1 + \cdots + n_r = \sum_{k=1}^{r} \dim \mathrm{F}_k = \dim \bigoplus_{k=1}^{r} \mathrm{F}_k$, et alors $\mathcal{B}'$ est une base de $\bigoplus_{k=1}^{r} \mathrm{F}_k$.

Objectif du chapitre : Pour une application linéaire $u \in \mathcal{L}(\mathrm{E})$ ou matrice $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ donnée, on souhaite trouver une base β ou une matrice $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$ où la matrice de u (resp. $\mathrm{P}^{-1}\mathrm{AP}$) est simple, *i.e.*
$$\mathrm{P}^{-1}\mathrm{AP} = [u]_\beta = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ ou } \mathrm{P}^{-1}\mathrm{AP} = [u]_\beta = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$
avec $\lambda_1,\dots,\lambda_r \in \mathbb{K}$, ou alors si u est nilpotente,
$$\mathrm{P}^{-1}\mathrm{AP} = [u]_\beta = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix},$$
ou alors
$$\mathrm{P}^{-1}\mathrm{AP} = [u]_\beta = \begin{pmatrix} \mathrm{J}_1 & & 0 \\ & \ddots & \\ 0 & & \mathrm{J}_k \end{pmatrix},$$

(cc) BY-NC-SA

tel que pour tout $k \in [\![1;r]\!]$, J_k un bloc de Jordan, *i.e.*

$$\mathrm{J}_k = \begin{pmatrix} a_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & a_k \end{pmatrix},$$

avec $a_k \in \mathbb{K}$, ou alors sous forme de matrice compagnon

$$\mathrm{P}^{-1}\mathrm{AP} = [u]_\beta = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix},$$

avec $a_0, \ldots, a_{n-1} \in \mathbb{K}$.

I Stabilité

Dans cette partie, on considère u un endomorphisme de E.

Définition I.1.

Un sous-espace vectoriel F de E est stable par u si $u(\mathrm{F}) \subset \mathrm{F}$.

Proposition I.2.

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

1. Si F et G sont stables par u, alors $\mathrm{F} + \mathrm{G}$ et $\mathrm{F} \cap \mathrm{G}$ le sont aussi.
2. Si F est de dimension finie, est stable par u et $\mathrm{F} \cap \operatorname{Ker} u = \{0\}$, alors $u(\mathrm{F}) = \mathrm{F}$.
3. u est une homothétie $\Longleftrightarrow$ u laisse stable tout sous-espace de E

Démonstration.

1. Supposons que F et G sont stables par u et soient $x, y \in \mathrm{E}$.
   - → Supposons que $x \in \mathrm{F}$ et $y \in \mathrm{G}$. Montrons que $u(x+y) \in \mathrm{F} + \mathrm{G}$.
     Par stabilité de F et G, on a $u(x) \in \mathrm{F}$ et $u(y) \in \mathrm{G}$ donc $u(x+y) = u(x) + u(y) \in \mathrm{F} + \mathrm{G}$, donc $\mathrm{F} + \mathrm{G}$ est stable par u.
   - → Supposons que $x \in \mathrm{F} \cap \mathrm{G}$.
     On a $x \in \mathrm{F}$ et $x \in \mathrm{G}$, donc $u(x) \in \mathrm{F}$ et $u(x) \in \mathrm{G}$ et alors $u(x) \in \mathrm{F} \cap \mathrm{G}$. On en déduit que $\mathrm{F} \cap \mathrm{G}$ est stable par u.
2. On considère l'application $\tilde{u} : \begin{cases} \mathrm{F} & \longrightarrow \mathrm{F} \\ x & \longmapsto u(x) \end{cases}$. On a $\operatorname{Ker} \tilde{u} = \operatorname{Ker} u \cap \mathrm{F}$ et en utilisant le théorème du rang,

$$\dim \operatorname{Ker} \tilde{u} + \operatorname{rg} \tilde{u} = \dim \mathrm{F},$$

donc

$$\dim \operatorname{Ker} u \cap \mathrm{F} + \dim u(\mathrm{F}) = \dim \mathrm{F},$$

et finalement

$$\dim u(\mathrm{F}) = \dim \mathrm{F}.$$

De plus, $u(\mathrm{F}) \subset \mathrm{F}$, donc $u(\mathrm{F}) = \mathrm{F}$.

BY-NC-SA

3. Montrons le résultat par double implication. Remarquons tout d'abord que si $\mathrm{E} = \{0\}$, alors le résultat est évident. Supposons donc que $\mathrm{E} \neq \{0\}$.

   → ($\Rightarrow$) Le sens direct est une conséquence de la stabilité de F par multiplication par un scalaire.

   → ($\Leftarrow$) Supposons que u laisse stable tout sous-espace de E.
   Pour tout $x, y \in \mathrm{E} \setminus \{0\}$, u laisse stable $\mathrm{Vect}(x)$ et $\mathrm{Vect}(y)$, donc il existe $\lambda_x, \lambda_y \in \mathbb{K}$ vérifiant

$$u(x) = \lambda_x x \text{ et } u(y) = \lambda_y y.$$

   u laisse aussi $\mathrm{Vect}(x+y)$ stable, donc il existe λ_{x+y} vérifiant $u(x+y) = \lambda_{x+y}(x+y)$. On a alors

$$\lambda_{x+y}(x+y) = u(x+y) = u(x) + u(y) = \lambda_x x + \lambda_y y,$$

   et alors

$$(\lambda_{x+y} - \lambda_x)x + (\lambda_{x+y} - \lambda_y)y = 0.$$

   - Si (x, y) est libre, alors $\lambda_{x+y} - \lambda_x = 0$ et $\lambda_{x+y} - \lambda_y = 0$, et alors $\lambda_x = \lambda_{x+y} = \lambda_y$.
   - Si (x, y) est liée, *i.e.* il existe $\lambda \in \mathbb{K}$ tel que $x = \lambda y$, alors

$$\lambda_y y = u(y) = \lambda u(x) = \lambda \lambda_x x = \lambda_x y,$$

     et donc, puisque $y \neq 0$, $\lambda_x = \lambda_y$.

   $x \longmapsto \lambda_x$ est alors constante sur $\mathbb{E} \setminus \{0\}$. Il existe donc $\lambda \in \mathbb{K}$ tel que pour tout $x \in \mathrm{E} \setminus \{0\}$, $u(x) = \lambda x$. Cette propriété est en particulier vraie pour $x = 0$, donc pour tout $x \in \mathrm{E}$, $u(x) = \lambda x$, *i.e.* u est une homothétie.

Exemple. Les sous espaces $\{0\}$, $\mathrm{Im}\, u^k$ et $\mathrm{Ker}\, u^k$ avec $k \in \mathbb{N}$ sont stables par u.

Proposition I.3.

Soit v un endomorphisme de E. Si u et v commutent, alors v laisse $\mathrm{Ker}\, u$ et $\mathrm{Im}\, u$ stables.
Plus généralement, $\forall \mathrm{P} \in \mathbb{K}[\mathrm{X}]$, $\mathrm{Ker}\, \mathrm{P}(u)$ et $\mathrm{Im}\, \mathrm{P}(u)$ sont stables par v.

Démonstration. Supposons que u et v commutent.

→ Si $x \in \mathrm{Ker}\, u$, alors $u(v(x)) = u \circ v(x) = v \circ u(x) = v(0) = 0$, donc $v(x) \in \mathrm{Ker}\, u$.

→ Si $y = u(x) \in \mathrm{Im}\, u$, alors $v(y) = u(v(x))$ et donc $v(y) \in \mathrm{Im}\, u$.

→ Étant donné que u et v commutent, une récurrence simple permet de montrer que pour tout $k \in \mathbb{N}$, u^k et v commutent également. Tout polynôme P en u est une combinaison linéaire des u^k qui commutent tous avec v, donc $\mathrm{P}(u)$ et v commutent, et alors le point précédent nous permet d'affirmer que v laisse $\mathrm{Ker}\, \mathrm{P}(u)$ et $\mathrm{Im}\, \mathrm{P}(u)$ stables.

Remarque. La réciproque est fausse en général. Prenons un contre-exemple. Soit (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$. On définit u et v comme suit :

$$\begin{cases} u(e_1) = 0 \\ u(e_2) = e_3 \\ u(e_3) = e_3 \end{cases} \quad \text{et} \quad \begin{cases} v(e_1) = e_1 \\ v(e_2) = 0 \\ v(e_3) = e_3. \end{cases}$$

On peut facilement voir que v laisse stable $\mathrm{Ker}\, u$ et $\mathrm{Im}\, u$, mais

$$\begin{cases} u \circ v(e_2) = 0 \\ v \circ u(e_2) = e_3. \end{cases}$$

(cc) BY-NC-SA

Proposition I.4.

Soit F un sous-espace vectoriel de E, β_1 une base de F et β_2 une base d'un supplémentaire de F. En posant $\beta = (\beta_1, \beta_2)$ la concaténation des deux bases et $p = \dim \mathrm{F} = |\beta_1|$, pour tout $u \in \mathcal{L}(\mathrm{E})$, on a équivalence entre les propositions suivantes.

1. u laisse stable F.
2. $\exists(\mathrm{A}, \mathrm{B}, \mathrm{C}) \in \mathcal{M}_p(\mathbb{K}) \times \mathcal{M}_{n-p}(\mathbb{K}) \times \mathcal{M}_{p,n-p}(\mathbb{K}),\ [u]_\beta = \begin{pmatrix} \mathrm{A} & \mathrm{C} \\ 0 & \mathrm{B} \end{pmatrix}$.

Démonstration.

→ (1) ⇒ (2) Supposons que u laisse stable F. Pour tout $e \in \beta_1$, $u(e) \in \mathrm{F} = \mathrm{Vect}(\beta_1)$ car u stabilise F, donc les coefficients dans la famille β_2 de e sont nuls.

→ (2) ⇒ (1) Pour tout $e \in \beta_1$, $[u(e)]_\beta$ a des coefficients nuls dans les $|\beta_2|$ dernières coordonnées, donc $u(e) \in \mathrm{Vect}(\beta_1) = \mathrm{F}$ pour tout $e \in \beta_1$. Par combinaison linéaire, on peut conclure que $u(\mathrm{F}) \subset \mathrm{F}$.

Proposition I.5.

Soit $r \in \mathbb{N}^*$ et $\mathrm{F}_1, \dots, \mathrm{F}_r$ des sous-espaces vectoriels de E non réduits à $\{0\}$ de dimensions respectives $n_1, \dots, n_r$. On suppose que $\mathrm{E} = \bigoplus_{k=1}^{r} \mathrm{F}_k$. Pour tout $j \in [\![1; r]\!]$, on considère β_j une base de F_j. Soit $\beta = (\beta_1, \dots, \beta_r)$ la concaténation des β_j. On a alors pour tout $u \in \mathcal{L}(\mathrm{E})$, on a équivalence entre les deux propositions suivantes.

1. u laisse stable F_j pour tout j.
2. $\exists(\mathrm{A}_1, \dots, \mathrm{A}_r) \in \mathcal{M}_{n_1}(\mathbb{K}) \times \cdots \times \mathcal{M}_{n_r}(\mathbb{K}),\ [u]_\beta = \begin{pmatrix} \mathrm{A}_1 & & 0 \\ & \ddots & \\ 0 & & \mathrm{A}_r \end{pmatrix}$.

Démonstration.

→ (1) ⇒ (2) Soit $k \in [\![1; r]\!]$. Pour tout $e \in \beta_k$, $u(e) \in \mathrm{Vect}(\beta_k)$, donc la sous-matrice de $[u]_\beta$ contenant des colonnes de rang entre $|\beta_1| + \cdots + |\beta_{k-1}| + 1$ et $|\beta_1| + \cdots + |\beta_k|$, *i.e.* la représentation de l'image de la base β_k par u dans la base β s'écrit $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathrm{A}_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ avec $\mathrm{A}_k \in \mathcal{M}_{n_k}(\mathbb{K})$. On en déduit donc directement que la matrice de u dans la base β s'écrit comme $\begin{pmatrix} \mathrm{A}_1 & & 0 \\ & \ddots & \\ 0 & & \mathrm{A}_r \end{pmatrix}$ avec $(\mathrm{A}_1, \dots, \mathrm{A}_r) \in \mathcal{M}_{n_1}(\mathbb{K}) \times \cdots \times \mathcal{M}_{n_r}(\mathbb{K})$.

→ (2) ⇒ (1) De la même manière, pour tout $k \in [\![1; r]\!]$, pour tout $e \in \beta_k$, étant donné la forme de la matrice, on voit que tous les coefficients de $u(e)$ dans $\beta \setminus \beta_k$ sont nuls, donc $u(e) \in \mathrm{Vect}(\mathrm{F}_k)$ et donc finalement $u(\mathrm{Vect}(\beta_k)) \subset \mathrm{F}_k$ *i.e.* $u(\mathrm{F}_k) \subset \mathrm{F}_k$.

(cc) BY-NC-SA

II Éléments propres

Définition II.1.

Soit $u \in \mathcal{L}(\mathrm{E})$. Un vecteur $x \in \mathrm{E}$ est appelé vecteur propre de u lorsque $x \neq 0$ et qu'il existe $\lambda \in \mathbb{K}$ tel que $u(x) = \lambda x$. Un tel scalaire λ est unique et est appelé valeur propre de u.

Proposition II.2.

Pour tout $\lambda \in \mathbb{K}$, l'équivalence suivante est vérifiée.

$$\lambda \text{ valeur propre de } u \Longleftrightarrow \operatorname{Ker}(u - \lambda \operatorname{Id}) \neq \{0\}.$$

Démonstration. La démonstration de ce résultat ne présente pas de difficulté et est laissée comme exercice au lecteur.

Vocabulaire. Lorsque $u \in \mathcal{L}(\mathrm{E})$ et $\lambda \in \mathbb{K}$, $\mathrm{E}_{\lambda,u} = \operatorname{Ker}(u - \lambda \operatorname{Id})$ est appelé espace propre de u associé à λ. Lorsqu'il n'y a pas ambiguïté sur l'endomorphisme associé à cet espace, on notera simplement $\mathrm{E}_{\lambda,u} = \mathrm{E}_\lambda$.

Proposition II.3.

Soit $u \in \mathcal{L}(\mathrm{E})$.

1. Soit $x \in \mathrm{E} \setminus \{0\}$. L'équivalence suivante est vérifiée.

$$x \text{ est un vecteur propre de } u \Longleftrightarrow \text{La droite } \mathbb{K}x \text{ est stable par } u.$$

2. Pour tout $x \in \mathrm{E}$ l'application $\tilde{u} : \begin{cases} \mathrm{E}_{\lambda,u} & \longrightarrow \mathrm{E}_{\lambda,u} \\ x & \longmapsto \lambda x \end{cases}$ commute avec tout élément de $\mathcal{L}(\mathrm{E}_{\lambda,u})$.

Démonstration. Même chose pour ces deux propositions.

Proposition II.4.

Si E est de dimension finie non réduit à $\{0\}$ et que $\mathbb{K} = \mathbb{C}$, alors u possède au moins une valeur propre.

Démonstration. Soit β une base de E. En posant $[u]_\beta = (a_{i,j})_{i,j \in [\![1;n]\!]}$, on a

$$\begin{aligned}
\lambda \text{ est une valeur propre de } u &\Longleftrightarrow \operatorname{Ker}(u - \lambda \operatorname{Id}) \neq \{0\} \\
&\Longleftrightarrow \det([u]_\beta - \lambda \mathrm{I}) = 0 \\
&\Longleftrightarrow \begin{vmatrix} a_{1,1} - \lambda & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - \lambda & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n-1} & a_{n,n} - \lambda \end{vmatrix} = 0.
\end{aligned}$$

L'application $\lambda \longmapsto \det([u]_\beta - \lambda \mathrm{I})$ est donc une application polynomiale dans $\mathbb{C}$ de degré n, elle admet donc au moins une racine λ, donc il existe au moins un $\lambda \in \mathbb{C}$ vérifiant $\operatorname{Ker}(u - \lambda \operatorname{Id}) \neq \{0\}$, *i.e.* λ est une valeur propre.

(cc) BY-NC-SA

Exercice II.5.

Supposons que $\mathbb{K}$ soit égal à $\mathbb{R}$ ou $\mathbb{C}$. Soit $n \geq 2$. Existe-t-il un plan P de $\mathcal{L}(\mathbb{K}^n)$ tel que $\mathrm{P} \setminus \{0\} \subset \mathrm{GL}(\mathbb{K})$?

Notation. Pour tout $u, v \in \mathcal{L}(\mathrm{E})$, on note $[u, v] = u \circ v - v \circ u$.

Proposition II.6.

Soit $u, v \in \mathcal{L}(\mathrm{E})$ tel que $[u, v] = 0$. Si λ est une valeur propre de u, alors v laisse stable $\mathrm{E}_{\lambda,u} = \mathrm{Ker}(u - \lambda\,\mathrm{Id})$.

Démonstration. On a $[u, v] = 0$, donc $[u - \lambda\,\mathrm{Id}, v] = 0$, *i.e.* $u - \lambda\,\mathrm{Id}$ et v commutent, donc d'après la proposition I.3, v laisse stable $\mathrm{Ker}(u - \lambda\,\mathrm{Id})$.

Exercice II.7.

On dit qu'une partie L de $\mathcal{L}(\mathbb{C}^n)$ est irréductible lorsque les seuls sous-espaces vectoriels de E qui sont stables par tous les éléments de L sont $\{0\}$ et E. Soit L une partie irréductible de $\mathcal{L}(\mathbb{C}^n)$ et u un élément de $\mathcal{L}(\mathbb{C}^n)$ qui commute avec tous les éléments de L. Montrer que u est une homothétie.

Proposition II.8.

Supposons que E est un espace vectoriel muni de $\|.\|$ et $u \in \mathcal{L}_c(\mathrm{E})$. Pour toute valeur propre λ de u, on a $|\lambda| \leq |||u|||$.

Démonstration. Soit λ une valeur propre de u et x un vecteur propre pour λ, *i.e.* $u(x) = \lambda x$ et $x \neq 0$. On a alors

$$|\lambda|\,\|x\| = \|u(x)\| \leq |||u||| \times \|x\|\,.$$

x est non nul, on peut donc simplifier par $\|x\|$ pour obtenir $|\lambda| \leq |||u|||$.

Remarque. Soit $u \in \mathcal{L}(\mathrm{E})$ et F un sous-espace vectoriel de E. Si F est stable par u, en considérant l'endomorphisme $v : \begin{cases} \mathrm{F} & \longrightarrow \mathrm{F} \\ x & \longmapsto u(x) \end{cases}$, on a $\mathrm{Ker}(v - \lambda\,\mathrm{Id}) = \mathrm{F} \cap \mathrm{Ker}(u - \lambda\,\mathrm{Id})$.

III Endomorphismes et matrices diagonalisables

Définition III.1.

On dit que $u \in \mathcal{L}(\mathrm{E})$ est diagonalisable lorsqu'il existe $r \in \mathbb{N}^*$ et $\lambda_1 \ldots \lambda_r \in \mathbb{K}$ tels que $\mathrm{E} = \mathrm{E}_{\lambda_1,u} + \cdots + \mathrm{E}_{\lambda_r,u}$.

Lemme III.2.

Soit $u \in \mathcal{L}(\mathrm{E})$, $p \in \mathbb{N}^*$, $\lambda_1, \ldots, \lambda_p$ des valeurs propres deux à deux distinctes de u et $x_1, \ldots, x_p \in \mathrm{E}$ des vecteurs propres respectivement associés à $\lambda_1, \ldots, \lambda_p$. La famille $(x_1, \ldots, x_p)$ est libre.

Démonstration. Nous allons montrer ce lemme par récurrence sur p.

BY-NC-SA

→ Pour le cas $p = 1$, $x_1$est un vecteur propre, il est donc non nul et alors la famille (x_1) est libre.

→ Soit $p \in \mathbb{N}^*$. Supposons que la propriété est vraie pour p. Soit $\mu_1, \dots, \mu_{p+1} \in \mathbb{K}$ tels que

$$\mu_1 x_1 + \cdots + \mu_{p+1} x_{p+1} = 0. \tag{17}$$

En appliquant u est deux côtés, on obtient

$$\mu_1 \lambda_1 x_1 + \cdots + \mu_{p+1} \lambda_{p+1} x_{p+1} = 0. \tag{18}$$

En faisant (17) $\times \lambda_{p+1}$ − (18), on obtient

$$\mu_1(\lambda_{p+1} - \lambda_1)x_1 + \cdots + \mu_p(\lambda_{p+1} - \lambda_p)x_p = 0.$$

Par hypothèse de récurrence, que la famille $(x_1, \dots, x_p)$ est libre, donc pour tout $k \in [\![1;p]\!]$, $\mu_k(\lambda_{p+1} - \lambda_k) = 0$. $\lambda_1, \dots, \lambda_{p+1}$ sont deux à deux distinctes par hypothèse donc pour tout $k \in [\![1;p]\!]$, $\mu_k = 0$ et alors (17) devient $\mu_{p+1}x_{p+1} = 0$, ce qui finalement implique que pour tout $k \in [\![1;p+1]\!]$, $\mu_k = 0$ *i.e.* $(x_1, \dots, x_{p+1})$ est libre, d'où le résultat voulu.

Application. La famille $(f_\alpha)_{\alpha\in\mathbb{R}} = (x \longmapsto x^\alpha)_{\alpha\in\mathbb{R}}$ est libre dans $\mathcal{C}^\infty(\mathbb{R}^*, \mathbb{R})$. En effet, en posant $u : f \longmapsto (x \longmapsto xf'(x))$, (f_α) est une famille de vecteurs propres associés à des valeurs propres deux à deux distinctes de u, elle est donc libre. On peut dire la même chose de la famille $(g_\beta)_{\beta\in\mathbb{C}}$, avec pour tout $\beta \in \mathbb{C}$, $g_\beta : x \longmapsto e^{\beta x}$.

Proposition III.3.

On suppose que E est de dimension finie. Soit $u \in \mathcal{L}(\mathrm{E})$. Les propositions suivantes sont équivalentes.

1. u est diagonalisable.
2. Il existe une base de E composée de vecteurs propres de u.
3. Il existe une base de E telle que $[u]_\beta$ est diagonale.
4. $\displaystyle\sum_{k=1}^{p} \mathrm{E}_{\lambda_k,u} = \mathrm{E}$, avec $\lambda_1, \dots, \lambda_p$ les valeurs propres de u.

Démonstration.

→ (1) ⇒ (2) Prenons $\lambda_1, \dots, \lambda_p$ comme dans la définition III.1. Quitte à éliminer les λ_i tels que $\mathrm{E}_{\lambda_i,u} = \{0\}$, on peut supposer sans perte de généralité que $\forall i \in [\![1;p]\!]$ $\mathrm{E}_{\lambda_i,u} \neq \{0\}$ et que les λ_i sont distincts. Le lemme III.2 nous permet de dire que toute famille de vecteurs $(x_1, \dots, x_p) \in \mathrm{E}_{\lambda_1,u} \times \cdots \times \mathrm{E}_{\lambda_p,u}$ est libre et que donc la somme $\mathrm{E} = \displaystyle\sum_{i=1}^{p} \mathrm{E}_{\lambda_i,u}$ est directe.

Chacun de ces espaces étant de dimension finie (non nulle), on peut alors prendre, pour chaque i, une base $(b_i^1, \dots, b_i^{p_i})$ de E_i avec $p_i = \dim \mathrm{E}_{\lambda_i,u}$. Il est clair que la concaténation de ces bases est une base de $\mathrm{E} = \displaystyle\bigoplus_{i=1}^{p} \mathrm{E}_{\lambda_i,u}$ et que chaque élément de cette dernière est un vecteur propre de u.

→ (2) ⇒ (3) Si $(b_1, \dots, b_n)$ est une base de E formée de vecteurs propres de u, et que pour tout $i \in [\![1;p]\!]$, b_i est associé à la valeur propre λ_i, alors la matrice de u dans cette base est la matrice diagonale $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

→ (3) ⇒ (4) Prenons une base $\beta = (\beta_1, \dots, \beta_n)$ qui rend $[u]_\beta$ diagonale. Il est alors clair que pour tout $i \in [\![1;p]\!]$, il existe $\lambda_i \in \mathbb{K}$ tel que β_i est un vecteur propre de u, associé à λ_i (les λ_i ne sont pas forcément distincts). On a alors en notant $\mathrm{VP}(u)$ l'ensemble (fini) des valeurs propres de u,

$$\mathrm{E} = \mathrm{Vect}(\beta_1, \dots, \beta_n) \subset \mathrm{Vect}(\mathrm{E}_{\lambda_1}, \dots, \mathrm{E}_{\lambda_n}) = \sum_{i=1}^{n} \mathrm{E}_{\lambda_i} \subset \sum_{\lambda\in\mathrm{VP}(u)} \mathrm{E}_\lambda,$$

(cc) BY-NC-SA

d'où le résultat.

→ (4) ⇒ (1) Cette implication ne présente pas de difficulté majeure.

Remarque 1. Si u possède $n = \dim \mathrm{E}$ valeurs propres différentes, alors u est diagonalisable et chaque espace propre est de dimension exactement une.
Ceci étant car, en notant $\lambda_1, \dots, \lambda_n$ ces n valeurs propres différentes, que

$$n \geq \dim \left(\sum_{i=1}^{n} \mathrm{E}_{\lambda_i} \right) \underset{\text{lemme III.2}}{=} \dim \left(\bigoplus_{i=1}^{n} \mathrm{E}_{\lambda_i} \right) = \sum_{i=1}^{n} \dim(\mathrm{E}_{\lambda_i}) \geq n.$$

On a alors $\dim \left(\bigoplus_{i=1}^{n} \mathrm{E}_{\lambda_i} \right) = n = \dim \mathrm{E}$, *i.e.* $\mathrm{E} = \bigoplus_{i=1}^{n} \mathrm{E}_{\lambda_i}$ et tous les $\mathrm{E}_{\lambda_i,u}$ ne peuvent pas avoir une dimension strictement supérieure à 1 ou égale à 0, ils sont donc de dimension 1. On a donc bien les deux résultats voulus.

Remarque 2. Lorsque u est diagonalisable, il est facile de retrouver son image et son noyau. En effet, considérons un base de vecteurs propres de u $\beta = (b_1, \dots, b_n)$ puis ordonnons la de manière à ce que les valeurs propres associées à $b_1, \dots, b_p$ soient non nulles et celles associées à $b_{p+1}, \dots, b_n$ soient nulles. Alors $\mathrm{Im}(u) = \mathrm{Vect}(b_1, \dots, b_p)$ et $\mathrm{Ker}(u) = \mathrm{Vect}(b_{p+1}, \dots, b_n)$.

Démonstration.
Une matrice diagonale à coefficients diagonaux tous non nuls dans $\mathbb{K}$ est inversible. Ainsi, dans la base β, $[u]_\beta$ s'écrit :

$$\begin{pmatrix} \mathrm{A} & 0_{\mathcal{M}_{p,n-p}(\mathbb{K})} \\ 0_{\mathcal{M}_{n-p,p}(\mathbb{K})} & 0_{\mathcal{M}_{n-p}(\mathbb{K})} \end{pmatrix},$$

avec A une matrice diagonale à coefficients diagonaux non nuls. On en déduit donc que

$$\mathrm{Ker}(u) = \left\{ a \in \mathrm{E},\ \exists v \in \mathbb{K}^{n-p},\ [a]_\beta = \begin{pmatrix} 0_{\mathbb{K}^p} \\ v \end{pmatrix} \right\} = \mathrm{Vect}(e_{p+1}, \dots, e_n),$$

$$\mathrm{Im}(u) = \left\{ a \in \mathrm{E},\ \exists v \in \mathbb{K}^{p},\ [a]_\beta = \begin{pmatrix} v \\ 0_{\mathbb{K}^{n-p}} \end{pmatrix} \right\} = \mathrm{Vect}(e_1, \dots, e_p).$$

Proposition III.4.

Soit $u \in \mathcal{L}(\mathrm{E})$, $w \in \mathrm{GL}(\mathrm{E})$ et $v = w \circ u \circ w^{-1}$.

1. $\lambda \in \mathbb{K}$ est une valeur propre de u si et seulement si c'est une valeur propre de v et $\mathrm{E}_{\lambda,v} = w(\mathrm{E}_{\lambda,u})$.
2. u est diagonalisable si et seulement si v l'est.

1. Soit $\lambda \in \mathbb{K}$. On a

$$\begin{aligned} x \in \mathrm{E}_{\lambda,v} &\iff w \circ u \circ w^{-1}(x) = \lambda x \\ &\iff u \circ w^{-1}(x) = \lambda w^{-1}(x) \\ &\iff w^{-1}(x) \in \mathrm{E}_{\lambda,u} \iff x \in w(\mathrm{E}_{\lambda,u}), \end{aligned}$$

d'où $\mathrm{E}_{\lambda,v} = w(\mathrm{E}_{\lambda,u})$ et donc

$$\lambda \in \mathrm{VP(v)} \iff \mathrm{E}_{\lambda,v} = w(\mathrm{E}_{\lambda,u}) \neq \{0\} \underset{w \in \mathrm{GL(E)}}{\iff} \mathrm{E}_{\lambda,u} \neq \{0\} \iff \lambda \in \mathrm{VP}(u).$$

(cc) BY-NC-SA

2. Soit $\lambda_1, \ldots, \lambda_r$ les valeurs propres de u et v. On a

$$u \text{ est diagonalisable } \iff \sum_{i=1}^{r} \mathrm{E}_{\lambda_i,u} = \mathrm{E}$$
$$\underset{w\in \mathrm{GL(E)}}{\iff} \sum_{i=1}^{r} w(\mathrm{E}_{\lambda_i,u}) = \sum_{i=1}^{r} \mathrm{E}_{\lambda_i,v} = \mathrm{E} \iff v \text{ est diagonalisable .}$$

Exercice III.5.

Soit $u \in \mathcal{L}(\mathrm{E})$ ayant n valeurs propres distinctes et v un endomorphisme de E qui commute avec u. Montrer que v est diagonalisable et que v est un polynôme en u.

Proposition III.6.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ et $\mathrm{E} = \mathbb{K}^n$. Les propriétés suivantes sont équivalentes.

1. $\exists \mathrm{P} \in \mathrm{GL}_n(\mathbb{K}),\ \exists \Delta \in \mathrm{D}_n(\mathbb{K}),\ \mathrm{P}^{-1}\mathrm{AP} = \Delta.$
2. $f_\mathrm{A} : \begin{cases} \mathbb{K}^n & \longrightarrow \mathbb{K}^n \\ \mathrm{X} & \longmapsto \mathrm{AX} \end{cases}$ est diagonalisable.
3. Il existe $u \in \mathcal{L}(\mathrm{E})$ et β une base de E tels que u diagonalisable et $[u]_\beta = \mathrm{A}$.
4. Pour tout $u \in \mathcal{L}(\mathrm{E})$ et pour tout β base de E, l'implication suivante est vérifiée.

$$[u]_\beta = \mathrm{A} \Longrightarrow u \text{ est diagonalisable.}$$

Démonstration.

→ (1) ⇒ (2) Soit β la base associée à la matrice P. Alors $[f_\mathrm{A}]_\beta = \mathrm{P}^{-1}\mathrm{AP} = \Delta$ est diagonale.

→ (2) ⇒ (3) Soit $\beta = (e_1, \ldots, e_n)$ une base de E et considérons l'application

$$u : \begin{cases} \mathrm{E} & \longrightarrow \mathrm{E} \\ \displaystyle\sum_{i=1}^{n} x_i e_i & \longmapsto \displaystyle\sum_{i=1}^{n} y_i e_i, \end{cases}$$

où $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = f_\mathrm{A}(\mathrm{X})$ avec $\mathrm{X} = (x_1, \ldots, x_n)^\mathrm{T}$. Une vérification rapide nous permet de voir que $[u]_\beta = \mathrm{A}$.

De plus, en considérant $\big((f_{1,i})_{i\in[\![1;n]\!]}, \ldots, (f_{n,i})_{i\in[\![1;n]\!]}\big)$ une base de vecteurs propres de f_A, il est facile de voir que la famille $(f_1, \ldots, f_n)$ telle que pour tout $k \in [\![1;n]\!]$, $f_k = f_{k,1}e_1 + \cdots + f_{k,n}e_n$ est une base de vecteurs propres de u, ce qui nous permet d'affirmer que u est diagonalisable.

→ (3) ⇒ (1) Soit $u \in \mathcal{L}(\mathrm{E})$ et β une base tels que $[u]_\beta = \mathrm{A}$. Soit γ une base de diagonalisation de u et P la matrice de passage de γ à β. On a alors

$$[u]_\gamma = \mathrm{P}[u]_\beta\mathrm{P}^{-1} = \mathrm{P}^{-1}\mathrm{AP}.$$

De plus, pour tout $i \in [\![1;n]\!]$, si e est le $i-$ème vecteur de la base canonique de $\mathcal{M}_n(\mathbb{K})$ et b_i est le $i-$ème vecteur de la base γ, alors il existe $\lambda_i \in \mathbb{K}$ tel que $u(b_i) = \lambda_i b_i$, *i.e.*, en passant aux matrices dans la base γ, $[u]_\gamma e_i = \lambda e_i$, donc il existe une matrice diagonale Δ telle que $[u]_\gamma = \Delta$. On en déduit donc finalement que $\mathrm{A} = \mathrm{P}^{-1}\Delta\mathrm{P}$.

→ (4) ⇔ (1) Laissé comme exercice au lecteur.

Vocabulaire. Lorsque A vérifie l'une de ces propriétés, on dit que A est diagonalisable.

CPGE paradise

(cc) BY-NC-SA

IV Action des polynômes

1. Rappels

L'algèbre des polynômes $\mathbb{K}[X]$ agit naturellement sur $\mathcal{L}(E)$. Plus exactement, si $P = \sum_0^p a_k X^k$ et $u \in \mathcal{L}(E)$ alors on définit

$$\phi_u : \begin{cases} \mathbb{K}[X] & \longrightarrow \mathcal{L}(E) \\ P & \longmapsto P(u) = \sum_{k=0}^{p} a_k u^k. \end{cases}$$

ϕ_u est alors un morphisme d'algèbre unitaire. En particulier, $I_u := \mathrm{Ker}(\phi_u)$ est un idéal.

Rappel IV.1.

Soit $(\mathcal{A}, +, .)$ un anneau commutatif. Un sous ensemble I de $\mathcal{A}$ est appelé idéal lorsque

- $\rightarrow$ $(I, +)$ est un sous groupe de $(\mathcal{A}, +)$.
- $\rightarrow$ Pour tout $x \in \mathcal{A}$ et $i \in I$, $ix \in I$.

I est dit principal lorsqu'il est engendré par un seul élément *i.e.* il existe $x \in \mathcal{A}$ tel que $I = \langle x \rangle := x\mathcal{A}$.
Un anneau commutatif $\mathcal{A}$ est dit principal si tout idéal de $\mathcal{A}$ est principal.
On notera en particulier que $\mathbb{K}[X]$ est principal lorsque $\mathbb{K}$ est un corps.

Si ce dernier est non trivial, chose toujours vraie en dimension finie, alors il est engendré par un unique polynôme unitaire μ_u, appelé polynôme minimal de u. Autrement dit si $P \in \mathbb{K}[X]$ et $P(u) = 0$, alors $P \in I_u$, c'est à dire qu'il existe $Q \in \mathbb{K}[X]$ tel que $P(X) = Q(X)\mu_u(X)$, *i.e.* $\mu_u | P$.

$\mathbb{K}[X]$ agit similairement sur les matrices carrées et commute avec les automorphismes intérieurs (*i.e.* les automorphismes de la forme $M \longmapsto PMP^{-1}$ avec $P \in GL_n(\mathbb{K})$) de conjugaison *i.e.* pour tout $P \in GL_n(\mathbb{K})$ et $Q \in \mathbb{K}[X]$, $Q(PAP^{-1}) = PQ(A)P^{-1}$.

Vocabulaire. Un élément de I_u, *i.e.* un polynôme P tel que $P(u) = 0_{\mathcal{L}(E)}$ est appelé un polynôme annulateur de u.

Proposition IV.2.

Soit x un vecteur propre de u de valeur propre associée λ. Les proposition suivantes sont vraies.

1. $\forall P \in \mathbb{K}[X],\ P(u)(x) = P(\lambda)x$.
2. Si $P \in I_u$ alors $P(\lambda) = 0$.
3. Soit $\lambda \in \mathbb{K}$. Alors $\mu_u(\lambda) = 0 \Longleftrightarrow \lambda$ valeur propre de u.
   En particulier, lorsque $\mathbb{K} = \mathbb{C}$, il existe $(\ell_\lambda)_{\lambda \in VP(u)} \in \mathbb{N}^{*VP(u)}$ telle que

$$\mu_u(X) = \prod_{\lambda \in VP(u)} (X - \lambda)^{\ell_\lambda}.$$

Démonstration.

1. On peut montrer par une récurrence rapide que $\forall k \geq 0\ u^k(x) = \lambda^k x$ (avec par convention $0^0 = 1$).
   Ainsi, si $P(X) = \sum_{k=0}^{d} a_k X^k$, alors $P(u)(x) = \sum_{k=0}^{d} a_k \cdot u^k(x) = \sum_{k=0}^{d} a_k \cdot \lambda^k x = P(\lambda)x$.
2. On sait que $0 = P(u)(x) = P(\lambda)x$ et donc $P(\lambda) = 0$ car $x \neq 0$.
3. Soit $\lambda \in \mathbb{K}$

(cc) BY-NC-SA

→ (⇐) Il suffit d'appliquer le point (2) pour $\mathrm{P} = \mu_u$.

→ (⇒) $\mu_u(\lambda) = 0$, on peut donc écrire $\mu_u(\mathrm{X}) = (\mathrm{X} - \lambda)\mathrm{Q}(\mathrm{X})$. On a alors

$$0 = \mu_u(u) = (u - \lambda \,\mathrm{Id}) \circ \mathrm{Q}(u).$$

Si λ n'est pas une valeur propre de u, la proposition II.2 nous permet d'affirmer que $u - \lambda \,\mathrm{Id}$ est inversible et que donc $\mathrm{Q}(u) = 0$, ce qui contredit la minimalité de μ_u. Donc λ est bien une valeur propre de u.

Exemple. Soit p un projecteur différent de 0 et Id. On a $\mu_p(\mathrm{X}) = \mathrm{X}^2 - \mathrm{X}$.

Exercice IV.3.

Soit $u \in \mathcal{L}(\mathrm{E})$ tel que $u \circ u = \mathrm{Id}$ et $u \neq \pm \,\mathrm{Id}$. Montrer que $\mu_u = \mathrm{X}^2 - 1$.

2. Décomposition des noyaux

Proposition (Lemme de décomposition des noyaux) IV.4.

Soit $\mathrm{P}_1, \ldots, \mathrm{P}_r$ des éléments de $\mathbb{K}[\mathrm{X}]$, deux à deux premiers entre eux et $u \in \mathcal{L}(\mathrm{E})$. On a $\operatorname{Ker} \mathrm{P}_1 \ldots \mathrm{P}_r(u) = \bigoplus_{i=1}^{r} \operatorname{Ker} \mathrm{P}_i(u)$. En particulier, en considérant le cas où $\mathrm{P}_i(\mathrm{X}) = \mathrm{X} - \lambda_i$ avec pour tout i, λ_i une valeur propre de u, on retrouve que les espaces propres sont en somme directe.

Démonstration. Montrons le résultat pour $r = 2$.
Soit P et Q deux polynômes premiers entre eux et soit $x \in \operatorname{Ker} \mathrm{P}(u) \cap \operatorname{Ker} \mathrm{Q}(u)$.
P et Q sont premiers entre eux et donc par Bézout, on dispose de $\mathrm{A}, \mathrm{B} \in \mathbb{K}[\mathrm{X}]$ tel que $\mathrm{AP} + \mathrm{BQ} = 1$.
On a alors $x = \mathrm{AP}(u)(x) + \mathrm{BQ}(u)(x) = 0 + 0 = 0$ et alors $\operatorname{Ker} \mathrm{P}(u) \cap \operatorname{Ker} \mathrm{Q}(u) = \{0\}$. $\operatorname{Ker} \mathrm{P}(u)$ et $\operatorname{Ker} \mathrm{Q}(u)$ sont donc bien en somme directe.

Montrons à présent que $\operatorname{Ker} \mathrm{PQ}(u) = \operatorname{Ker} \mathrm{P}(u) \oplus \operatorname{Ker} \mathrm{Q}(u)$.

→ (⊃) Si $x \in \operatorname{Ker} \mathrm{P}(u) \oplus \operatorname{Ker} \mathrm{Q}(u)$, alors en posant $x = x_{\operatorname{Ker} \mathrm{P}(u)} + x_{\operatorname{Ker} \mathrm{Q}(u)}$ on retrouve

$$\mathrm{PQ}(u)(x) = \mathrm{Q}(u) \circ \mathrm{P}(u) \left(x_{\operatorname{Ker} \mathrm{P}(u)}\right) + \mathrm{P}(u) \circ \mathrm{Q}(u) \left(x_{\operatorname{Ker} \mathrm{Q}(u)}\right) = 0.$$

→ (⊂) Soit $x \in \operatorname{Ker} \mathrm{PQ}(u)$. La relation obtenue via Bézout nous permet d'écrire

$$x = \underbrace{\mathrm{AP}(u)(x)}_{\in \operatorname{Ker} \mathrm{Q}(u)} + \underbrace{\mathrm{BQ}(u)(x)}_{\in \operatorname{Ker} \mathrm{P}(u)}.$$

Enfin, pour généraliser pour $r \geq 2$, on peut le faire simplement de la manière suivante. Soit $\mathrm{P}_1, \ldots, \mathrm{P}_r \in \mathbb{K}[\mathrm{X}]$ premiers entre eux. On a

$$\operatorname{Ker} \mathrm{P}_1 \ldots \mathrm{P}_r(u) = \operatorname{Ker} \mathrm{P}_1(u) \oplus \operatorname{Ker} \mathrm{P}_2 \ldots \mathrm{P}_r(u) = \cdots = \bigoplus_{i=1}^{r} \operatorname{Ker} \mathrm{P}_i(u).$$

(cc) BY-NC-SA

Proposition IV.5.

Soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ (non constant) irréductible et prenons l'unique $\ell \in \mathbb{N}$ tel que $\mathrm{P}^\ell | \mu_u$ et $\mathrm{P}^{\ell+1} \nmid \mu_u$, alors

$$\operatorname{Ker}(\mathrm{P}^0(u)) \subsetneq \operatorname{Ker}(\mathrm{P}^1(u)) \subsetneq \cdots \subsetneq \operatorname{Ker} \mathrm{P}^\ell(u) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u) = \ldots .$$

Démonstration. Les inclusions sont faciles à voir. On veut donc surtout prouver le fait que ces inclusions sont strictes. Supposons sans perte de généralité que P est unitaire.

Montrons que pour tout $k \in \mathbb{N}$, si $\operatorname{Ker} \mathrm{P}^k(u) = \operatorname{Ker} \mathrm{P}^{k+1}(u)$ alors $\forall i \geq k \;\; \operatorname{Ker} \mathrm{P}^k(u) = \operatorname{Ker} \mathrm{P}^i(u)$.

Soit $k \in \mathbb{N}$. Le résultat est vrai pour $i = k$ et $i = k+1$. Montrons donc le résultat voulu par récurrence. Supposons que le résultat est vrai pour $i-1$ avec $i \geq k+2$ et montrons qu'il est vrai pour i.
Soit $x \in \operatorname{Ker} \mathrm{P}^i(u)$. On a $\mathrm{P}^{i-1} \circ \mathrm{P}(u)(x) = 0$ et donc

$$\mathrm{P}(u)(x) \in \operatorname{Ker} \mathrm{P}^{i-1}(u) \underset{\substack{\uparrow \\ \text{hypothèse de récurrence}}}{=} \operatorname{Ker} \mathrm{P}^k(u).$$

Ainsi,

$$\mathrm{P}^{k+1}(u)(x) = \mathrm{P}^k \circ \mathrm{P}(u)(x) = 0,$$

et donc $x \in \operatorname{Ker} \mathrm{P}^{k+1}(u) = \operatorname{Ker} \mathrm{P}^k(u)$. En utilisant le résultat ci-dessus, il suffit de vérifier que $\operatorname{Ker} \mathrm{P}^{\ell-1}(u) \neq \operatorname{Ker} \mathrm{P}^\ell(u) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u)$ pour finir la démonstration.

$\rightarrow$ Montrons que $\operatorname{Ker} \mathrm{P}^\ell(u) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u)$.
Écrivons $\mu_u = \mathrm{P}^\ell \mathrm{Q}$ avec $\mathrm{P} \wedge \mathrm{Q} = 1$ (P est irréductible). On a alors, d'après le lemme des noyaux,

$$\mathrm{E} = \operatorname{Ker}(\mu_u(u)) = \operatorname{Ker}(\mathrm{P}^\ell \mathrm{Q}(u)) = \operatorname{Ker} \mathrm{P}^\ell(u) \oplus \operatorname{Ker} \mathrm{Q}(u)$$

Similairement, étant donné que $\mathrm{P}^{\ell+1}\mathrm{Q}(u) = \mathrm{P}\mu_u(u) = 0$, par le lemme des noyaux

$$\mathrm{E} = \operatorname{Ker}(\mathrm{P}^{\ell+1}(u) \circ \mathrm{Q}(u)) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u) \oplus \operatorname{Ker} \mathrm{Q}(u).$$

Ainsi,

$$\operatorname{Ker} \mathrm{P}^\ell(u) \oplus \operatorname{Ker} \mathrm{Q}(u) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u) \oplus \operatorname{Ker} \mathrm{Q}(u),$$

ce qui donne $\dim \operatorname{Ker} \mathrm{P}^{\ell+1}(u) = \dim \operatorname{Ker} \mathrm{P}^\ell(u)$. Étant donné que $\operatorname{Ker} \mathrm{P}^\ell(u) \subset \operatorname{Ker} \mathrm{P}^{\ell+1}(u)$, on obtient $\operatorname{Ker} \mathrm{P}^\ell(u) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u)$.

$\rightarrow$ Montrons que $\operatorname{Ker} \mathrm{P}^{\ell-1}(u) \neq \operatorname{Ker} \mathrm{P}^\ell(u)$.

Supposons par l'absurde que $\operatorname{Ker} \mathrm{P}^{\ell-1}(u) = \operatorname{Ker} \mathrm{P}^\ell(u)$. Alors

$$\mathrm{E} = \operatorname{Ker} \mathrm{P}^\ell(u) \oplus \operatorname{Ker} \mathrm{Q}(u) = \operatorname{Ker} \mathrm{P}^{\ell-1}(u) \oplus \operatorname{Ker} \mathrm{Q}(u) = \operatorname{Ker}(\mathrm{P}^{\ell-1}\mathrm{Q}(u)),$$

et donc $\mathrm{P}^{\ell-1}\mathrm{Q}(u) = 0$ *i.e.* $\mathrm{P}^\ell \mathrm{Q} = \mu_u | \mathrm{P}^{\ell-1}\mathrm{Q}$, ce qui est faux par hypothèse.

Remarque. En décomposant P en irréductibles, on obtient comme conséquence la généralisation suivante :
Soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ non constant et $\ell \in \mathbb{N}$ le plus petit entier tel que $\mathrm{P}^\ell \wedge \mu_u = \mathrm{P}^{\ell+1} \wedge \mu_u$, alors

$$\operatorname{Ker}(\mathrm{P}^0(u)) \subsetneq \operatorname{Ker}(\mathrm{P}^1(u)) \subsetneq \cdots \subsetneq \operatorname{Ker} \mathrm{P}^\ell(u) = \operatorname{Ker} \mathrm{P}^{\ell+1}(u) = \ldots$$

(cc) BY-NC-SA

Proposition IV.6.

Soit $u \in \mathcal{L}(\mathrm{E})$ avec $\dim \mathrm{E} < +\infty$. Les propriétés suivantes sont équivalentes.

1. u est diagonalisable.
2. u est annulé par un polynôme scindé à racines simples (dans $\mathbb{K}$).
3. μ_u est scindé à racines simples.

Bien entendu, on possède un théorème similaire pour les matrices.

Proposition IV.7.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$. Les propriétés suivantes sont équivalentes.

1. A est diagonalisable.
2. A est annulé par un polynôme scindé à racines simples (dans $\mathbb{K}$).
3. μ_{A} est scindé à racines simples.

→ (1) ⇒ (2) A est diagonalisable, il existe donc $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tel que que A est semblable à $\mathrm{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Notons $\delta_1, \ldots, \delta_r$ tous les scalaires λ_i en enlevant les doublons et considérons le polynôme scindé à racines simples $\mathrm{P}(x) = \prod_{i=1}^{r}(\mathrm{X} - \delta_i)$. On peut facilement vérifier que

$$\mathrm{P}(\mathrm{D}) = \mathrm{diag}(\mathrm{P}(\lambda_1), \ldots, \mathrm{P}(\lambda_n)) = 0.$$

→ (2) ⇒ (3) Soit P un polynôme scindé à racines simples annulant A. On sait que $\mu_{\mathrm{A}} | \mathrm{P}$ ($\mathrm{P} \neq 0$ et μ_{A} non constant), donc μ_{A} est bien scindé à racines simples.

→ (3) ⇒ (1) Écrivons $\mu_{\mathrm{A}} = \prod_{i=1}^{r}(\mathrm{X} - \lambda_i)$ avec les λ_i distincts. On a alors d'après le lemme de décomposition des noyaux

$$\mathrm{E} = \mathrm{Ker}\, \mu_u(u) = \bigoplus_{i=1}^{r} \mathrm{Ker}(u - \lambda_i \,\mathrm{Id}).$$

Considérons pour tout i, β_i une base de $\mathrm{Ker}(u - \lambda_i \,\mathrm{Id})$ puis posons $\beta = (\beta_1, \ldots, \beta_r)$ la concaténation de ces bases. β est une base de E. On a alors $[\mathrm{A}]_\beta = \begin{pmatrix} \lambda_1 \mathrm{I}_{\ell_1} & & \\ & \ddots & \\ & & \lambda_r \mathrm{I}_{\ell_r} \end{pmatrix}$ et donc A est bien diagonalisable.

Exemple. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ tel que $\mathrm{A}^m = \mathrm{I}_n$ alors A est diagonalisable étant donné que

$$\mathrm{X}^m - 1 = \prod_{\omega \in \mathrm{U}_m} (\mathrm{X} - \omega)$$

est un polynôme annulateur de u scindé à racines simples sur $\mathbb{C}$.

En pratique, pour des corps algébriquement clos tels que $\mathbb{C}$, tout polynôme est scindé, donc dire qu'un polynôme P non constant est scindé à racines simples est équivalent à ce que $\mathrm{P} \wedge \mathrm{P}' = 1$ ou à ne pas avoir de racines multiples *i.e.* ne pas avoir de racine commune à P et P'. Dans l'exemple précédant, si $m \geq 2$, alors $\mathrm{Z}(m\mathrm{X}^{m-1}) = \{0\}$ et 0 n'est pas racine de X^{m-1}. De même,

$$\mathrm{X}^m - 1 \wedge m\mathrm{X}^{m-1} = (\mathrm{X}^m - 1) \wedge \mathrm{X}^{m-1} = (\mathrm{X}^m - 1 - \mathrm{X} \times \mathrm{X}^{m-1}) \wedge \mathrm{X}^{m-1} = 1.$$

BY-NC-SA

Finalement, la notion de diagonalisabilité passe aux sous-espaces. Plus exactement :

Proposition IV.8.

Soit $u \in \mathcal{L}(\mathrm{E})$ diagonalisable et F un sous-espace vectoriel de E. Si u stabilise F (*i.e.* $u(\mathrm{F}) \subset \mathrm{F}$) alors $v := u|_{\mathrm{F}}^{\mathrm{F}}$ est diagonalisable.

Démonstration. Soit P un polynôme scindé à racines simples tel que $\mathrm{P}(u) = 0$. On a alors $\mathrm{P}(v) = 0$, v est annulé par un polynôme scindé à racines simples, c'est donc un endomorphisme diagonalisable.

Remarque. Pour les endomorphismes diagonalisables, les espaces stables et

$$\mathrm{Com}(u) = \{v \in \mathcal{L}(\mathrm{E}),\ v \circ u = u \circ v\}$$

sont faciles à caractériser. On peut voir cela à partir des deux exercices ci-dessous.

Exercice IV.9.

Soit $u \in \mathcal{L}(\mathrm{E})$ diagonalisable de valeurs propres $\lambda_1, \dots, \lambda_s$ distinctes. Montrer que les propositions suivantes sont équivalentes.

1. $\mathrm{F} \subset \mathrm{E}$ est un sous-espace vectoriel de E stable par u.
2. $\mathrm{F} = \bigoplus_{i=1}^{s} \mathrm{F} \cap \mathrm{E}_{\lambda_i,u}$.
3. Il existe une famille $(\mathrm{F}_i)_{i\in[\![1;s]\!]}$ de sous-espaces vectoriels de E telle que pour tout $i \in [\![1;s]\!]$ $\mathrm{F}_i \subset \mathrm{E}_{\lambda_i,u}$ et $\mathrm{F} = \bigoplus_{i=1}^{s} \mathrm{F}_i$.

Exercice IV.10.

Soit $u \in \mathcal{L}(\mathrm{E})$ diagonalisable de valeurs propres $\lambda_1, \dots, \lambda_s$ distinctes et soit β une base de E telle que $[u]_\beta = \begin{pmatrix} \lambda_1 \mathrm{I}_{\ell_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_s \mathrm{I}_{\ell_s} \end{pmatrix}$, avec $(\ell_1, \dots, \ell_s) \in [\![1;n]\!]$ tel que $\ell_1 + \dots \ell_s = n$.
Montrer que

$$\mathrm{Com}(u) = \left\{ v \in \mathcal{L}(\mathrm{E}),\ \exists(\mathrm{A}_1, \dots, \mathrm{A}_s) \in \mathcal{M}_{\ell_1}(\mathbb{K}) \times \dots \times \mathcal{M}_{\ell_s}(\mathbb{K}),\ [v]_\beta = \begin{pmatrix} \mathrm{A}_1 & & 0 \\ & \ddots & \\ 0 & & \mathrm{A}_s \end{pmatrix} \right\}.$$

En déduire une autre solution de l'exercice III.5.

(cc) BY-NC-SA

3. Compléments : projecteurs spectraux et codiagonalisation

Proposition IV.11.

Soit $u \in \mathcal{L}(\mathrm{E})$ diagonalisable de valeurs propres $\lambda_1, \dots, \lambda_r$ distinctes. Les propositions suivantes sont vraies.

1. $\mu_u = \prod_{i=1}^{r}(\mathrm{X} - \lambda_i)$.
2. Soit (L_i) les interpolateurs de Lagrange des (λ_i), *i.e.* pour tout i, L_i est l'unique polynôme de degré $r-1$ tel que pour tout $k \neq i$ $\mathrm{L}_i(\lambda_k) = 0$ et $\mathrm{L}_i(\lambda_i) = 1$. On a $\mathrm{L}_1 + \cdots + \mathrm{L}_r = 1$ et $\forall i \neq j$, $\mu_u | \mathrm{L}_i\mathrm{L}_j$.
3. En posant $p_i = \mathrm{L}_i(u)$. Alors p_i est le projecteur sur $\mathrm{E}_{\lambda_i,u}$ parallèlement à $\bigoplus_{j\neq i} \mathrm{E}_{\lambda_j,u}$.

1. La proprosition IV.5 donne le fait que μ_u est scindé à racines simples car u est diagonalisable. Ainsi, il existe $\delta_1, \dots, \delta_r \in \mathrm{K}$ deux à deux distincts tels que $\mu_u = \prod_{i=1}^{s}(\mathrm{X} - \lambda_i)$ où les λ_i sont distincts. De plus, d'après la proposition IV.2
$$\forall \lambda \in \mathbb{K},\ \mu_u(\lambda) = 0 \iff \lambda \in \mathrm{VP}(u),$$
et donc finalement $\mu_u = \prod_{i=1}^{r}(\mathrm{X} - \lambda_i)$
2. Posons $\mathrm{P} = \mathrm{L}_1 + \cdots + \mathrm{L}_r - 1$. Alors pour tout $i \in [\![1;r]\!]$ $\mathrm{P}(\lambda_i) = 1 - 1 = 0$. Ainsi, étant donne que $\mu_u(\mathrm{X}) = \prod_{i=1}^{s}(\mathrm{X} - \lambda_i)$, $\mu_u|\mathrm{P}$ et $\deg \mathrm{P} \leq r-1$ et alors $\mathrm{P} = 0$ *i.e.* $\mathrm{L}_1 + \cdots + \mathrm{L}_r = 1$. Finalement, pour tous $i \neq j$ et tout $k \in [\![1;r]\!]$, $\mathrm{L}_i\mathrm{L}_j(\lambda_k) = 0$ et donc pour tous $i \neq j$ et tout $k \in [\![1;r]\!]$, $(\mathrm{X} - \lambda_k)|\mathrm{L}_i\mathrm{L}_j$ *i.e.* $\mu_u|\mathrm{L}_i\mathrm{L}_j$ si $i \neq j$.
3. Soit $j \in [\![1;r]\!]$ et $x \in \mathrm{E}_{\lambda_j,u}$. On a pour tout $i \in [\![1;r]\!]$, $\mathrm{L}_i(u)(x) = \mathrm{L}_i(\lambda_j)x = \delta_{i,j}x$. En considérant pour tout $i \in [\![1;r]\!]$, β_i une base de $\mathrm{E}_{\lambda_i,u}$, on voit que pour tout $j \in [\![1;r]\!]$, pour tout $i \in [\![1;r]\!]$ et $e \in \beta_j$, $\mathrm{L}_i(u)(e) = e$ si $i = j$ et 0 sinon. On a aussi clairement $\mathrm{L}_i(u) \circ \mathrm{L}_i(u) = \mathrm{L}_i(u)$, d'où le résultat.

Exercice IV.12.

Soit $\mathrm{S} \subset \mathcal{L}(\mathrm{E})$ un ensemble non vide d'endomorphismes diagonalisables de E. Montrer que les éléments de S sont codiagonalisables (c'est-à-dire que leurs matrices sont toutes diagonales dans une même base) si et seulement s'ils commutent tous deux à deux, *i.e.*
$$\forall (u,v) \in \mathrm{S}^2,\ u \circ v = v \circ u.$$

Exercice IV.13.

Soit $\mathbb{K}$ un corps de caractéristique différente de 2 et $m, n \geq 1$. Montrer que s'il existe $\varphi : \mathrm{GL}_m(\mathbb{K}) \longrightarrow \mathrm{GL}_n(\mathbb{K})$ morphisme de groupes injectif alors $m \leq n$. En particulier, si $\mathrm{GL}_m(\mathbb{K})$ et $\mathrm{GL}_n(\mathbb{K})$ sont isomorphes alors $m = n$.

Exercice IV.14.

Dans cet exercice, on suppose que $\mathbb{K} = \mathbb{C}$. Soit $u \in \mathcal{L}(\mathrm{E})$ tel que u^2 soit diagonalisable. Montrer que u est diagonalisable si et seulement si $\operatorname{Ker} u = \operatorname{Ker} u^2$.

(cc) BY-NC-SA

V Polynôme caractéristique

1. Généralités

Soit E un espace vectoriel **de dimension finie** $n \geq 1$ et $u \in \mathcal{L}(\mathrm{E})$.

Proposition V.1.

1. Soit $\lambda \in \mathbb{K}$. λ est une valeur propre de u si et seulement si $\det(u - \lambda \,\mathrm{Id}) = 0$.
2. Si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$ semblables, alors $\det(\mathrm{A} - \mathrm{XI}_n) = \det(\mathrm{B} - \mathrm{XI}_n)$. En particulier, la quantité $\det([u]_\beta - \mathrm{XI}_n)$ est indépendante de la base β. On appelle alors polynôme caractéristique de u le polynôme

$$\chi_u(\mathrm{X}) = \det(\mathrm{XI}_n - [u]_\beta) = (-1)^n \det([u]_\beta - \mathrm{XI}_n).$$

Démonstration.

1. $\lambda \in \mathrm{VP}(u) \Longleftrightarrow \mathrm{Ker}(u - \lambda \,\mathrm{Id}) \neq \{0\} \underset{\text{dimension finie}}{\Longleftrightarrow} u - \lambda \,\mathrm{Id} \notin \mathrm{GL(E)} \Longleftrightarrow \det(u - \lambda \,\mathrm{Id}) = 0$.
2. $\det(\mathrm{PAP}^{-1} - \mathrm{X}\,\mathrm{Id}) = \det(\mathrm{P}(\mathrm{A} - \lambda \,\mathrm{Id})\mathrm{P}^{-1}) = \det(\mathrm{P})\det(\mathrm{P}^{-1})\det(\mathrm{A} - \mathrm{X}\,\mathrm{Id}) = \det(\mathrm{A} - \mathrm{X}\,\mathrm{Id})$.

Remarque. le point 1 se reformule de la manière suivante.

$$\lambda \in \mathbb{K} \text{ est une valeur propre de } u \Longleftrightarrow \chi_u(\lambda) = 0.$$

En particulier, lorsque $\mathbb{K} = \mathbb{C}$, $\chi_u(\mathrm{X}) = \displaystyle\prod_{\delta \in \mathrm{VP}(u)} (\mathrm{X} - \delta)^{\ell_\delta}$ où $\forall \delta \in \mathrm{VP}(u)$, $\ell_\delta \geq 1$.

Attention : χ_u peut avoir des racines non contenues dans $\mathbb{K}$. En effet, si on considère une matrice de rotation d'angle $\frac{\pi}{2}$, $\mathrm{M} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, on a $\chi_\mathrm{M}(\mathrm{X}) = \mathrm{X}^2 + 1$. Ce polynôme admet deux racines complexes mais aucune racine réelle.

Astuce utile : Un calcul explicite nous permet d'affirmer que

$$\det(\mathrm{XI}_n - \mathrm{A}) = \mathrm{X}^n - tr(\mathrm{A})\mathrm{X}^{n-1} + \cdots + (-1)^n \det \mathrm{A}.$$

En particulier, il est bon de retenir les formules suivantes

→ Lorsque $n = 2$, $\chi_\mathrm{A}(\mathrm{X}) = \mathrm{X}^2 - \mathrm{Tr(A)X} + \det \mathrm{A}$.

→ Lorsque $n = 3$, $\chi_\mathrm{A}(\mathrm{X}) = \mathrm{X}^3 - \mathrm{Tr(A)X}^2 + \mathrm{C}_2(\mathrm{A})\mathrm{X} - \det \mathrm{A}$
où $\mathrm{C}_2(\mathrm{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{12} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$ lorsque $\mathrm{A} = (a_{i,j})_{i,j \in [\![1;3]\!]}$.

Notation. Pour tout $u \in \mathcal{L}(\mathrm{E})$, on désigne par $\mathrm{Spec}_\mathbb{K}(u)$ la liste non ordonnée des valeurs propres de u qui sont contenues dans $\mathbb{K}$ en prenant compte de leurs multiplicités dans χ_u, c'est-à-dire que $\mathrm{Spec}_\mathbb{K}(u)$ peut être vu comme un ensemble mais contrairement à un ensemble usuel, un élément peut être compté plus d'une fois et contrairement aux $k-$uplets l'ordre n'est pas pris en compte. Par exemple, lorsque $\mathrm{Spec}_\mathbb{K}(u)$ contient deux fois 1 et une fois 2, nous noterons $\mathrm{Spec}_\mathbb{K}(u) = [1, 1, 2]$.
Parfois, par abus de notation, on pourra voir $\mathrm{Spec}_\mathbb{K}(u)$ comme un ensemble normal aussi.

Exemple. Si $\mathrm{A} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ (pas forcément distincts) alors

$$\chi_\mathrm{A} = \prod_{i=1}^{n} (\mathrm{X} - \lambda_i) \text{ et } \mathrm{Spec}_\mathbb{K}(\mathrm{A}) = [\lambda_1, \ldots, \lambda_n].$$

Attention : En général, $\mathrm{Spec}_\mathbb{R}(u) \subsetneq \mathrm{Spec}_\mathbb{C}(u)$.

CPGE paradise
(cc) BY-NC-SA

Proposition V.2.

Lorsque $\mathbb{K} = \mathbb{C}$, pour tout $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, si $\mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) = [\lambda_1, \dots, \lambda_n]$ alors $\mathrm{Tr}(\mathrm{A}) = \lambda_1 + \dots + \lambda_n$ et $\det \mathrm{A} = \lambda_1 \dots \lambda_n$.

Démonstration. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$. On a

$$\chi_{\mathrm{A}}(\mathrm{X}) = \det(\mathrm{XI}_n - \mathrm{A}) = \prod_{i=1}^{n} (\mathrm{X} - \lambda_i) = \mathrm{X}^n - \mathrm{Tr}(\mathrm{A})\mathrm{X}^{n-1} + \dots + (-1)^n \det(\mathrm{A}).$$

En identifiant les coefficients de l'avant-dernier terme et du dernier terme, on trouve bien que $\mathrm{Tr}(\mathrm{A}) = \lambda_1 + \dots + \lambda_n$ et $\det \mathrm{A} = \lambda_1 \dots \lambda_n$.

Exemples.

1. Si $\mathrm{A} = (a_{i,j})_{i,j \in [\![1;n]\!]}$ est une matrice triangulaire supérieure (resp. inférieure), alors

$$\chi_{\mathrm{A}}(\mathrm{X}) = \prod_{i=1}^{n} (\mathrm{X} - a_{ii}).$$

2. Soit $\mathrm{P}(\mathrm{X}) = \mathrm{X}^n + a_{n-1}\mathrm{X}^{n-1} + \dots + a_0 \in \mathbb{K}[\mathrm{X}]$ et $\mathrm{C}_{\mathrm{P}} = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -a_{n-2} \\ & & 1 & -a_{n-1} \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$.

   On a $\chi_{\mathrm{C}_{\mathrm{P}}} = \mathrm{P}$. On appelle C_{P} la matrice compagnon de P.

La démonstration (classique) de ces deux points se fait par récurrence en développant suivant la première colonne.

Exercice V.3.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$. En utilisant la base $(\mathrm{E}_{k,\ell})_{k,j \in [\![1;n]\!]} = ((\delta_{i,k}\delta_{j,\ell})_{i,j \in [\![1;n]\!]})$, trouver le polynôme caractéristique de

$$\phi_{\mathrm{A}} : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ \mathrm{X} & \longmapsto \mathrm{XA}. \end{cases}$$

Proposition V.4.

Soit F un sous-espace vectoriel de E stable par u. Posons $v = u\Big|_{\mathrm{F}}^{\mathrm{F}}$. Montrer que $\chi_v | \chi_u$.

Démonstration. Considérons β_1 une base de F qu'on complète en une base $\beta = (\beta_1, \beta_2)$ de E. On a alors

$$[\mathrm{X}\,\mathrm{Id} - u]_\beta = \begin{pmatrix} [\mathrm{X}\,\mathrm{Id} - v]_{\beta_1} & * \\ 0 & \mathrm{X}\,\mathrm{Id} - \mathrm{B} \end{pmatrix},$$

où $\mathrm{B} \in \mathcal{M}_{|\beta_2|}(\mathbb{K})$. On en déduit alors que $\chi_u(\mathrm{X}) = \chi_v(\mathrm{X})\chi_{\mathrm{B}}(\mathrm{X})$ (on rappelle que le déterminant d'une matrice triangulaire supérieure par bloc est le produit des déterminants des blocs diagonaux). Ceci donne bien le résultat voulu.

2. Caractérisation des endomorphismes diagonalisables

Dans cette sous-partie, nous allons présenter des moyens de dire si un endomorphisme est diagonalisable ou non.

(cc) BY-NC-SA

Définition V.5.

Soit $u \in \mathcal{L}(\mathrm{E})$ et $\lambda \in \mathrm{Spec}_{\mathbb{K}}(u)$.

1. La multiplicité algébrique de λ est sa multiplicité en tant que racine de χ_u, notée $\alpha_u(\lambda)$.
2. La multiplicité géométrique de λ est $\beta_u(\lambda) := \dim \mathrm{E}_{\lambda,u}$.

S'il n'y a pas d'ambiguïté sur l'endomorphisme u, on pourra les noter $\alpha(\lambda)$ (ou α_λ) et $\beta(\lambda)$ (ou β_λ) respectivement.

Proposition V.6.

Soit $u \in \mathcal{L}(\mathrm{E})$. Les propositions suivantes sont vraies.

1. Pour toute valeur propre λ de u, $\alpha_u(\lambda) \geq \beta_u(\lambda)$.
2. u est diagonalisable $\Longleftrightarrow$ χ_u scindé et $\forall \lambda \in \mathrm{Spec}_{\mathbb{K}}(u),\ \alpha_u(\lambda) = \beta_u(\lambda)$.

Démonstration.

1. $\mathrm{F} = \mathrm{Ker}(u - \lambda\,\mathrm{Id})$ est stable par u. Considérons $v = u\Big|_{\mathrm{F}}^{\mathrm{F}}$. $\chi_v = (\mathrm{X} - \lambda\,\mathrm{Id})^{\beta(\lambda)} | \chi_u$ d'où $\beta_\lambda \leq \alpha_\lambda$.
2. Soit $u \in \mathcal{L}(\mathrm{E})$.

   → ($\Rightarrow$) Soit β une base de E et $\lambda_1, \ldots, \lambda_r$ deux à deux distincts tels que $[u]_\beta = \begin{pmatrix} \lambda_1 \mathrm{I}_{\ell_1} & & \\ & \ddots & \\ & & \ell b_r \mathrm{I}_{\ell_r} \end{pmatrix}$ avec $\lambda_1, \ldots, \lambda_r \in \mathbb{N}^*$. On a alors

$$\chi_u(\mathrm{X}) = \det(\mathrm{X}\mathrm{I}_n - [u]_\beta) = \det \begin{pmatrix} (\mathrm{X}-\lambda_1)\mathrm{I}_{\ell_1} & & \\ & \ddots & \\ & & (\mathrm{X}-\lambda_n)\mathrm{I}_{\ell_r} \end{pmatrix}$$
$$= \prod_{i=1}^{r} (\mathrm{X} - \lambda_i)^{\ell_i}.$$

   De là, on voit bien que χ_u et scindé et pour tout $i \in [\![1;r]\!]$, $\ell_i = \alpha_u(\lambda_i) = \beta_u(\lambda_i)$.

   → ($\Leftarrow$) On a

$$n = \sum_{\lambda \in \mathrm{VP}(u)} \alpha_u(\lambda) = \sum_{\lambda \in \mathrm{VP}(u)} \beta_u(\lambda) = \dim \left(\bigoplus_{\lambda \in \mathrm{VP}(u)} \mathrm{E}_{\lambda,u} \right).$$

   On en déduit donc que $\mathrm{E} = \bigoplus_{\lambda \in \mathrm{VP}(u)} \mathrm{E}_{\lambda,u}$, *i.e.* u est diagonalisable.

Remarque. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$. Si χ_A possède n racines distinctes dans $\mathbb{K}$ alors

→ A est diagonalisable dans $\mathcal{M}_n(\mathbb{K})$.

→ χ_A est scindé à racines simples.

→ $\forall \lambda \in \mathrm{Spec}_{\mathbb{K}}(\mathrm{A}),\ \alpha_\lambda = \beta_\lambda = 1$.

Remarque. En général, pour deux matrices $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{K})$, $\chi_\mathrm{A} = \chi_\mathrm{B}$ n'implique pas forcément que A est semblable à B. Toutefois, si A et B sont toutes les deux diagonalisables, alors l'implication devient vraie. Ceci car, dans ce cas,

$$\mathrm{Spec}_{\mathbb{K}}(\mathrm{A}) = \mathrm{Spec}_{\mathbb{K}}(\mathrm{B}) = [\lambda_1, \ldots, \lambda_n],$$

donc A et B sont toutes deux semblables à $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ et alors A et B sont donc semblables par transitivité de la relation d'équivalence matricielle. Notons aussi que le fait d'avoir la diagonalisabilité

BY-NC-SA

pour uniquement une seule d'elles ne suffit pas. En effet, pour le voir, il suffit de prendre $A = 0$ et B une matrice nilpotente non nulle quelconque (par exemple strictement triangulaire supérieure non nulle), ces deux matrices admettent X^n comme polynôme caractéristique mais ne sont clairement pas équivalentes.

Le théorème suivant est l'un des plus utiles en algèbre linéaire

Théorème (Cayley-Hamilton) V.7.

Pour tout $A \in \mathcal{M}_n(\mathbb{K})$, $\chi_A(A) = 0$.

Remarque. Le théorème ci-dessus est équivalent à dire que pour tout $A \in \mathcal{M}_n(\mathbb{K})$, $\mu_A | \chi_A$.

Démonstration. Plusieurs démonstrations sont possibles, on rencontrera notamment plusieurs dans ce cours. La démonstration présentée ici une basée sur le changement de corps. On considérera dans cette démonstration que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et on posera $\mathbb{L} := \mathbb{C}$. Noter que cette démonstration est généralisable à d'autres corps si on connaît de théorie d'extension de corps.

Remarquons d'abord que χ_A est indépendant du corps où on le calcule. On peut donc travailler dans $M_n(\mathbb{L})$ et montrer que dans $\mathbb{L}$, $\chi_A(A) = 0$. Montrons donc le résultat pour les matrices dans $M_n(\mathbb{L})$.

Dans $\mathbb{L}$, $\mu_{A,\mathbb{L}}$ (le polynôme minimal de u dans $\mathbb{L}[X]$) et χ_A sont scindés (car $\mathbb{L}$ est algébriquement clos, *i.e.* tous les polynômes de $\mathbb{L}[X]$ sont scindés) et leurs racines sont exactement les mêmes (il s'agit des valeurs propres de A dans $\mathbb{L}$). Notons alors $\lambda_1, \dots, \lambda_s$ les valeurs propres distinctes de A. On a alors

$$\mu_{A,\mathbb{L}} = \prod_{i=1}^{s} (X - \lambda_i)^{\gamma_i} \text{ et } \chi_A = \prod_{i=1}^{s} (X - \lambda_i)^{\alpha_i}.$$

Il suffit maintenant de montrer que pour tout i $\gamma_i \leq \alpha_i$, car une fois cela fait, on aurait $\mu_{A,\mathbb{L}} | \chi_A$ et donc $\chi_A(A) = 0$. Par symétrie des rôles des γ_i, il suffit de montrer que $\gamma_1 \leq \alpha_1$. Soit $u \in \mathcal{L}(\mathbb{L}^n)$ l'endomorphisme associé à A dans la base canonique. De même, E désignera désormais $\mathbb{L}^n$. Posons

$$\begin{aligned} F &= \operatorname{Ker}(u - \lambda_1 \operatorname{Id})^{\gamma_1} & H &= \bigoplus_{i=2}^{s} \operatorname{Ker}(u - \lambda_i \operatorname{Id})^{\gamma_i} \\ d &= \dim(F) & v &= u\Big|_F^F. \end{aligned}$$

Remarquons que v est bien défini car F est stable par u. Soit β_1 une base de F et β_2 une base de H. $\beta = (\beta_1, \beta_2)$ est une base de E et

$$[u]_\beta = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}.$$

Soit μ_B et μ_C les polynômes minimaux de B et C respectivement. Remarquons tout d'abord que B est la matrice de v dans la base β.

v est annulé par $(X - \lambda_1)^{\gamma_1}$, donc $\mu_B | (X - \lambda_1)^{\gamma_1}$, *i.e.* il existe $p \in [\![1; \gamma_1]\!]$ tel que $\mu_B = (X - \lambda_1)^p$. De même, $\prod_{i=2}^{s} (X - \lambda_i)^{\gamma_i}$ annule C, donc $\mu_C \,\Big|\, \prod_{i=2}^{s} (X - \lambda_i)^{\gamma_i}$. De plus, on a pour tout $P \in \mathbb{L}[X]$, P annule u si et seulement si P annule B et C, *i.e.* $P \in \langle \mu_B \rangle \cap \langle \mu_C \rangle = \langle \mu_B \vee \mu_C \rangle$. Le polynôme unitaire de plus petit degré vérifiant cette propriété est $\mu_C \vee \mu_B$, donc

$$\prod_{i=1}^{s} (X - \lambda_i)^{\gamma_i} = \mu_u = \mu_C \vee \mu_B \underset{\mu_B \wedge \mu_C = 1}{\overset{\uparrow}{=}} \mu_B \mu_C = (X - \lambda_1)^p \mu_C.$$

On a alors $(X - \lambda_1)^{\gamma_1} | (X - \lambda_1)^p \mu_C$ et $\mu_C \wedge (X - \lambda_1)^{\gamma_1} = 1$, donc $(X - \lambda_1)^{\gamma_1} | (X - \lambda_1)^p$, donc $\gamma_1 \leq p$ (et aussi $\gamma_1 \geq p$), donc $\gamma_1 = p$, et alors $\mu_B = (X - \lambda_1)^{\gamma_1}$.

CC BY-NC-SA

La matrice $\mathrm{N} = \mathrm{B} - \lambda_1 \mathrm{Id}$ est donc nilpotente d'indice de nilpotence γ_1. En particulier, on dispose de $\mathrm{X} \in \mathbb{L}^d$ tel que $\mathrm{N}^{\gamma_1 - 1}\mathrm{X} \neq 0$. Ainsi, $\gamma_1 \leq d$ car $(\mathrm{X}, \mathrm{NX}, \dots, \mathrm{N}^{\gamma_1-1}\mathrm{X})$ est libre et ne peut donc pas contenir plus de d éléments (pour le montrer, il suffit de considérer une combinaison linéaire de cette famille non triviale et d'appliquer un nombre suffisant de fois N).

La seule valeur propre de N est 0 car si $\lambda \neq 0$ est une valeur propre de N et X un vecteur propre associé, $0 = \mathrm{N}^{\gamma_1}\mathrm{X} = \lambda^{\gamma_1}\mathrm{X}$ ce qui est absurde. De plus, χ_{N} est de degré d, unitaire, scindé dans $\mathbb{L}$ et admet uniquement 0 comme racine donc $\chi_{\mathrm{N}}(\mathrm{X}) = \mathrm{X}^d$. Ceci nous permet d'écrire

$$\chi_{\mathrm{B}}(\mathrm{X}) = \det(\mathrm{XI} - \mathrm{B}) = \det((\mathrm{X} - \lambda_1)\mathrm{I} - \mathrm{N}) = \chi_{\mathrm{N}}(\mathrm{X} - \lambda_1) = (\mathrm{X} - \lambda_1)^d.$$

Pour conclure, on a

$$(\mathrm{X} - \lambda_1)^d = \chi_{\mathrm{B}}(\mathrm{X}) | \chi_{\mathrm{A}}(\mathrm{X}) = \prod_{i=1}^{s} (\mathrm{X} - \lambda_i)^{\alpha_i},$$

et alors $d \leq \alpha_1$ et finalement $\gamma_1 \leq d \leq \alpha_1$, ce qui est bien le résultat voulu.

Corollaire V.8.

Soit $u \in \mathcal{L}(\mathrm{E})$. Les propositions suivantes sont vraies.

1. $\deg(\mu_u) \leq n$.
2. Les facteurs irréductibles de μ_u et χ_u sont exactement les mêmes à puissance près.
3. $\chi_u | \mu_u^n$ et $\mathrm{Z}_{\mathbb{K}}(\mu_u) = \mathrm{Z}_{\mathbb{K}}(\chi_u) = \mathrm{Spec}_{\mathbb{K}}(u)$.

Remarque.

→ Dans le cas particulier où $\mathbb{K} = \mathbb{C}$, le point (2) est équivalent à dire

$$\chi_u(\mathrm{X}) = \prod_{i=1}^{r} (\mathrm{X} - \lambda_i)^{\alpha_i} \text{ et } \mu_u(\mathrm{X}) = \prod_{i=1}^{r} (\mathrm{X} - \lambda_i)^{\gamma_i},$$

où λ_i distincts et pour tout $i \in [\![1; r]\!]$, $1 \leq \gamma_i \leq \alpha_i$.

→ Ce résultat est vrai même pour $\mathbb{K}$ égal à un corps quelconque, pas forcément égal à $\mathbb{R}$ ou $\mathbb{C}$.

Applications.

→ Soit $\mathrm{A} \in \mathcal{M}_2(\mathbb{K})$. Si $\mathrm{Tr}(\mathrm{A}) = 0$, alors A^2 est une homothétie. En effet, $0 = \chi_{\mathrm{A}}(\mathrm{A}) = \mathrm{A}^2 - \mathrm{Tr}(\mathrm{A})\mathrm{A} + \det(\mathrm{A})\mathrm{I}_2 = \mathrm{A}^2 + \det(\mathrm{A})\mathrm{I}_2$ et alors $\mathrm{A}^2 = -\det(\mathrm{A})\mathrm{I}_2$.

→ Soit $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{Z})$ tel que $\det \mathrm{A} \wedge \det \mathrm{B} = 1$. Alors $\exists \mathrm{U}, \mathrm{V} \in \mathcal{M}_n(\mathbb{Z})$ tel que $\mathrm{UA} + \mathrm{VB} = \mathrm{I}_n$.
En effet, $\chi_{\mathrm{A}}(\mathrm{A}) = 0$ nous donne que

$$\mathrm{A}^n + (\mathrm{Tr}\,\mathrm{A})\mathrm{A}^{n-1} + \cdots + c_1\mathrm{A} + (-1)^n \det \mathrm{A} \cdot \mathrm{I}_n = 0,$$

i.e.

$$\det \mathrm{A} \cdot \mathrm{I}_n = \mathrm{A} \times \underbrace{(-1)^{n+1}(\mathrm{A}^{n-1} + (\mathrm{Tr}\,\mathrm{A})\mathrm{A}^{n-2} + \cdots + c_1\mathrm{I}_n)}_{\mathrm{U} \in \mathcal{M}_n(\mathbb{Z})}.$$

De la même manière, il existe $\mathrm{V} \in \mathcal{M}_n(\mathbb{Z})$ polynôme en B tel que $\det \mathrm{B} \cdot \mathrm{I}_n = \mathrm{VB}$ et finalement d'après Bézout, on dispose de $a, b \in \mathbb{Z}$ tel que $a \cdot \det \mathrm{A} + b \cdot \det \mathrm{B} = 1$ ce qui permet de dire, en multipliant par I_n des deux côtés que $(a\mathrm{U})\mathrm{A} + (b\mathrm{V})\mathrm{B} = \mathrm{I}_n$.

Dans ce qui suit, on aura besoin du lemme suivant, qui est souvent assez utile.

(cc) BY-NC-SA

Lemme V.9.

Soit $u \in \mathcal{L}(\mathrm{E})$, $\lambda \in \mathbb{K}$ et $k \geq 2$, on a

$$\operatorname{Ker}(u - \lambda \operatorname{Id})^k = \operatorname{Ker}(u - \lambda \operatorname{Id}) \iff \operatorname{Ker}(u - \lambda \operatorname{Id})^2 = \operatorname{Ker}(u - \lambda \operatorname{Id}).$$

Démonstration.

→ ($\Leftarrow$) Soit $k \geq 2$ et $x \in \mathrm{E}$, on a

$$\begin{aligned}(u - \lambda \operatorname{Id})^k(x) = 0 &\implies (u - \lambda \operatorname{Id})^2 \circ (u - \lambda \operatorname{Id})^{k-2}(x) = 0 \\ &\implies (u - \lambda \operatorname{Id}) \circ (\mathrm{A} - \lambda \operatorname{Id})^{k-2}(x) = 0 \\ &\implies (u - \lambda \operatorname{Id})^{k-1}(x) = 0.\end{aligned}$$

En itérant ce procédé $k-1$ fois, on obtient que

$$(u - \lambda \operatorname{Id})^k(x) = 0 \implies (u - \lambda \operatorname{Id})^{k-1}(x) = 0 \implies \cdots \implies (u - \lambda \operatorname{Id})(x) = 0.$$

On en déduit donc que $\operatorname{Ker}(u - \lambda \operatorname{Id})^k \subset \operatorname{Ker}(u - \lambda \operatorname{Id})$, *i.e.* $\operatorname{Ker}(u - \lambda \operatorname{Id})^k = \operatorname{Ker}(u - \lambda \operatorname{Id})$.

→ ($\Rightarrow$) On a clairement

$$\operatorname{Ker}(u - \lambda \operatorname{Id}) \subset \operatorname{Ker}(u - \lambda \operatorname{Id})^2 \subset \cdots \subset \operatorname{Ker}(u - \lambda \operatorname{Id})^k.$$

Le fait que $\operatorname{Ker}(u - \lambda \operatorname{Id})^k = \operatorname{Ker}(u - \lambda \operatorname{Id})$ implique que toutes ces inclusions sont des égalités, et en particulier que $\operatorname{Ker}(u - \lambda \operatorname{Id}) = \operatorname{Ker}(u - \lambda \operatorname{Id})^2$.

Exercice V.10.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ avec χ_{A} scindé. Montrer que

$$\mathrm{A} \text{ est diagonalisable } \iff \forall \lambda \in \operatorname{Spec}_{\mathbb{K}}(\mathrm{A}),\ \ \dim \operatorname{Ker}(\mathrm{A} - \lambda \mathrm{I}_n) = \dim \operatorname{Ker}(\mathrm{A} - \lambda \mathrm{I}_n)^2.$$

VI Trigonalisation

1. Généralités

Soit E un $\mathbb{K}-$espace vectoriel de dimension finie $n \in \mathbb{N}^*$.

Définition VI.1.

On dit que $u \in \mathcal{L}(\mathrm{E})$ est trigonalisable s'il existe β base de E tel que $[u]_\beta$ est triangulaire supérieure.

Remarque. Soit $\beta = (e_1, \ldots, e_n)$ est une base de E. $[u]_\beta$ est triangulaire supérieure si et seulement si u stabilise le drapeau associé à β *i.e.* $\forall k \in [\![1; n]\!]$, $u(\operatorname{Vect}(e_1, \ldots, e_k)) \subset \operatorname{Vect}(e_1, \ldots, e_k)$.

Définition VI.2.

On dit que $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable s'il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$ tel que $\mathrm{P}^{-1}\mathrm{AP}$ soit triangulaire supérieure.

(cc) BY-NC-SA

Remarque. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$. Lorsque A est triangulaire supérieure, notons $\mathrm{A} = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$, on a $\mathrm{Spec}_{\mathbb{K}}(\mathrm{A}) = [a_{11}, \dots, a_{nn}]$. La démonstration de ce résultat est laissée comme exercice au lecteur.

Proposition VI.3.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$ et $u \in \mathcal{L}(\mathrm{E})$.

1. A est trigonalisable $\iff f_{\mathrm{A}} : \mathrm{X} \longmapsto \mathrm{AX}$ est trigonalisable.
2. u est trigonalisable $\iff$ Il existe β une base de E tel que $[u]_\beta$ est trigonalisable $\iff$ pour tout base β base de E, $[u]_\beta$ est trigonalisable.
3. Soit $\mathrm{A} = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$ et $\mathrm{B} = \begin{pmatrix} b_{11} & & * \\ & \ddots & \\ 0 & & b_{nn} \end{pmatrix}$ Si A et B sont semblables, alors $(a_{11}, \dots, a_{nn})$ et $(b_{11}, \dots, b_{nn})$ sont égaux à permutation près, soit, avec nos notations, $[a_{11}, \dots, a_{nn}] = [b_{11}, \dots, b_{nn}] (= \mathrm{Spec}_{\mathbb{K}}(\mathrm{A}) = \mathrm{Spec}_{\mathbb{K}}(\mathrm{B}))$.

Démonstration. La démonstration est facile et est laissée au lecteur (elle est très similaire à celle sur la diagonalisabilité).

Proposition VI.4.

Soit $u \in \mathcal{L}(\mathrm{E})$. Les propriétés suivantes sont équivalentes.

1. u est trigonalisable.
2. χ_u est scindé.
3. μ_u est scindé.
4. $\exists \mathrm{P} \in \mathbb{K}[\mathrm{X}] \setminus \{0\},\ \mathrm{P}(u) = 0$ et P est scindé.

En particulier, sur un corps algébriquement clos tel que $\mathbb{C}$, tout endomorphisme est trigonalisable.

Démonstration.

→ (1) ⇒ (2) Si u est trigonalisable donc on dispose de β base de E tel que $[u]_\beta = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ et alors

$$\chi_u(\mathrm{X}) = \det([\mathrm{X}\,\mathrm{Id} - u]_\beta) = \det \begin{pmatrix} \mathrm{X} - \lambda_1 & & * \\ & \ddots & \\ 0 & & \mathrm{X} - \lambda_n \end{pmatrix} = \prod_{i=1}^{n} (\mathrm{X} - \lambda_i).$$

χ_u est donc bien scindé.

→ (2) ⇒ (3) $\mu_u | \chi_u$, donc μ_u est aussi scindé.

→ (3) ⇒ (4) Il suffit de prendre $\mathrm{P} = \mu_u$.

→ (4) ⇒ (1) Procédons par récurrence forte sur n, la dimension de E.

- Le cas $n = 1$ est évident.
- Soit $n \geq 2$. Supposons que pour tout $k \in [\![1; n]\!]$, la propriété soit vraie. $\mu_u | \mathrm{P}$, donc μ_u est aussi scindé. μ_u admet donc une racine λ. Considérons $\mathrm{F} = \mathrm{Ker}(u - \lambda\,\mathrm{Id}) \neq \{0\}$ et β_1 une base de

(cc) BY-NC-SA

F qu'on complète en une base $\beta = (\beta_1, \beta_2)$ de E. On a alors

$$[u]_\beta = \begin{pmatrix} \lambda\,\mathrm{Id} & * \\ 0 & \mathrm{B} \end{pmatrix} \text{ et donc } 0 = [\mathrm{P}(u)]_\beta = \begin{pmatrix} \mathrm{P}(\lambda)\,\mathrm{Id} & * \\ 0 & \mathrm{P}(\mathrm{B}) \end{pmatrix},$$

donc $\mathrm{P}(\mathrm{B}) = 0$. $\mathrm{B} \in \mathcal{M}_\ell(\mathbb{K})$ (avec $\ell \leq n-1$) est annulé par un polynôme scindé, on peut donc appliquer l'hypothèse de récurrence et dire que B est trigonalisable. Il existe donc $\mathrm{Q} \in \mathrm{GL}_\ell(\mathbb{K})$ tel que $\mathrm{T} := \mathrm{QBQ}^{-1}$ soit triangulaire supérieure. Finalement,

$$[u]_\beta = \begin{pmatrix} \lambda\,\mathrm{Id} & * \\ 0 & \mathrm{B} \end{pmatrix} \simeq \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Q} \end{pmatrix} \begin{pmatrix} \lambda\,\mathrm{Id} & * \\ 0 & \mathrm{B} \end{pmatrix} \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Q}^{-1} \end{pmatrix} = \begin{pmatrix} \lambda\,\mathrm{Id} & * \\ 0 & \mathrm{T} \end{pmatrix}.$$

Cette matrice est triangulaire supérieure, donc u est trigonalisable.

Applications.

1. Soit $u \in \mathcal{L}(\mathrm{E})$. Les propriétés suivantes sont équivalentes
   - → u est nilpotente.
   - → Il existe une base β de E telle que $[u]_\beta$ est strictement triangulaire supérieure.
   - → $\chi_u = \mathrm{X}^n$;

   Pour montrer cela, il suffit de remarquer que le seul facteur irréductible de X^d est X.
2. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ et $\mathbb{P} \in \mathrm{K}[\mathrm{X}]$.

$$\mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) = [\lambda_1, \ldots, \lambda_n] \Longrightarrow \mathrm{Spec}_{\mathbb{C}}(\mathrm{P}(\mathrm{A})) = [\mathrm{P}(\lambda_1), \ldots, \mathrm{P}(\lambda_n)].$$

   Cette propriété se démontre aisément en trigonalisant la matrice A et en utilisant le fait que le spectre d'une matrice triangulaire est égal aux coefficients diagonaux de cette matrice.

Exercice VI.5.

Soit $u \in \mathcal{L}(\mathbb{C}^n)$. Montrer que les propriétés suivantes sont équivalentes.

1. u est nilpotent.
2. $\forall k \in [\![1; n]\!],\ \mathrm{Tr}(u^k) = 0$.
3. $\forall \mathrm{P} \in \mathbb{K}[\mathrm{X}],\ \mathrm{Tr}(\mathrm{P}(u)) = n\mathrm{P}(0)$.

L'exercice suivant est assez classique. Nous recommandons fortement au lecteur d'au moins mémoriser les résultats qu'il fournit, en particulier les classes de similitude de $\mathcal{M}_2(\mathbb{C})$ et $\mathcal{M}_3(\mathbb{C})$.

Exercice VI.6.

Supposons que $\mathbb{K} = \mathbb{C}$.

1. Trouver toutes les classes de similitudes des matrices suivant les valeurs de μ_{A} et χ_{A} pour $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ lorsque $n = 2$ et $n = 3$.
2. En déduire que pour tout $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{C})$,
   - → Pour $n = 2$, $\mu_{\mathrm{A}} = \mu_{\mathrm{B}} \Longleftrightarrow \mathrm{A} \simeq \mathrm{B}$.
   - → Pour $n = 3$, $(\mu_{\mathrm{A}} = \mu_{\mathrm{B}}$ et $\chi_{\mathrm{A}} = \chi_{\mathrm{B}}) \Longleftrightarrow \mathrm{A} \simeq \mathrm{B}$.
3. Trouver un contre-exemple des propriétés précédentes dans $n = 4$.

(cc) BY-NC-SA

Exercice VI.7.

Soit $\mathrm{S} \subset \mathrm{M}_n(\mathbb{K})$ un sous-ensemble non vide de matrices trigonalisables qui commutent deux à deux. Montrer qu'il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$ telle que pour tout $\mathrm{A} \in \mathrm{S}$, PAP^{-1} est triangulaire supérieure.

2. Décomposition de Jordan-Dunford

Proposition VI.8.

Soit $u \in \mathcal{L}(\mathrm{E})$. Supposons que χ_u soit scindé, *i.e.* qu'on peut écrire $\mu_u(\mathrm{X}) = \prod_{i=1}^{r}(\mathrm{X}-\lambda_i)^{\gamma_i}$ et $\chi_u(\mathrm{X}) = \prod_{i=1}^{r}(\mathrm{X}-\lambda_i)^{\alpha_i}$ avec les λ_i distincts. posons $\mathrm{F}_{\lambda_i} = \mathrm{Ker}(u-\lambda_i \mathrm{Id})^{\gamma_i} = \mathrm{Ker}(u-\lambda_i \mathrm{Id})^{\alpha_i}$.
Les propriétés suivantes sont vraies.

1. F_{λ_i} est stable par u et est de dimension α_i.
2. $\mathrm{E} = \bigoplus_{i=1}^{r} \mathrm{F}_{\lambda_i}$.
3. $u\big|_{\mathrm{F}_{\lambda_i}}^{\mathrm{F}_{\lambda_i}} = \lambda_i \mathrm{Id}_{\mathrm{F}_{\lambda_i}} + \mathrm{N}_{\lambda_i}$ où $\mathrm{N}_{\lambda_i} \in \mathcal{L}(\mathrm{F}_{\lambda_i})$ est nilpotent.

Vocabulaire. On appelle F_{λ_i} l'espace caractéristique de u associé à λ_i.

En considérant pour tout i, β_i une base de F_{λ_i} où $\left[u\big|_{\mathrm{F}_{\lambda_i}}^{\mathrm{F}_{\lambda_i}}\right]_{\beta_i}$ est triangulaire supérieure et β la concaténation de ces bases (qui est une base de E), on obtient que

$$[u]_\beta = \begin{pmatrix} \mathrm{A}_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{A}_{\lambda_r} \end{pmatrix} \text{ où } \mathrm{A}_{\lambda_i} = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} \in \mathcal{M}_{\alpha_i}(\mathbb{K}).$$

Remarquons que lorsque χ_u est scindé que $\forall i \in [\![1;r]\!],\ \mathrm{F}_{\lambda_i,u} = \mathrm{E}_{\lambda_i,u}$ si et seulement si u est diagonalisable. Ceci étant vrai car

$$\begin{aligned} \mathrm{F}_{\lambda_i,u} = \mathrm{E}_{\lambda_i,u} &\iff \mathrm{Ker}(u-\lambda \mathrm{Id}) = \mathrm{Ker}((u-\lambda \mathrm{Id})^{\alpha_i}) \\ &\iff \mathrm{Ker}(u-\lambda \mathrm{Id}) = \mathrm{Ker}((u-\lambda \mathrm{Id})^2) \iff u \text{ est diagonalisable.} \end{aligned}$$

Le cas $\alpha_i = 1$ est facile et peut être traité séparément par le lecteur. La dernière équivalence est vraie d'après l'exercice V.10. et l'avant-dernière d'après le lemme V.9.
Rassemblons maintenant tous les blocs diagonaux en une seule matrice D et de même pour ceux nilpotents (la partie strictement triangulaire supérieure) en N. En considérant $\delta, \nu \in \mathcal{L}(\mathrm{E})$ tel que

$$[\delta]_\beta = \mathrm{D} = \begin{pmatrix} \lambda_1 \cdot \mathrm{I}_{\alpha_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_r \cdot \mathrm{I}_{\alpha_r} \end{pmatrix} \text{ et } [\nu]_\beta = \mathrm{N} = \begin{pmatrix} \mathrm{N}_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{N}_{\lambda_r} \end{pmatrix},$$

on obtient la décomposition suivante, appelée décomposition de Dunford.

BY-NC-SA

Proposition (Décomposition de Dunford) VI.9.

Soit $u \in \mathcal{L}(\mathrm{E})$ tel que χ_u soit scindé. Il existe un unique couple d'endomorphismes $(\delta, \nu) \in \mathcal{L}(\mathrm{E})^2$ telle que

→ $u = \delta + \nu$.

→ δ est diagonalisable et ν et nilpotente.

→ δ et ν commutent.

De plus, δ et ν sont polynomiales en u *i.e.* il existe $\mathrm{P}, \mathrm{Q} \in \mathbb{K}[\mathrm{X}]$ tels que $\delta = \mathrm{P}(u)$ et $\nu = \mathrm{Q}(u)$.

Démonstration. L'existence a déjà été établie avant, il reste à montrer le caractère polynomial et l'unicité.
Remarquons d'abord que $\delta \in \mathbb{K}[u] \iff \nu \in \mathbb{K}[u]$. Il suffit donc de le montrer pour δ ou ν, disons δ.
Reprenons les notations de la proposition, ainsi que celles considérées juste avant. Remarquons que $\delta = \sum_1^r \lambda_i \pi_i$ où π_i est la projection sur $\mathrm{F}_{\lambda_i,u}$ parallèlement à $\bigoplus_{j\neq i} \mathrm{F}_{\lambda_j,u}$ et donc il suffit de vérifier que $\pi_i \in \mathbb{K}[u]$ pour tout i. Par symétrie des rôles des π_i, il suffit de le montrer par exemple pour $i = 1$. Posons

$$\mathrm{P}_1 = \prod_{j=2}^{r} (\mathrm{X} - \lambda_j)^{\alpha_j},\ \lambda = \mathrm{P}(\lambda_1) \neq 0,\ \mathrm{P}_2 = \mathrm{P}_1 - \lambda$$

$$\mathrm{P}_3 = (\mathrm{P}_2)^n,\ \mathrm{P}_4 = \mathrm{P}_3 - (-\lambda)^n,\ \mathrm{P}_5 = \frac{-1}{(-\lambda)^n}\mathrm{P}_5.$$

Pour voir que ces polynômes sont naturels à considérer, il suffit de voir la succession d'égalités ci-dessous.

$$[\mathrm{P}_1(u)]_\beta = \begin{pmatrix} \lambda\,\mathrm{Id} + \mathrm{N} & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & 0 \end{pmatrix} \text{ donc } [\mathrm{P}_2(u)]_\beta = \begin{pmatrix} \mathrm{N} & & 0 \\ & -\lambda\,\mathrm{Id} & \\ & & \ddots \\ 0 & & -\lambda\,\mathrm{Id} \end{pmatrix},$$

et alors

$$[\mathrm{P}_3(u)]_\beta = \begin{pmatrix} \mathrm{N}^n = 0 & & 0 \\ & (-\lambda)^n\,\mathrm{Id} & \\ & & \ddots \\ 0 & & (-\lambda)^n\,\mathrm{Id} \end{pmatrix} \text{ ainsi } [\mathrm{P}_4(u)]_\beta = \begin{pmatrix} -(-\lambda)^n\,\mathrm{Id} & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & 0 \end{pmatrix}.$$

Finalement,

$$[\mathrm{P}_5(u)]_\beta = \begin{pmatrix} \mathrm{Id} & & 0 \\ & 0 & \\ & & \ddots \\ 0 & & 0 \end{pmatrix} = [\pi_1]_\beta,$$

ce qui nous donne bien le résultat voulu.

Montrons enfin l'unicité. Considérons (δ', ν') vérifiant les mêmes propriétés que (δ, ν). On a

$$u = \delta + \nu = \delta' + \nu' \ \textit{i.e.}\ \delta - \delta' = \nu - \nu'.$$

ν et ν' commutent, car $\nu \in \mathbb{K}[u]$ et ν' commute avec δ' et alors avec $\delta' + \nu' = u$. En posant $v = \delta - \delta' = \nu - \nu'$, on obtient

$$v^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} \nu^k \nu'^{2n-k}.$$

BY-NC-SA

Sachant que $\forall k \in [\![1;2n]\!]$, soit $k \geq n$ soit $2n-k \geq n$ (on rappelle que l'indice de nilpotence de ν et ν' est forcément inférieur à n), tous les termes de cette somme sont nuls. Ainsi, les seules valeurs propres de v dans $\mathbb{C}$ sont 0, mais v est diagonalisable car il est égal à la somme de δ et δ' qui sont diagonalisables et commutent entre eux car δ est dans $\mathbb{K}[u]$ et δ' commute avec ν' et donc avec $\delta' + \nu' = u$ et sont donc codiagonalisables. On en déduit donc directement que $v = 0$, *i.e.* $(\delta, \nu) = (\delta', \nu')$, d'où l'unicité de δ et ν.

Exercice VI.10.

Résoudre les équations suivantes.

1. $\mathrm{X}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}$ dans $\mathcal{M}_3(\mathbb{R})$.
2. $\mathrm{X}^2 = \begin{pmatrix} a^2 & 1 \\ 0 & a^2 \end{pmatrix}$ dans $\mathcal{M}_2(\mathbb{C})$ avec $a \in \mathbb{C}$.
3. $\mathrm{X}^2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ dans $\mathcal{M}_3(\mathbb{R})$.

Exercice VI.11.

Soit $u \in \mathcal{L}(\mathbb{C}^n)$. On considère l'application

$$\phi_u : \begin{cases} \mathcal{L}(\mathbb{C}^n) & \longrightarrow \mathcal{L}(\mathrm{C}^n) \\ v & \longmapsto u \circ v - v \circ u. \end{cases}$$

Montrer que u est diagonalisable si et seulement si ϕ_u est diagonalisable.

Exercice VI.12.

Soit $u \in \mathrm{GL}_n(\mathbb{C})$. Montrer que u admet une racine, *i.e.* qu'il existe $v \in \mathrm{GL}_n(\mathbb{C})$, $v^2 = u$.

Proposition (Décomposition de Jordan 1) VI.13.

Soit $v \in \mathcal{L}(\mathrm{E})$ nilpotent. Il existe une base β de E tel que

$$[v]_\beta = \begin{pmatrix} \mathrm{J}_{\ell_1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{J}_{\ell_s} \end{pmatrix} \text{ où } \forall k \in [\![1;s]\!],\ \mathrm{J}_k = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in \mathcal{M}_k(\mathbb{K}),$$

avec $\ell_1, \ldots, \ell_s \in [\![1;n]\!]$ et $\ell_1 + \cdots + \ell_s = n$.

Démonstration. Soit $u \in \mathcal{L}(\mathrm{E})$ nilpotent, il est clair qu'il existe $p \geq 1$ tel que $\mu_u = \mathrm{X}^p$ pour un certain $p \geq 1$ (en effet, u est annulé par X^ℓ pour ℓ assez grand, donc $\mu_u | \mathrm{X}^\ell$).
Montrons le résultat par récurrence forte sur la dimension n de E.

→ Le cas $n = 1$ est évident.

→ Soit $n \geq 2$ tel que $\dim \mathrm{E} = n$. Supposons que la propriété soit vraie pour tout $k \in [\![1;n]\!]$. Soit $x_1 \in \mathrm{E} \setminus \{0\}$ tel que $u^{p-1}(x_1) \neq 0$. En posant $\mathrm{F}_{x_1} = \mathrm{Vect}(x_1, \ldots, u^{p-1}(x_1))$ (remarquons que cette

(cc) BY-NC-SA

famille est libre, c'est donc une base de F_{x_1}), on voit que F_{x_1} est stable et

$$\left[u\Big|_{\mathrm{F}_{x_1}}^{\mathrm{F}_{x_1}}\right] = \mathrm{J}_p.$$

Notre intuition est de trouver un supplémentaire de F_{x_1} stable par u et lui appliquer l'hypothèse de récurrence.

La famille $(x_1, u(x_1), \dots, u^{p-1}(x_1))$ est libre, on peut donc considérer $\phi : \mathrm{E} \longrightarrow \mathbb{K}$ une forme linéaire telle que $\varphi(u^{p-1}(x_1)) = 1$ et pour tout $k \in [\![0; p-2]\!]$, $\phi(u^k(x_1)) = 0$.

À partir de ϕ on construit l'application linéaire suivante

$$\varphi : \begin{cases} \mathrm{E} & \longrightarrow \mathbb{K}^p \\ y & \longmapsto (\phi(u^{p-1}(y), \dots, \phi(u(y)), \phi(y)), \end{cases}$$

et on pose $\mathrm{F} = \operatorname{Ker} \varphi$. F est stable par u. En effet, on a pour tout $y \in \mathrm{E}$,

$$\begin{aligned} \varphi(y) = 0 &\Longrightarrow \forall k \in [\![0; p-1]\!],\ u^k(y) = 0 \\ &\Longrightarrow \forall k \in [\![0; p-1]\!],\ u^k(u(y)) = 0 \text{ (car } u^p = 0) \\ &\Longrightarrow u(y) \in \mathrm{F}. \end{aligned}$$

φ est surjective. en effet, pour tout $k \in [\![0; p-1]\!]$, $\varphi(u^k(x_1)) = (\delta_{ik})_{i \in [\![0;p-1]\!]}$ donc pour tout $(a_0, \dots, a_{p-1}) \in \mathbb{K}^p$,

$$\varphi(a_0 u^0(x_1) + \dots + a_{p-1} u^{p-1}(x_1)) = (a_0, \dots, a_{p-1}).$$

$\mathrm{F} \cap \mathrm{F}_{x_1} = \{0\}$. En effet, pour tout $y \in \mathrm{F}_{x_1}$, en posant $y = y := \sum\limits_0^{p-1} a_k u^k(x_1)$, on a

$$\varphi(y) = 0 \Longrightarrow 0 = \varphi\left(\sum_0^{p-1} a_k u^k(x_1)\right) = (a_0, \dots, a_{p-1}) \Longrightarrow y = 0,$$

donc F et F_{x_1} sont en somme directe et par la formule du rang

$$\dim \mathrm{F} = \dim \operatorname{Ker} \varphi = \dim \mathrm{E} - \dim \operatorname{Im} \varphi = n - p,$$

et alors $\dim \mathrm{F}_{x_1} + \dim \mathrm{F} = n = \dim \mathrm{E}$. F et F_{x_1} sont donc supplémentaires.

Pour finir, on voit que $w = u\Big|_{\mathrm{F}}^{\mathrm{F}}$ est aussi nilpotente, on peut donc appliquer l'hypothèse de récurrence à w sur F ce qui nous donne bien le résultat voulu.

Remarques.

→ Lorsqu'on ajoute la condition $\ell_1 \geq \dots \geq \ell_s \geq 1$, cette décomposition est unique, *i.e.* s'il existe une autre base β' de E, où n est de la même forme avec $m_1 \geq \dots \geq m_r \geq 1$ les tailles respectives de ses blocs $(\mathrm{J}_k)_{k \in [\![1;r]\!]}$, alors $r = s$ et $\forall i \in [\![1; s]\!],\ \ell_i = m_i$. Pour s'en convaincre, notons

$$\forall i \in \mathbb{N}^*,\ \mathrm{F}(i) = |\{k \in [\![1; s]\!],\ \ell_k = i\}| \text{ et } \mathrm{G}(i) = |\{k \in [\![1; r]\!],\ m_k = i\}|.$$

Il est aisé de vérifier que

$$\forall r \in \mathbb{N}^*,\ \forall k \in \mathbb{N},\ \dim \operatorname{Ker}(\mathrm{J}_r^k) = \min(k, r).$$

Ceci permet donc d'affirmer que

$$\forall k \in \mathbb{N}^*,\ \dim \operatorname{Ker}(u^k) = \sum_{i=1}^{\infty} \mathrm{F}(i) \min(k, i) = \sum_{i=1}^{\infty} \mathrm{G}(i) \min(k, i).$$

(cc) BY-NC-SA

Ainsi, $F = G$ (Il suffit de considérer par absurde le premier indice $i \in \mathbb{N}^*$ où $F(i) \neq G(i)$ pour tomber sur une contradiction) et donc on a l'unicité.

→ Remarquons que le bloc associé au sous-espace F_{x_1} a la plus grande taille des blocs. En effet, la taille du bloc correspond à la dimension de ce sous-espace, $\dim F_{x_1} = p$ et la dimension de tout espace défini de la même manière (en itérant u sur un élément de E) est de dimension au plus p.

Proposition (Décomposition de Jordan 2) VI.14.

Soit $u \in \mathcal{L}(E)$ tel que χ_u est scindé dans $\mathbb{K}$. Notons $\lambda_1, \ldots, \lambda_r$ les valeurs propres distinctes (comptées sans multiplicité) de u. Il existe une base β de E telle que

$$[u]_\beta = \begin{pmatrix} \lambda_1 \cdot I_{\ell_1^i} + A_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r \cdot I_{\ell_{s_i}^i} + A_r \end{pmatrix} \text{ avec } A_i = \begin{pmatrix} J_{\ell_1^i} & & 0 \\ & \ddots & \\ 0 & & J_{\ell_{s_i}^i} \end{pmatrix},$$

où pour tout $i \in [\![1; r]\!]$, $s_i \in \mathbb{N}^*$ et $\ell_1^i \geq \cdots \geq \ell_{s_i}^i \geq 1$. De plus, cette écriture est unique à permutation des blocs $\lambda_i I_{\ell_{s_i}^i} + A_i$ près.
On appelle cette décomposition réduction de Jordan de u.

Remarque. Certains auteurs n'exigent pas les inégalités sur les tailles des blocs dans la réduction de Jordan. Cela ne change que la partie unicité de ce théorème.

Exercice VI.15.

Soit $A \in \mathcal{M}_n(\mathbb{C})$. Montrer que $\dim \mathrm{Com}(A) \geq n$.

VII Endomorphismes cycliques

Cette partie hors programme est assez classique. En effet, elle revient dans un nombre considérable d'exercices d'oraux et de sujets d'écrits, voir par exemple Centrale Math 1 2019, où l'intégralité du sujet portait sur les endomorphismes cycliques. Les élèves (dont ceux des MP* de Louis-le-Grand) ayant vu cette notion étaient très avantagés par rapport aux autres cette année-là.
Soit E un $\mathbb{K}-$espace vectoriel de dimension finie $n \geq 1$.

Définition VII.1.

Soit $u \in \mathcal{L}(E)$ et $x \in E \setminus \{0\}$. On appelle espace cyclique engendré par x pour u le sous-espace vectoriel défini par

$$F_{x,u} = \mathrm{Vect}\{u^k(x), k \in \mathbb{N}\}.$$

On le notera aussi F_x lorsqu'il n'y a pas d'ambiguïté sur u.

BY-NC-SA

Proposition VII.2.

Soit $u \in \mathcal{L}(\mathrm{E})$ et $x \in \mathrm{E} \setminus \{0\}$. Les propositions suivantes sont vraies.

1. $\mathrm{F}_{x,u}$ est le plus petit sous-espace vectoriel de E stable par u contenant x.
2. Soit $p = \max\{k \in \mathbb{N}^*,\ (x, u(x), \dots, u^{k-1}(x)) \text{ est libre}\}$. $\beta = (x, \dots, u^{p-1}(x))$ est une base de $\mathrm{F}_{x,u}$.
3. Il existe $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ unitaire tel que $\left[u\Big|_{\mathrm{F}_{x,u}}^{\mathrm{F}_{x,u}}\right]_\beta = \mathrm{C_P}$ la matrice compagnon associée à P.
4. P est le générateur normalisé de $\mathcal{H}_{u,x} = \{\mathrm{Q} \in \mathbb{K}[\mathrm{X}],\ \mathrm{Q}(u)(x) = 0\}$, *i.e.* P est l'unique polynôme unitaire tel que $\mathcal{H}_{u,x} = \langle \mathrm{P} \rangle := \mathrm{P} \cdot \mathbb{K}[\mathrm{X}]$. On appellera P le polynôme minimal (de u) en x et on le notera $\mu_{x,u}$ ou μ_x s'il n'y a pas d'ambiguïté sur u. De plus, $\mu_{x,u} | \mu_u$.

Démonstration.

1. Clair (il suffit de l'écrire).
2. β est libre par définition. Montrons par récurrence forte que $\forall \ell \in \mathbb{N}\ u^\ell(x) \in \mathrm{Vect}(\beta)$.
   - → La propriété est évidente pour $\ell \in [\![0; p-1]\!]$.
   - → Soit $\ell \geq 1$. Supposons que la propriété soit vraie pour tout $k \in [\![0; \ell - 1]\!]$. Si$\ell \leq p$, la propriété est vraie. Supposons donc que $\ell > p$. Par définition, $(x, \dots, u^\ell(x))$ est liée, il existe donc $m \in [\![0; \ell]\!]$ et $(\lambda_0, \dots, \lambda_m) \in \mathbb{K}^n \setminus \{(0, \dots, 0)\}$ tels que
   $$u^m(x) = \sum_0^{m-1} \lambda_k u^k(x).$$
   On a donc, par hypothèse de récurrence,
   $$u^\ell(x) = u^{\ell-m}\left(\sum_0^{m-1} \lambda_k u^k(x)\right) = \sum_0^{m-1} \lambda_k \underbrace{u^{k+\ell-m}(x)}_{\in \mathrm{Vect}(\beta)} \in \mathrm{Vect}(\beta),$$
   ce qui est bien le résultat voulu.
3. β est libre mais $(\beta, u^p(x))$ ne l'est pas, il existe donc $\lambda_0, \dots, \lambda_{p-1} \in \mathbb{K}$ tels que $u^p(x) = \sum_0^{p-1} \lambda_k u^k(x)$
   En Posant $\mathrm{P} := \mathrm{X}^p - \sum_{k=0}^{p-1} \lambda_k \mathrm{X}^k$, il est aisé de vérifier que $\left[u\Big|_{\mathrm{F}_{x,u}}^{\mathrm{F}_{x,u}}\right]_\beta = \mathrm{C_P}$.
4. Remarquons que le polynôme P retrouvé au point (3) est unitaire, de degré p et annule u en x. Remarquons de plus que $(x, \dots, u^{p-1}(x))$ est libre *i.e.* tout élément $\mathrm{Q} \in \mathcal{H}_{u,x}$ non nul doit être de degré supérieur à p. En particulier, si Q est le générateur unitaire de $\mathcal{H}_{u,x}$ (existe car $\mathcal{H}_{u,x}$ est un idéal non trivial), alors Q|P et Q et P unitaires (non nuls) et $\deg(\mathrm{P}) = p \leq \deg(\mathrm{Q})$ d'où $\mathrm{P} = \mathrm{Q}$ et donc $\mathcal{H}_{u,x} = \langle \mathrm{P} \rangle$.

Définition VII.3.

On dit que u est cyclique lorsqu'il existe $x \in \mathrm{E} \setminus \{0\}$ tel que $\mathrm{F}_x = \mathrm{E}$ *i.e.* il existe $x \in \mathrm{E}$ tel que $(x, \dots, u^{n-1}(x))$ est libre.

Proposition VII.4.

Soit u cyclique et $x \in \mathrm{E}\setminus\{0\}$ tel que $\mathrm{F}_x = \mathrm{E}$. Alors $\mu_x = \mu_u = \chi_u$.

(cc) BY-NC-SA

Démonstration. Les trois polynômes sont tous unitaires et $\mu_x|\mu_u|\chi_u$ et donc il suffit de montrer que $\deg(\mu_x) \geq n$, ce qui est vrai car $(x, u(x), \dots, u^{n-1}(x))$ est libre et donc aucun polynôme de degré $n-1$ ou moins ne peut annuler u.
Remarquons qu'en prenant u quelconque, $x \in \mathrm{E} \setminus \{0\}$ et en considérant $v = u\Big|_{\mathrm{F}_x}^{\mathrm{F}_x}$, on obtient que $\mu_{u,x} = \mu_{v,x} = \chi_v$ et donc, en particulier, $\deg(\mu_x) = \dim(\mathrm{F}_x)$.

Proposition VII.5.

Soit $u \in \mathcal{L}(\mathrm{E})$. Notons μ le polynôme minimal unitaire de u. Si $d \in \mathbb{N}$ est le degré de μ, alors $(\mathrm{Id}_\mathrm{E}, \dots, u^{d-1})$ est une base de $\mathrm{K}[u]$.

Démonstration. Notons $\mathrm{F} = \mathrm{Vect}(\mathrm{Id}_\mathrm{E}, \dots, u^{d-1})$

→ Montrons que la famille $(\mathrm{Id}_\mathrm{E}, \dots, u^{d-1})$ est génératrice de F.
Soit $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$. Faisons la division euclidienne de P par μ :

$$\mathrm{P} = \mathrm{Q}\mu + \mathrm{R} \text{ avec } \deg \mathrm{R} \leq d-1.$$

On obtient alors

$$\mathrm{P}(u) = \mathrm{Q}(u) \circ \underbrace{\mu(u)}_{0} + \underbrace{\mathrm{R}(u)}_{\in \mathrm{F}} = \mathrm{R}(u) \in \mathrm{F},$$

car R ne contient que des termes de degré inférieur ou égal à $d-1$.

→ Montrons que la famille $(\mathrm{Id}_\mathrm{E}, \dots, u^{d-1})$ est libre.
Soit $(\lambda_0, \dots, \lambda_{d-1}) \in \mathbb{R}^d$ tel que $\sum_{k=0}^{d-1} \lambda_k u^k = 0$. Posons $\mathrm{P} = \sum_{k=0}^{d-1} \lambda_k \mathrm{X}^k \in \mathbb{K}_{d-1}[\mathrm{X}]$, on a $\mathrm{P}(u) = 0$ et donc $\mu|\mathrm{P}$. De plus, $\deg \mathrm{P} < \deg \mu = d$, d'où $\mathrm{P} = 0$ et la famille est donc libre.

Remarque. Cette proposition donne un corollaire assez souvent utile : si $u \in \mathcal{L}(\mathrm{E})$ est nilpotente d'indice de nilpotence d, alors $(\mathrm{Id}_\mathrm{E}, \dots, u^{d-1})$ est une base de $\mathbb{K}[u]$. En particulier, elle est libre.

Exercice VII.6.

Soit $u \in \mathcal{L}(\mathrm{E})$. On admettra qu'il existe $x \in \mathrm{E} - \{0\}$ tel que $\mu_{x,u} = \mu_u$ (démontré dans un exercice ultérieur).

1. Montrer que $\mu_u = \chi_u$ si et seulement si u est cyclique.
2. Supposons que u est cyclique. Montrer que $\mathrm{Com}(u) = \mathbb{K}[u] = \{\mathrm{P}(u),\ \mathrm{P} \in \mathbb{K}[\mathrm{X}]\}$ et en déduire que $\dim \mathrm{Com}(u) = n$.
3. On ne suppose plus que u est cyclique. Déduire de la question précédente que $\dim \mathrm{Com}(u) \geq n$.
4. Montrer que $\mu_u = \chi_u$ si et seulement si $\mathrm{Com}(u) = \mathbb{K}[u]$.

Exercice VII.7.

Soit $u \in \mathcal{L}(\mathrm{E})$. Montrer que χ_u est irréductible si et seulement si les seuls sous-espaces vectoriels stables par u sont $\{0\}$ et E.
Au passage remarquer que, dans ce cas, tout élément non nul est cyclique pour u.

CC BY-NC-SA

Exercice VII.8.

Soit $A \in \mathcal{M}_n(\mathbb{C})$.

1. Montrer que la dimension de chaque espace propre est égale à 1 si A est semblable à une matrice compagnon.
2. En déduire que si A est diagonalisable, alors A est cyclique si et seulement si A possède n valeurs propres distinctes.

Exercice VII.9.

Soit $P \in \mathbb{Z}[X]$ un polynôme unitaire de degré n et de racines (dans $\mathbb{C}$, possiblement avec répétition) $\lambda_1, \dots, \lambda_n$

1. Montrer que $\forall k \geq 1$, $\lambda_1^k, \dots, \lambda_n^k$ sont les racines d'un polynôme unitaire de $\mathbb{Z}[X]$ de degré n.
2. En déduire que si $|\lambda_1| = \cdots = |\lambda_n| = 1$ alors les λ_j sont des racines de l'unité.

Exercice VII.10.

Dans cet exercice, on notera $\forall P \in \mathbb{K}[X]$ et $x \in E$, $P \cdot x = P(u)(x)$ et $\mu_0 = 1$.

1. Soit $x \in E \setminus \{0\}$. Supposons qu'il existe $P, Q \in \mathbb{K}[X]$ unitaires tels que $\mu_x = PQ$ et posons $y = Q \cdot x$. Vérifier que $\mu_y = P$.
2. Soit $x, y \in E$ non nuls tels que $\mu_x \wedge \mu_y = 1$. Montrer que $\mu_{x+y} = \mu_x \mu_y$.
3. On ne suppose plus que $\mu_x \wedge \mu_y = 1$. Montrer que qu'il existe $z \neq 0$ tel que $\mu_z = \mu_x \vee \mu_y$.
4. Montrer qu'il existe $a \neq 0$ tel que pour tout $x \neq 0$, $\mu_x | \mu_a$. En déduire que $\mu_a = \mu_u$, le polynôme minimal de u.

Application. En utilisant le résultat de la dernière question, on peut démontrer le théorème de Cayley-Hamilton. En effet, considérons $u \in \mathcal{L}(E)$ et $x \in E \setminus \{0\}$ tel que $\mu_x = \mu_u$. F_x étant stable, $\mu_u = \mu_x = \chi_v | \chi_u$ où $v = u\Big|_{F_x}^{F_x}$. On en déduit donc directement que $\chi_u(u) = 0$.
Une autre manière similaire de démontrer le théorème de Cayley Hamilton sans le résultat de la dernière question est de voir que pour tout $x \in E \setminus \{0\}$, F_x étant stable, $\mu_x = \chi_v | \chi_u$ où $v = u\Big|_{F_x}^{F_x}$. On a alors pour tout $x \in E \setminus \{0\}$, $\chi_u(u)(x) = 0$, *i.e.* $\chi_u(u) = 0$.

VIII Réduction et topologie

1. Normes, valeurs propres

On suppose dans cette partie que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Rappel VIII.1.

On suppose que $E = \mathcal{M}_n(\mathbb{K})$. Soit $||.||$ une norme sur $\mathbb{K}^n$. On peut définir une norme sur E, appelée norme d'opérateur, par

$$|||A||| = \sup_{||X||=1} ||AX|| = \sup_{||X|| \leq 1} ||AX|| = \sup_{X \in E \setminus \{0\}} \frac{||AX||}{||X||}.$$

(cc) BY-NC-SA

Remarque. cette norme vérifie les propriétés suivantes

→ $\|\!|\mathrm{I}_n\|\!| = 1$.

→ $\forall \mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{K}),\ \|\!|\mathrm{AB}\|\!| \leq \|\!|\mathrm{A}\|\!| \times \|\!|\mathrm{B}\|\!|$.

→ $\forall \mathrm{A} \in \mathcal{M}_n(\mathbb{K}),\quad \sup\limits_{\lambda \in \mathrm{Spec}_{\mathbb{K}}(\mathrm{A})} |\lambda| \leq \|\!|\mathrm{A}\|\!|$.

Proposition VIII.2.

Soit $\mathrm{A} = (a_{ij})_{i,j \in [\![1;n]\!]} \in \mathcal{M}_n(\mathbb{K})$. On a $\mathrm{VP}(\mathrm{A}) \subset \bigcup\limits_{k=1}^{n} \mathrm{B}_f\left(a_{kk}, \sum\limits_{j=1, j\neq k}^{n} |a_{kj}|\right)$.

Démonstration. Posons $\mathrm{H} = \bigcup\limits_{k=1}^{n} \mathrm{B}_f\left(a_{kk}, \sum\limits_{j=1}^{n} |a_{kj}|\right)$. Soit $\lambda \notin \mathrm{H}$, alors

$$\forall k \in [\![1;n]\!],\ |\lambda - a_{kk}| > \sum_{j=1,\ j\neq k}^{n} |a_{kj}|\,.$$

D'après Hadamard (voir le chapitre des systèmes linéaires), ceci est équivalent à dire que la matrice $\mathrm{A}-\lambda\mathrm{I}$ est inversible et que donc λ n'est pas une valeur propre de A. On en déduit donc bien que $\mathrm{VP}(\mathrm{A}) \subset \mathrm{H}$.

Application. En utilisant ce résultat, on peut redémontrer le résultat sur les racines d'un polynôme vu au chapitre 1. En effet, considérons $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ et soit $\mathrm{A} = \mathrm{C}_{\mathrm{P}}$. On a bien entendu

$$\mathrm{P}(\mathrm{X}) = \chi_{\mathrm{A}}(\mathrm{X}) = \mathrm{X}^n + \alpha_{n-1}\mathrm{X}^{n-1} + \cdots + \alpha_1\mathrm{X} + \alpha_0.$$

On a donc, en utilisant le résultat ci-dessus

$$\mathrm{Z}(\mathrm{P}) = \mathrm{Spec}_{\mathbb{K}}(\mathrm{A}) \subset \bigcup_{k=0}^{n-2} \mathrm{B}_f(0, 1 + |\alpha_k|) \cup \mathrm{B}_f(\alpha_{n-1}, 1),$$

i.e. pour toute racine λ de P, on a l'inégalité

$$|\lambda| \leq 1 + \max_{k\in[\![0;n-1]\!]} |a_k|\,.$$

Exercice VIII.3.

Soit F un fermé de $\mathbb{C}$. Montrer que $\widetilde{\mathrm{F}} = \{\mathrm{A} \in \mathcal{M}_n(\mathbb{C}),\ \mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) \subset \mathrm{F}\}$ est fermé dans $\mathcal{M}_n(\mathbb{C})$.

Application. $\mathcal{M}_n(\mathbb{R})$ est fermé dans $\mathcal{M}_n(\mathbb{C})$, donc en appliquant le résultat de cet exercice, on peut affirmer que $\{\mathrm{A} \in \mathcal{M}_n(\mathbb{C}),\ \mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) \subset \mathbb{R}\}$ est fermé dans $\mathcal{M}_n(\mathbb{C})$. De plus, on a

$$\begin{aligned}\{\mathrm{A} \in \mathcal{M}_n(\mathbb{R}),\ \mathrm{A} \text{ est trigonalisable}\} &\underset{\substack{\uparrow\\ \text{proposition VI.4}}}{=} \{\mathrm{A} \in \mathcal{M}_n(\mathbb{R}),\ \chi_{\mathrm{A}} \text{ est scindé}\}\\ &= \{\mathrm{A} \in \mathcal{M}_n(\mathbb{C}),\ \mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) \subset \mathbb{R}\} \cap \mathcal{M}_n(\mathbb{R}).\end{aligned}$$

On en déduit donc que l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ trigonalisables s'écrit sous forme d'une intersection de deux fermés, c'est donc un fermé dans $\mathcal{M}_n(\mathbb{R})$ et $\mathcal{M}_n(\mathbb{C})$.

(cc) BY-NC-SA

2. Diagonalisation à ε près

Proposition VIII.4.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ et $\lambda_1, \ldots, \lambda_n$ les coefficients diagonaux d'une trigonalisation de A. Pour tout $\varepsilon > 0$, il existe $P \in GL_n(\mathbb{C})$ tel que $P^{-1}AP = \begin{pmatrix} \lambda_1 & & b_{ij} \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ avec $\forall i \neq j\ |b_{ij}| \leq \varepsilon$.

Démonstration. Soit $\beta = (b_1, \ldots, b_n)$ une base de trigonalisation de A dans $\mathcal{M}_n(\mathbb{C})$ *i.e.* une base telle que

$$A_\beta = QAQ^{-1} = \begin{pmatrix} \lambda_1 & & a_{ij} \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

avec Q la matrice de la base β. Soit $p > 0$, posons $\beta_p = \left(b_1, \frac{b_2}{p}, \ldots, \frac{b_n}{p^{n-1}}\right)$. β_p est aussi une base de $\mathcal{M}_n(\mathbb{C})$. La matrice de A dans cette base s'écrit

$$V_p QAQ^{-1} V_p^{-1} = \begin{pmatrix} \lambda_1 & & \frac{a_{ij}}{p^{|i-j|}} \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

avec $V_p = \operatorname{diag}\left(1, \frac{1}{p}, \ldots, \frac{1}{p^{n-1}}\right)$. Il suffit ensuite de prendre p assez grand pour qu'on ait pour tout $i \neq j$, $\frac{|a_{i,j}|}{p} \leq \varepsilon$.

Exercice VIII.5.

Soit $A \in \mathcal{M}_n(\mathbb{C})$ tel que $\operatorname{Spec}(A) \subset B(0,1)$. Montrer que $A^p \xrightarrow[p \to +\infty]{} 0$.

Remarque. La propriété connue $|||A^p|||^{1/p} \xrightarrow[p \to \infty]{} \rho(A) = \sup\{|\lambda|,\ \lambda \in \operatorname{Spec}_{\mathbb{C}}(A)\}$ dans $\mathbb{C}$ permet de rapidement en conclure aussi.

Exercice VIII.6.

Soit $A \in \mathcal{M}_n(\mathbb{Z})$ tel que $\operatorname{Spec}_{\mathbb{C}}(A) \subset B(0,1)$. Montrer que A est nilpotente

3. Densité des matrices diagonalisables

Proposition VIII.7.

$\Omega = \{A \in \mathcal{M}_n(\mathbb{C}),\ A \text{ possède } n \text{ valeurs propres distinctes}\}$ est dense dans $\mathcal{M}_n(\mathbb{C})$.

Démonstration.

Soit β une base de $\mathbb{C}^n$ tel que $[A]_\beta = PAP^{-1} = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ avec $P \in GL_n(\mathbb{K})$. Considérons alors (A_p)

(cc) BY-NC-SA

la suite de matrices vérifiant

$$\forall p \in \mathbb{N}, \mathrm{A}_p = \mathrm{A} + \mathrm{P}^{-1} \begin{pmatrix} \dfrac{1}{p+1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{1}{p+n} \end{pmatrix} \mathrm{P}.$$

On a alors

$$\forall p \in \mathbb{N},\ [\mathrm{A}_p]_\beta = [\mathrm{A}]_\beta + \begin{pmatrix} \dfrac{1}{p+1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{1}{p+n} \end{pmatrix}.$$

On remarque que $\mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) = [\lambda_1, \ldots, \lambda_n]$. A partir de là, il y a deux moyens de conclure.

Méthode 1 : Considérons le polynôme non nul

$$\mathrm{P}(\mathrm{X}) = \prod_{i \neq j} (\mathrm{X}+i)(\mathrm{X}+j)\left(\lambda_i + \frac{1}{\mathrm{X}+i} - \lambda_j - \frac{1}{\mathrm{X}+j}\right) = \prod_{i \neq j} ((\mathrm{X}+i)(\mathrm{X}+j)(\lambda_i - \lambda_j) + j - i)).$$

Il est clair que si A_p n'admet pas n valeurs propres distinctes alors $\mathrm{P}(p) = 0$. P n'a qu'un nombre fini de racines et donc à partir d'un certain rang A_p admet toujours n valeurs propres distinctes. Finalement

$$\mathrm{A}_p = \mathrm{A} + \underbrace{\mathrm{P}^{-1} \begin{pmatrix} \dfrac{1}{p+1} & & 0 \\ & \ddots & \\ 0 & & \dfrac{1}{p+n} \end{pmatrix} \mathrm{P}}_{\xrightarrow[p \to +\infty]{} 0} \xrightarrow[p \to +\infty]{} \mathrm{A},$$

d'où le résultat recherché.

Méthode 2 : Posons $\delta = \dfrac{1}{2} \min\{|\lambda_i - \lambda_j|,\ i \neq j \text{ et } \lambda_i \neq \lambda_j\} > 0$ et soit $k \in \mathbb{N}^*$ tel que $\dfrac{1}{k} < \delta$. Supposons par l'absurde qu'il existe deux termes diagonaux de $[\mathrm{A}_p]_\beta$ égaux, il existe donc $i \neq j$ tel que $\lambda_i + \dfrac{1}{k+i} = \lambda_j + \dfrac{1}{k+j}$. Deux cas se présentent.

→ $\lambda_i = \lambda_j$ et donc $0 = |\lambda_i - \lambda_j| = \left|\dfrac{1}{i+k} - \dfrac{1}{j+k}\right| \neq 0$, ce qui est absurde.

→ $\lambda_i \neq \lambda_j$ et donc $\delta < |\lambda_i - \lambda_j| = \left|\dfrac{1}{i+k} - \dfrac{1}{j+k}\right| = \dfrac{|i-j|}{(k+i)(k+j)} \leq \frac{1}{k} < \delta$, ce qui est absurde.

En en déduit donc que pour tout $p \geq k$, A_p admet n valeurs propres distinctes, et $\mathrm{A}_p \xrightarrow[p \to +\infty]{} \mathrm{A}$, d'où le résultat voulu.

Conséquence. D'après la question 1 de l'exercice VII.4, pour tout $u \in \mathcal{L}(\mathrm{E})$, $\chi_u = \mu_u$ implique que u est cyclique. Or si u admet n valeurs propres deux à deux différentes, $\chi_u = \mu_u$ et donc u est cyclique. L'ensemble des endomorphismes ayant n valeurs propres deux à deux différentes (qui est dense) est inclus dans l'ensemble des endomorphismes cycliques. On en déduit donc que l'ensemble des les endomorphismes cycliques est dense dans $\mathcal{M}_n(\mathbb{C})$.

Application. On peut utiliser ce résultat pour redémontrer le théorème de Cayley-Hamilton.
Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ et (A_p) une suite de matrices dont chacune possède n valeurs propres distinctes (et donc, a fortiori, est diagonalisable) tel que $\mathrm{A}_p \xrightarrow[p \to +\infty]{} \mathrm{A}$.

(cc) BY-NC-SA

Lorsque $D \in \mathcal{M}_n(\mathbb{C})$ est diagonale, il est aisé d'établir que $\chi_D(D) = 0$. Idem pour le cas où D est diagonalisable. Ainsi, par continuité de $(A, B) \longmapsto \chi_A(B)$ (l'image de (A, B) est une matrice dont les coordonnées sont produit et somme des coefficients de A et B),

$$\chi_A(A) = \lim_{p\to\infty} \chi_{A_p}(A_p) = \lim_{p\to\infty} 0 = 0.$$

Remarque. On peut également montrer que Ω est ouvert. En effet,

$$A \in \Omega \iff \chi_A \wedge \chi_A' = 1 \iff \det \chi_A'(A) \neq 0.$$

$h : A \longmapsto \det(\chi_A'(A))$ est continue et $\mathbb{K}^*$ est ouvert, donc $\Omega = h^{-1}(\mathbb{K}^*)$ est ouvert.

Exercice VIII.8.

Soit $A \in \mathcal{M}_n(\mathbb{C})$, montrer que

$$\text{A est diagonalisable } \iff \{P^{-1}AP,\ P \in GL_n(\mathbb{C})\} \text{ est fermé.}$$

4. Valeurs propres pures

Définition VIII.9.

$\lambda \in \mathrm{Spec}_{\mathbb{K}}(A)$ est dite pure lorsque

$$E_{\lambda,u} = \mathrm{Ker}(A - \lambda I_n) = \mathrm{Ker}(A - \lambda I_n)^{\alpha_\lambda} = F_{\lambda,u},$$

ou alors d'une manière équivalente $\mathrm{Ker}(A - \lambda I_n) = \mathrm{Ker}(A - \lambda I_n)^2$.

Remarque. Cette propriété est aussi équivalente au fait que la composante nilpotente dans la décomposition de Dunford associée à l'espace caractéristique de λ est nulle (*i.e.* u est simplement une homothétie sur cet espace).

Exercice VIII.10.

Soit λ une valeur propre de $A \in \mathcal{M}_n(\mathbb{C})$ et $\|.\|$ une norme sur $\mathbb{C}^n$. Montrer que si $|\lambda| = |||A|||$, avec $|||.|||$ la norme d'opérateur associée à $\|.\|$, alors λ est pure.

Application. Si $u \in \mathcal{M}_n(\mathbb{C})$ est une isométrie pour $\|.\|$ alors u est diagonalisable. En effet, $\forall \lambda \in \mathrm{Spec}_{\mathbb{C}}(u)$, $|\lambda| = |||u||| = 1$ et donc $E_{\lambda,u} = F_{\lambda,u}$.

(cc) BY-NC-SA

Correction de l'exercice II.5. :

→ Lorsque $\mathbb{K} = \mathbb{R}$, un tel plan existe.
En effet, le plan suivant privé de 0 est inclus dans $\mathrm{GL}_2(\mathbb{R})$

$$\mathrm{P} = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, (a,b) \in \mathbb{R}^2 \right\}.$$

→ Lorsque $\mathbb{K} = \mathbb{C}$, il n'existe pas de tel plan.
En effet, supposons qu'il existe un plan P vérifiant ces conditions et soit (u,v) une base de P et $\lambda \in \mathbb{C}$. On a alors

$$u - \lambda v \in \mathrm{GL}_n(\mathbb{C}) \iff \det(u - \lambda v) \neq 0 \underset{v \in \mathrm{GL}_n(\mathbb{C})}{\iff} \det(uv^{-1} - \lambda \,\mathrm{Id}) \neq 0,$$

ce qui ne peut pas être vrai pour tout $\lambda \in \mathbb{C}$ car uv^{-1} admet une valeur propre dans $\mathbb{C}$.

Correction de l'exercice II.7. :

Soit λ une valeur propre de u. $u - \lambda \,\mathrm{Id}$ commute avec tous les éléments de L et donc $\mathrm{F} = \mathrm{Ker}(u - \lambda id) \neq \{0\}$ est stable par tous les éléments L. L étant irréductible, on a $\mathrm{F} = \mathrm{E}$ et $u = \lambda \,\mathrm{Id}$.

Correction de l'exercice III.5. :

Soit $\beta = (b_1, \dots, b_n)$ une base telle que $[u]_\beta = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ où les λ_i sont les n valeurs propres distinctes de u. v commute avec u et donc pour tout $i \in [\![1;n]\!]$, $\mathrm{Ker}(u - \lambda_i \,\mathrm{Id}) = \mathrm{Vect}(b_i)$ est stable par v et donc pour tout $i \in [\![1;n]\!]$, il existe $\mu_i \in \mathbb{K}$ tel que $v(b_i) = \mu_i b_i$ et donc finalement $[v]_\beta = \mathrm{diag}(\mu_1, \dots, \mu_n)$. On en déduit directement que v est diagonalisable.
En considérant un polynôme interpolateur $\mathrm{P} \in \mathbb{K}[\mathrm{X}]$ tel que $\forall i \in [\![1;n]\!]$, $\mathrm{P}(\lambda_i) = \mu_i$ (chose possible vu que les λ_i sont distincts), on voit que

$$[\mathrm{P}(u)]_\beta = \mathrm{diag}(\mathrm{P}(\lambda_1), \dots, \mathrm{P}(\lambda_n)) = \mathrm{diag}(\mu_1, \dots, \mu_n) = [v]_\beta,$$

d'où le résultat.

Correction de l'exercice IV.3. :

$\mathrm{X}^2 - 1$ est un polynôme annulateur de u, donc $\mu_u | \mathrm{X}^2 - 1$. μ_u n'est pas constant, il y a donc 2 possibilités.

→ $\mu_u = \mathrm{X} - 1$ ou $\mu_u = \mathrm{X} + 1$, ce qui est impossible car $u \neq \pm \,\mathrm{Id}$.

→ $\mu_u(\mathrm{X}) = \mathrm{X}^2 - 1$ est la seule possibilité qui reste, d'où la résultat voulu.

Correction de l'exercice IV.8. :

→ (1) ⇒ (2) Considérons $v = u\Big|_{\mathrm{F}}^{\mathrm{F}}$ (bien défini car F stable par u). $\mathrm{P}(\mathrm{X}) = \prod_{i=1}^{s} (\mathrm{X} - \lambda_s)$ annule v, on a donc par le lemme de décomposition des noyaux,

$$\mathrm{F} = \mathrm{Ker}(\mathrm{P}(v)) = \bigoplus_{i=1}^{s} \mathrm{Ker}(v - \lambda_i \,\mathrm{Id}) = \bigoplus_{i=1}^{s} \mathrm{F} \cap \mathrm{Ker}(u - \lambda_i \,\mathrm{Id}) = \bigoplus_{i=1}^{s} \mathrm{F} \cap \mathrm{E}_{\lambda_i, u}.$$

→ (2) ⇒ (3) Il suffit de prendre pour tout $i \in [\![1;s]\!]$, $\mathrm{F}_i = \mathrm{F} \cap \mathrm{E}_{\lambda_i, u}$.

→ (3) ⇒ (1) Chaque F_i est stable par u, donc F, qui est égal à la somme directe des F_i, est également stable par u.

(cc) BY-NC-SA

Correction de l'exercice IV.9. :

Méthode 1 :

→ Soit $v \in \mathrm{Com}(u)$. Posons
$$[v]_\beta = \begin{pmatrix} \mathrm{A}_{1,1} & \dots & \mathrm{A}_{1,s} \\ \vdots & \ddots & \vdots \\ \mathrm{A}_{s,1} & \dots & \mathrm{A}_{s,s} \end{pmatrix},$$
telle que pour tout $i, j \in [\![1; s]\!]$, $\mathrm{A}_{i,j} \in \mathcal{M}_{\ell_i,\ell_j}(\mathbb{K})$. On a alors
$$\begin{pmatrix} \lambda_1\mathrm{A}_{1,1} & \dots & \lambda_1\mathrm{A}_{1,s} \\ \vdots & \ddots & \vdots \\ \lambda_s\mathrm{A}_{s,1} & \dots & \lambda_s\mathrm{A}_{s,s} \end{pmatrix} = [u]_\beta \times [v]_\beta = [u \circ v]_\beta = [v \circ u]_\beta = [v]_\beta \times [u]_\beta = \begin{pmatrix} \lambda_1\mathrm{A}_{1,1} & \dots & \lambda_s\mathrm{A}_{1,s} \\ \vdots & \ddots & \vdots \\ \lambda_1\mathrm{A}_{s,1} & \dots & \lambda_s\mathrm{A}_{s,s} \end{pmatrix}.$$
On a alors $(\lambda_i \mathrm{A}_{i,j})_{i,j\in[\![1;s]\!]} = (\lambda_j \mathrm{A}_{i,j})_{i,j\in[\![1;s]\!]}$, *i.e.* pour tout $i \neq j$,
$$\lambda_i \mathrm{A}_{i,j} = \lambda_j \mathrm{A}_{i,j},$$
ce qui implique que $\mathrm{A}_{i,j} = 0$ car les λ_i sont deux à deux distincts. v s'écrit donc bien dans la base β sous la forme
$$v = \begin{pmatrix} \mathrm{A}_{1,1} & & 0 \\ & \ddots & \\ 0 & & \mathrm{A}_{s,s} \end{pmatrix}.$$

→ Réciproquement, si on suppose que v s'écrit comme ci-dessus, alors
$$[v \circ u]_\beta = [v]_\beta \times [u]_\beta = \begin{pmatrix} \lambda_1\mathrm{A}_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_s\mathrm{A}_s \end{pmatrix} = [u]_\beta \times [v]_\beta = [u \circ v]_\beta,$$
donc $v \in \mathrm{Com}(u)$.

Méthode 2 :

→ Notons S l'ensemble de droite. Il est alors clair que $\mathrm{S} \subset \mathrm{Com}(u)$ (pour s'en convaincre, voir le dernier point de la méthode 1).

→ Réciproquement, soit $v \in \mathrm{Com}(u)$. Pour tout $i \in [\![1; s]\!]$, $\mathrm{E}_{\lambda_i,u}$ est stable par v. La proposition I.5 nous permet d'affirmer que dans la base β, v s'écrit bien de la manière voulue.

Une autre solution de l'exercice III.5 consiste à utiliser ce résultat pour $s = n$ et $\ell_i = 1$ pour tout $i \in [\![1; s]\!]$.

Remarque. une conséquence de ce résultat est que $\dim \mathrm{Com}(u) = \sum_{i=1}^{s} \ell_i^2 \geq n$ avec égalité si et seulement si $n = s$ et pour tout $i \in [\![1; s]\!]$, $\ell_i = 1$. L'inégalité $\dim_{\mathbb{C}} \mathrm{Com}(u) \geq n$ est en fait toujours vraie même sans diagonalisabilité lorsque $u \in \mathcal{L}(\mathbb{C}^n)$. On le prouvera dans un exercice ultérieur.

Correction de l'exercice IV.11. :

→ (⇐) Dans une base β où les matrices de tous les endomorphismes de S sont de matrice diagonale, toutes ces matrices commutent, ce qui nous donne bien le résultat voulu.

→ (⇒) Procédons par récurrence forte sur la dimension n de E.

- Le cas $n = 1$ est évident, car dans ce cas tous les endomorphismes de E commutent et sont diagonaux.
- Soit $n \geq 1$. Supposons que la propriété est vraie pour tout $k \in [\![1; n]\!]$. Le cas où tous les éléments de S sont des homothéties est évident, on suppose donc que ce n'est pas le cas. On

(cc) BY-NC-SA

dispose donc de $u \in \mathrm{S}$, $s \geq 2$ et $\lambda_1, \ldots, \lambda_s$ ses valeurs propres comptées sans multiplicité. Tout élément $v \in \mathrm{S}$ commute avec u, donc pour tout $i \in [\![1; s]\!]$, $\mathrm{E}_{\lambda_i,u}$ est stable par tout élément de S. En posant $\mathrm{F} = \bigoplus_{i=2}^{s} \mathrm{E}_{\lambda_i,u}$, on voit clairement que F est stable par tout élément de S. Considérons les deux ensembles

$$\mathrm{S}_1 = \left\{ v\Big|_{\mathrm{E}_{\lambda_1,u}}^{\mathrm{E}_{\lambda_1,u}},\ v \in \mathrm{S} \right\} \text{ et } \mathrm{S}_2 = \left\{ v\Big|_{\mathrm{F}}^{\mathrm{F}},\ v \in \mathrm{S} \right\}.$$

On peut appliquer l'hypothèse de récurrence à $\mathrm{E}_{\lambda_1,u}$ pour S_1 et à F pour S_2. Il existe une base β_1 (resp. β_2) de $\mathrm{E}_{\lambda_1,u}$ (resp. F) dans laquelle les matrices de tous les éléments de S_1 (resp. S_2) sont diagonales. On en déduit donc que les matrices de tous les éléments de S dans la base $\beta = (\beta_1, \beta_2)$ sont diagonales, ce qui est bien le résultat recherché.

Correction de l'exercice IV.12. :

Considérons l'ensemble

$$\mathrm{S} = \{\mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_m),\ (\varepsilon_1, \ldots, \varepsilon_m) \in \{-1, 1\}^m\}.$$

S est clairement un sous-groupe de $\mathrm{GL}_m(\mathbb{K})$ de cardinal 2^m (car car $\mathbb{K} \neq 2$).
Posons $\mathrm{S}' := \varphi(\mathrm{S})$. φ est injective, donc S' est un sous-groupe de $\mathrm{GL}_n(\mathbb{K})$ de cardinal 2^m.

Remarquons que pour tout $\mathrm{A} \in \mathrm{S}$, $\mathrm{A}^2 = \mathrm{I}_m$ donc pour tout $\mathrm{A}' \in \mathrm{S}'$, il existe $\mathrm{A} \in \mathrm{S}$ tel que $\mathrm{A}' = \varphi(\mathrm{A})$ et donc $\mathrm{A}'^2 = \varphi(\mathrm{A})^2 = \mathrm{I}_n$. $\mathrm{X}^2 - 1$ annule donc tout élément de S' et est scindé à racines simples dans $\mathbb{K}$ (car car $\mathbb{K} \neq 2$) et donc chaque élément de S' est diagonalisable. De plus, les éléments de S commutent entre eux, donc ceux de S' aussi.

Ainsi, en utilisant l'exercice précédent, on dispose d'une base β de $\mathrm{M}_n(\mathbb{K})$ où la matrice de tout élément de S' est diagonale. Par construction, le carré de chacune de ces matrices est égal à l'identité, ce qui impose que $[\mathrm{S}']_\beta = \{[v]_\beta,\ v \in \mathrm{S}'\} \subset \{\mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_n),\ (\varepsilon_1, \ldots, \varepsilon_n) \in \{-1, 1\}^n\}$ On a alors

$$2^m = |\mathrm{S}'| = |[\mathrm{S}']_\beta| \leq 2^n,$$

i.e. $m \leq n$.

Correction de l'exercice IV.13. :

→ ($\Rightarrow$) Supposons que u^2 et u soient diagonalisables. Il existe une base β et une matrice diagonale $\mathrm{D} \in \mathcal{M}_n(\mathbb{C})$ telle que $[u]_\beta = \mathrm{D}$. De plus $\mathrm{rg}\,\mathrm{D} = \mathrm{rg}\,\mathrm{D}^2$ (il s'agit du nombre de coefficients non nuls sur la diagonale) et donc $\mathrm{rg}\, u = \mathrm{rg}\, u^2$. La formule du rang nous donne donc $\dim \mathrm{Ker}\, u = \dim \mathrm{Ker}\, u^2$. En combinant ce résultat avec le fait que $\mathrm{Ker}\, u \subset \mathrm{Ker}\, u^2$, on obtient $\mathrm{Ker}\, u = \mathrm{Ker}\, u^2$.

→ ($\Leftarrow$) Deux cas se présentent

- u est inversible (*i.e.* $\mathrm{Ker}\, u^2 = \mathrm{Ker}\, u = \{0\}$).
  Considérons $\lambda_1, \ldots, \lambda_s$ les valeurs propres comptées sans multiplicité non nulles de u^2 et $\alpha_1, -\alpha_1, \ldots, \alpha_s, -\alpha_s$ leurs racines complexes distinctes respectives. Le polynôme $\mathrm{P}(\mathrm{X}) = \prod_{i=1}^{s}(\mathrm{X} - \lambda_i)$ annule u^2 et donc le polynôme

$$\mathrm{P}(\mathrm{X}^2) = \prod_{i=1}^{s}(\mathrm{X}^2 - \lambda_i) = \prod_{i=1}^{s}(\mathrm{X} - \alpha_i)(\mathrm{X} + \alpha_i)$$

  scindé à racines simples annule u, donc u est diagonalisable.

- u n'est pas inversible, donc u^2 non plus.

(cc) BY-NC-SA

Reprenons les notations du point précédent et considérons que $\lambda_1 = 0$. Le polynôme

$$\mathrm{P}(\mathrm{X}^2) = \mathrm{X}^2 \prod_{i=2}^{s} (\mathrm{X} - \alpha_i)(\mathrm{X} + \alpha_i)$$

annule u, et donc par le lemme des noyaux

$$\begin{aligned}\mathrm{E} = \operatorname{Ker} \mathrm{P}(u^2) &= \operatorname{Ker} u^2 \oplus \bigoplus_{i=2}^{s} (\operatorname{Ker}(u - \alpha_i \operatorname{Id}) \oplus \operatorname{Ker}(u + \alpha_1 \operatorname{Id})) \\ &= \operatorname{Ker} u \oplus \bigoplus_{i=2}^{s} (\operatorname{Ker}(u - \alpha_i \operatorname{Id}) \oplus \operatorname{Ker}(u + \alpha_1 \operatorname{Id}))\,.\end{aligned}$$

On en déduit que E est somme directe d'espaces propres de u, *i.e.* u est diagonalisable.

Correction de l'exercice V.3. :

Posons $\mathrm{A} = (a_{i,j})_{i,j \in [\![1;n]\!]}$. Pour calculer le polynôme caractéristique de ϕ_A, nous allons calculer sa matrice dans la base

$$\mathcal{B} = (\mathrm{E}_{1,1}, \ldots, \mathrm{E}_{1,n}, \ldots, \mathrm{E}_{n,1}, \ldots, \mathrm{E}_{n,n}).$$

Posons pour tout $i \in [\![1;n]\!]$, $\mathcal{B}_i = (\mathrm{E}_{i,j})_{j \in [\![1;n]\!]}$ et remarquons que $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_n)$.
Soit $i \in [\![1;n]\!]$. On a pour tout $j \in [\![1;n]\!]$, $\phi_\mathrm{A}(\mathrm{E}_{i,j})$ est égal à une matrice de coefficients nuls partout, sauf à la ligne i où se trouve la ligne j de la matrice A. On a donc

$$\phi_\mathrm{A}(\mathrm{E}_{i,j}) = a_{j,1}\mathrm{E}_{i,1} + \cdots + a_{j,n}\mathrm{E}_{i,n} \in \operatorname{Vect} \mathcal{B}_i.$$

On en déduit que pour tout $i \in [\![1;n]\!]$, $\operatorname{Vect} \mathcal{B}_i$ est stable par ϕ_A. Il existe donc des matrices $\mathrm{B}_1, \ldots, \mathrm{B}_n \in \mathcal{M}_n(\mathbb{C})$ telles que

$$[\phi_\mathrm{A}]_\mathcal{B} = \begin{pmatrix} \mathrm{B}_1 & & 0 \\ & \ddots & \\ 0 & & \mathrm{B}_n \end{pmatrix}.$$

Soit $k \in [\![1;n]\!]$. Pour tout $i, j \in [\![1;n]\!]$, le coefficient de position i, j de B_k est égal à le coefficient de $\mathrm{E}_{k,i}$ dans la décomposition de $\phi_\mathrm{A}(\mathrm{E}_{k,j})$ dans la base $\mathcal{B}_k$, qui est égal à $a_{j,i}$. On en déduit donc que pour tout $k \in [\![1;n]\!]$, $\mathrm{B}_k = \mathrm{A}^\mathrm{T}$ et alors

$$\chi_{\phi_\mathrm{A}}(\mathrm{X}) = \det(\mathrm{XI}_{n^2} - [\phi_\mathrm{A}]_\mathcal{B}) = \det \begin{pmatrix} \mathrm{XI}_n - \mathrm{A}^\mathrm{T} & & 0 \\ & \ddots & \\ 0 & & \mathrm{XI}_n - \mathrm{A}^\mathrm{T} \end{pmatrix} = \det \left(\mathrm{XI}_n - \mathrm{A}^\mathrm{T}\right)^n = \chi_\mathrm{A}(\mathrm{X})^n.$$

Correction de l'exercice V.10. :

→ (⇒) Supposons que A soit diagonalisable. Il existe donc $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$ et D une matrice diagonale $\mathrm{D} \in \mathcal{M}_n(\mathbb{K})$ telle que $\mathrm{PAP}^{-1} = \mathrm{D}$. On en déduit donc que pour tout $\lambda \in \mathrm{VP}(\mathrm{A})$,

$$\operatorname{rg}(\mathrm{A} - \lambda \operatorname{Id})^2 = \operatorname{rg} \mathrm{P}(\mathrm{A} - \lambda \operatorname{Id})^2\mathrm{P}^{-1} = \operatorname{rg}(\mathrm{D} - \lambda \operatorname{Id})^2.$$

$\operatorname{rg}(\mathrm{D} - \lambda \operatorname{Id})^2$ est égal au nombre de coefficients diagonaux non nuls dans $(\mathrm{D} - \lambda \operatorname{Id})^2$. Ce nombre est le même que le nombre de coefficients non nuls dans la matrice diagonale $\mathrm{D} - \lambda \operatorname{Id}$ (les deux matrices sont diagonales). On a alors

$$\operatorname{rg}(\mathrm{A} - \lambda\mathrm{I})^2 = \operatorname{rg}(\mathrm{A} - \lambda\mathrm{I}),$$

(cc) BY-NC-SA

et finalement par la formule du rang

$$\dim \operatorname{Ker}(\mathrm{A}-\lambda \mathrm{I})^2 = \dim \operatorname{Ker}(\mathrm{A}-\lambda \mathrm{I}).$$

→ ($\Leftarrow$) Supposons que χ_{A} est scindé et que pour tout $\lambda \in \mathrm{VP}(\mathrm{A})$,

$$\dim \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id}) = \dim \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})^2.$$

Soit $\lambda \in \mathrm{VP}(\mathrm{A})$. On a $\operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id}) \subset \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})^2$, et ces deux sous espaces sont de même dimension, donc $\operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id}) = \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})^2$. nous allons montrer que pour tout $k \in \mathbb{N}^*$, $\operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id}) = \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})^k$. Soit $k \geq 2$ et $\mathrm{X} \in \mathbb{K}^n$, on a

$$\begin{aligned}(\mathrm{A}-\lambda\,\mathrm{Id})^k\mathrm{X} = 0 &\Longrightarrow (\mathrm{A}-\lambda\,\mathrm{Id})^2(\mathrm{A}-\lambda\,\mathrm{Id})^{k-2}\mathrm{X} = 0\\ &\Longrightarrow (\mathrm{A}-\lambda\,\mathrm{Id})(\mathrm{A}-\lambda\,\mathrm{Id})^{k-2}\mathrm{X} = 0\\ &\Longrightarrow (\mathrm{A}-\lambda\,\mathrm{Id})^{k-1}\mathrm{X} = 0.\end{aligned}$$

En itérant ce procédé $k-1$ fois, on obtient que

$$(\mathrm{A}-\lambda\,\mathrm{Id})^k\mathrm{X} = 0 \Longrightarrow (\mathrm{A}-\lambda\,\mathrm{Id})^{k-1}\mathrm{X} = 0 \Longrightarrow \cdots \Longrightarrow (\mathrm{A}-\lambda\,\mathrm{Id})\mathrm{X} = 0.$$

On en déduit donc que $\operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})^k \subset \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})$, *i.e.* $\operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})^k = \operatorname{Ker}(\mathrm{A}-\lambda\,\mathrm{Id})$. Remarquons que ce raisonnement peut aussi être fait par récurrence.
χ_{A} est scindé, on peut donc écrire

$$\chi_{\mathrm{A}}(\mathrm{X}) = \prod_{i=1}^{s}(\mathrm{X}-\lambda_i)^{\alpha_i},$$

où $\lambda_1, \ldots, \lambda_s \in \mathbb{K}$ sont les valeurs propres de A. On a donc par le lemme des noyaux

$$\mathbb{K}^n = \operatorname{Ker}\chi_{\mathrm{A}}(\mathrm{A}) = \bigoplus_{i=1}^{s}\operatorname{Ker}(\mathrm{A}-\lambda_i\,\mathrm{Id})^{\alpha_i} = \bigoplus_{i=1}^{s}\operatorname{Ker}(\mathrm{A}-\lambda_i\,\mathrm{Id}).$$

$\mathbb{K}^n$ est donc somme directe des espaces propres de A *i.e.* A est diagonalisable.

Correction de l'exercice VI.5. :

→ (1) ⇒ (2) Supposons que u est nilpotent. Pour tout $k \geq 1$, la seule valeur propre de u^k dans $\mathbb{C}$ est 0 et $\operatorname{Tr}(u^k)$ est égal à la somme des valeurs propres de u^k, et donc pour tout $k \in [\![1;n]\!]$, $\operatorname{Tr}(u^k) = 0$.

→ (2) ⇒ (1) Supposons que pour tout $k \in [\![1;n]\!]$, $\operatorname{Tr}(u^k) = 0$. Soit $\lambda_1, \ldots, \lambda_s$ les coefficients diagonaux sans répétition et sans coefficients nuls de la matrice trigonalisée de u et soit $n_1, \ldots, n_s \in [\![1;n]\!]$ le nombre respectifs d'occurrences de $\lambda_1, \ldots, \lambda_s$ dans cette diagonale. supposons que $s \geq 1$. La condition $\operatorname{Tr}(u^k) = 0$ pour tout $k \in [\![1;s]\!]$ donne

$$\begin{cases} n_1\lambda_1 + \cdots + n_s\lambda_s &= 0\\ n_1\lambda_1^2 + \cdots + n_s\lambda_s^2 &= 0\\ \quad\vdots & \\ n_1\lambda_1^s + \cdots + n_s\lambda_s^s &= 0,\end{cases}$$

i.e.

$$\begin{pmatrix}\lambda_1 & \ldots & \lambda_s\\ \lambda_1^2 & \ldots & \lambda_s^2\\ \vdots & \ddots & \vdots\\ \lambda_1^s & \ldots & \lambda_s^s\end{pmatrix}\begin{pmatrix}n_1\\ n_2\\ \vdots\\ n_s\end{pmatrix} = \begin{pmatrix}0\\0\\ \vdots\\0\end{pmatrix}.$$

(cc) BY-NC-SA

En posant $\mathrm{M} = \begin{pmatrix} \lambda_1 & \dots & \lambda_s \\ \lambda_1^2 & \dots & \lambda_s^2 \\ \vdots & \ddots & \vdots \\ \lambda_1^s & \dots & \lambda_s^s \end{pmatrix}$, on voit que

$$\begin{aligned}\det \mathrm{M} = \lambda_1 \dots \lambda_s \det \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_s \\ \vdots & \ddots & \vdots \\ \lambda_1^{s-1} & \dots & \lambda_s^{s-1} \end{pmatrix} &= \lambda_1 \dots \lambda_s \det \mathrm{V}(\lambda_1, \dots, \lambda_s)^{\mathrm{T}} \\ &= \lambda_1 \dots \lambda_s \prod_{i,j \in [\![1;n]\!],\ i \neq j} (\lambda_i - \lambda_j) \neq 0,\end{aligned}$$

où $\mathrm{V}(\lambda_1, \dots, \lambda_s)$ est la matrice de Vandermonde associée à $\lambda_1, \dots, \lambda_s$. Ce dernier nombre est non nul car les λ_i sont tous non nuls et distincts. La formule de ce déterminant est assez classique, nous ne le démontrerons donc pas (une récurrence suffit pour le faire).
M est donc inversible et alors $n_1 = \dots = n_s = 0$, ce qui est absurde. On a alors $s = 1$, *i.e.* les coefficients diagonaux de la matrice trigonalisée de u sont donc tous égaux à 0.

→ (2) ⇒ (3) Supposons que (2) est vérifiée. u est alors nilpotente et le coefficient de nilpotence de u est inférieur à n, donc pour tout $k \in \mathbb{N}^*$, $\mathrm{Tr}(u^k) = 0$. Soit $m \in \mathbb{N}$ et $\mathrm{P}(\mathrm{X}) = \sum_{i=0}^{m} a_i \mathrm{X}^i \in \mathbb{K}[\mathrm{X}]$. On a

$$\mathrm{Tr}(\mathrm{P}(u)) = \sum_{i=0}^{m} a_i \,\mathrm{Tr}(u^i) = a_0 \,\mathrm{Tr}(\mathrm{Id}) = n\mathrm{P}(0).$$

→ (3) ⇒ (2) Supposons que (3) est vérifiée, on a alors pour tout $k \in \mathbb{N}^*$, en posant $\mathrm{P}(\mathrm{X}) = \mathrm{X}^k$,

$$\mathrm{Tr}(u^k) = \mathrm{Tr}(\mathrm{P}(u)) = n\mathrm{P}(0) = 0.$$

Correction de l'exercice VI.6. :

Dans cet exercice, on suppose que $\mathbb{K} = \mathbb{C}$.

1. Faisons une discussion des cas pour $n = 2$ puis $n = 3$.

   → Pour $n = 2$.

   - Si $\deg \mu_{\mathrm{A}} = 1$, alors il existe $\lambda \in \mathbb{K}$, tel que $\mu_{\mathrm{A}}(\mathrm{X}) = \mathrm{X} - \lambda$, et donc $\mathrm{A} = \lambda \mathrm{I}$.
   - Si $\deg \mu_{\mathrm{A}} = 2$, alors il y a deux possibilités.
     - ▷ Il existe $\lambda, \delta \in \mathbb{K}$ différents tels que $\chi_{\mathrm{A}}(\mathrm{X}) = \mu_{\mathrm{A}}(\mathrm{X}) = (\mathrm{X} - \lambda)(\mathrm{X} - \delta)$. χ_{A} est scindé à racines simples, donc A est diagonalisable et $\mathrm{A} \simeq \mathrm{diag}(\lambda, \delta)$.
     - ▷ Il existe $\lambda \in \mathrm{K}$ tel que $\chi_{\mathrm{A}}(\mathrm{X}) = \mu_{\mathrm{A}}(\mathrm{X}) = (\mathrm{X} - \lambda)^2$. En trigonalisant, on voit que $\mathrm{A} \simeq \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix}$ dans une base $\beta = (b_1, b_2)$. De plus $a \neq 0$, car sinon on aurait $\mu_{\mathrm{A}}(\mathrm{X}) = \mathrm{X} - \lambda$.
       Enfin, on voit qu'en regardant la matrice A dans la base (ab_1, b_2), $\mathrm{A} \simeq \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

   On en déduit donc que les classes de similitudes de $\mathcal{M}_2(\mathbb{K})$ sont

$$\mathcal{C}_2 = \left\{ \left\langle \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \right\rangle,\ \lambda \in \mathbb{K} \right\} \cup \left\{ \left\langle \begin{pmatrix} \lambda & 0 \\ 0 & \delta \end{pmatrix} \right\rangle,\ \lambda, \delta \in \mathbb{K} \right\}.$$

   → Pour $n = 3$.

(cc) BY-NC-SA

- S'il existe $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{K}$ distincts tels que $\chi_A(X) = (X - \lambda_1)(X - \lambda_2)(X - \lambda_3)$, alors A est diagonalisable, $\mu_A = \chi_A$ et $A \simeq \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$.
- S'il existe $\lambda, \delta \in \mathbb{K}$ différents tels que $\chi_A(X) = (X-\lambda)^2(X-\delta)$, alors deux cas se présentent.
  - ▷ Si $\mu_A(X) = (X - \lambda)(X - \delta)$, alors A est diagonalisable et $A \simeq \operatorname{diag}(\lambda, \lambda, \delta)$.
  - ▷ Si $\mu_A = \chi_A$, alors on a
  $$\mathbb{K}^n = \operatorname{Ker}(A - \lambda I)^2 \oplus \operatorname{Ker}(A - \delta I).$$
  Il existe donc une base où $A \simeq \begin{pmatrix} B & 0 \\ 0 & \delta \end{pmatrix}$ où $B \in \mathcal{M}_2(\mathbb{K})$ est la matrice de la corestriction de A à $\operatorname{Ker}(A - \delta \operatorname{Id})^2$. On a $\mu_B(X) = (X - \lambda)^2$, donc d'après la discussion pour $n = 2$,
  $$B \simeq \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \text{ et finalement } A \simeq \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \delta \end{pmatrix}.$$
- S'il existe $\lambda \in \mathbb{K}$ tel que $\chi_A(X) = (X - \lambda)^3$, alors on envisage trois cas.
  - ▷ Si $\mu_A(X) = X - \lambda$, alors $A \simeq \lambda I$.
  - ▷ Si $\mu_A(X) = (X - \lambda)^2$ alors on peut écrire $A = \lambda I + N$ avec N nilpotente d'indice de nilpotence égal à 2. Soit $X \in \mathbb{K}^3$ tel que $NX \neq 0$. la famille (NX, X) est libre (raisonnement déjà fait avant). Complétons cette famille en une base $\beta = (NX, X, Y)$ avc $Y \in \mathbb{K}^n$. On a alors, dans cette base
  $$[N]_\beta = \begin{pmatrix} 0 & 1 & a \\ 0 & 0 & b \\ 0 & 0 & c \end{pmatrix},$$
  avec $a, b, c \in \mathbb{K}$. c est une valeur propre de N, donc on a nécessairement $c = 0$. De plus, on a
  $$0 = N^2Y = N(NY) = N(aNX + bX) = bNX,$$
  donc $b = 0$. Enfin, en regardant N dans la base $\beta' = (NX, X, Y - aX)$, on voit que
  $$[N]_{\beta'} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et donc finalement } A \simeq \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}.$$
  - ▷ Si $\mu_A(X) = \chi_A(X) = (X - \lambda)^3$, alors on peut écrire $A = \lambda I + N$ avec N d'ordre de nilpotence égal à 3. En prenant donc $X \in \mathbb{K}^3$ tel que $N^2X \neq 0$, on voit que $\beta = (N^2X, NX, X)$ est une base de $\mathbb{K}^3$ et dans cette base
  $$[N]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ et donc finalement } A \simeq \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

On en déduit donc que les classes de similitude dans $\mathcal{M}_3(\mathbb{K})$ sont

$$\mathcal{C}_3 = \left\{ \left\langle \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \right\rangle,\ \lambda_1, \lambda_2, \lambda_3 \in \mathbb{K} \right\} \cup \left\{ \left\langle \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \delta \end{pmatrix} \right\rangle,\ \lambda, \delta \in \mathbb{K} \right\} \cup \left\{ \left\langle \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \right\rangle,\ \lambda \in \mathbb{K} \right\}.$$

2. Il suffit de revoir la discussion ci-dessus pour répondre à la question.

(cc) BY-NC-SA

3. On peut trouver le contre-exemple suivant pour $n = 4$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ et } B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

En effet, $\mu_A = \mu_B = X^2$ et $\chi_A = \chi_B = X^4$ mais $A \not\simeq B$ car $\mathrm{rg}(A) = 1$ et $\mathrm{rg}(B) = 2$.

Correction de l'exercice VI.7. :

Procédons par récurrence forte sur la dimension n de E.

→ Le cas $n = 1$ est évident car dans ce cas tous les endomorphismes de E commutent et sont triangulaires supérieurs dans toute base.

→ Soit $n \geq 1$. Supposons que la propriété est vraie pour tout $k \in [\![1; n]\!]$.
Soit $u \in S$ non homothétie (le cas $S \subset \mathbb{K} \cdot \mathrm{Id}$ étant trivial) et soit λ_1 une de ses valeurs propres. Posons β_1 une base de $E_{\lambda_1,u}$, de dimension $a < n$, qui est stable par tout élément de S. Prenons maintenant F un supplémentaire quelconque de $E_{\lambda_1,u}$ dans E de dimension b et de base β_2. Tous les éléments de S sont triangulaires supérieurs par blocs dans $\beta = (\beta_1, \beta_2)$. En particulier, pour tout $u \in S$, il existe $(A_u, B_u, C_u) \in \mathcal{M}_a(\mathbb{K}) \times \mathcal{M}_{a\times b}(\mathbb{K}) \times \mathcal{M}_b(\mathbb{K})$ tels que

$$[u]_\beta = \begin{pmatrix} A_u & B_u \\ 0 & C_u \end{pmatrix}.$$

On considère alors les deux ensembles

$$S_1 = \{A_v,\ v \in S\} \text{ et } S_2 = \{C_s,\ v \in S\}.$$

On peut appliquer l'hypothèse de récurrence à S_1 et S_2 (les matrices commutent toujours et sont toujours trigonalisables et de plus $1 \leq a \leq n-1$ et $1 \leq b \leq n-1$). Il existe donc une matrice inversible $P \in GL_a(\mathbb{K})$ (resp. $Q \in GL_b(\mathbb{K})$) telle que pour tout $s \in S$, $P^{-1}A_sP \in \mathcal{T}_a^+(\mathbb{K})$ (resp. $Q^{-1}C_sQ \in \mathcal{T}_b^+(\mathbb{K})$). Finalement, pour la base γ correspondante au changement de base (depuis β) induit par

$$R = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}.$$

Il est clair que $[u]_\gamma = R^{-1}[u]_\beta R$ est triangulaire supérieure pour tout $u \in S$, ce qui est bien le résultat voulu.

Correction de l'exercice VI.10. :

1. $X^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}$ dans $\mathcal{M}_3(\mathbb{R})$.

Soit X une solution de l'équation. X^2 et X commutent, donc X laisse stable les espaces propres de X^2, *i.e.* X laisse stable $\mathrm{Vect}(e_1), \mathrm{Vect}(e_2), \mathrm{Vect}(e_3)$ avec (e_1, e_2, e_3) la base canonique de $\mathbb{R}^3$. On en déduit donc que X est diagonale, *i.e.* il existe $\alpha, \beta, \gamma \in \mathbb{R}$ tels que

$$X = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}.$$

En réinjectant dans l'équation, on voit qu'on a nécessairement $(\alpha^2, \beta^2, \gamma^2) = (1, 4, 9)$ et donc $(\alpha, \beta, \gamma) = (\pm 1, \pm 2, \pm 3)$. Réciproquement, une matrice de cette forme vérifie bien l'équation, donc

(cc) BY-NC-SA

l'ensemble des solutions est bien

$$\mathrm{S} = \left\{ \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix},\ (\alpha,\beta,\gamma) \in \{1,-1\} \times \{2,-2\} \times \{3,-3\} \right\}.$$

2. $\mathrm{X}^2 = \begin{pmatrix} a^2 & 1 \\ 0 & a^2 \end{pmatrix}$ dans $\mathcal{M}_2(\mathbb{C})$ avec $a \in \mathbb{C}$.

Soit X une solution de l'équation. $\mathrm{Vect}(e_1)$ est un espace propre de X^2. De plus X et X^2 commutent, donc X laisse stable $\mathrm{Vect}(e_1)$. Il existe donc $\alpha, \beta, \gamma \in \mathbb{C}$ tels que

$$\mathrm{X} = \begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix}.$$

On a alors

$$\begin{pmatrix} \alpha^2 & \alpha\beta + \gamma\beta \\ 0 & \gamma^2 \end{pmatrix} = \mathrm{X}^2 = \begin{pmatrix} a^2 & 1 \\ 0 & a^2 \end{pmatrix}.$$

En utilisant cette inégalité, on voit que $a \neq 0$, car sinon $\alpha = \gamma = \beta = 0$ et alors $\mathrm{X} = 0$. En elevant les cas où l'égalité est fausse, on obtient que $(\alpha, \beta, \gamma) \in \left\{ \left(a, \frac{1}{2a}, a\right), \left(-a, -\frac{1}{2a}, -a\right) \right\}$. Réciproquement, en réinjectant ces deux matrices possibles dans l'équation, on voit que elles sont bien solutions. On en déduit que l'ensemble des solutions est bien

$$\mathrm{S} = \left\{ \begin{pmatrix} a & \frac{1}{2a} \\ 0 & a \end{pmatrix}, \begin{pmatrix} -a & \frac{-1}{2a} \\ 0 & -a \end{pmatrix} \right\}.$$

3. $\mathrm{X}^2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ dans $\mathcal{M}_3(\mathbb{R})$.

Soit X une solution de l'équation. Posons $\beta = (e_3, e_1, e_2)$ une permutation de la base canonique et regardons la matrice X dans cette base. Posons $\mathrm{Y} = [\mathrm{X}]_\beta$. On a

$$\mathrm{Y}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Y commute avec Y^2 et laisse donc ses espaces propres stables, donc Y laisse stable

$$\mathrm{Vect}\left((1,0,0)^{\mathrm{T}}, (0,1,0)^{\mathrm{T}}\right).$$

Il existe donc $x, y, z, t, c, d, e \in \mathbb{R}$ tels que

$$\mathrm{Y} = \begin{pmatrix} x & y & c \\ z & t & d \\ 0 & 0 & e \end{pmatrix}.$$

En posant $\mathrm{A} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$, on voit que $\mathrm{A}^2 = \mathrm{I}$, donc A est inversible. On a alors

$$\begin{pmatrix} \mathrm{A}^2 & (\mathrm{A} + e\mathrm{I}) \times \begin{pmatrix} c \\ d \end{pmatrix} \\ 0 & e^2 \end{pmatrix} = \mathrm{Y}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

(cc) BY-NC-SA

donc

$$\mathrm{A}^2 = \mathrm{I} \qquad (\mathrm{A} + e\mathrm{I}) \times \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad e^2 = 1.$$

La suite de l'exercice est laissée au lecteur : il suffit d'évaluer les différents cas possible pour trouver toutes les solutions de l'équation.

Correction de l'exercice VI.11. :

→ ($\Rightarrow$) Supposons que u est diagonalisable. Soit $\beta = (e_1, \ldots, e_n)$ une base de diagonalisation de u. Il existe donc $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tels que $[u]_\beta = \mathrm{D} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Pour tout $i, j \in [\![1;n]\!]$, on considère $v_{i,j} \in \mathcal{L}(\mathrm{E})$ définie par

$$\forall k \in [\![1;n]\!],\ v_{i,j}(e_k) = \begin{cases} e_i & \text{si } k = j \\ 0 & \text{sinon.} \end{cases}$$

On a alors, pour tout $i, j \in [\![1;n]\!]$,

$$\begin{aligned} [\phi_u(v_{i,j})]_\beta &= [u \circ v_{i,j} - v_{i,j} \circ u]_\beta \\ &= \mathrm{DE}_{i,j} - \mathrm{E}_{i,j}\mathrm{D} \\ &= (\lambda_i - \lambda_j)\mathrm{E}_{i,j} = [(\lambda_i - \lambda_j)v_{i,j}]_\beta \end{aligned}$$

On en déduit donc que $(v_{i,j})_{i,j\in[\![1;n]\!]}$ est une base de vecteurs propres de ϕ_u, donc ϕ_u est diagonalisable.

→ ($\Leftarrow$) Supposons que ϕ_u est diagonalisable. Utilisons la décomposition de Dunford. Sois $\delta, \nu \in \mathcal{L}(\mathrm{E})$ tels que δ est diagonalisable, ν est nilpotent et $u = \delta + \nu$. δ et ν commutent, donc il est facile de vérifier que ϕ_δ et ϕ_ν aussi. De plus, d'après le point précédent, ϕ_δ est diagonalisable et pour tout $p \in \mathbb{N}$ et $v \in \mathcal{L}(\mathrm{E})$,

$$\phi_\nu^p(v) = \sum_{k=0}^{p} \binom{p}{k} (-1)^{p-k} \nu^k \circ v \circ \nu^{p-k},$$

donc ϕ_ν est nilpotent. ϕ_u est diagonalisable, donc par unicité de la décomposition de Dunford, $\phi_\nu = 0$, *i.e.* $\nu = 0$ et donc u est diagonalisable.

Correction de l'exercice VI.12. :

En décomposant sur chaque espace caractérisique (proposition VI.8), on sait que E s'écrit sous forme de somme directe de sous espaces (espaces caractéristiques) $\mathrm{F}_1, \ldots, \mathrm{F}_s$ stables par u et où pour tout i, il existe ν_i nilpotent et $\lambda_i \in \mathbb{K}^*$ tels que

$$u\Big|_{\mathrm{F}_i}^{\mathrm{F}_i} = \lambda_i \operatorname{Id} + \nu_i.$$

Il suffit donc de montrer ce résultat à l'application linéaire ci-dessus. Nous allons nous inspirer du développement limité en 0 de $x \longmapsto \sqrt{1+x}$. On a

$$\sqrt{1+x} = 1 + \frac{x}{2} + \cdots + (-1)^{n-1} \frac{(2n)!}{2^{2n}(n!)^2} x^n + o(x^n) = \mathrm{P}_n(x) + o(x^n).$$

On a alors

$$\mathrm{P}_n(x)^2 = 1 + x + \underbrace{o(x^n)}_{\mathrm{Q}(x)},$$

où Q est un polynôme de terme de plus petit degré égal à au moins n. On peut donc écrire $\mathrm{Q(X)} = \mathrm{X}^n\mathrm{R(X)}$ avec $\mathrm{R} \in \mathbb{K}[\mathrm{X}]$. On a alors, si α_i est une racine de λ_i, alors

$$\left(\alpha_i \mathrm{P}_n\left(\frac{1}{\lambda_i}\nu_i\right)\right)^2 = \lambda_i \operatorname{Id} + \nu_i + \frac{1}{\lambda_i^{n-1}} \nu_i^n \mathrm{R}\left(\frac{1}{\lambda_i}\nu_i\right) = \lambda_i \operatorname{Id} + \nu_i,$$

(cc) BY-NC-SA

et donc $\alpha_i \mathrm{P}_n\left(\frac{1}{\lambda_i}\nu_i\right)$ est une racine de $\lambda_i \operatorname{Id} + \nu_i$, d'où le résultat voulu.

Correction de l'exercice VI.15. :

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$. En utilisant la décomposition de Jordan, on sait qu'il existe une base β, $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ non nécessairement distincts et $s_1, \dots, s_r \in [\![1; n]\!]$ tels que $s_1 + \cdots + s_r = n$ et

$$[\mathrm{A}]_\beta = \begin{pmatrix} \lambda_1 \cdot \mathrm{I}_{s_1} + \mathrm{J}_{s_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_r \cdot \mathrm{I}_{s_i} + \mathrm{J}_{s_i} \end{pmatrix}.$$

Pour tout $i \in [\![1; r]\!]$, on considère l'application linéaire injective $\phi_i : \mathcal{M}_{s_i}(\mathbb{C}) \longrightarrow \mathcal{M}_n(\mathbb{C})$ définie par

$$\forall \mathrm{B} \in \mathcal{M}_{s_i}(\mathbb{C}),\ \phi_i(\mathrm{B}) = \begin{pmatrix} 0_{\mathcal{M}_{s_1}(\mathbb{C})} & & & & & & 0 \\ & \ddots & & & & & \\ & & 0_{\mathcal{M}_{s_{i-1}}(\mathbb{C})} & & & & \\ & & & \mathrm{B} & & & \\ & & & & 0_{\mathrm{M}_{s_{i+1}}(\mathbb{C})} & & \\ & 0 & & & & \ddots & \\ & & & & & & 0_{\mathcal{M}_{s_r}(\mathbb{C})} \end{pmatrix}.$$

On a alors

$$\operatorname{Com}(\mathrm{A}) \approx \operatorname{Com}([\mathrm{A}]_\beta) \supset \bigoplus_{i=1}^{r} \phi_i(\operatorname{Com}(\lambda_i \mathrm{I}_{s_i} + \mathrm{J}_{s_i})) = \bigoplus_{i=1}^{r} \phi_i(\operatorname{Com}(\mathrm{J}_{s_i})).$$

Il suffit de montrer que pour tout $s \in [\![1; n]\!]$, $\dim \operatorname{Com}(\mathrm{J}_s) \geq s$ car une fois cela fait, on aurait

$$\dim \operatorname{Com}(\mathrm{A}) = \dim \operatorname{Com}([\mathrm{A}]_\beta) \geq \sum_{i=1}^{r} \dim \operatorname{Com}(\mathrm{J}_{s_i}) \geq s_1 + \cdots + s_r = n.$$

Soit $s \in [\![1; n]\!]$. En posant $\mathrm{X} = (1, 0, \dots, 0)^{\mathrm{T}} \in \mathbb{K}^s$, on a

$$\mathrm{J}_s^{s-1}\mathrm{X} = (0, \dots, 0, 1)^{\mathrm{T}} \neq 0,$$

donc l'indice de nilpotence de J_s est égal à s, *i.e.* $\mu_{\mathrm{J}_s}(\mathrm{X}) = \mathrm{X}^s$. En utilisant ce résultat, il est aisé de montrer que $\dim \mathbb{K}[\mathrm{J}_s] = s$ ($(\mathrm{I}, \mathrm{J}_s, \dots, \mathrm{J}_s^{s-1})$ est une base de cet espace). Enfin, puisque $\mathbb{K}[\mathrm{J}_s] \subset \operatorname{Com}(\mathrm{J}_s)$, on a

$$\dim \operatorname{Com}(\mathrm{J}_s) \geq \dim \mathbb{K}[\mathrm{J}_s] = s,$$

d'où le résultat voulu.

Correction de l'exercice VII.5. :

1. Soit $u \in \mathcal{L}(\mathrm{E})$.
   - $\rightarrow$ ($\Leftarrow$) Supposons que u est cyclique. Il existe donc $x_0 \in \mathrm{E}$ tel que $(x_0, u(x_0), \dots, u^{n-1}(x_0))$ est une base de E et alors $\deg \mu_u \geq n$ (car tout polynôme de degré inférieur ou égal à $n-1$ ne peut pas annuler u). De plus, $\mu_u | \chi_u$, $\deg \chi_u = n$ et les deux polynômes sont unitaires, ce qui nous permet d'affirmer que $\chi_u = \mu_u$.
   - $\rightarrow$ ($\Rightarrow$) Soit $x_0 \in \mathrm{E}$ tel que $\mu_{x_0,u} = \mu_u$ (existe d'après l'énoncé). On a alors $\chi_u = \mu_u = \mu_{x_0,u}$, donc $\deg \mu_{x_0,u} = \deg \chi_u = n$. On a alors pour tout $\lambda_0, \dots, \lambda_{n-1} \in \mathbb{K}^n$,

   $$\lambda_0 x_0 + \lambda_1 u(x_0) + \cdots + \lambda_{n-1} u^{n-1}(x_0) = 0 \Longrightarrow \lambda_0 = \cdots = \lambda_{n-1} = 0,$$

   car sinon x_0 serait annulé en u par un polynôme de degré strictement inférieur à n. On en déduit donc que $(x_0, u(x_0), \dots, u^{n-1}(x_0))$ est une base de E, *i.e.* u est cyclique. On aurait aussi

(cc) BY-NC-SA

pu conclure en utilisant la proposition VII.4 (la remarque en dessous en particulier).

2. Supposons que $u \in \mathcal{L}(\mathrm{E})$ est cyclique. Montrons par double inclusion que $\mathrm{Com}(u) = \mathbb{K}[u]$.
   → ($\supset$) Cette inclusion est évidente car tout polynôme en u commute avec u.
   → ($\subset$) Soit $v \in \mathrm{Com}(u)$ et x_0 tel que $\beta = (x_0, u(x_0), \dots, u^{n-1}(x_0))$ est une base de E. Ecrivons la décomposition de $v(x_0)$ dans la base β

$$v(x_0) = a_0 x_0 + a_1 u(x_1) + \cdots + a_{n-1} u^{n-1}(x_0),$$

et considérons l'application linéaire w définie par

$$w = v - \sum_{k=0}^{n-1} a_k u^k.$$

w commute avec u et $w(x_0) = 0$. De plus, on a pour tout $k \in [\![1; n-1]\!]$,

$$w(u^k(x_0)) = w \circ u^k(x_0) = u^k(w(x_0)) = 0.$$

w est nul sur la base β, donc $w = 0$ et donc $v \in \mathbb{K}[u]$ et alors $\mathrm{Com}(u) = \mathbb{K}[u]$. De plus, $(\mathrm{Id}, u, \dots, u^{n-1})$ est une base de $\mathbb{K}[u]$ (pour le montrer, il suffit d'effectuer la division euclidienne de tout polynôme de $\mathbb{K}[\mathrm{X}]$ en u par μ_u) et donc

$$\dim \mathrm{Com}(u) = \dim \mathbb{K}[u] = n.$$

3. Déjà fait dans l'exercice VI.15.

4. Soit $u \in \mathcal{L}(\mathrm{E})$. Procédons par double implication.
   → ($\Rightarrow$) Si $\chi_u = \mu_u$, alors d'après la question 1, u est cyclique et donc d'après la question 2, $\mathrm{Com}(u) = \mathbb{K}[u]$.
   → ($\Leftarrow$) Si $\mathrm{Com}(u) = \mathbb{K}[u]$, alors

$$\deg \mu_u = \dim \mathbb{K}[u] = \dim \mathrm{Com}(u) \geq n.$$

La dernière inégalité est vraie d'après la question précédente. On a de plus $\mu_u | \chi_u$ et ces deux polynômes sont unitaires de même degré, donc $\chi_u = \mu_u$.

Correction de l'exercice VII.6. :

L'énoncé est équivalent à

$$\chi_u \text{ non irréductible} \iff \text{il existe un sous-espace vectoriel de E non trivial stable par } u.$$

→ ($\Rightarrow$) Supposons que χ_u n'est pas irréductible. Soit P un terme irréductible de la décomposition en facteurs irréductibles de χ_u tel que $\deg \mathrm{P} \in [\![1; n-1]\!]$. D'après le corollaire V.8, on sait que $\mathrm{P}|\mu_u$. Soit $x \in \mathrm{Ker}\, \mathrm{P} \setminus \{0\} \neq \varnothing$. $\mathrm{F}_{x,u}$ est stable par u et

$$1 \underset{x \in \mathrm{F}_{x,u}}{\leq} \dim \mathrm{F}_{x,u} \underset{(*)}{=} \deg \mu_{x,u} \underset{\substack{\uparrow \\ \text{Proprosition VII.2}}}{\leq} \deg \mathrm{P} \leq n-1.$$

L'inégalité $(*)$ est vraie d'après la proposition VII.4. $\mathrm{F}_{x,u}$ est donc un sous-espace vectoriel de E non trivial stable par u.

→ ($\Leftarrow$) Soit F un sous-espace vectoriel de E non trivial stable par u et posons $v = u\Big|_{\mathrm{F}}^{\mathrm{F}}$. On a alors $\chi_v | \chi_u$ et $\deg \chi_v \in [\![1; n-1]\!]$.

(cc) BY-NC-SA

Correction de l'exercice VII.7. :

1. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$. Supposons que A est semblable à une matrice compagnon. Il existe donc une base β de $\mathbb{K}^n$ et $a_0, \ldots, a_{n-1} \in \mathbb{C}$ tels que
$$[\mathrm{A}]_\beta = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}.$$
On a alors pour tout $\lambda \in \mathbb{C}$,
$$[\mathrm{A} - \lambda \operatorname{Id}]_\beta = \begin{pmatrix} -\lambda & & & -a_0 \\ 1 & \ddots & & \vdots \\ & \ddots & -\lambda & \vdots \\ & & 1 & -a_{n-1} - \lambda \end{pmatrix}.$$
Le sous bloc de $[\mathrm{A} - \lambda \operatorname{Id}]_\beta$ regrouppant les colonnes de position 1 à $n-1$ est égal à
$$\mathrm{B} = \begin{pmatrix} -\lambda & & 0 \\ 1 & \ddots & \\ & \ddots & -\lambda \\ & & 1 \end{pmatrix}.$$
Les colonnes de B sont libres, donc $\operatorname{rg}(\mathrm{A} - \lambda \operatorname{Id}) \geq n-1$, *i.e.* par la formule du rang $\dim \operatorname{Ker}(\mathrm{A} - \lambda \operatorname{Id}) \leq 1$. En particulier, lorsque λ est une valeur propre de A, $\dim \operatorname{Ker}(\mathrm{A} - \lambda \operatorname{Id}) = 1$.
2. Procédons par double implication. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$ diagonalisable.
   - → ($\Rightarrow$) Si A est cyclique, alors A est semblable à une matrice compagnon et donc d'après la question précédente, tous les espaces propres de A sont de dimension 1, *i.e.* toutes les valeurs propres de A sont de multiplicité 1, il y en a donc n.
   - → ($\Leftarrow$) Supposons que A admet n valeurs propres deux à deux distinctes. Soit $\beta = (e_1, \ldots, e_n)$ une base de diagonalisation de A. Il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ deux à deux distincts tels que
   $$[\mathrm{A}]_\beta = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$
   Soit $x = e_1 + \cdots + e_n$. On a pour tout $k \in \mathbb{N}$,
   $$u^k(x) = \lambda_1^k e_1 + \cdots + \lambda_n^k e_n.$$
   On a alors
   $$\left[(x, u(x), \ldots, u^{n-1}(x))\right]_\beta = \begin{pmatrix} 1 & \lambda_1 & \ldots & \lambda_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \ldots & \lambda_n^{n-1} \end{pmatrix}.$$
   Cette matrice est une matrice de Vandermonde associée aux $\lambda_1, \ldots, \lambda_n$ qui sont deux à deux distincts, elle est donc inversible. La famille $(x, u(x), \ldots, u^{n-1}(x))$ est donc libre et alors c'est une base de $\mathbb{K}^n$. Finalement, A est cyclique, *i.e.* semblable à une matrice compagnon.

(cc) BY-NC-SA

Correction de l'exercice VII.8. :

1. Considérons $C_P \in M_n(\mathbb{Z})$ la matrice compagnon associée à P. On sait que

$$\mathrm{Spec}_{\mathbb{C}}(C_P) = [\lambda_1, \dots, \lambda_n] \text{ et donc } \mathrm{Spec}_{\mathbb{C}}(C_P^k) = [\lambda_1^k, \dots, \lambda_n^k].$$

$\lambda_1^k, \dots, \lambda_n^k$ sont exactement les racines (avec multiplicité) de $\chi_{C_P^k}$. De plus $C_P \in \mathcal{M}_n(\mathbb{Z})$, donc $C_P^k \in \mathcal{M}_n(\mathbb{Z})$ et alors $\chi_{C_P^k} \in \mathbb{Z}[X]$ car $k \geq 1$. On a alors $\chi_{C_P^k} \in \mathbb{Z}_n[X]$ et est unitaire. Le polynôme $\chi_{C_P^k}$ convient donc.

2. Supposons que $|\lambda_1| = \cdots = |\lambda_n| = 1$ et posons pour tout $k \geq 1$,

$$P_k = \prod_1^n (X - \lambda_i^k) = X^n + \sum_{i=0}^{n-1} a_{k,i} X^k \in \mathbb{Z}[X].$$

(existe d'après la question précédente). On veut montrer que les coefficients des polynômes P_k sont uniformément bornés. Deux méthodes sont possibles.

→ **Méthode 1 (Topologie) :** On pour tout $k \in \mathbb{N}^*$,

$$\forall x \in [0,1],\ |P_k(x)| = \prod_{i=1}^n \left| x - \lambda_i^k \right| \leq \prod_{i=1}^n \left(|x| + |\lambda_i|^k \right) \leq \prod_{i=1}^n (1+1) = 2^n.$$

Considérons les deux normes suivantes sur $\mathbb{R}_n[X]$

$$\|.\|_\infty : \begin{cases} \mathbb{R}_n[X] & \longrightarrow \mathbb{R}^+ \\ P & \longmapsto \sup\limits_{x \in [0,1]} |P(x)|, \end{cases}$$

ainsi que

$$\|.\|_{\infty,\mathrm{coef}} : \begin{cases} \mathbb{R}_n[X] & \longrightarrow \mathbb{R}^+ \\ P & \longmapsto \sup\limits_{k \in [\![0;n]\!]} |a_k| \text{ lorsque } P(X) = \sum\limits_{k=0}^n a_k X^k. \end{cases}$$

$\mathbb{R}_n[X]$ est de dimension finie, donc toutes les normes y sont équivalentes (voir chapitre sur l'équivalence des normes). Il existe donc une constante $C > 0$ telle que

$$\forall P \in \mathbb{R}_n[X],\ \|P\|_{\infty,\mathrm{coef}} \leq C \|P\|_\infty.$$

On a donc pour tout $k \in \mathbb{N}^*$,

$$\|P_k\|_{\infty,\mathrm{coef}} \leq C 2^n.$$

→ **Méthode 2 (Vieta) :** On souhaite borner les coefficients de P_k en utilisant une majoration des racines de ce dernier, il est donc naturel de chercher à utiliser des relations coefficients racines. Introduisons donc le lemme assez connu suivant.

Lemme (Formules de Vieta) VIII.11.

Soit $n \in \mathbb{N}^*$, $P(X) = a_0 + a_1 X + \cdots + a_n X^n$ et $\lambda_1, \dots, \lambda_n$ les racines (non nécessairement différentes) de P. Supposons que $a_n \neq 0$. On a pour tout $\ell \in [\![0;n]\!]$

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_\ell \leq n} \left(\prod_{j=1}^{\ell} \lambda_{i_j} \right) = (-1)^\ell \frac{a_{n-\ell}}{a_n}.$$

Nous ne démontrerons pas ce lemme, mais nous encourageons le lecteur à aller regarder la

(cc) BY-NC-SA

démonstration de ce lemme qui peut être quelques fois assez utile. Appliquons ce lemme, on a pour tout $k \in \mathbb{N}$ et $\ell \in [\![1; n]\!]$

$$\begin{aligned} |a_{k,n-\ell}| &= \left| \sum_{1 \le i_1 < i_2 < \cdots < i_\ell \le n} \left(\prod_{j=1}^{\ell} \lambda_{i_j} \right) \right| \\ &\le \sum_{1 \le i_1 < i_2 < \cdots < i_\ell \le n} \left| \prod_{j=1}^{\ell} \lambda_{i_j} \right| \\ &= \sum_{1 \le i_1 < i_2 < \cdots < i_\ell \le n} 1 = \binom{n}{\ell}. \end{aligned}$$

Ceci nous permet donc d'affirmer que pour tout $i \in [\![0; n]\!]$ les coefficients $|a_{k,i}|$ sont bornés par une constante indépendante de k.

On a donc pour tout $i \in [\![0; n-1]\!]$, l'ensemble $\{a_{k,i},\ k \in \mathbb{N}^*\}$ est une partie bornée de $\mathbb{Z}$, elle est donc finie. L'ensemble $\{\mathrm{P}_k,\ k \in \mathbb{N}\}$ est alors fini, ce qui implique que pour tout $i \in [\![1; n]\!]$, l'ensemble $\{\lambda_i^k,\ k \in \mathrm{N}\}$ est fini. On en déduit que pour tout $i \in [\![1; n]\!]$, il existe $k, \ell \in \mathbb{N}$ différents (on suppose sans perte de généralité que $k < \ell$) tels que $\lambda_i^k = \lambda_i^\ell$, *i.e.* $\lambda_i^{k-\ell} = 1$ et enfin pour tout $i \in [\![1; n]\!]$, λ_i est une racine de l'unité.

Correction de l'exercice VII.9. :

1. On a

$$\begin{aligned} \mu_y \cdot y = 0 &\iff \mu_y \cdot (\mathrm{Q} \cdot x) = 0 \iff (\mu_y \mathrm{Q}) \cdot x = 0 \\ &\iff \mu_x | \mu_y \mathrm{Q} \iff \mathrm{PQ} | \mu_y \mathrm{Q} \iff \mathrm{P} | \mu_y, \end{aligned}$$

mais $\mathrm{P} \cdot y = 0$, donc $\mu_y | \mathrm{P}$ et ces deux polynômes sont unitaires, donc $\mu_y = \mathrm{P}$.

2. Il est clair que $\mu_x \mu_y \cdot (x+y) = 0$ donc $\mu_{x+y} | \mu_x \mu_y$. De plus, on a

$$\mu_{x+y} \cdot (x+y) = 0 \Longrightarrow \mu_{x+y} \mu_x \cdot y = 0 \Longrightarrow \mu_y | \mu_{x+y} \mu_x,$$

mais $\mu_y \wedge \mu_x = 1$, donc d'après Gauß, $\mu_y | \mu_{x+y}$. Par le même raisonnement, on peut montrer également que $\mu_x | \mu_{x+y}$. Encore une fois, puisque $\mu_y \wedge \mu_x = 1$, on a $\mu_x \mu_y | \mu_{x+y}$. On en déduit donc que $\mu_x \mu_y = \mu_{x+y}$. $\mu_x \mu_y$ et μ_{x+y} sont unitaires ce qui nous permet d'affirmer que $\mu_x \mu_y = \mu_{x+y}$.

3. Posons $\mu_x = \mathrm{P}_1^{\alpha_1} \dots \mathrm{P}_r^{\alpha_r}$ et $\mu_y = \mathrm{P}_1^{\beta_1} \dots \mathrm{P}_r^{\beta_r}$. Posons également

$$\alpha_i' = \begin{cases} \alpha_i & \text{si } \alpha_i \ge \beta_i \\ 0 & \text{sinon} \end{cases} \text{ et } \beta_i' = \begin{cases} \beta_i & \text{si } \beta_i > \alpha_i \\ 0 & \text{sinon} \end{cases}$$

$$\mathrm{P}_x = \mathrm{P}_1^{\alpha_1'} \dots \mathrm{P}_r^{\alpha_r'} \text{ et } \mathrm{Q}_y = \mathrm{P}_1^{\beta_1'} \dots \mathrm{P}_r^{\beta_r'}.$$

Posons également $x' = \dfrac{\mu_x}{\mathrm{P}_x} x$ et $y' = \dfrac{\mu_y}{\mathrm{Q}_y} y$. D'après la question 1, on sait que $\mu_{x'} = \mathrm{P}_x$ et $\mu_{y'} = \mathrm{Q}_y$. De plus, on a $\mathrm{P}_x \wedge \mathrm{Q}_y = 1$, donc d'après la question précédente,

$$\mu_{x'+y'} = \mathrm{P}_x \mathrm{Q}_y = \mu_x \vee \mu_y.$$

On en déduit donc que $z = x' + y'$ convient.

4. Soit $(e_1, \dots, e_n)$ une base de E. Par une récurrence rapide, on sait qu'il existe d'après la question précédente $a \in \mathrm{E} \backslash \{0\}$ tel que $\mu_a = \mu_{e_1} \vee \cdots \vee \mu_{e_n}$. On a donc pour tout $x = x_1 e_1 + \cdots + x_n e_n \in \mathrm{E} \backslash \{0\}$,

$$\mu_a \cdot x = x_1 \mu_a \cdot e_1 + \cdots + x_n \mu_a \cdot e_n = 0,$$

(cc) BY-NC-SA

donc $\mu_x|\mu_a$. On en déduit que $\mu_a(u) = 0$ *i.e.* $\mu_u|\mu_a$ et de plus $\mu_u \cdot a = 0$ donc $\mu_a|\mu_u$ et ces deux polynômes sont unitaires, donc $\mu_a = \mu_u$.

Correction de l'exercice VIII.3. :

Soit $(\mathrm{A}_p) \in \widetilde{\mathrm{F}}^{\mathbb{N}}$ telle que $\mathrm{A}_p \xrightarrow[p\to+\infty]{} \mathrm{A}$ et $z \in \mathbb{C} \setminus \mathrm{F}$. Posons pour tout $p \in \mathbb{N}$, $\mathrm{Spec}_{\mathbb{C}}(\mathrm{A}_p) = [\lambda_{1,p}, \dots, \lambda_{n,p}]$. On a

$$\left|\chi_{\mathrm{A}_p}(z)\right| = \prod_{i=1}^{n} |z - \lambda_{i,p}| \geq d(z, \mathrm{F})^n > 0.$$

En passant donc à la limite, on a par continuité

$$|\chi_{\mathrm{A}}(z)| \geq d(z, \mathrm{F})^n > 0,$$

et donc $z \notin \mathrm{Spec}_{\mathbb{C}}(\mathrm{A})$, *i.e.* $\mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) \subset \mathrm{F}$. On a donc bien que $\mathrm{A} \in \widetilde{\mathrm{F}}$, *i.e.* $\widetilde{\mathrm{F}}$ est fermé.

Correction de l'exercice VIII.5. :

Soit $\varepsilon > 0$. D'après la proposition VIII.4, Il existe $\mathrm{P} \in \mathrm{GL}_n(\mathbb{C})$ tel que

$$\mathrm{A} = \mathrm{P} \underbrace{\begin{pmatrix} \lambda_1 & & b_{ij} \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}_{\mathrm{B}} \mathrm{P}^{-1}$$

et $\forall i, j \in [\![1; n]\!]$ tels que $j > i$, $|b_{i,j}| < \varepsilon$. Munissons $\mathbb{C}^n$ de la norme $\|.\|_1$, dont on rappelle la définition

$$\|.\|_1 : \begin{cases} \mathbb{C}^n & \longrightarrow \mathbb{R}^+ \\ (x_1, \dots, x_n)^{\mathrm{T}} & \longmapsto |x_1| + \cdots + |x_n|. \end{cases}$$

En posant $(e_1, \dots, e_n)$ la base canonique de $\mathbb{C}^n$, on a pour tout $\mathrm{X} = x_1e_1 + \cdots + x_ne_n \in \mathbb{C}^n$,

$$\|\mathrm{BX}\|_1 = \left\|\sum_{i=1}^{n} x_i \mathrm{B}e_i\right\|_1 \leq \sum_{i=1}^{n} |x_i| \, \|\mathrm{B}e_i\|_1 \leq \mathrm{C} \, \|\mathrm{X}\|_1,$$

avec $\mathrm{C} = \sup\limits_{i\in[\![1;n]\!]} \|\mathrm{B}e_i\|$. On a donc, en considérant $|\!|\!|.|\!|\!|$ la norme d'opérateur associée à $\|.\|_1$ que

$$|\!|\!|\mathrm{B}|\!|\!| \leq \mathrm{C} = \sup_{i\in[\![1;n]\!]} \|\mathrm{B}e_i\|_1 = \sup_{i\in[\![1;n]\!]} \left|\lambda_i + \sum_{j=i+1}^{n} b_{ij}\right| \leq \sup_{i\in[\![1;n]\!]} |\lambda_i| + (n-1)\varepsilon.$$

On choisit donc ε tel que $\sup_{i\in[\![1;n]\!]} |\lambda_i| + (n-1)\varepsilon < 1$. Ceci nous permet d'avoir $|\!|\!|\mathrm{B}|\!|\!| < 1$ et donc

$$\begin{aligned} |\!|\!|\mathrm{A}^p|\!|\!| = |\!|\!|\mathrm{P}\mathrm{B}^p\mathrm{P}^{-1}|\!|\!| &\leq |\!|\!|\mathrm{P}|\!|\!| \times |\!|\!|\mathrm{B}^p|\!|\!| \times |\!|\!|\mathrm{P}^{-1}|\!|\!| \\ &\leq |\!|\!|\mathrm{P}|\!|\!| \times |\!|\!|\mathrm{P}^{-1}|\!|\!| \times |\!|\!|\mathrm{B}|\!|\!|^p \xrightarrow[p\to+\infty]{} 0 \end{aligned}$$

et donc finalement $\mathrm{A}^p \xrightarrow[p\to+\infty]{} 0$.

Correction de l'exercice VIII.6. :

Munissons $\mathcal{M}_n(\mathbb{C})$ de la norme $\|.\|$ définie par

$$\|.\| : \begin{cases} \mathcal{M}_n(\mathbb{C}) & \longrightarrow \mathbb{R}^+ \\ (a_{i,j})_{i,j\in[\![1;n]\!]} & \longmapsto \sup\limits_{i,j\in[\![1;n]\!]} |a_{i,j}|. \end{cases}$$

(cc) BY-NC-SA

Pour tout $p \in \mathbb{N}$, $\mathrm{A}^p \in \mathcal{M}_n(\mathbb{Z})$, donc $\|\mathrm{A}^p\| \in \mathbb{Z}$. De plus, $\mathrm{Spec}_{\mathbb{C}}(\mathrm{A}) \subset \mathrm{B}(0,1)$, donc d'après l'exercice précédent et par équivalence des normes en dimension finie, $\mathrm{A}^p \xrightarrow[p\to+\infty]{} 0$. $(\|\mathrm{A}\|^p)_{p\in\mathbb{N}}$ est une suite à valeurs dans $\mathbb{Z}$ qui converge vers 0, elle est donc stationnaire en 0. On en déduit donc que il existe $r \in \mathbb{N}$ tel que pour tout $p \geq r$, $\|\mathrm{A}\|^p = 0$ *i.e.* $\mathrm{A}^p = 0$. A est alors bien nilpotente.

Correction de l'exercice VIII.8. :

Posons $\mathrm{S} = \{\mathrm{P}^{-1}\mathrm{AP},\ \mathrm{P} \in \mathrm{GL}_n(\mathbb{C})\}$.

→ ($\Leftarrow$) Supposons que S soit fermé et montrons qu'il existe une matrice diagonale D dans S. D'après la proposition VIII.4, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ et une suite de matrices $(\mathrm{P}_k)_{k\in\mathbb{N}} \in \mathrm{GL}_n(\mathbb{C})^{\mathbb{N}}$ telle que pour tout $k \in \mathbb{N}$,

$$\mathrm{P}_k^{-1}\mathrm{AP}_k = \begin{pmatrix} \lambda_1 & & b_{k,i,j} \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

et pour tout $j > i$, $|b_{k,i,j}| \leq \dfrac{1}{k+1}$. Cet argument nous permet d'affirmer que

$$\mathrm{P}_k^{-1}\mathrm{AP}_k \xrightarrow[p\to+\infty]{} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

S est fermé, donc cette matrice appartient à S. On en déduit donc que A est semblable à une matrice diagonale, *i.e.* A est diagonalisable.

→ ($\Rightarrow$) Supposons que A soit diagonalisable. Montrons que $\bar{\mathrm{S}} = \mathrm{S}$. Soit $\mathrm{B} \in \bar{\mathrm{S}}$. On veut montrer que $\mathrm{B} \in \mathrm{S}$. L'application

$$\phi : \begin{cases} \mathcal{M}_n(\mathbb{C}) & \longrightarrow \mathbb{C}_n[\mathrm{X}] \\ \mathrm{M} & \longmapsto \chi_{\mathrm{M}} \end{cases}$$

est continue, car pour tout $\mathrm{M} \in \mathcal{M}_n(\mathbb{C})$, les coefficients de $\phi(\mathrm{M})$ dans $\mathbb{C}_n[\mathrm{X}]$ sont polynomiaux en les coefficients de M dans $\mathcal{M}_n(\mathbb{C})$. De plus, ϕ est constante sur S égale à χ_{A}, elle est donc constante aussi sur $\bar{\mathrm{S}}$ égale à χ_{A} par continuité. Ceci nous permet de dire que $\chi_{\mathrm{A}} = \chi_{\mathrm{B}}$ et que A et B ont les mêmes valeurs propres et même multiplicité.

$$\psi : \begin{cases} \mathcal{M}_n(\mathbb{C}) & \longrightarrow \mathcal{M}_n(\mathbb{C}) \\ \mathrm{M} & \longmapsto \mu_{\mathrm{A}}(\mathrm{M}) \end{cases}$$

est continue car polynomiale et constante sur S égale à 0, donc par continuité elle est aussi égale à 0 sur l'adhérence de S et alors $\mu_{\mathrm{A}}(\mathrm{B}) = 0$. A est diagonalisable, donc μ_{A} est scindé à racines simples et alors B est diagonalisable. D'après ce qui précède, A et B ont même valeurs propres et même multiplicités, donc A et B sont semblables diagonalisables et finalement $\mathrm{B} \in \mathrm{S}$.

Correction de l'exercice VIII.10. :

Le cas $\mathrm{A} = 0$ est trivial. Supposons que $\mathrm{A} \neq 0$. Soit $\lambda \in \mathbb{C}$ une valeur propre de A. Supposons que $|\lambda| = \|\!|\mathrm{A}|\!\|$. Quitte à diviser les deux côtés de l'égalité par $|\lambda|$, on suppose que $\|\!|\mathrm{A}|\!\| = 1 = |\lambda|$. On a alors

$$\forall p \in \mathbb{N}^*,\ \|\!|\mathrm{A}^p|\!\| \leq \|\!|\mathrm{A}|\!\|^p = 1.$$

Soit $\mathrm{F} = \mathrm{Ker}(\mathrm{A} - \lambda\mathrm{I})^{\alpha_{\mathrm{A}}(\lambda)}$. En posant $u : \mathrm{X} \longmapsto \mathrm{AX}$ et $v = u\Big|_{\mathrm{F}}^{\mathrm{F}}$, on sait que il existe N nilpotente telle que $v = \lambda\mathrm{I} + \mathrm{N}$. On veut montrer donc que $\mathrm{N} = 0$.

(cc) BY-NC-SA

Idée : On sait que pour tout $p \in \mathbb{N}^*$, $|||v^p||| \leq |||\mathrm{A}^p||| \leq 1$, donc v^p est borné. On va donc montrer que si $\mathrm{N} \neq 0$, alors v^p est non bornée.

Supposons que $\mathrm{N} \neq 0$. Soit $m > 1$ l'indice de nilpotence de N. Il existe donc $\mathrm{X} \in \mathbb{C}^n$ tel que $\mathrm{N}^{m-1}\mathrm{X} \neq 0$. Posons $\mathrm{Y} = \mathrm{N}^{m-2}\mathrm{X}$. On a alors $\mathrm{NY} \neq 0$ et $\mathrm{N}^2\mathrm{Y} = 0$. On a alors pour tout $p \in \mathbb{N}^*$

$$v^p\mathrm{Y} = (\lambda\mathrm{I} + \mathrm{N})^p\mathrm{Y} = \sum_{k=0}^{p} \binom{p}{k} \lambda^k \mathrm{N}^{n-k}\mathrm{Y} = \lambda^p\mathrm{Y} + p\lambda^{p-1}\mathrm{NY}.$$

Or $\lambda^p\mathrm{Y}$ est borné ca $|\lambda| \leq 1$ et

$$\left\| p\lambda^{p-1}\mathrm{NY} \right\| = p\,\|\mathrm{NY}\| \xrightarrow[p\to+\infty]{} +\infty,$$

donc v^p est non bornée, ce qui est absurde. On en déduit donc que $\mathrm{N} = 0$ et alors $\mathrm{Ker}(\mathrm{A} - \lambda\mathrm{I})^{\alpha_{\mathrm{A}}(\lambda)} = \mathrm{Ker}(\mathrm{A} - \lambda\mathrm{I})$ *i.e.* λ est pure.

(cc) BY-NC-SA

Chapitre 30.1

Probabilités

Dans tout ce chapitre, E désigne un ensemble.
Préliminaires

→ Soient P et Q deux assertions. Si $\mathrm{A} = \{\mathrm{P}\}$, $\mathrm{B} = \{\mathrm{Q}\}$, alors $\mathrm{A}^a = \{\neg \mathrm{P}\}$, $\mathrm{A} \cap \mathrm{B} = \{\mathrm{P} \wedge \mathrm{Q}\}$, et $\mathrm{A} \cup \mathrm{B} = \{\mathrm{P} \vee \mathrm{Q}\}$.

→ Soit Ω un ensemble. Soit $(\mathrm{A}_n) \in \mathcal{P}(\Omega)^{\mathbb{N}}$. Alors,

$$\bigcap_{\mathrm{N}=0}^{+\infty} \left(\bigcup_{n=\mathrm{N}}^{+\infty} \mathrm{A}_n \right) = \{\omega \in \Omega \mid \forall \mathrm{N} \in \mathbb{N},\ \exists n \in [\![\mathrm{N}, +\infty[\![,\ \omega \in \mathrm{A}_n\}.$$

Si $\omega \in \Omega$, alors ω appartient cet ensemble si, et seulement si, ω appartient à une infinité de termes de la suite $(\mathrm{A}_n)_{n\in\mathbb{N}}$.

I Tribus

On appelle tribu sur E toute partie $\mathcal{T}$ de $\mathcal{P}(\mathrm{E})$ qui vérifie les propriétés suivantes :

→ $\varnothing$ et $\mathrm{E} \in \mathcal{T}$.

→ Si $(\mathrm{A}_n) \in \mathcal{T}^{\mathbb{N}}$, $\displaystyle\bigcup_{n=0}^{+\infty} \mathrm{A}_n \in \mathcal{T}$.

→ Si $\mathrm{A} \in \mathcal{T}$, $\mathrm{A}^a \in \mathcal{T}$.

Exemples. $\{\varnothing, \mathrm{E}\}$ est une tribu sur E, et si E est dénombrable, $\mathcal{P}(\mathrm{E})$ est une tribu sur E.

Proposition I.1.

Soit $\mathcal{T}$ une tribu sur E. Soit $(\mathrm{A}_n) \in \mathcal{T}^{\mathbb{N}}$.

→ Pour tout $\mathrm{N} \in \mathbb{N}$, $\displaystyle\bigcup_{n=0}^{\mathrm{N}} \mathrm{A}_n = \left(\bigcup_{n=0}^{\mathrm{N}} \mathrm{A}_n \right) \cup \varnothing \cup \varnothing \cup \ldots \in \mathcal{T}$.

→ Pour tout $\mathrm{N} \in \mathbb{N}$, $\displaystyle\bigcap_{n=0}^{\mathrm{N}} \mathrm{A}_n = \left(\bigcup_{n=0}^{\mathrm{N}} \mathrm{A}_n^a \right)^a \in \mathcal{T}$.

→ Il existe une suite $(\mathrm{B}_n) \in \mathcal{T}^{\mathbb{N}}$ d'ensembles deux à deux disjoints telle que $\displaystyle\bigcup_{n=0}^{+\infty} \mathrm{A}_n = \bigcup_{n=0}^{+\infty} \mathrm{B}_n$, et, pour tout $\mathrm{N} \in \mathbb{N}$, $\displaystyle\bigcup_{n=0}^{\mathrm{N}} \mathrm{A}_n = \bigcup_{n=0}^{\mathrm{N}} \mathrm{B}_n$.

→ Si Λ est un ensemble non vide, et $(\mathcal{T}_\lambda)_{\lambda\in\Lambda}$ une famille de tribus sur E, alors $\displaystyle\bigcap_{\lambda\in\Lambda} \mathcal{T}_\lambda$ est aussi une tribu sur E.

Démonstration de la troisième propriété. En posant, pour tout $n \in \mathbb{N}$, $\mathrm{B}_n = \mathrm{A}_n \setminus \left(\displaystyle\bigcup_{k=0}^{n-1} \mathrm{A}_k \right)$, on construit bien une suite d'ensembles deux à deux disjoints qui vérifie les deux propriétés.

CPGE paradise
(cc) BY-NC-SA

Définition (Tribu engendrée) I.2.

Soit X une partie de $\mathcal{P}(\mathrm{E})$, et notons $\mathcal{T}_{\mathrm{X}}$ l'ensemble des tribus sur E contenant X. On appelle *tribu engendrée par* X, la tribu

$$\mathcal{T}(\mathrm{X}) = \bigcap_{\mathcal{T} \in \mathcal{T}_{\mathrm{X}}} \mathcal{T}.$$

C'est la plus petite tribu sur E contenant X.

Exemples.

→ On appelle tribu borélienne sur $\mathbb{R}$, et on note $\mathcal{B}(\mathbb{R})$, la tribu engendrée par l'ensemble des ouverts de $\mathbb{R}$. Par complémentarité, cette tribu contient également tous les fermés de $\mathbb{R}$. Donc cette tribu contient tous les points de $\mathbb{R}$, et par union dénombrable, $\mathbb{Q} \subset \mathcal{B}(\mathbb{R})$ puis, par complémentarité, $\mathbb{R} \setminus \mathbb{Q} \subset \mathcal{B}(\mathbb{R})$.

→ Soit $n \in \mathbb{N}^*$. Soient $\mathrm{A}_1, \ldots, \mathrm{A}_n$ des parties de E incluant $\varnothing$ et E. Posons

$$\mathcal{T}_a = \left\{ \bigcap_{k=1}^{n} \mathrm{B}_k \;\middle|\; \forall k \in [\![1;n]\!],\ \mathrm{B}_k \in \{\mathrm{A}_k, \mathrm{A}_k^a\} \right\}.$$

Les éléments de $\mathcal{T}_a$ sont deux à deux disjoints et la tribu engendrée par $\{\mathrm{A}_1, \ldots, \mathrm{A}_n\}$ est l'ensemble de réunions d'ensembles appartenant à $\mathcal{T}_a$.

→ Soit $(\mathrm{A}_n) \in \mathcal{P}(\mathrm{E})^{\mathbb{N}}$ une suite de parties de E deux à deux disjointes telles que $\bigcup_{n=0}^{+\infty} \mathrm{A}_n = \mathrm{E}$. La tribu engendrée par l'ensemble des termes de $(\mathrm{A}_n)_{n\in\mathbb{N}}$ est $\mathcal{T} = \left\{ \bigcup_{n \in \mathrm{I}} \mathrm{A}_n \;\middle|\; \mathrm{I} \subset \mathbb{N} \right\}$.

En effet, $\mathrm{E} = \bigcup_{n=0}^{+\infty} \mathrm{A}_n \in \mathcal{T}$, $\varnothing = \bigcup_{n \in \varnothing} \mathrm{A}_n \in \mathcal{T}$, et $\mathcal{T}$ est stable par union dénombrable.

Enfin, si $\mathrm{I} \subset \mathbb{N}$, $\left(\bigcup_{n \in \mathrm{I}} \mathrm{A}_n \right)^a = \bigcup_{n \in \mathbb{N}\setminus\mathrm{I}} \mathrm{A}_n \in \mathcal{T}$.

Notation. Soient A et B deux parties de E. Si l'union $\mathrm{A} \cup \mathrm{B}$ est disjointe, on la note $\mathrm{A} \sqcup \mathrm{B}$.

Proposition I.3.

Soit F un ensemble et f une application de E dans F. Si $\mathcal{T}'$ est une tribu sur F, l'ensemble des images réciproques des éléments de $\mathcal{T}'$ par f, noté $f^{-1}\langle \mathcal{T}' \rangle$, est une tribu sur E, appelée tribu image réciproque de $\mathcal{T}'$ sous f.
Si, de plus, on munit E d'une tribu $\mathcal{T}$ sur E, et F de la tribu $\mathcal{T}'$, alors $(\mathrm{E}, \mathcal{T})$ et $(\mathrm{F}, \mathcal{T}')$ sont qualifiés d'espaces mesurables, et $f : (\mathrm{E}, \mathcal{T}) \longrightarrow (\mathrm{F}, \mathcal{T}')$ est dite mesurable lorsque $f^{-1}\langle \mathcal{T}' \rangle \subset \mathcal{T}$

BY-NC-SA

II Espaces probabilisés

Dans la suite du chapitre, on munit E d'une tribu $\mathcal{T}$ sur E si bien que $(\mathrm{E},\mathcal{T})$ est un espace mesurable, ou encore un espace probabilisable. Les éléments de $\mathcal{T}$ sont appelés *événements*.

Définition II.1.

On dit qu'une application $\mathbb{P} : \mathcal{T} \longrightarrow [0,1]$ est une probabilité sur l'espace probabilisable $(\mathrm{E},\mathcal{T})$ lorsqu'elle vérifie les propriétés suivantes :

→ $\mathbb{P}(\varnothing) = 0$ et $\mathbb{P}(\mathrm{E}) = 1$.

→ Si $(\mathrm{A}_n) \in \mathcal{T}^{\mathbb{N}}$ est une suite d'événements deux à deux incompatibles, *i.e.* une suite d'éléments de $\mathcal{T}$ deux à deux disjoints, alors $\mathbb{P}\left(\bigcup_{n=0}^{+\infty} \mathrm{A}_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{A}_n)$.

Proposition II.2.

Soit $\mathbb{P}$ une probabilité sur $(\mathrm{E},\mathcal{T})$. Alors $\mathbb{P}$ vérifie les propriétés suivantes :

→ Soit $n \in \mathbb{N}$. Si $\mathrm{A}_0,\ldots,\mathrm{A}_n$ sont des événements deux à deux incompatibles, alors

$$\mathbb{P}\left(\bigcup_{k=0}^{n} \mathrm{A}_k\right) = \sum_{k=0}^{n} \mathbb{P}(\mathrm{A}_k).$$

→ *Croissance :* Soient A et B deux événements. Si $\mathrm{A} \subset \mathrm{B}$, alors $\mathbb{P}(\mathrm{A}) \leq \mathbb{P}(\mathrm{B})$.

→ Soit $n \in \mathbb{N}^*$. Soient $\mathrm{A}_1,\ldots,\mathrm{A}_n \in \mathcal{T}$. Alors

$$\mathbb{P}\left(\bigcup_{k=1}^{n} \mathrm{A}_k\right) \leq \sum_{k=1}^{n} \mathbb{P}(\mathrm{A}_k) \qquad \text{et} \qquad \mathbb{P}\left(\bigcup_{k=1}^{n} \mathrm{A}_k\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq j_1 < \cdots < j_k \leq n} \mathbb{P}\left(\bigcap_{\ell=1}^{k} \mathrm{A}_{j_\ell}\right).$$

L'égalité se nomme *la formule de Poincaré.*

→ *Continuité croissante :* Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite croissante d'événements *i.e.* pour tout $n \in \mathbb{N}$, $\mathrm{A}_n \subset \mathrm{A}_{n+1}$. Alors

$$\mathbb{P}(\mathrm{A}_n) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\bigcup_{k=0}^{+\infty} \mathrm{A}_k\right).$$

→ *Continuité décroissante :* Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite décroissante d'événements *i.e.* pour tout $n \in \mathbb{N}$, $\mathrm{A}_{n+1} \subset \mathrm{A}_n$. Alors

$$\mathbb{P}(\mathrm{A}_n) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\bigcap_{k=0}^{+\infty} \mathrm{A}_k\right).$$

Démonstration de la continuité décroissante à partir de la continuité croissante

$$\left(\bigcup_{k=0}^{n} \mathrm{A}_k^a\right)_{n\in\mathbb{N}} \text{ est croissante donc } \underbrace{\mathbb{P}\left(\bigcup_{k=0}^{n} \mathrm{A}_k^a\right)}_{=1-\mathbb{P}\left(\underbrace{\bigcup_{k=0}^{n} \mathrm{A}_k}_{=\mathrm{A}_n}\right)} \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\bigcup_{k=0}^{+\infty} \mathrm{A}_k^a\right) = 1 - \mathbb{P}\left(\bigcap_{k=0}^{+\infty} \mathrm{A}_k\right).$$

(cc) BY-NC-SA

Conséquence. Si $(\mathrm{A}_n)_{n\in\mathbb{N}}$ est une suite d'événements telle que $\sum_{n\geq 0}\mathbb{P}(\mathrm{A}_n)$ converge, alors

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty}\mathrm{A}_n\right)\leq\sum_{n=0}^{+\infty}\mathbb{P}(\mathrm{A}_n).$$

Démonstration. $\left(\bigcup_{k=0}^{n}\mathrm{A}_k\right)_{n\in\mathbb{N}}$ est une suite croissante d'événements, donc par continuité croissante

$$\mathbb{P}\left(\bigcup_{k=0}^{+\infty}\mathrm{A}_k\right)=\lim_{n\to+\infty}\mathbb{P}\left(\bigcup_{k=0}^{n}\mathrm{A}_k\right)\leq\lim_{n\to+\infty}\sum_{k=0}^{n}\mathbb{P}(\mathrm{A}_k)=\sum_{n=0}^{+\infty}\mathbb{P}(\mathrm{A}_n).$$

III Événement presque sûr, événement négligeable

Dans la suite du chapitre, on munit l'espace probabilisable $(\mathrm{E},\mathcal{T})$ d'une probabilité $\mathbb{P}$ sur $(\mathrm{E},\mathcal{T})$, si bien que $(\mathrm{E},\mathcal{T},\mathbb{P})$ est un espace probabilisé.

Définition III.1.

On dit qu'un événement A est presque sûr lorsque $\mathbb{P}(\mathrm{A})=1$. On dit qu'un événement A est négligeable lorsque $\mathbb{P}(\mathrm{A})=0$.

Proposition III.2.

- → Soient A et B sont deux événements tels que $\mathrm{A}\subset\mathrm{B}$. Si A est presque sûr, alors B est presque sûr.
- → Si $(\mathrm{A}_n)_{n\in\mathbb{N}}$ est une suite d'événements négligeables, alors $\bigcup_{n=0}^{+\infty}\mathrm{A}_n$ est négligeable.

Démonstration.

- → $1=\mathbb{P}(\mathrm{A})\leq\mathbb{P}(\mathrm{B})\leq 1$, donc B est presque sûr.
- → La série $\sum_{n\geq 0}\mathbb{P}(\mathrm{A}_n)$ converge car elle est nulle donc $0\leq\mathbb{P}\left(\bigcup_{n=0}^{+\infty}\mathrm{A}_n\right)\leq\sum_{n\geq 0}\mathbb{P}(\mathrm{A}_n)=0$, donc $\bigcup_{n=0}^{+\infty}\mathrm{A}_n$ est négligeable.

Exercice III.3.

Soit $(\mathrm{E},\mathcal{T},\mathbb{P})$ un espace probabilisé. Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite d'événements telle que $\sum_{n\geq 0}\mathbb{P}(\mathrm{A}_n)$ converge.

Que dire que l'événement $\bigcap_{\mathrm{N}=0}^{+\infty}\left(\bigcup_{n=\mathrm{N}}^{+\infty}\mathrm{A}_n\right)$?

IV Exemples d'espaces probabilisés

1. Ensembles finis

Espace avec équiprobabilité : Si E est non vide et fini, $\mathcal{T}=\mathcal{P}(\mathrm{E})$, et pour tout $\mathrm{A}\in\mathcal{T}$, $\mathbb{P}(\mathrm{A})=\dfrac{|\mathrm{A}|}{|\mathrm{X}|}$ alors $(\mathrm{E},\mathcal{T},\mathbb{P})$ est un espace probabilisé.

BY-NC-SA

Espace image d'une loi binomiale : Soient $n \in \mathbb{N}^*$ et $p \in]0,1[$. Si $\mathrm{E} = [\![0;n]\!]$, $\mathcal{T} = \mathcal{P}(\mathrm{E})$, et pour tout $k \in \mathrm{E}$, $\mathbb{P}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}$, alors $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ est un espace probabilisé.

2. Ensembles dénombrables

Supposons dans cette partie que E soit dénombrable. Soit $(x_n) \in \mathrm{E}^{\mathbb{N}}$ une énumeration bijective de E. Soit $(a_n) \in \ell^1(\mathbb{N}, \mathbb{R}^+)$ telle que $\sum_{n=0}^{+\infty} a_n = 1$.

Posons $\mathcal{T} = \mathcal{P}(\mathrm{E})$. Pour tout $\mathrm{A} \in \mathcal{T}$, posons $\mathrm{I_A} = \{n \in \mathbb{N} \mid x_n \in \mathrm{A}\}$, puis $\mathbb{P}(\mathrm{A}) = \sum_{n \in \mathrm{I_A}} a_n$.

On a bien $\mathbb{P}(\varnothing) = 0$, $\mathbb{P}(\mathrm{E}) = 1$, et si (A_n) est une suite d'événements deux à deux incompatibles, alors

$$\mathbb{P}\Bigg(\underbrace{\bigcup_{n=0}^{+\infty} \mathrm{A}_n}_{\text{noté A}}\Bigg) = \sum_{k \in \mathrm{I_A}} a_k = \sum_{n=0}^{+\infty} \left(\sum_{k \in \mathrm{I}_{\mathrm{A}_n}} a_k \right) = \sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{A}_n),$$

donc $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ est un espace probabilisé.

Exemples.

→ Soit $a > 1$. Si $\mathrm{E} = \mathbb{N}^*$, $\mathcal{T} = \mathcal{P}(\mathrm{E})$ et, pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(\{n\}) = \dfrac{1}{\zeta(a) n^a}$, alors $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ est un espace probabilisé. C'est l'univers image d'une loi ζ de paramètre a.

→ Soit $p \in]0,1[$. Si $\mathrm{E} = \mathbb{N}^*$, $\mathcal{T} = \mathcal{P}(\mathrm{E})$ et, pour tout $n \in \mathbb{N}^*$, $\mathbb{P}(\{n\}) = (1-p)^{n-1} p$, alors $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ est un espace probabilisé. C'est l'univers image d'une loi géométrique de paramètre p.

→ Soit $\lambda > 0$. Si $\mathrm{E} = \mathbb{N}$, $\mathcal{T} = \mathcal{P}(\mathrm{E})$ et, pour tout $n \in \mathbb{N}$, $\mathbb{P}(\{n\}) = \dfrac{\lambda^n e^{-\lambda}}{n!}$, alors $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ est un espace probabilisé. C'est l'univers image d'une loi de Poisson de paramètre λ.

→ Soit F un ensemble dénombrable et $(y_n)_{n \in \mathbb{N}}$ une énumeration bijective de F. Soit $(b_n) \in \ell^1(\mathbb{N}, \mathbb{R}^+)$ tel que $\sum_{n=0}^{+\infty} b_n = 1$. Alors en posant, pour tout $n \in \mathbb{N}$, $\mathbb{P}(\{y_n\}) = b_n$, on définit bien une probabilité.
On définit la probabilité produit $\mathbb{P}_\times$ sur l'espace probabilisable $(\mathrm{E} \times \mathrm{F}, \mathcal{P}(\mathrm{E} \times \mathrm{F}))$ par

$$\forall (n,m) \in \mathbb{N}^2, \qquad \mathbb{P}_\times(\{x_n, y_m\}) = a_n b_m.$$

On vérifie que $\mathbb{P}_\times(\mathrm{E} \times \mathrm{F}) = \sum_{(n,m) \in \mathbb{N}^2} a_n b_m = \left(\sum_{n=0}^{+\infty} a_n \right) \left(\sum_{m=0}^{+\infty} a_m \right) = 1$.

(cc) BY-NC-SA

Correction de l'exercice II.3. :

Pour tout $\mathrm{M} \in \mathbb{N}$,

$$\mathbb{P}\left(\bigcap_{\mathrm{N}=0}^{+\infty}\left(\bigcup_{n \geq \mathrm{N}} \mathrm{A}_n\right)\right) \underset{\substack{\uparrow \\ \text{croissance de la probabilité}}}{\leq} \mathbb{P}\left(\bigcup_{n \geq \mathrm{M}} \mathrm{A}_n\right) \leq \sum_{n=\mathrm{M}}^{+\infty} \mathbb{P}(\mathrm{A}_n) \xrightarrow[\mathrm{M} \to +\infty]{} 0,$$

donc $\displaystyle\bigcap_{\mathrm{N}=0}^{+\infty}\left(\bigcup_{n \geq \mathrm{N}} \mathrm{A}_n\right)$ est négligeable.

(cc) BY-NC-SA

Chapitre 30.2

Indépendance, conditionnement

Dans tout ce chapitre, $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé.

I Événements indépendants

Définition I.1.

Deux événements A et B sont indépendants lorsque $\mathbb{P}(\mathrm{A} \cap \mathrm{B}) = \mathbb{P}(\mathrm{A})\mathbb{P}(\mathrm{B})$.

Attention : Si $\mathrm{A} \cap \mathrm{B} = \varnothing$, alors A et B sont indépendants si, et seulement si, A est négligeable ou B est négligeable.

Exemples.

- → Un événement négligeable est indépendant de tout autre.
- → Un événement presque sûr est indépendant de tout autre. En effet, si A est presque sûr, et si B est un événement quelconque, alors $\mathrm{B} = (\mathrm{A} \cap \mathrm{B}) \sqcup (\underbrace{\mathrm{A}^a \cap \mathrm{B}}_{\subset \mathrm{A}^a \text{ qui est négligeable}})$, donc $\mathbb{P}(\mathrm{A})\mathbb{P}(\mathrm{B}) = \mathbb{P}(\mathrm{B}) = \mathbb{P}(\mathrm{A} \cap \mathrm{B})$.

Proposition I.2.

Si A et B sont deux événements indépendants, alors A^a et B sont indépendants.

Démonstration. $\mathbb{P}(\mathrm{B}) = \mathbb{P}(\mathrm{A}^a \cap \mathrm{B}) + \mathbb{P}(\mathrm{A} \cap \mathrm{B}) = \mathbb{P}(\mathrm{A}^a \cap \mathrm{B}) + \mathbb{P}(\mathrm{A})\mathbb{P}(\mathrm{B})$. Donc $\mathbb{P}(\mathrm{A}^a \cap \mathrm{B}) = \mathbb{P}(\mathrm{B})(\underbrace{1 - \mathbb{P}(\mathrm{A})}_{= \mathbb{P}(\mathrm{A}^a)})$. **Conséquence.** Dans ce cas, B^a et A sont indépendants, ainsi que A^a et B^a.

Définition I.3.

Soit n un entier supérieur ou égal à 3. Des événements $\mathrm{A}_1, \ldots, \mathrm{A}_n$ sont dits mutuellement indépendants lorsque pour tout $\mathrm{I} \subset [\![1; n]\!]$, $\mathbb{P}\left(\bigcap_{i \in \mathrm{I}} \mathrm{A}_i\right) = \prod_{i \in \mathrm{I}} \mathbb{P}(\mathrm{A}_i)$.

Proposition I.4.

Soit n un entier supérieur ou égal à 3. Soient $\mathrm{A}_1, \ldots, \mathrm{A}_n$ des événements mutuellement indépendants.

- → Si I et $\mathrm{J} \subset [\![1; n]\!]$ et $\mathrm{I} \cap \mathrm{J} = \varnothing$, alors $\bigcap_{i \in \mathrm{I}} \mathrm{A}_i$ et $\bigcap_{j \in \mathrm{J}} \mathrm{A}_j$ sont indépendants.
- → $\mathrm{A}_1^a, \mathrm{A}_2, \ldots, \mathrm{A}_n$ sont mutuellement indépendants.

Conséquence. Les événements de la tribu engendrée par $\{\mathrm{A}_1, \ldots, \mathrm{A}_n\}$ sont mutuellement indépendants.

CC BY-NC-SA

Exercice I.5.

Soit $(\mathrm{E}, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Soit n un entier supérieur ou égal à 3. Soient $\mathrm{A}_1, \ldots, \mathrm{A}_n$ des événements mutuellement indépendants. Montrer que la probabilité qu'aucun d'eux ne se réalise est majorée par $\exp\left(-\sum_{i=1}^{n} \mathbb{P}(\mathrm{A}_i)\right)$.

II Conditionnement

Loi conditionnelle

Soit A un événement tel que $\mathbb{P}(\mathrm{A}) \neq 0$.

Pour tout $\mathrm{B} \in \mathcal{T}$, posons $\mathbb{P}(\mathrm{B} \mid \mathrm{A}) = \dfrac{\mathbb{P}(\mathrm{B} \cap \mathrm{A})}{\mathbb{P}(\mathrm{A})}$. Alors $\mathbb{P}(\,\cdot \mid \mathrm{A})$ est une probabilité sur $(\mathrm{E}, \mathcal{T})$. Elle vérifie les propriétés suivantes :

- → $\mathbb{P}(\mathrm{A} \cap \mathrm{B}) = \mathbb{P}(\mathrm{B} \mid \mathrm{A})\mathbb{P}(\mathrm{A})$.
- → Si A et B sont deux événements indépendants (au sens de $\mathbb{P}$), alors $\mathbb{P}(\mathrm{B} \mid \mathrm{A}) = \mathbb{P}(\mathrm{B})$.
- → Si n est un entier supérieur ou égal à 3, et $\mathrm{A}_1, \ldots, \mathrm{A}_n$ sont des événements tels que, pour tout $i \in [\![1; n-1]\!]$, $\mathbb{P}\left(\bigcap_{k=1}^{i} \mathrm{A}_k\right) > 0$, alors,

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \mathrm{A}_i\right) = \mathbb{P}(\mathrm{A}_1)\mathbb{P}(\mathrm{A}_2 \mid \mathrm{A}_1)\mathbb{P}(\mathrm{A}_3 \mid \mathrm{A}_1 \cap \mathrm{A}_2) \ldots \mathbb{P}(\mathrm{A}_n \mid \mathrm{A}_1 \cap \mathrm{A}_2 \cap \ldots \cap \mathrm{A}_n)$$

- → Si $(\mathrm{A}_n)_{n \in \mathbb{N}}$ est une partition de E telle que pour tout $n \in \mathbb{N}$, $\mathbb{P}(\mathrm{A}_n) > 0$, alors pour tout événement B non négligeable, on a

$$\mathbb{P}(\mathrm{B}) = \sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{B} \mid \mathrm{A}_n)\mathbb{P}(\mathrm{A}_n) \qquad \text{et} \qquad \mathbb{P}(\mathrm{A}_n \mid \mathrm{B}) = \frac{\mathbb{P}(\mathrm{B} \mid \mathrm{A}_n)\mathbb{P}(\mathrm{A}_n)}{\sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{B} \mid \mathrm{A}_n)\mathbb{P}(\mathrm{A}_n)}.$$

III Tribus indépendantes

Définition III.1.

Soit I un ensemble contenant au moins deux éléments.
Une famille $(\mathcal{T}_i)_{i \in \mathrm{I}}$ de tribus sur E est dite indépendante lorsque, pour tout ensemble fini $\mathrm{J} \subset \mathrm{I}$, toute famille $(\mathrm{A}_j)_{j \in \mathrm{J}} \in \prod_{j \in \mathrm{J}} \mathcal{T}_j$ est une famille d'événements mutuellement indépendants.

Indépendance de tribus engendrées par des partitions indépendantes

Soient $(\mathrm{A}_n)_{n \in \mathbb{N}}$ et $(\mathrm{B}_n)_{n \in \mathbb{N}}$ deux partitions de E telles que, pour tout $(n, p) \in \mathbb{N}^2$, A_n et B_n soient indépendants. Les tribus engendrées par ces partitions, $\mathcal{T}_{\mathrm{A}} = \left\{ \bigcup_{n \in \mathrm{I}} \mathrm{A}_n \;\middle|\; \mathrm{I} \in \mathcal{P}(\mathbb{N}) \right\}$ et $\mathcal{T}_{\mathrm{B}} = \left\{ \bigcup_{n \in \mathrm{I}} \mathrm{B}_n \;\middle|\; \mathrm{I} \in \mathcal{P}(\mathbb{N}) \right\}$, sont indépendantes.

(cc) BY-NC-SA

En effet, pour tous I et J $\in \mathcal{P}(\mathbb{N})$, en utilisant le théorème d'associativité pour les familles sommables,

$$\mathbb{P}\left(\left(\bigcup_{n\in \mathrm{I}} \mathrm{A}_n\right) \cap \left(\bigcup_{p\in \mathrm{J}} \mathrm{B}_n\right)\right) = \mathbb{P}\left(\bigsqcup_{(n,p)\in \mathrm{I}\times \mathrm{J}} (\mathrm{A}_n \cap \mathrm{B}_p)\right) = \sum_{(n,p)\in \mathrm{I}\times\mathrm{J}} \underbrace{\mathbb{P}(\mathrm{A}_n \cap \mathrm{B}_p)}_{\mathbb{P}(\mathrm{A}_n)\mathbb{P}(\mathrm{B}_p)} = \mathbb{P}\left(\bigcup_{n\in \mathrm{I}} \mathrm{A}_n\right) \mathbb{P}\left(\bigcup_{n\in \mathrm{J}} \mathrm{B}_n\right).$$

IV Borel-Cantelli

Théorème IV.1.

Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite d'événements mutuellement indépendants telle que $\sum_{n\geq 0} \mathbb{P}(\mathrm{A}_n)$ diverge. Alors,

$$\mathbb{P}\left(\bigcap_{\mathrm{N}=0}^{+\infty}\left(\bigcup_{n=\mathrm{N}}^{+\infty} \mathrm{A}_n\right)\right) = 1.$$

Vocabulaire. Si $\omega \in \mathrm{E}$ réalise $\bigcap_{\mathrm{N}=0}^{+\infty}\left(\bigcup_{n=\mathrm{N}}^{+\infty} \mathrm{A}_n\right)$ *i.e.* ω appartient à cet ensemble, alors ω réalise une infinité d'événements de la suite $(\mathrm{A}_n)_{n\in\mathbb{N}}$. La conclusion de ce théorème assure alors que, presque sûrement, une infinité d'événements de $(\mathrm{A}_n)_{n\in\mathbb{N}}$ sont réalisés, ou alors que A_n **est réalisé infiniment souvent.**

Démonstration. Posons, pour tout $\mathrm{N} \in \mathbb{N}$, $\mathrm{B}_\mathrm{N} = \bigcup_{n=\mathrm{N}}^{+\infty} \mathrm{A}_n$. Alors pour tout $\mathrm{N} \in \mathbb{N}$, $\mathrm{B}_\mathrm{N}^a = \bigcap_{n=\mathrm{N}}^{+\infty} \mathrm{A}_n^a$ et en introduisant la suite décroissante d'événements $\left(\mathrm{C}_\mathrm{M}^\mathrm{N}\right)_{\mathrm{M}\in\mathbb{N}} = \left(\bigcap_{n=\mathrm{N}}^{\mathrm{M}} \mathrm{A}_n^a\right)_{\mathrm{M}\in\mathbb{N}}$, on a, par continuité décroissante

$$\mathbb{P}\left(\mathrm{C}_\mathrm{M}^\mathrm{N}\right) \xrightarrow[\mathrm{M}\to+\infty]{} \mathbb{P}\left(\mathrm{B}_\mathrm{N}^a\right)$$

D'après l'exercice, par indépendance, pour tout $\mathrm{N} \in \mathbb{N}$ et pour tout entier $\mathrm{M} \geq \mathrm{N}$,

$$\mathbb{P}\left(\mathrm{C}_\mathrm{M}^\mathrm{N}\right) \leq \exp\left(-\sum_{n=\mathrm{N}}^{\mathrm{M}} \underbrace{\mathbb{P}(\mathrm{A}_n)}_{=\mathbb{P}((\mathrm{A}_n^a)^a)}\right).$$

donc en passant à la limite $\mathrm{M} \longrightarrow +\infty$, il vient $\mathbb{P}(\mathrm{B}_\mathrm{N}^a) = 0$.
Donc $\mathbb{P}\left(\bigcup_{\mathrm{N}=0}^{+\infty} \mathrm{B}_\mathrm{N}^a\right) = 0$ car $\mathbb{P}\left(\bigcup_{\mathrm{N}=0}^{+\infty} \mathrm{B}_\mathrm{N}^a\right) \leq \underbrace{\sum_{\mathrm{N}=0}^{+\infty} \mathbb{P}\left(\mathrm{B}_\mathrm{N}^a\right)}_{\text{somme d'une série convergente}}$. En passant au complémentaire, on obtient la conclusion du théorème.

Théorème IV.2.

Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite d'événements telle que $\sum_{n\geq 0} \mathbb{P}(\mathrm{A}_n)$ converge. Alors,

$$\mathbb{P}\left(\bigcap_{\mathrm{N}=0}^{+\infty}\left(\bigcup_{n=\mathrm{N}}^{+\infty} \mathrm{A}_n\right)\right) = 0.$$

Démonstration. En reprenant les notations de la démonstration précédente, $(\mathrm{B}_\mathrm{N})_{\mathrm{N}\in\mathbb{N}}$ est une suite décroissante d'événements, donc par continuité décroissante

$$\mathbb{P}(\mathrm{B}_\mathrm{N}) \xrightarrow[\mathrm{N}\to+\infty]{} \mathbb{P}\left(\bigcap_{\mathrm{N}=0}^{+\infty} \mathrm{B}_\mathrm{N}\right).$$

(cc) BY-NC-SA

Par ailleurs, pour tout $\mathrm{N} \in \mathbb{N}$, $\mathbb{P}(\mathrm{B}_\mathrm{N}) \leq \sum_{n=\mathrm{N}}^{+\infty} \mathbb{P}(\mathrm{A}_n) \xrightarrow[\mathrm{N}\to+\infty]{} 0$ donc par unicité de la limite, $\mathbb{P}\left(\bigcap_{\mathrm{N}=0}^{+\infty} \mathrm{B}_\mathrm{N}\right) = 0$, ce qu'on voulait.

(cc) BY-NC-SA

Correction de l'exercice I.5. :

Par indépendance mutuelle,

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \mathrm{A}_i^a\right) = \prod_{i=1}^{n}(1 - \mathbb{P}(\mathrm{A}_i)) \leq \prod_{i=1}^{n} e^{-\mathbb{P}(\mathrm{A}_i)} = \exp\left(-\sum_{i=1}^{n}\mathbb{P}(\mathrm{A}_i)\right).$$

Il est loisible de retenir que pour tout $x \in \mathbb{R}$, $e^x \geq 1 + x$.

(cc) BY-NC-SA

Chapitre 30.3

Variables aléatoires

Dans tout ce chapitre, $(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probabilisé, et $(\mathrm{E}, \mathcal{U})$ un espace probabilisable.

I Généralités

Une variable aléatoire à valeurs dans E est une application $\mathrm{X} : (\Omega, \mathcal{T}, \mathbb{P}) \longrightarrow (\mathrm{E}, \mathcal{U})$ telle que pour tout $\mathrm{U} \in \mathcal{U}$, $\mathrm{X}^{-1}(\mathrm{U}) \in \mathcal{T}$. Autrement dit, $\mathrm{X} : (\Omega, \mathcal{T}) \longrightarrow (\mathrm{E}, \mathcal{U})$ est une fonction mesurable.
Loi de probabilité : La loi de X, notée $\mathcal{L}_{\mathrm{X}}$, est une application de $\mathcal{U}$ dans $[0,1]$ définie par

$$\forall \mathrm{U} \in \mathcal{U}, \qquad \mathcal{L}_{\mathrm{X}}(\mathrm{U}) = \mathbb{P}(\mathrm{X}^{-1}(\mathrm{U})).$$

Ainsi, $\mathcal{L}_{\mathrm{X}}(\mathrm{E}) = \mathbb{P}(\Omega) = 1$, $\mathcal{L}_{\mathrm{X}}(\varnothing) = \mathbb{P}(\varnothing) = 0$ et, pour toute suite $(\mathrm{U}_n) \in \mathcal{U}^{\mathbb{N}}$ d'événements deux à deux disjoints,

$$\mathcal{L}_{\mathrm{X}}\left(\bigcup_{n=0}^{+\infty} \mathrm{U}_n\right) = \mathbb{P}\left(\bigsqcup_{n=0}^{+\infty} \mathrm{X}^{-1}(\mathrm{U}_n)\right) = \sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{X}^{-1}(\mathrm{U}_n)) = \sum_{n=0}^{+\infty} \mathcal{L}_{\mathrm{X}}(\mathrm{U}_n).$$

Notons que sur tout espace probabilisable, on peut définir une loi de probabilité $\mathcal{L}$ sans se donner de variable aléatoire.

II Variables discrètes

Définition II.1.

On dit qu'une variable aléatoire $\mathrm{X} : (\Omega, \mathcal{T}, \mathbb{P}) \longrightarrow (\mathrm{E}, \mathcal{U})$ est discrète lorsque $\mathrm{X}(\Omega)$ est dénombrable.

Notation II.2.

Soit $d \in \mathbb{N}^*$. Soit $\mathrm{X} : (\Omega, \mathcal{T}, \mathbb{P}) \longrightarrow (\mathrm{E}, \mathcal{U})$ une variable aléatoire. Soient $x, y \in \mathrm{E}$.

- → L'événement $\mathrm{X}^{-1}(\{x\})$ est noté $[\mathrm{X} = x]$.
- → Pour tout $\mathrm{A} \in \mathcal{U}$, l'événement $\mathrm{X}^{-1}(\mathrm{A})$ est noté $[\mathrm{X} \in \mathrm{A}]$.
- → Si $(\mathrm{E}, \mathcal{U}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, l'événement $\mathrm{X}^{-1}(]-\infty, x])$ (respectivement $\mathrm{X}^{-1}(]-\infty, x[)$) est noté $[\mathrm{X} \leq x]$ (respectivement $[\mathrm{X} < x]$).
- → Si $(\mathrm{E}, \mathcal{U}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, l'événement $\mathrm{X}^{-1}([x, +\infty[)$ (respectivement $\mathrm{X}^{-1}(]x, +\infty[)$) est noté $[\mathrm{X} \geq x]$ (respectivement $[\mathrm{X} > x]]$).
- → Si $(\mathrm{E}, \mathcal{U}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, et $x \leq y$, l'événement $\mathrm{X}^{-1}([x, y])$ est noté $[x \leq \mathrm{X} \leq y]$, et on note de manière analogue les événements $\mathrm{X}^{-1}(]x, y])$, $\mathrm{X}^{-1}([x, y[)$, et $\mathrm{X}^{-1}(]x, y[)$.

Proposition II.3.

Soit $d \in \mathbb{N}^*$. Soit $\lambda \in \mathbb{R}$. Si X et Y sont des variables aléatoires discrètes définies sur $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, alors $\mathrm{X} + \mathrm{Y}$ et $\lambda \mathrm{X}$ sont des variables aléatoires discrètes.

Démonstration. L'application

$$\begin{aligned} \mathrm{X}(\Omega) \times \mathrm{Y}(\Omega) &\longrightarrow \mathbb{R}^d \\ (x, y) &\longmapsto x + y. \end{aligned}$$

a pour image une partie dénombrable de $\mathbb{R}^d$ contenant $(\mathrm{X}+\mathrm{Y})(\Omega)$.
De plus, pour tout $z \in \mathbb{R}^d$, $(\mathrm{X}+\mathrm{Y})^{-1}(z) = \bigcup\limits_{\substack{(x,y)\in \mathrm{X}(\Omega)\times \mathrm{Y}(\Omega)\\ x+y=z}} ([\mathrm{X}=x] \cap [\mathrm{Y}=y]) \in \mathcal{T}$ en tant qu'union dénombrable d'événements.

Remarque II.4.

L'union étant disjointe, la loi de $\mathrm{X}+\mathrm{Y}$ est donnée par

$$\forall z \in (\mathrm{X}+\mathrm{Y})(\Omega), \qquad \mathcal{L}_{\mathrm{X}+\mathrm{Y}}(z) = \sum_{\substack{(x,y)\in \mathrm{X}(\Omega)\times \mathrm{Y}(\Omega)\\ x+y=z}} \mathbb{P}\left([\mathrm{X}=x] \cap [\mathrm{Y}=y]\right).$$

III Variables aléatoires discrètes indépendantes

Dans la suite, on considère que les variables aléatoires sont définies sur $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans $(\mathrm{E}, \mathcal{U})$.

Définition III.1.

Soit $n \in \mathbb{N}^*$. Soient $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des variables aléatoires.
On dit que $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes lorsque, pour tous $\mathrm{A}_1, \ldots, \mathrm{A}_n \in \mathcal{U}$, les événements $[\mathrm{X}_1 \in \mathrm{A}_1], \ldots, [\mathrm{X}_n \in \mathrm{A}_n]$ sont mutuellement indépendants.

Suites : On dit que $(\mathrm{X}_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires indépendantes lorsque pour toute partie finie I de $\mathbb{N}$, les variables aléatoires de $(\mathrm{X}_i)_{i\in \mathrm{I}}$ sont indépendantes.

Remarque III.2.

Soit $n \in \mathbb{N}^*$. Soient $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des variables aléatoires. Alors $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes si, et seulement si, les tribus images réciproques $\mathrm{X}_1^{-1}\langle\mathcal{U}\rangle, \ldots, \mathrm{X}_n^{-1}\langle\mathcal{U}\rangle$ sont indépendantes.

Proposition III.3.

Soient $n, d \in \mathbb{N}^*$. On suppose que $(\mathrm{E}, \mathcal{U}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Soient $\mathrm{X}_1, \ldots, \mathrm{X}_n$ des variables aléatoires discrètes. Alors :

- → $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes si, et seulement si, pour tout $(x_1, \ldots, x_n) \in \mathrm{E}^n$, les événements $[\mathrm{X}_1 = x_1]$, …, $[\mathrm{X}_n = x_n]$ sont mutuellement indépendants.
- → *Lemme des coalitions :* Soient $p, q \in \mathbb{N}$ tels que $p + q = n$. Soient $f : (\mathbb{R}^d)^p \longrightarrow \mathbb{R}^d$ et $g : (\mathbb{R}^d)^q \longrightarrow \mathbb{R}^d$ deux fonctions.
  Si $\mathrm{X}_1, \ldots, \mathrm{X}_n$ sont indépendantes alors, $f(\mathrm{X}_1, \ldots, \mathrm{X}_p)$ est une variable aléatoire indépendante de $g(\mathrm{X}_{p+1}, \ldots, \mathrm{X}_n)$.

Exemples. $\mathrm{X}_1 + \cdots + \mathrm{X}_p$ et $\mathrm{X}_{p+1} + \cdots + \mathrm{X}_n$ sont indépendantes.

BY-NC-SA

Pour tout $i \in \mathbb{N}$ tel que $1 \leq i < i+3 \leq n$, X_iX_{i+1} et $X_{i+2}X_{i+3}$ sont indépendantes.

Démonstration des propositions.

→ $(\Leftarrow)$ Soient $A_1, \ldots, A_n \subset E$. En utilisant le théorème d'associativité pour les familles sommables,

$$\begin{aligned}\mathbb{P}\left([X_1 \in A_1] \cap \ldots \cap [X_n \in A_n]\right) &= \sum_{(x_1,\ldots,x_n)\in A_1\times\cdots\times A_n} \mathbb{P}\left([X_1 = x_1] \cap \ldots \cap [X_n = x_n]\right) \\ &= \sum_{(x_1,\ldots,x_n)\in A_1\times\cdots\times A_n} \mathbb{P}\left(X_1 = x_1\right) \ldots \mathbb{P}\left(X_n = x_n\right) \\ &= \mathbb{P}(X_1 \in A_1) \times \cdots \times \mathbb{P}(X_n \in A_n).\end{aligned}$$

→ Soient $x, y \in \mathbb{R}^d$. Par indépendance des variables $X_1, \ldots, X_n$,

$$\begin{aligned}&\mathbb{P}([f(X_1, \ldots, X_p) = x] \cap g(X_{p+1}, \ldots, X_n) = y) \\ &= \sum_{\substack{(x_1,\ldots,x_n)\in E^n \\ f(x_1,\ldots,x_p)=x \\ g(x_{p+1},\ldots,x_n)=y}} \mathbb{P}\left([X_1 = x_1] \cap \ldots \cap [X_n = x_n]\right) \\ &= \sum_{\substack{(x_1,\ldots,x_n)\in E^n \\ f(x_1,\ldots,x_p)=x \\ g(x_{p+1},\ldots,x_n)=y}} \mathbb{P}\left([X_1 = x_1]\right) \ldots \mathbb{P}\left([X_n = x_n]\right) \\ &= \sum_{\substack{(x_1,\ldots,x_p)\in E^p \\ f(x_1,\ldots,x_p)=x}} \mathbb{P}\left([X_1 = x_1]\right) \ldots \mathbb{P}\left([X_p = x_p]\right) \mathbb{P}\left(g\left(X_{p+1}, \ldots, X_n\right) = y\right) \\ &= \mathbb{P}\left(f(X_1, \ldots, X_p) = x\right) \mathbb{P}\left(g\left(X_{p+1}, \ldots, X_n\right) = y\right).\end{aligned}$$

Exercice III.4.

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Soit $X : (\Omega, \mathcal{T}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire discrète. Donner une condition nécessaire et suffisante pour que X soit indépendante d'elle-même.

Exercice III.5.

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Soient X, Y deux variables aléatoires de $(\Omega, \mathcal{T}, \mathbb{P})$ à valeurs dans $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$, indépendantes, et telles que, pour tout $i \in \mathbb{N}^*$, $\mathbb{P}(X = i) = \mathbb{P}(Y = i) = \dfrac{1}{2^i}$.

1. Calculer $\mathbb{P}(X = Y)$.
2. Calculer $\mathbb{P}(X > Y)$.
3. Soit $N \in \mathbb{N}^*$. Calculer $\mathbb{P}(\min(X, Y) \leq N)$.

IV Lois usuelles

Dans cette partie, toutes les variables aléatoires sont discrètes.

(cc) BY-NC-SA

Schéma de Bernoulli[1] **:** Soit $p \in]0,1[$. Un schéma de Bernoulli est une suite finie d'expériences identiques, deux à deux indépendantes et ayant, chacune, exactement deux issues possibles : un succès ou un échec. La probabilité qu'une expérience soit un succès vaut p.
Cette suite d'expériences est modélisée par les termes d'une suite $(\mathrm{X}_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes, à valeurs dans $\{0,1\}$, et suivant la même loi de Bernoulli $\mathcal{B}(p)$ *i.e.* pour tout $k \in \mathbb{N}$, $\mathbb{P}(\mathrm{X}_k = 1) = p$.
Loi binomiale : Soient $n \in \mathbb{N}$, et $p \in]0,1[$. Considérons un schéma de Bernoulli d'ordre n (*i.e.* où n expériences sont menées) où la probabilité de succès vaut p. Le nombre de succès est aléatoire et est compris entre 0 et n, et la probabilité que k succès aient lieu vaut la probabilité que l'une des situations où $k \in [\![0;n]\!]$ succès sont disséminés parmi les n expériences ait lieu : l'une de ces situations a une probabilité $p^k(1-p)^{n-k}$ d'avoir lieu (par indépendance), et il y a $\binom{n}{k}$ combinaisons de k succès parmi n expériences à deux issues.
La variable aléatoire $\mathrm{S} = \sum_{k=1}^{n} \mathrm{X}_k$ qui compte le nombre de succès d'un schéma de Bernoulli d'ordre n et de probabilité de succès p est à valeurs dans $[\![0;n]\!]$ et suit la loi binomiale $\mathcal{B}(n,p)$:

$$\forall k \in [\![0;n]\!], \qquad \mathbb{P}(\mathrm{S} = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Loi de Rademacher : Il est possible de modéliser un schéma de Bernoulli d'ordre $n \in \mathbb{N}$ et de probabilité de succès $p \in]0,1[$ par les termes d'une suite $(\mathrm{Y}_n)_{n\geq 0}$ de variables aléatoires à valeurs dans $\{-1,1\}$ indépendantes et identiquement distribuées suivant la loi de Rademacher *i.e.* pour tout $k \in \mathbb{N}$, $\mathbb{P}(\mathrm{Y}_k = 1) = p$.
Dans un tel contexte, il est loisible de se ramener à la suite $(\mathrm{X}_n)_{n\in\mathbb{N}}$ de variables de Bernoulli correspondante en remarquant que, presque sûrement,

$$\forall k \in \mathbb{N}, \qquad \mathrm{Y}_k = 2\mathrm{X}_k - 1.$$

Loi géométrique : Considérons un schéma de Bernoulli infini *i.e.* où l'on réalise une infinité dénombrable d'expériences identiques et indépendantes, et où la probabilité de succès vaut $p \in]0,1[$.
Le temps d'attente du premier succès est une variable aléatoire T à valeurs dans $\mathbb{N}^*$ suivant une loi géométrique $\mathcal{G}(p)$.
Soit $n \in \mathbb{N}^*$. Pour que le premier succès ait lieu à la n-ième expérience, il faut et il suffit que les expériences antérieures se soldent toutes par des échecs et que la n-ième soit un succès, c'est-à-dire que

$$\mathbb{P}(\mathrm{T} = n) = (1-p)^{n-1} p.$$

Théorème IV.1.

Soit $p \in]0,1[$. Soit $(\mathrm{X}_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées suivant la loi de Bernoulli $\mathcal{B}(p)$. Alors

- → Presque sûrement, il existe $k \in \mathbb{N}^*$ tel que $\mathrm{X}_k = 1$.
- → $\mathrm{Y} = \min\{k \in \mathbb{N}^* \mid \mathrm{X}_k = 1\}$ est une variable aléatoire à valeurs dans $\mathbb{N}^*$ suivant la loi géométrique $\mathcal{G}(p)$.

Démonstration.

→ La suite d'événements $\left(\bigcap_{k=1}^{n} [\mathrm{X}_k = 0]\right)_{n\in\mathbb{N}^*}$ est décroissante, donc par continuité décroissante et uni-

1. Rappelons que Jacques Bernoulli n'est pas une nouille.

(cc) BY-NC-SA

cité de la limite :

$$\underbrace{\mathbb{P}\left(\bigcap_{k=1}^{n}[\mathrm{X}_k=0]\right)}_{=(1-p)^n} \xrightarrow[n\to+\infty]{} \mathbb{P}\left(\neg[\exists k\in\mathbb{N}^*,\quad \mathrm{X}_k=1]\right)=0.$$

→ Soit $n\in\mathbb{N}^*$. $[\mathrm{Y}=n]=[\mathrm{X}_0=0]\cap\ldots\cap[\mathrm{X}_{n-1}=0]\cap[\mathrm{X}_n=1]\in\mathcal{T}$, donc Y est mesurable, c'est donc une variable aléatoire à valeurs dans $(\mathbb{N}^*,\mathcal{P}(\mathbb{N}^*))$.
Par indépendance,

$$\mathbb{P}(\mathrm{Y}=n)=(1-p)^{n-1}p.$$

Remarque IV.2.

Souvent, le premier succès arrive assez tôt. En effet, il existe $\mu>0$ tel que $1-p=e^{-\mu}$, donc $(1-p)^{n-1}p=e^{-(n-1)\mu}(1-e^{-\mu})$, ce qui exhibe une décroissance exponentielle de la probabilité du premier succès avec le rang de ce succès.

Proposition IV.3.

Soit $p\in\,]0,1[$. Soit X une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$. Alors

$$\forall(n,k)\in\mathbb{N}^2,\qquad \mathbb{P}(\mathrm{X}>n+k\,|\,\mathrm{X}>n)=\mathbb{P}(\mathrm{X}>k).$$

Vocabulaire : On dit qu'une loi géométrique est une loi sans mémoire.

Démonstration. Soit $(n,k)\in\mathbb{N}^2$.

$$\begin{aligned}\mathbb{P}(\mathrm{X}>n+k\,|\,\mathrm{X}>n)&=\frac{\mathbb{P}(\mathrm{X}>n+k)}{\mathbb{P}(\mathrm{X}>n)}\\&=\frac{\displaystyle\sum_{m=n+k+1}^{+\infty}(1-p)^{m-1}\cancel{p}}{\displaystyle\sum_{m=n+1}^{+\infty}(1-p)^{m-1}\cancel{p}}\\&=\frac{(1-p)^{n+k}\times\cancel{\dfrac{1}{1-(1-p)}}}{(1-p)^{n}\times\cancel{\dfrac{1}{1-(1-p)}}}\\&=(1-p)^k=\mathbb{P}(\mathrm{X}>k).\end{aligned}$$

formule utile pour une loi géométrique

Loi de Poisson : Soit $\lambda>0$. On dit qu'une variable aléatoire X à valeurs dans $\mathbb{N}$ suit une loi de Poisson $\mathcal{P}(\lambda)$ lorsque pour tout $n\in\mathbb{N}$, $\mathbb{P}(\mathrm{X}=n)=e^{-\lambda}\dfrac{\lambda^n}{n!}$.

Théorème IV.4.

Soit $\lambda>0$. Soit $(p_n)\in\,]0,1[^{\mathbb{N}}$ telle que $np_n\xrightarrow[n\to+\infty]{}\lambda$. Soit $(\mathrm{X}_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour tout $n\in\mathbb{N}$, X_n suit la loi $\mathcal{B}(n,p_n)$.
Alors $(\mathrm{X}_n)_{n\in\mathbb{N}}$ converge en loi vers une loi de Poisson $\mathcal{P}(\lambda)$ *i.e.*

$$\forall m\in\mathbb{N},\qquad \mathbb{P}(\mathrm{X}_n=m)\xrightarrow[n\to+\infty]{}e^{-\lambda}\frac{\lambda^m}{m!}.$$

BY-NC-SA

Démonstration. Soit $m \in \mathbb{N}$. Pour tout entier $n > m$,

$$\begin{aligned}\mathbb{P}(\mathrm{X}_n = m) &= \binom{n}{m} p_n^m (1-p_n)^{n-m} \\ &= \frac{n \times \cdots \times (n-m+1)}{m!} (1-p_n)^{-m} p_n^m \, (1-p_n)^n .\end{aligned}$$

Or $(1-p_n)^n \xrightarrow[n\to+\infty]{} e^{-\lambda}$ car $p_n \underset{n\to+\infty}{\sim} \dfrac{\lambda}{n}$. De plus,

$$n \times \cdots \times (n-m+1)(1-p_n)^{-m} p_n^m = \underbrace{(np_n)^m}_{\xrightarrow[n\to+\infty]{}\lambda^m} \underbrace{\left(1-\frac{1}{n}\right) \times \cdots \times \left(1-\frac{m-1}{n}\right)}_{\xrightarrow[n\to+\infty]{}1} \underbrace{(1-p_n)^{-m}}_{\xrightarrow[n\to+\infty]{}1} .$$

ce qu'il fallait démontrer.

Exercice IV.5.

Soient $\lambda > 0$ et $p \in]0,1[$. On modélise le nombre d'œufs pondus par un poisson par une loi de Poisson $\mathcal{P}(\lambda)$. On modélise l'éclosion d'un œuf indépendamment des autres par une loi de Bernoulli $\mathcal{B}(p)$. Déterminer la loi du nombre d'œufs éclos.

Exercice IV.6.

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Soit $p \in]0,1[$. Soit $(\mathrm{X}_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires définies sur $(\Omega, \mathcal{T})$, indépendantes et identiquement distribuées suivant la loi de Rademacher de paramètre p.
On pose $\mathrm{S}_0 = 0$ et, pour tout $n \in \mathbb{N}^*$, $\mathrm{S}_n = \mathrm{X}_1 + \cdots + \mathrm{X}_n$. Pour tout $m \in \mathbb{N}^*$, on pose $\mathrm{A}_m = \bigcap_{n=0}^{+\infty} [|\mathrm{S}_n| \leq m]$.

1. On suppose que $p = \dfrac{1}{2}$. Montrer que pour tout $m \in \mathbb{N}^*$, $\mathbb{P}(\mathrm{A}_m) = 0$.
2. On suppose que $p \neq \dfrac{1}{2}$. Montrer que $\mathbb{P}([\mathrm{S}_n = 0]$ infiniment souvent$) = 0$.

V Compléments

Fonction de répartition : Soit X une variable aléatoire à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On introduit la fonction de répartition de X, notée F_X, définie par

$$\begin{aligned}\mathrm{F}_\mathrm{X} : \mathbb{R} &\longrightarrow [0,1] \\ x &\longmapsto \mathbb{P}(\mathrm{X} \leq x).\end{aligned}$$

Remarque. Cette fonction ne caractérise que la loi de X.

BY-NC-SA

Proposition V.1.

Soit X une variable aléatoire à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Alors :

→ F_X est croissante.

→ $F_X(x) \xrightarrow[x \to -\infty]{} 0$ et $F_X(x) \xrightarrow[x \to +\infty]{} 1$.

→ F_X est continue à droite en tout point de $\mathbb{R}$.

→ Soit $a \in \mathbb{R}$. F_X n'est pas continue en a si, et seulement si, $\mathbb{P}(X = a) > 0$.

Démonstration.

→ La croissance de F_X découle de la croissance de $\mathbb{P}$.

→ Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle décroissante divergeant vers $-\infty$. $([X \leq x_n])_{n\in\mathbb{N}^*}$ est une suite décroissante d'événements donc par continuité décroissante,

$$F_X(x_n) \xrightarrow[n \to +\infty]{} \mathbb{P}\left(\bigcap_{n=0}^{+\infty} [X \leq x_n]\right) = \mathbb{P}(\varnothing) = 0.$$

Par un raisonnement analogue, on prouve de même que $F_X(x) \xrightarrow[n \to +\infty]{} 1$

→ Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle décroissante convergeant vers $a \in \mathbb{R}$. $([X \leq x_n])_{n\in\mathbb{N}^*}$ est une suite décroissante d'événements donc par continuité décroissante, $F_X(x_n) \xrightarrow[n \to +\infty]{} F_X(a)$. Donc F_X est continue à droite en a.

→ $(\Rightarrow)$ Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle décroissante convergeant vers a et $(y_n)_{n\in\mathbb{N}}$ une suite réelle croissante convergeant vers a sans valoir a. Par continuité croissante et décroissante, on montre que $F_X(x_n) \xrightarrow[n \to +\infty]{} F_X(a)$ et $F_X(y_n) \xrightarrow[n \to +\infty]{} F_X(a^-)$.
Donc $\mathbb{P}(X = a) = F_X(a) - F_X(a^-) > 0$ car F_X est discontinue en a.
$(\Leftarrow)$ Si F_X est continue en a, alors $\mathbb{P}(X = a) = F_X(a) - F_X(a^-) = 0$.

Variables à densité : Soit X une variable aléatoire à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On dit que X est une variable aléatoire à densité s'il existe une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}^+$ intégrable et telle que

$$\forall a \in \mathbb{R}, \qquad F_X(a) = \int_{-\infty}^{a} f(t)\mathrm{d}t.$$

Une telle fonction détermine entièrement la loi de X.

Exemples.

→ Pour tout $t \in \mathbb{R}$, $f(t) = e^{-t}\mathbb{1}_{[0,+\infty[}(t)$. f est la densité d'une loi exponentielle $\mathcal{E}(1)$.

→ Pour tout $t \in \mathbb{R}$, $f(t) = \dfrac{1}{\sqrt{2\pi}} e^{-t^2}$. f est la densité d'une loi normale centrée réduite $\mathcal{N}(0, 1)$.

→ Pour tout $t \in \mathbb{R}$, $f(t) = \dfrac{1}{(1+t^2)\pi}$. f est la densité d'une loi de Cauchy.

Il est alors possible de définir, sous réserve d'existence, l'espérance et le moment d'ordre 2 de telles variables par $\mathbb{E}(X) = \int_{-\infty}^{+\infty} tf(t)\mathrm{d}t$, et $\mathbb{E}(X^2) = \int_{-\infty}^{+\infty} t^2 f(t)\mathrm{d}t$.

Proposition V.2.

Si X est une variable aléatoire à densité, alors pour tout $a \in \mathbb{R}$, $\mathbb{P}(X = a) = 0$.

Démonstration. Soit $a \in \mathbb{R}$. $\mathbb{P}(X = a) = F_X(a) - F_X(a^-) = 0$ car la fonction de répartition d'une variable aléatoire à densité est continue. En effet, soit $\varepsilon > 0$. Pour tout $x \in [a - \varepsilon, a]$,

$$|F_X(a) - F_X(x)| \leq \|f\|_{\infty,[a-\varepsilon,a]} |x - a|,$$

BY-NC-SA

ce qui démontre à la continuité à gauche, donc la continuité.

(cc) BY-NC-SA

Correction de l'exercice III.4. :

Supposons que X soit indépendante d'elle-même. Soit $x \in \mathbb{R}$ tel que $\mathbb{P}(X = x) > 0$. Alors par indépendance des événements $[X = x]$ et $[X = x]$, $\mathbb{P}(X = x) = \mathbb{P}([X = x] \cap [X = x]) = \mathbb{P}(X = x)^2$, donc $\mathbb{P}(X = x) = 1$. Donc X est presque sûrement constante.
Réciproquement, toute variable aléatoire constante est indépendante d'elle-même.
En conclusion, une variable aléatoire est indépendante d'elle-même si, et seulement si, elle est presque sûrement constante.

Correction de l'exercice III.5. :

1. $\mathbb{P}(X = Y) = \mathbb{P}\left(\bigsqcup_{i=1}^{+\infty} [X = i] \cap [Y = i]\right) \underset{\substack{\uparrow \\ \text{indépendance}}}{=} \sum_{i=1}^{+\infty} \frac{1}{4^i} = \frac{1}{3}.$

2. $\mathbb{P}(X > Y) = \mathbb{P}\left(\bigsqcup_{i=1}^{+\infty} [Y = i] \cap [X > i]\right) \underset{\substack{\uparrow \\ \text{indépendance}}}{=} \sum_{i=1}^{+\infty} \frac{1}{2^i} \sum_{j=1}^{+\infty} \frac{1}{2^{i+j}} = \sum_{i=1}^{+\infty} \frac{1}{4^i} = \frac{1}{3}.$

3. $\mathbb{P}(\min(X, Y) \leq N) = 1 - \mathbb{P}(\underbrace{\min(X, Y) > N}_{=[X>N]\cap[Y>N]}) \underset{\substack{\uparrow \\ \text{indépendance}}}{=} 1 - \mathbb{P}(X > N)\mathbb{P}(Y > N) \underset{\substack{\uparrow \\ \text{calcul aisé}}}{=} 1 - \frac{1}{4^N}$

Correction de l'exercice IV.5. :

Notons X la variable aléatoire désignant le nombre d'œufs pondus, et Y le nombre d'œufs ayant éclos.
Pour tout $n \in \mathbb{N}$, $\mathbb{P}(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$ et pour tout $k \in [\![0; n]\!]$, $\mathbb{P}(Y = k \mid X = n) = \binom{n}{k} p^k (1-p)^{n-k}$ car lorsque $[X = n]$ est réalisé, chaque éclosion est considérée, indépendamment des autres comme un succès dans un schéma de Bernoulli d'ordre n et de probabilité de succès p.
Soit $k \in \mathbb{N}$.

$$\begin{aligned}
\mathbb{P}(Y = k) &= \sum_{n=k}^{+\infty} \mathbb{P}([Y = k] \cap [X = n]) = \sum_{n=k}^{+\infty} \mathbb{P}(Y = k \mid X = n)\mathbb{P}(X = n) \\
&= \sum_{n=k}^{+\infty} \binom{n}{k} p^k (1-p)^{n-k} \times e^{-\lambda} \frac{\lambda^n}{n!} \\
&\underset{\substack{\uparrow \\ m \leftarrow n-k}}{=} e^{-\lambda} \sum_{m=0}^{+\infty} \frac{\cancel{(m+k)!}}{k! m!} p^k (1-p)^m \frac{\lambda^{m+k}}{\cancel{(m+k)!}} \\
&= e^{-\lambda} \frac{(p\lambda)^k}{k!} \sum_{m=0}^{+\infty} \frac{(1-p)^m \lambda^m}{m!} = e^{-\lambda} \frac{(p\lambda)^k}{k!} e^{(1-p)\lambda} \\
&= e^{-p\lambda} \frac{(p\lambda)^k}{k!},
\end{aligned}$$

donc Y suit la loi de Poisson $\mathcal{P}(p\lambda)$.

Correction de l'exercice IV.6. :

1. Soit $m \in \mathbb{N}^*$. Soit $k \in \mathbb{N}^*$. Posons $B_k = \bigcap_{i=0}^{2m+1} [X_{km+i} = 1]$. Par indépendance, $\mathbb{P}(B_k) = \frac{1}{2^{2m+2}}$ et par construction, si B_k est réalisé, alors $|S_{mk+2m+1} - S_{mk-1}| \geq 2m + 2$ est réalisé, donc $[S_{mk+2m+1} \notin [\![-m; m]\!]] \cup [S_{mk-1} \notin [\![-m; m]\!]]$ est réalisé, et ceci est valable pour tout $k \in \mathbb{N}^*$. Pour tout $\ell \in \mathbb{N}^*$

BY-NC-SA

posons $k_\ell = 4\ell m$. Par le lemme des coalitions, pour tous $\ell, \ell' \in \mathbb{N}^*$ tels que $\ell \neq \ell'$, les événements B_ℓ et $\mathrm{B}_{\ell'}$ sont indépendants. De plus $\sum_{\ell \geq 1} \mathbb{P}(\mathrm{B}_{k_\ell})$ est divergente (son terme général est non nul et indépendant de ℓ).

Le premier théorème de Borel-Cantelli assure alors que $\mathbb{P}\left(\bigcap_{\ell=1}^{+\infty}\left(\bigcup_{\ell'=\ell+1}^{+\infty} \mathrm{B}_{k'_\ell}\right)\right) = 1$

Autrement dit, presque sûrement, B_{k_ℓ} est réalisé infiniment souvent, donc l'événement A_m est négligeable : il existe une partie infinie J de $\mathbb{N}$ telle que pour tout $n \in \mathrm{J}$, $\mathrm{S}_n \notin [\![-m; m]\!]$.

2. Pour tout $m \in \mathbb{N}^*$, posons $u_m = \mathbb{P}(\mathrm{S}_{2m} = 0)$.
Soit $m \in \mathbb{N}^*$. Remarquons que le retour à l'origine se fait si, et seulement si, autant de déplacements à gauche qu'à droite ont lieu et, sur $2m$ déplacements il y a $\binom{2m}{m}$ telles combinaisons et par indépendance des déplacements, $u_m = \binom{2m}{m} p^m (1-p)^m$. Alors

$$\frac{u_{m+1}}{u_m} = p(1-p)\frac{(2m+2)(2m+1)}{(m+1)^2} \xrightarrow[m\to+\infty]{} 4p(1-p) < 1.$$

En effet, $y = x(1-x)$ est une parabole tournée vers le bas atteignant son maximum entre les deux racines de $\mathrm{X}(\mathrm{X}-1)$, donc en $x = \frac{1}{2}$, et celui-ci vaut $\frac{1}{4}$. $p \neq \frac{1}{2}$ donc l'inégalité est stricte. D'après la règle de d'Alembert, $\sum_{m \geq 1} u_m$ converge.
Le deuxième théorème de Borel-Cantelli assure alors que $\mathbb{P}([\mathrm{S}_{2m} = 0]$ infiniment souvent$) = 0$.

(cc) BY-NC-SA

Chapitre 30.4

Couples de variables aléatoires

Soit $d \in \mathbb{N}^*$. Dans tout ce chapitre, les variables aléatoires sont définies sur l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans l'espace mesurable $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

I Loi conjointe, loi marginale

Définition I.1.

Soient X, Y deux variables aléatoires discrètes. La loi conjointe de (X, Y) est définie par

$$\forall (x,y) \in (\mathbb{R}^d)^2, \qquad \mathcal{L}_{(\mathrm{X},\mathrm{Y})}(\{(x,y)\}) = \mathbb{P}([\mathrm{X}=x] \cap [\mathrm{Y}=y]).$$

Les lois $\mathcal{L}_\mathrm{X}$ et $\mathcal{L}_\mathrm{Y}$ sont appelées lois marginales de $\mathcal{L}_{(\mathrm{X},\mathrm{Y})}$.

Remarque I.2.

Les lois marginales se calculent à partir de la loi conjointe :

$$\forall x \in \mathbb{R}^d, \qquad \mathbb{P}(\mathrm{X}=x) = \mathbb{P}\left(\bigsqcup_{y \in \mathrm{Y}(\Omega)} [\mathrm{X}=x] \cap [\mathrm{Y}=y]\right) = \sum_{y \in \mathrm{Y}(\Omega)} \mathcal{L}_{(\mathrm{X},\mathrm{Y})}\left(\{(x,y)\}\right).$$

Exercice I.3.

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé.
Soient X et Y deux variables aléatoires à valeurs dans $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ telles que pour tout $(i,j) \in \mathbb{N}^2$

$$\mathcal{L}_{(\mathrm{X},\mathrm{Y})}(\{(i,j)\}) = \frac{1}{2^{i+j+2}}.$$

1. Vérifier que $\mathcal{L}_{(\mathrm{X},\mathrm{Y})}$ est bien une loi de probabilité.
2. Calculer les lois marginales $\mathcal{L}_\mathrm{X}$ et $\mathcal{L}_\mathrm{Y}$.
3. X et Y sont-elles indépendantes ?

II Produit de convolution

Soient X et Y deux variables aléatoires à valeurs dans $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. On définit le produit de convolution de $\mathcal{L}_\mathrm{X}$ et $\mathcal{L}_\mathrm{Y}$, et note $\mathcal{L}_\mathrm{X} \star \mathcal{L}_\mathrm{Y}$, par :

$$\forall n \in \mathbb{N}, (\mathcal{L}_\mathrm{X} \star \mathcal{L}_\mathrm{Y})(\{n\}) = \sum_{\substack{(i,j) \in \mathbb{N}^2 \\ i+j=n}} \mathcal{L}_\mathrm{X}(\{i\}) \mathcal{L}_\mathrm{Y}(\{j\}).$$

(cc) BY-NC-SA

Si, de plus, X et Y sont indépendantes, alors $\mathcal{L}_{X+Y} = \mathcal{L}_X \star \mathcal{L}_Y$.En effet, pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(X+Y=n) = \sum_{\substack{(i,j)\in\mathbb{N}^2\\ i+j=n}} \mathbb{P}([X=i]\cap[Y=j]) \underset{\substack{\uparrow\\ \text{indépendance}}}{=} \sum_{\substack{(i,j)\in\mathbb{N}^2\\ i+j=n}} \mathbb{P}(X=i)\mathbb{P}(Y=j) = (\mathcal{L}_X \star \mathcal{L}_Y)(\{n\}).$$

Exemples. On suppose que $d = 1$.

→ **Produit de convolution de lois binomiales :** Soient $n, m \in \mathbb{N}^*$ et $p \in]0,1[$. Soient $X_1, \dots, X_{n+m}$ des variables aléatoires indépendantes et identiquement distribuées suivant la loi $\mathcal{B}(p)$. Posons $X = X_1 + \cdots + X_n$ et $Y = X_{n+1} + \cdots + X_{n+m}$. Alors $\mathcal{L}_{X+Y}$ est une loi binomiale $\mathcal{B}(n+m, p)$.

Produit de convolution de lois de Poisson : Soient $\lambda, \mu > 0$. Posons $\mathcal{L}_1 = \mathcal{P}(\lambda)$ $\mathcal{L}_2 = \mathcal{P}(\mu)$. Alors, pour tout $n \in \mathbb{N}$,

$$\begin{aligned}(\mathcal{L}_1 \star \mathcal{L}_2)(\{n\}) &= \sum_{\substack{(i,j)\in\mathbb{N}^2\\ i+j=n}} \frac{\lambda^i}{i!}e^{-\lambda}\cdot\frac{\mu^j}{j!}e^{-\mu} \\ &= e^{-(\lambda+\mu)} \sum_{\substack{(i,j)\in\mathbb{N}^2\\ i+j=n}} \frac{1}{n!}\binom{n}{i}\lambda^i\mu^{n-i} \\ &= e^{-(\lambda+\mu)}\sum_{i=0}^{n}\frac{1}{n!}\binom{n}{i}\lambda^i\mu^{n-i} = \frac{(\lambda+\mu)^n}{n!}e^{-(\lambda+\mu)}.\end{aligned}$$

On reconnaît une loi de Poisson $\mathcal{P}(\lambda+\mu)$.

(cc) BY-NC-SA

Correction de l'exercice I.3. :

1. Le nombre de manières d'écrire un entier naturel n non nul sous la forme d'une somme de deux entiers naturels vaut le nombre de manières de placer un bâton parmi n objets alignés de sorte à les scinder en deux groupes d'objets en autorisant toutefois qu'au moins l'un des groupes n'en contienne aucun. De tels emplacements pour le bâton, il y en a $n+1$. Ainsi, dans la somme calculée ci-dessous, il y a $n+1$ termes égaux à $\dfrac{1}{2^{i+j+2}}$ lorsque $i+j=n$.
   Ce raisonnement est aisément généralisable si l'on veut écrire un entier naturel non nul sous la forme d'une somme à plus de deux termes d'entier naturel.
   $$\sum_{(i,j)\in\mathbb{N}^2} \mathcal{L}_{(\mathrm{X},\mathrm{Y})}(\{(i,j)\}) = \sum_{n=0}^{+\infty}\frac{n+1}{2^{n+2}} = \frac{1}{4}\sum_{n=0}^{+\infty}\frac{n+1}{2^n} = \frac{1}{4}\cdot\frac{1}{\left(1-\frac{1}{2}\right)^2} = 1,$$
   donc $\mathcal{L}_{(\mathrm{X},\mathrm{Y})}$ est bien une loi de probabilité.
   Il est loisible de retenir les sommes des séries dérivées d'une série géométrique convergente. Soit $q\in\mathbb{C}$ tel que $|q|<1$. Alors
   $$\sum_{n=1}^{+\infty} nq^{n-1} = \frac{1}{(1-q)^2} \qquad \text{et} \qquad \sum_{n=2}^{+\infty} n(n-1)q^{n-2} = \frac{2}{(1-q)^3}.$$
2. Soit $i\in\mathbb{N}$.
   $$\mathcal{L}_{\mathrm{X}}(\{i\}) = \sum_{j=0}^{+\infty}\mathcal{L}_{(\mathrm{X},\mathrm{Y})}(\{(i,j)\}) = \frac{1}{2^{i+2}}\sum_{j=0}^{+\infty}\frac{1}{2^j} = \frac{1}{2^{i+1}} = \mathcal{L}_{\mathrm{Y}}(\{i\}).$$
3. Oui, car pour tout $(i,j)\in\mathbb{N}^2$, $\mathcal{L}_{\mathrm{X}}(\{i\})\mathcal{L}_{\mathrm{Y}}(\{j\}) = \dfrac{1}{2^{i+j+2}} = \mathcal{L}_{(\mathrm{X},\mathrm{Y})}(\{i,j\})$.

(cc) BY-NC-SA

CHAPITRE 32

Exponentielle d'une matrice

Dans tout le chapitre $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $(\mathrm{E}, \|.\|)$ est un $\mathbb{K}$ espace vectoriel normé de dimension finie n, et $\mathcal{L}(\mathrm{E})$ désigne l'espace vectoriel (de dimension finie n^2) des endomorphismes linéaires de E, implicitement muni de la norme d'opérateur $|||.|||$ induite par $\|.\|$.

Notations

→ On note pour tout $h \in \mathcal{L}(\mathrm{E})$ et $\mathrm{R} \geq 0$,

$$\mathrm{B}_{|||.|||}(h, \mathrm{R}) = \{u \in \mathcal{L}(\mathrm{E}),\ |||u - h||| < \mathrm{R}\},$$

et lorsqu'il n'y a pas ambiguïté, on note tout simplement $\mathrm{B}(h, \mathrm{R})$.

→ Pour toute suite $(a_k) \in \mathbb{K}^{\mathbb{N}}$, on dit que $\sum a_k z^k$ est une série entière de rayon au moins $\mathrm{R} \in]0, +\infty[$ lorsque pour tout $z \in \mathbb{C}$ tel que $|z| < \mathrm{R}$, la série $\sum\limits_{k \geq 0} a_k z^k$ est convergente, *i.e.* la suite $\left(\sum\limits_{k=0}^{n} a_k z\right)_{n \in \mathbb{N}}$ est convergente.

→ Pour toute base β de E et tous $u \in \mathcal{L}(\mathrm{E})$ et $x \in \mathrm{E}$, on note $[u]_\beta$ la matrice de u dans la base β et $[x]_\beta$ le vecteur colonne des coordonnées de x dans la base β.

→ Pour tout $u \in \mathcal{L}(\mathrm{E})$, on note $\mathbb{K}[u] = \mathrm{Vect}(\{u^k,\ k \in \mathbb{N}\})$.

→ Pour toute matrice $\mathrm{M} \in \mathcal{M}_n(\mathbb{C})$, on note $\mathbb{K}[\mathrm{M}] = \mathrm{Vect}(\{\mathrm{M}^k,\ k \in \mathbb{N}\})$.

→ Pour toute matrice $\mathrm{M} \in \mathcal{M}_n(\mathbb{C})$, on note $\langle \mathrm{M} \rangle_{\mathbb{K}} = \{\mathrm{PMP}^{-1},\ \mathrm{P} \in \mathrm{GL}_n(\mathbb{K})\}$.

→ Soit X, Y deux ensembles munis de lois de compositions internes, on note $\mathrm{Hom}(\mathrm{X}, \mathrm{Y})$ l'ensemble des homomorphismes de X dans Y.

I Généralités et propriétés fonctionnelles

1. Calcul opérationnel

Soit $\sum a_k z^k$ une série entière de rayon $\mathrm{R} \in]0, +\infty]$.

Proposition I.1.

1. Pour tout $u \in \mathcal{L}(\mathrm{E})$ tel que $|||u||| < \mathrm{R}$, $\sum a_k u^k$ est une série absolument convergente.
2. La fonction

$$f : \begin{cases} \mathrm{B}_{|||.|||}(0, \mathrm{R}) & \longrightarrow \mathcal{L}(\mathrm{E}) \\ u & \longmapsto f(u) = \sum\limits_{k=0}^{\infty} a_k u^k \end{cases}$$

est continue.

1. On rappelle qu'une norme d'opérateur est une norme d'algèbre (sous-multiplicative) et que donc

(cc) BY-NC-SA

$\left\|u^k\right\| \leq \|u\|^k$. Par conséquent, on a

$$\forall k \in \mathbb{N},\ \left\|a_k u^k\right\| \leq |a_k| \cdot \|u\|^k$$

et la série $\sum |a_k| \, \|u\|^k$ est convergente car $\|u\| < \mathrm{R}$. Ainsi, la série est absolument convergente. De plus, $\mathcal{L}(\mathrm{E})$ étant complet (E est de dimension finie sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et donc $\mathcal{L}(\mathrm{E})$ aussi), la série converge bien dans $\mathcal{L}(\mathrm{E})$ aussi (d'après le corolaire III.4 du chapitre 11.2).

Remarque. On rappelle que, par équivalence des normes en dimension finie sur $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, la convergence normale ne dépend pas de la norme choisie (qu'elle soit triple ou pas), donc ce résultat est toujours valable si on avait choisi une autre norme N sur E et sa norme d'opérateur associée $\|.\|_{\mathrm{N}}$ (ici par contre, on oblige qu'elle soit triple pour assurer sa sous-multiplicativité). En particulier, pour montrer que la série $\sum a_k u^k$ converge, il suffit de trouver une norme $\|.\|$ sur E dont la norme triple vérifie $\|u\|_{\|.\|} < \mathrm{R}$ où R est le rayon de convergence de $\sum a_k z^k$.

2. Soit ρ tel que $0 < \rho < \mathrm{R}$. La série $\sum a_k z^k$ est normalement convergente sur $\mathrm{B}(0, \rho)$, ceci étant car

$$\forall u \in \mathrm{B}(0, \rho),\ \forall k \in \mathbb{N},\ \left\|a_k u^k\right\| \leq |a_k| \cdot \|u\|^k \leq |a_k|\, \rho^k,$$

et $\sum_{k=0}^{\infty} |a_k| \cdot \rho^k$ est convergente. En particulier, $\mathrm{B}(0, \rho)$ étant ouvert, f est donc continue sur ce dernier.

On déduit donc que f est continue sur $\bigcup_{0<\rho<\mathrm{R}} \mathrm{B}(0, \rho) = \mathrm{B}(0, \mathrm{R})$.

Notation. Par abus de notation, parfois on confondra série entière $\sum a_k z^k$ et son application associée sur $\mathcal{L}(\mathrm{E})$, $u \longmapsto \sum_{k=0}^{\infty} a_k u^k$. Par exemple, on dira que la série entière $\sum a_k u^k$ est convergente sur $\mathrm{B}_{\|.\|}(h, \mathrm{R})$ (avec $(h, \mathrm{R}) \in \mathcal{L}(\mathrm{E}) \times \mathbb{R}^+$) si la suite $\left(\sum_{k=0}^{n} a_k u^k\right)_{n \in \mathbb{N}}$ est convergente pour tout $u \in \mathrm{B}_{\|.\|}(h, \mathrm{R})$.

Exercice I.2.

1. On suppose que $\mathbb{K} = \mathbb{C}$, $u \in \mathcal{L}(\mathrm{E})$ et $\sum_{k=0}^{\infty} a_k z^k$ une série entière de rayon $\mathrm{R} > 0$. Montrer que si
$$\rho(u) := \sup\{|\lambda|,\ \lambda \in \mathrm{VP}(u)\} < \mathrm{R},$$
alors $\sum a_k u^k$ est convergente.

2. Soit $u \in \mathcal{L}(\mathrm{E})$, $f = \sum_{k=0}^{\infty} a_k z^k$ et $g = \sum_{k=0}^{\infty} b_k z^k$ tel que $\|u\| < \min(\rho(f), \rho(g))$. Montrer que $f(u) \circ g(u) = (fg)(u)$.

(cc) BY-NC-SA

Définition I.3.

Soit $u \in \mathcal{L}(\mathrm{E})$, on appelle exponentielle de u l'endomorphisme $\exp(u) := \sum_{k=0}^{\infty} \frac{u^k}{k!} \in \mathcal{L}(\mathrm{E})$.

Par ce qui précède,

$$\exp : \begin{cases} \mathcal{L}(\mathrm{E}) & \longrightarrow \mathcal{L}(\mathrm{E}) \\ u & \longmapsto \exp(u) = \sum_{k=0}^{\infty} \frac{u^k}{k!} \end{cases}$$

est bien définie partout et est continue (et même vérifie $|||\exp(u)||| \leq e^{|||u|||}$).

Remarque. Pour montrer que l'application exp telle que définie ci-dessus est bien définie, il suffit d'appliquer le théorème précédent à $\sum_{k=0}^{\infty} \frac{1}{k!} z^k$ qui est de rayon de convergence infini.

Proposition I.4.

Notons pour tout $(u, v) \in \mathcal{L}(\mathrm{E})^2$, $[u, v] := u \circ v - v \circ u$. Pour tout, $(u, v) \in \mathcal{L}(\mathrm{E})^2$ on a

$$[u, v] = 0 \Longrightarrow \exp(u + v) = \exp(u) \circ \exp(v).$$

En particulier, $\exp(u) \circ \exp(-u) = \mathrm{Id}_{\mathrm{E}}$ et donc exp est à valeurs dans $\mathrm{GL}(\mathrm{E})$.

On fournira deux démonstrations : la première type série entière et une seconde, plus longue, moins calculatoire, et plus aspect differentiel/morphisme (pour anticiper sur le chapitre suivant).

Démonstration 1. Posons pour tout $n \in \mathbb{N}$,

$$\mathrm{S}_n := \left(\sum_{k=0}^{n} \frac{u^k}{k!} \right) \circ \left(\sum_{k=0}^{n} \frac{v^k}{k!} \right) - \left(\sum_{k=0}^{n} \frac{(u+v)^k}{k!} \right).$$

On a pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
|||\mathrm{S}_n||| &= \left|\left|\left| \sum_{i=0}^{n}\sum_{j=0}^{n} \frac{u^i \circ v^j}{i!j!} - \sum_{k=0}^{n} \sum_{i+j=k} \frac{\binom{k}{i} u^i \circ v^j}{k!} \right|\right|\right| \\
&= \left|\left|\left| \sum_{i=0}^{n}\sum_{j=0}^{n} \frac{u^i \circ v^j}{i!j!} - \sum_{\substack{i+j \leq n \\ i \geq 0\ j \geq 0}} \frac{u^i \circ v^j}{i!j!} \right|\right|\right| = \left|\left|\left| \sum_{\substack{i+j \geq n+1 \\ 0 \leq i,j \leq n}} \frac{u^i \circ v^j}{i!j!} \right|\right|\right| \\
&\leq \sum_{\substack{i+j \geq n+1 \\ 0 \leq i,j \leq n}} \frac{|||u|||^i \cdot |||v|||^j}{i!j!} \\
&= \left(\sum_{0}^{n} \frac{|||u|||^k}{k!} \right) \cdot \left(\sum_{0}^{n} \frac{|||v|||^k}{k!} \right) - \left(\sum_{0}^{n} \frac{(|||u||| + |||v|||)^k}{k!} \right) \\
&\xrightarrow[n \to +\infty]{} e^{|||u|||} \cdot e^{|||v|||} - e^{|||u||| + |||v|||} = 0.
\end{aligned}$$

Ainsi

$$\mathrm{S}_n \xrightarrow[n \to +\infty]{} 0.$$

Or, par continuité de la composition, on a aussi

$$\mathrm{S}_n \xrightarrow[n \to +\infty]{} \exp(u) \circ \exp(v) - \exp(u + v),$$

(cc) BY-NC-SA

d'où le résultat

Démonstration 2. Cette démonstration est facultative et a été ajoutée uniquement pour donner un avant-goût du chapitre suivant. Le lecteur peut la sauter s'il le souhaite.

Pour tout $t \in \mathbb{R}$, posons $\exp_t : z \longmapsto \sum_{k=0}^{\infty} \frac{t^k}{k!} z^k = e^{tz}$. On a pour $t, t' \in \mathbb{R}$ et $u \in \mathcal{L}(\mathrm{E})$,

$$\exp(tu) \circ \exp(t'u) = \exp_t(u) \circ \exp_{t'}(u) \underset{(*)}{=} (\exp_t \times \exp_{t'})(u) = (\exp_{t+t'})(u) = \exp((t+t')u).$$

L'égalité $(*)$ est vraie d'après l'exercice I.2. Ceci nous assure que $\exp(u) \in \mathrm{GL}(\mathrm{E})$ étant donné que le résultat ci-dessus donne $\exp(u) \circ \exp(-u) = \exp(0) = \mathrm{Id}_{\mathrm{E}}$.

Pour tout $u \in \mathcal{L}(\mathrm{E})$, notons

$$f_u : \begin{cases} \mathbb{R} & \longrightarrow \mathrm{GL}(\mathrm{E}) \\ t & \longmapsto \exp(tu) \end{cases}$$

qui est, par ce qui précède, un morphisme (continu) de groupes.

Remarquons maintenant que

$$\forall t \in \mathbb{R}^*, \quad \frac{f_u(t) - f_u(0)}{t} = u + t \underbrace{\left(\sum_{k=0}^{\infty} \frac{t^k}{(k+2)!} u^{k+2} \right)}_{g(t)}.$$

On sait que $g : \mathbb{R} \to \mathcal{L}(\mathrm{E})$ est continue et est donc en particulier bornée sur le compact $[-1, 1]$.
On déduit en particulier que f_u est dérivable en 0 et qu'en passant à la limite, étant donné que g est bornée sur un voisinage de 0, on a $f_u'(0) = u$. De plus, en utilisant l'égalité vue en début de cette démonstration, on a pour tout $t, t_0 \in \mathbb{R}$,

$$f_u(t + t_0) = f_u(t_0) \circ f_u(t) = f_u(t) \circ f_u(t_0),$$

et donc

$$\begin{aligned} \frac{f_u(t) - f_u(t_0)}{t - t_0} &= \frac{f_u(t - t_0) \circ f_u(t_0) - f_u(t_0)}{t - t_0} \\ &= \frac{f_u(t - t_0) - f_u(0)}{t - t_0} \circ f_u(t_0) \xrightarrow[t \to t_0]{} u \circ f_u(t_0), \end{aligned}$$

et alors f_u est dérivable sur $\mathbb{R}$ et pour tout $t \in \mathbb{R}$, $f_u'(t) = u \circ f_u(t)$.

Remarquons que ceci, en utilisant le théorème de Cauchy-Lipshitz, vu dans le cas de $\mathbb{R}$ ou $\mathbb{C}$, et sera vu d'une perspective plus générale dans le chapitre suivant, fournit une caractérisation différentielle de l'exponentielle, à savoir :

$$\exp(u) = f_u(1) \text{ où } \begin{cases} f_u \in \mathrm{C}^1(\mathbb{R}, \mathcal{L}(\mathrm{E})) \\ f_u'(t) = u \circ f_u(t) \\ f_u(0) = \mathrm{Id}_{\mathrm{E}}\,. \end{cases}$$

Montrons à présent la propriété voulue en utilisant ce qui précède. Soit $u, v \in \mathcal{L}(\mathrm{E})$ tel que $[u, v] = 0$. Il est clair que u, v, $\exp(tu)$ et $\exp(tv)$ commutent deux à deux (car u^k et v^ℓ commutent pour tout $k, \ell \in \mathbb{N}$).

Considérons $g(t) = f_u(t) \circ f_v(t)$ qui, par commutation, vérifie $g \in \mathrm{Hom}(\mathbb{R}, \mathrm{GL}(\mathrm{E}))$. Par bilinéarité de la composition, on obtient que $g \in \mathrm{C}^1(\mathbb{R}, \mathrm{GL}(\mathrm{E}))$ et que

$$g'(t) = f_u'(t) \circ f_v(t) + f_u(t) \circ f_v'(t) = u \circ f_u(t) \circ f_v(t) + f_u(t) \circ v \circ f_v(t) = (u+v) \circ f_u(t) \circ f_v(t) = (u+v) \circ g(t).$$

(cc) BY-NC-SA

En évaluant en 0 on obtient
$$g'(0) = f'_u(0) + f'_v(0) = u + v.$$
On en déduit donc que g vérifie les propriétés suivantes
$$\begin{cases} g \in \mathrm{C}^1(\mathbb{R}, \mathcal{L}(\mathrm{E})) \\ g'(t) = (u+v) \circ g(t) \\ g(0) = \mathrm{Id}_\mathrm{E}, \end{cases}$$
ce qui signifie d'après l'unicité fournie par le théorème de Cauchy-Lipschitz que $g = f_{u+v}$ et donc finalement
$$\exp(u) \circ \exp(v) = g(1) = f_{u+v}(1) = \exp(u+v).$$

Remarque culturelle. Nous citons dans cette remarque des éléments hors programmes et non nécessaires pour les concours.

→ En fait, en examinant la démonstration ci-dessus, on peut facilement déduire une démonstration du résultat suivant.
$$\begin{cases} f \in \mathrm{C}^0(\mathbb{R}, \mathcal{L}(\mathrm{E})) \cap \mathrm{Hom}(\mathbb{R}, \mathcal{L}(\mathrm{E})) \\ f \text{ dérivable en } 0 \end{cases} \implies \begin{cases} f \in \mathrm{C}^1(\mathbb{R}, \mathcal{L}(\mathrm{E})) \\ \forall t \in \mathbb{R},\ f'(t) = f'(0) \circ f(t) \\ f(0) = \mathrm{Id}_\mathrm{E}\,. \end{cases}$$
Les conditions de gauches sont donc suffisantes pour caractériser l'exponentielle (elle y sont même clairement équivalentes, étant donné que l'exponentiel est dans $\mathrm{Hom}(\mathbb{R}, \mathcal{L}(\mathrm{E})) \cap \mathrm{C}^1(\mathbb{R}, \mathcal{L}(\mathrm{E}))$). On en déduit qu'une autre version de la démonstration aurait pu être la suivante : g vérifie les deux conditions de gauche ainsi que les deux conditions avec $g'(0) = u+v$, ce qui nous permet directement de dire que $g = f_{u+v}$ et finalement $\exp(u) \circ \exp(v) = g(1) = \exp(u+v)$.

→ L'intérêt des deux approches (homomorphisme et différentielle) est qu'elles sont généralisables, ou plutôt sont les approches typiques, lorsqu'on travaille avec des groupes et algèbres de Lie abstraites. Pour le lecteur curieux, la manière moderne typique de définir l'exponentielle de $v \in \mathrm{Lie(G)}$ est de considérer le flot de l'unique champs de vecteur invariant à gauche (ou droite) associé à v et de l'évaluer en 1. C'est ce qui a été fait (de manière plus simplifiée) ci dessus.

Exercice I.5.

Soit $u \in \mathcal{L}(\mathrm{E})$. Montrer que
$$\left(\mathrm{Id}_\mathrm{E} + \frac{u}{k}\right)^k \xrightarrow[k \to +\infty]{} \exp(u).$$

Remarque. Bien évidemment, tout ce qui a été construit jusqu'à présent est valable lorsqu'on remplace $\mathcal{L}(\mathrm{E})$ par $\mathcal{M}_n(\mathbb{K})$ et la composition d'endomorphisme par le produit matriciel.

2. Logarithme

Cette partie est à but purement culturel, nous ne fournirons donc pas les démonstrations de résultats ci-dessous.

(cc) BY-NC-SA

Définition I.6.

Notons Unip(E) l'ensemble des endomorphismes unipotents de E *i.e.*

$$\mathrm{Unip(E)} = \{u \in \mathcal{L}(\mathrm{E}),\ u - \mathrm{Id_E} \text{ est nilpotent}\}.$$

On définit le logarithme dans $\mathcal{L}(\mathrm{E})$ de la manière suivante.

$$\log : \begin{cases} \mathrm{B}_{\|\!|.\|\!|}(\mathrm{Id_E}, 1) \cup \mathrm{Unip(E)} & \longrightarrow \mathcal{L}(\mathrm{E}) \\ \mathrm{Id_E} + u & \longmapsto \displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} u^k. \end{cases}$$

Proposition I.7.

Les propriétés suivantes sont vraies.

→ log est continue sur $\mathrm{B}_{\|\!|.\|\!|}(\mathrm{Id_E}, 1)$.

→ Pour $u \in \mathcal{L}(\mathrm{E})$, $\dfrac{d}{dt} \log(\mathrm{Id} + tu)\Big|_{t=0} = u$.

→ Pour $u \in \mathcal{L}(\mathrm{E})$ $\log(u) \in \mathbb{K}[u]$.

→ $\forall u \in \mathrm{B}(\mathrm{Id_E}, 1) \cup \mathrm{Unip(E)},\ \exp(\log(u)) = u$.

→ $\log(\exp(u)) = u$ pour u nilpotent ou assez proche de 0.

II Propriétés géométriques

Cette partie est hors-programme. Désormais $\mathbb{K} = \mathbb{C}$. On note $\|.\|$ la norme matricielle sur $\mathcal{M}_n(\mathbb{K})$. On note de plus pour $\mathrm{P} \in \mathrm{GL}_n(\mathbb{K})$

$$i_{\mathrm{P}} : \begin{cases} \mathcal{M}_n(\mathbb{K}) & \longrightarrow \mathcal{M}_n(\mathbb{K}) \\ \mathrm{A} & \longmapsto \mathrm{PAP}^{-1}. \end{cases}$$

Exercice II.1.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{K})$, montrer que

$$\exp(\mathrm{A}) = \sum_{k=0}^{\infty} \frac{\mathrm{A}^k}{k!} \in \mathbb{K}[\mathrm{A}].$$

Proposition II.2.

Pour tout $(\mathrm{A}, \mathrm{P}) \in \mathcal{M}_n(\mathbb{K}) \times \mathrm{GL}_n(\mathbb{K})$, on a

$$i_{\mathrm{P}} \circ \exp = \exp \circ i_{\mathrm{P}} \ \ i.e.\ \mathrm{P} \exp(\mathrm{A}) \mathrm{P}^{-1} = \exp(\mathrm{PAP}^{-1}).$$

Démonstration. On a pour $\mathrm{N} \in \mathbb{N}$,

$$\sum_{0}^{\mathrm{N}} \frac{(\mathrm{PAP}^{-1})^k}{k!} = \mathrm{P}\left(\sum_{0}^{\mathrm{N}} \frac{\mathrm{A}^k}{k!}\right) \mathrm{P}^{-1}.$$

(cc) BY-NC-SA

Par continuité de i_{P} et en passant à la limite, on obtient le résultat.

Lemme II.3.

Soit $\mathrm{A} \in \mathcal{T}_n^+(\mathbb{K})$, écrivons $\mathrm{A} = \mathrm{D} + \mathrm{N}$ où $\mathrm{D} = \mathrm{diag}(\lambda_1, \dots, \lambda_n) \in \mathrm{D}_n(\mathbb{K})$ est diagonale et $\mathrm{N} \in \mathcal{T}_n^{++}(\mathbb{K})$ est strictement triangulaire supérieure. Les propriétés suivantes sont vérifiées.

1. $\exp(\mathrm{D}) = \mathrm{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.
2. $\exp(\mathrm{N}) = \mathrm{I}_n + \mathrm{N} + \dfrac{\mathrm{N}^2}{2!} + \cdots + \dfrac{\mathrm{N}^{n-1}}{(n-1)!}$, formule vraie pour tout nilpotent.
3. L'égalité suivante est vraie.

$$\mathrm{A} = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ 0 & & \lambda_n \end{pmatrix} \text{ et } \exp(\mathrm{A}) = \begin{pmatrix} e^{\lambda_1} & * & * \\ & \ddots & * \\ 0 & & e^{\lambda_n} \end{pmatrix}.$$

Démonstration.

1. Clair en remarquant que pour tout $k \in \mathbb{N}$, $\mathrm{D}^k = \mathrm{diag}(\lambda_1^k, \dots, \lambda_n^k)$. Le lecteur ayant un doute est encouragé à faire la démonstration lui-même.
2. N est nilpotente et son degré de nilpotence est au plus n, donc tous les termes de degré strictement supérieur à n dans la somme $\displaystyle\sum_{k=0}^{\infty} \frac{\mathrm{N}^k}{k!}$ sont nuls.
3. Pour montrer ce résultat, il suffit de remarquer que pour tout $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \frac{\mathrm{A}^k}{k!} = \begin{pmatrix} \displaystyle\sum_{k=0}^{n} \frac{\lambda_1^k}{k!} & * & * \\ & \ddots & * \\ 0 & & \displaystyle\sum_{k=0}^{n} \frac{\lambda_n^k}{k!} \end{pmatrix}.$$

Ceci qui donne le résultat en passant à la limite.

Remarque. Une manière pratique de calculer l'exponentielle d'une matrice A est de considérer sa décomposition de Jordan-Dunford $\mathrm{A} = \mathrm{D} + \mathrm{N}$ où D est diagonalisable, N nilpotente et $[\mathrm{D}, \mathrm{N}] = 0$. Dans ce cas, l'exponentielle de D et N est facile à calculer et en utilisant le fait que D et N commutent, on peut voir que $\exp(\mathrm{A}) = \exp(\mathrm{D}) \exp(\mathrm{N})$.

Proposition II.4.

Soit $\mathrm{A} \in \mathrm{M}_n(\mathbb{C})$.

1. On note $\mathrm{Spec}(\mathrm{A}) = [\lambda_1, \dots, \lambda_n]$ (on rappelle que la notation [.] veut dire un $n-$uplet avec possible répétition et sans prise en compte de l'ordre). On a
$$\mathrm{Spec}(\exp(\mathrm{A})) = [e^{\lambda_1}, \dots, e^{\lambda_n}].$$
2. $\det(\exp(\mathrm{A})) = e^{\mathrm{Tr}(\mathrm{A})}$ (on retrouve que $\det \mathrm{A} \neq 0$, *i.e.* $\exp(\mathrm{A}) \in \mathrm{GL}_n(\mathbb{C})$).

Démonstration. Trigonalisons la matrice A, on pose

$$\mathrm{A} = \mathrm{PBP}^{-1} \text{ avec } \mathrm{P} \in \mathrm{GL}_n(\mathbb{C}) \text{ et } \mathrm{B} = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ 0 & & \lambda_n \end{pmatrix}.$$

(cc) BY-NC-SA

On rappelle que $\exp(A) = P\exp(B)P^{-1}$.

1. Le lemme II.3 nous donne que

$$\operatorname{Spec}(\exp(A)) = \operatorname{Spec}(\exp(B)) = [e^{\lambda_1}, \dots, e^{\lambda_n}].$$

2. Toujours par le lemme II.3,

$$\det(\exp(A)) = \det(\exp(B)) = e^{\lambda_1} \dots e^{\lambda_n} = e^{\lambda_1 + \dots + \lambda_n} = e^{\operatorname{Tr}(A)}.$$

Proposition II.5.

Soit $A \in \mathcal{M}_n(\mathbb{C})$. Les propositions suivantes sont vraies.

1. Si A est diagonalisable, alors $\exp(A)$ l'est aussi.
2. Si A est nilpotente, alors $\exp(A) = I_n + A + \dots + \dfrac{A^{n-1}}{(n-1)!}$ est unipotente, *i.e.* s'écrit de la forme $I + N$ où N est une matrice nilpotente.
3. Si $A \in \mathcal{M}_n(\mathbb{R})$, alors $\det(\exp(A)) > 0$.

Démonstration.

1. Écrivons $A = PDP^{-1}$ où D est diagonale. Ceci donne que $\exp(A) = P\exp(D)P^{-1}$. Le lemme permet de conclure vu que $\exp(D)$ est également diagonale.
2. Il suffit de voir que $A + \dots + \dfrac{A^{n-1}}{(n-1)!}$ est nilpotente. Le lecteur ayant un doute peut le vérifier à la main.
3. On rappelle que $\det(\exp(A)) = e^{\operatorname{Tr}(A)} \in \mathbb{R}^{+*}$.

Exercice II.6.

1. Trouver les matrices $A \in GL_2(\mathbb{R})$ tel que $\det(A) > 0$ et $A \notin \exp(\mathcal{M}_2(\mathbb{R}))$.
2. Soit $A \in \mathcal{M}_n(\mathbb{C})$. Montrer que A est diagonalisable si et seulement si $\exp(A)$ l'est.
3. Montrer que $\exp(\mathcal{M}_n(\mathbb{C})) = GL_n(\mathbb{C})$.

(cc) BY-NC-SA

Correction de l'exercice I.2. :

1. Prenons $\varepsilon > 0$. Sur $\mathbb{C}$, on peut diagonaliser u à $\varepsilon-$près (voir proposition VIII.4 du chapitre 28 sur la réduction d'endomorphismes) *i.e.* il existe une base $\beta = (e_{1,\varepsilon}, \dots, e_{n,\varepsilon})$ telle que
$$[u]_\beta = \begin{pmatrix} \lambda_1 & & b_{i,j} \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$
où pour tout $1 \leq i \leq j \leq n$, $|b_{i,j}| \leq \varepsilon$. Quitte à diminuer ε, on peut supposer que
$$n\varepsilon + \max(|\lambda_1|, \dots, |\lambda_n|) < \mathrm{R}.$$
On considère la norme suivante sur E.
$$\mathrm{N} : \begin{cases} \mathrm{E} & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto \|[x]_\beta\|_1 \,. \end{cases}$$
Considérons $|||.|||_{\mathrm{N}}$, la norme d'opérateur associée à N. On a pour tout $x \in \mathrm{E}$, en posant $[x]_\beta = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}^{\mathrm{T}}$,
$$\begin{aligned} \mathrm{N}(u(x)) = \|[u]_\beta [x]_\beta\|_1 &= \sum_{k=1}^{n} \left| x_k \lambda_k + \sum_{j=k+1}^{n} b_{k,j} x_j \right| \\ &\leq \sum_{k=1}^{n} \underbrace{|\lambda_k|}_{\leq \rho(u)} |x_k| + \sum_{k=1}^{n} \sum_{j=k+1}^{n} |b_{k,j}|\,|x_j| \\ &\leq \rho(u)\mathrm{N}(x) + \varepsilon \sum_{k=1}^{n} \sum_{j=k+1}^{n} |x_j| \leq (\rho(u) + n\varepsilon)\mathrm{N}(x). \end{aligned}$$
On en déduit donc que $|||u|||_{\mathrm{N}} \leq \rho(u) + n\varepsilon < \mathrm{R}$, d'où le résultat. On notera au passage que cette démonstration nous permet de déduire le résultat suivant.
Soit $\mathrm{A} \in \mathrm{M}_n(\mathbb{C})$. Pour toute norme N de $\mathbb{C}^n$ on notera $||| \cdot |||_{\mathrm{N}}$ la norme d'opérateur associée. Alors
$$\rho(\mathrm{A}) = \inf \underbrace{\left\{ |||\mathrm{A}|||_{\mathrm{N}},\ \mathrm{N} \text{ norme de } \mathbb{C}^n \right\}}_{\mathrm{H}}.$$
En effet, on a pour tout $\lambda \in \mathrm{Spec}(\mathrm{A})$, $|\lambda| \leq |||\mathrm{A}|||_{\mathrm{N}}$. Pour voir ce point, il suffit de considérer $\mathrm{X} \in \mathrm{Ker}(\mathrm{A} - \lambda \mathrm{I}) \setminus \{0\}$ et de remarquer que $|||\mathrm{A}|||_{\mathrm{N}} \mathrm{N}(\mathrm{X}) \geq \mathrm{N}(\mathrm{AX}) = |\lambda| \mathrm{N}(\mathrm{X})$. Ceci entraîne que $\rho(\mathrm{A}) \leq |||\mathrm{A}|||_{\mathrm{N}}$, *i.e.* $\rho(\mathrm{A})$ est un minorant de H. La démonstration ci-dessus montre que pour tout $\mathrm{R} > \rho(\mathrm{A})$, il existe N tel que $|||\mathrm{A}|||_{\mathrm{N}} \leq \mathrm{R}$, ce qui donne le résultat voulu.

2. Démonstration exactement similaire à celle du théorème I.4 (qui consiste à manipuler des séries entières). en majorant par la norme et utilisant le produit de convolution.

Correction de l'exercice I.5. :

Méthode 1 (calculatoire) : Posons $r = |||u|||$. Soit $\varepsilon > 0$. Considérons N assez grand fixé tel que
$$\sum_{i=\mathrm{N}}^{\infty} \frac{r^i}{i!} < \frac{\varepsilon}{4}.$$
On a pour tout $k \geq \mathrm{N} + 1$,

(cc) BY-NC-SA

$$\left\|\left(\mathrm{Id}_{\mathrm{E}}+\frac{u}{k}\right)^{k}-\exp(u)\right\|=\left\|\sum_{i=0}^{k}\binom{k}{i}\frac{u^{i}}{k^{i}}-\sum_{i=0}^{\infty}\frac{u^{i}}{i!}\right\|=\left\|\sum_{i=0}^{k}\frac{\frac{k!}{(k-i)!k^{i}}-1}{i!}u^{i}-\sum_{j=k+1}^{\infty}\frac{u^{j}}{j!}\right\|$$

$$\leq\sum_{i=0}^{\mathrm{N}}\frac{\left|\frac{k!}{(k-i)!k^{i}}-1\right|}{i!}r^{i}+\sum_{i=\mathrm{N}+1}^{k}\frac{\overbrace{\frac{k!}{(k-i)!k^{i}}}^{\leq 1}+1}{i!}r^{i}+\sum_{j=k+1}^{\infty}\frac{r^{j}}{j!}\leq\underbrace{\sum_{i=0}^{\mathrm{N}}\frac{\overbrace{\left|\frac{k!}{(k-i)!k^{i}}-1\right|}^{\xrightarrow[k\to+\infty]{}0}}{i!}r^{i}}_{\xrightarrow[k\to+\infty]{}0}+\underbrace{2\sum_{j=\mathrm{N}+1}^{\infty}\frac{r^{j}}{j!}}_{<\frac{\varepsilon}{2}}.$$

En particulier, pour k assez grand on a

$$\left\|\left(\mathrm{Id}_{\mathrm{E}}+\frac{u}{k}\right)^{k}-\exp(u)\right\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

et donc

$$\left(\mathrm{Id}_{\mathrm{E}}+\frac{u}{k}\right)^{k}\xrightarrow[k\to+\infty]{}\exp(u).$$

Méthode 2 (réduction d'endomorphisme) : Attention, cette démonstration n'est valable que dans le cas $\mathbb{K}=\mathbb{C}$ car on y utilise la décomposition de Jordan-Dunford qui n'est valable que dans $\mathbb{C}$. Nous utilisons ici des propriétés vues dans le chapitre 28 sur la réduction d'endomorphisme.
Soit β une base de E où on peut écrire $u=d+\nu$ (Jordan-Dunford, proposition VI.9 chapitre 28) avec

$$[d]_{\beta}=\operatorname{diag}(\lambda_{1},\ldots,\lambda_{n}),\ \nu \text{ nilpotent et } d\circ\nu=\nu\circ d.$$

On a alors

$$\left[\left(\mathrm{Id}_{\mathrm{E}}+\frac{u}{k}\right)^{k}\right]_{\beta}=\left[\left(\mathrm{Id}_{\mathrm{E}}+\frac{d}{k}+\frac{\nu}{k}\right)^{k}\right]_{\beta}\underset{(*)}{=}\left[\sum_{i=0}^{n-1}\binom{k}{i}\left(\mathrm{Id}_{\mathrm{E}}+\frac{d}{k}\right)^{k-i}\frac{\nu^{i}}{k^{i}}\right]_{\beta}$$

$$=\sum_{i=0}^{n-1}\underbrace{\frac{k(k-1)\ldots(k-i+1)}{k^{i}}}_{\xrightarrow[k\to+\infty]{}1}\underbrace{\operatorname{diag}\left(\underbrace{\left(1+\frac{\lambda_{1}}{k}\right)^{k-i}}_{\xrightarrow[k\to+\infty]{}e^{\lambda_{1}}},\ldots,\underbrace{\left(1+\frac{\lambda_{n}}{k}\right)^{k-i}}_{\xrightarrow[k\to+\infty]{}e^{\lambda_{n}}}\right)}_{\xrightarrow[k\to+\infty]{}\operatorname{diag}(e^{\lambda_{1}},\ldots,e^{\lambda_{n}})=[\exp(d)]_{\beta}}\frac{[\nu]_{\beta}^{i}}{i!}.$$

Remarquons que dans le terme à droite de l'égalité $(*)$, l'indice de sommation s'arrête à n et non pas k car ν est nilpotent et pour tout $i\geq n$, $\nu^{n}=0$. Ainsi par continuité de la composition/multiplication ($n=\dim\mathrm{E}$ est fixé), on obtient

$$\left[\left(\mathrm{Id}_{\mathrm{E}}+\frac{u}{k}\right)^{k}\right]_{\beta}\xrightarrow[k\to+\infty]{}\left[\exp(d)\circ\left(\sum_{0}^{n-1}\frac{\nu^{i}}{i!}\right)\right]_{\beta}=[exp(d)\circ\exp(\nu)]_{\beta}\underset{\substack{\uparrow\\ d\circ\nu=\nu\circ d}}{=}[\exp(d+n)]_{\beta}=[\exp(u)]_{\beta},$$

d'où le résultat voulu.

Correction de l'exercice II.1. :

On rappelle que $\mathbb{K}[\mathrm{A}]$ est un sous-espace vectoriel de l'espace vectoriel $\mathcal{M}_{n}(\mathbb{K})$, qui est de dimension finie.

(cc) BY-NC-SA

$\mathbb{K}[A]$ est donc fermé (d'après la proposition III.6 du chapitre 11.7). exp(A) est limite d'une suite à valeurs dans $\mathbb{K}[A]$. En effet, on a

$$\underbrace{\sum_{k=0}^{n} \frac{A^k}{k!}}_{\in \mathbb{K}[A]} \xrightarrow[n\to+\infty]{} \exp(A).$$

$\mathbb{K}[A]$ étant fermé, on a bien $\exp(A) \in \mathbb{K}[A]$.

Correction de l'exercice II.6. :

1. Remarquons tout d'abord que pour tout $A \in \mathcal{M}_n(\mathbb{R})$ et $P \in GL_n(\mathbb{R})$,

$$Pe^{A}P^{-1} = e^{PAP^{-1}},$$

et donc

$$\forall B \in GL_2(\mathbb{R}),\ B \in \exp(\mathcal{M}_2(\mathbb{R})) \iff \exists A \in \mathcal{M}_2(\mathbb{R}),\ \langle B\rangle_{\mathbb{R}} = \exp(\langle A\rangle_{\mathbb{R}}).$$

Il suffit donc de calculer l'image de chaque classe d'équivalence (de conjugaison) par exp.
Rappelons également ce résultat classique

Lemme II.7.

Deux matrices $A, B \in \mathcal{M}_n(\mathbb{R})$ réelles sont équivalences dans $\mathbb{C}$ si et seulement si elles le sont dans $\mathbb{R}$, *i.e.*

$$\exists P \in GL_n(\mathbb{R}),\ PAP^{-1} = B \iff \exists Q \in GL_n(\mathbb{C}),\ QAQ^{-1} = B.$$

Ceci signifie que toutes les équivalences qu'on manipulera ci-dessous qui sont valables dans $\mathbb{C}$ le sont également dans $\mathbb{R}$ dès lors que les matrices sont réelles.
Remarquons qu'il est facile de voir que l'exponentielle de toute matrice est de déterminant strictement positif et qu'il suffit de traiter le cas des matrices de déterminant égal à 1. En effet, pour tout $A \in GL_2(\mathbb{R})$ et $B \in \mathcal{M}_2(\mathbb{R})$ avec $\det A > 0$, on a l'équivalence

$$A = \exp(B) \iff \underbrace{\frac{A}{\sqrt{\det(A)}}}_{\text{déterminant égal à 1}} = \exp\left(B - \frac{1}{2}\log(\det(A))I_2\right).$$

En utilisant cette remarque, on peut voir que pour tout $A \in \mathcal{M}_2(\mathbb{R})$ de déterminant égal à 1, s'il existe $B \in GL_2(\mathbb{R})$ telle que $A = \exp(B)$, alors

$$1 = \det A = \det\exp(B) = \exp\operatorname{Tr}(B) \text{ et donc } \operatorname{Tr}(B) = 0.$$

Sachant que $\det(\exp(B)) = \exp(\operatorname{Tr}(B))$, nous nous intéresserons donc à l'exponentielle de matrices de traces nulles. Trouvons donc toutes les classes d'équivalence (de conjugaison) d'une matrice de la forme $M = \exp(A)$ avec $A \in \mathcal{M}_2(\mathbb{R})$.

→ **Cas 1 :** A diagonalisable dans $\mathbb{R}$.
A étant de trace nulle, on a nécessairement $A \simeq \operatorname{diag}(\lambda, -\lambda)$ avec $\lambda \in \mathbb{R}$. Dans ce cas,

$$\exp(A) \simeq \operatorname{diag}(e^{\lambda}, e^{-\lambda}) \text{ et } \lambda \in \mathbb{R}.$$

Ainsi, ce cas correspond aux classes de conjugaison de la forme

(cc) BY-NC-SA

$$\left\{ \left\langle \begin{pmatrix} t & 0 \\ 0 & \dfrac{1}{t} \end{pmatrix} \right\rangle_{\mathbb{R}}, \ t > 0 \right\}.$$

→ **Cas 2 :** A diagonalisable dans $\mathbb{C}$ mais pas $\mathbb{R}$ (ou est nulle).

Encore une fois, on a $\mathrm{A} \simeq \operatorname{diag}(\lambda, -\lambda)$ et les valeurs propres de cette matrice sont conjuguées, car elle est réelle, ce qui implique que son polynôme caractéristique l'est aussi, ce qui fait que ses racines sont conjuguées. On a alors

$$\overline{\lambda} = -\lambda \iff \lambda = \pm \mu i \text{ avec } \mu \in \mathbb{R}.$$

On en déduit donc que $\mathrm{A} \simeq \operatorname{diag}(\mu i, -\mu i)$ avec $\mu \in \mathbb{R}$. Ainsi,

$$\exp(\mathrm{A}) \simeq \begin{pmatrix} \cos(\mu) + i\sin(\mu) & 0 \\ 0 & \cos(\mu) - i\sin(\mu) \end{pmatrix} \underset{(*)}{\simeq} \begin{pmatrix} \cos(\mu) & -\sin(\mu) \\ \sin(\mu) & \cos(\mu) \end{pmatrix} = \mathrm{R}_{\mu}.$$

L'équivalence $(*)$ est vraie, car les deux matrices sont diagonalisables et ont les mêmes valeurs propres, comptés avec multiplicité (un autre argument aurait pu être le fait qu'ils sont diagonalisables, de taille 2, et ont la même trace et le même déterminant). Ce cas correspond donc aux classes de conjugaison de la forme

$$\left\{ \left\langle \begin{pmatrix} \cos(\mu) & -\sin(\mu) \\ \sin(\mu) & \cos(\mu) \end{pmatrix} \right\rangle_{\mathbb{R}}, \ \mu \in \mathbb{R} \right\}.$$

→ **Cas 3 :** A est non diagonalisable dans $\mathbb{C}$.

Dans ce cas, les valeurs propres de A sont égales (car si elles étaient distinctes A serait diagonalisable). On a alors, étant donné que $\operatorname{Tr}(\mathrm{A}) = 0$, les valeurs propres de A sont nulles et $\mathrm{A} \simeq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (voir exercice VI.6 du chapitre 28 sur la réduction d'endomorphisme). Par conséquent, on en déduit que ce cas correspond à la classe d'équivalence $\langle \exp(\mathrm{A}) \rangle_{\mathbb{R}} = \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\rangle_{\mathbb{R}}$.

Revenons au cas général. Soit A une matrice de $\mathrm{GL}_2(\mathbb{R})$. On a

$$\begin{aligned} \exists \mathrm{B} \in \mathcal{M}_2(\mathbb{R}),\ \mathrm{A} = \exp(\mathrm{B}) &\iff \exists \mathrm{B} \in \mathcal{M}_2(\mathbb{R}),\ \frac{\mathrm{A}}{\sqrt{\det(\mathrm{A})}} = \exp\left(\mathrm{B} - \frac{1}{2}\log(\det(\mathrm{A}))\mathrm{I}_2\right) \\ &\iff \frac{\mathrm{A}}{\sqrt{\det(\mathrm{A})}} \in \exp(\mathcal{M}_2(\mathbb{R})). \end{aligned}$$

$\sqrt{\det(\mathrm{A})}$ peut prendre n'importe quelle valeur dans $\mathbb{R}_+^*$, on en déduit donc que $\mathrm{A} \in \exp(\mathcal{M}_2(\mathbb{R}))$ si, et seulement si, A appartient aux classes d'équivalences trouvées multipliées par un réel strictement positif, c'est-à-dire une classe appartenant à l'ensemble suivant

$$\left\{ \begin{pmatrix} a & a \\ 0 & a \end{pmatrix}, \ a > 0 \right\} \cup \left\{ \begin{pmatrix} a\cos(\mu) & -a\sin(\mu) \\ a\sin(\mu) & a\cos(\mu) \end{pmatrix}, \ \mu \in \mathbb{R} \text{ et } a > 0 \right\} \cup \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \ a, b > 0 \right\}.$$

Étant donné que pour tout $a > 0$, on a

$$\begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} a & a \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}^{-1} = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}.$$

L'ensemble des classes d'équivalences possibles pour $\mathrm{A} \in \exp(\mathcal{M}_2(\mathbb{C}))$ sont les classes de l'ensemble

(cc) BY-NC-SA

suivant.

$$\left\{\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix},\ a > 0\right\} \cup \left\{\begin{pmatrix} a\cos(\mu) & -a\sin(\mu) \\ a\sin(\mu) & a\cos(\mu) \end{pmatrix},\ \mu \in \mathbb{R} \text{ et } a > 0\right\} \cup \left\{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},\ a, b > 0\right\}.$$

On souhaite maintenant trouver les classes d'équivalences des matrices $\mathrm{A} \notin \exp(\mathcal{M}_2(\mathbb{R}))$. En s'inspirant de l'exercice VI.6 du chapitre 28 sur la réduction d'endomorphisme, on peut voir que les classes d'équivalence de $\{\mathrm{A} \in \mathrm{GL}_2(\mathbb{R}),\ \det \mathrm{A} > 0\}$ sont les suivantes.

$$\left\{\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix},\ a \in \mathbb{R}^*\right\} \cup \left\{\begin{pmatrix} a\cos(\mu) & -a\sin(\mu) \\ a\sin(\mu) & a\cos(\mu) \end{pmatrix},\ \mu \in \mathbb{R} \text{ et } a > 0\right\} \cup \left\{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},\ a, b \in \mathbb{R}^* \text{ et } ab > 0\right\},$$

donc les classes restantes n'appartenant pas à $\exp(\mathcal{M}_2(\mathbb{R}))$ sont les suivantes

$$\left\{\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix},\ a < 0\right\} \cup \left\{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix},\ a < 0,\ b < 0 \text{ et } a \neq b\right\}.$$

2. Le sens direct à déjà été fait, faisons donc la réciproque. Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$ telle que $\exp(\mathrm{A})$ est diagonalisable. Écrivons la décomposition de Dunford de A.

$$\mathrm{A} = \mathrm{D} + \mathrm{N},\ \mathrm{D} \text{ diagonalisable},\ \mathrm{N} \text{ nilpotente et } \mathrm{DN} = \mathrm{ND}.$$

Notons d l'indice de nilpotence de N. On a alors

$$\exp(\mathrm{A}) = \exp(\mathrm{D})\left(\sum_{k=0}^{d-1}\frac{\mathrm{N}^k}{k!}\right) = \exp(\mathrm{D}) + \exp(\mathrm{D})\underbrace{\left(\sum_{k=1}^{d-1}\frac{\mathrm{N}^k}{k!}\right)}_{\mathrm{N}'}.$$

Le fait que $\mathrm{DN} = \mathrm{ND}$ nous donne $\mathrm{DN}' = \mathrm{N}'\mathrm{D}$. De plus, par continuité de la multiplication matricielle, on a également, $\exp(\mathrm{D})\mathrm{N}' = \mathrm{N}'\exp(\mathrm{D})$ car pour tout $r \in \mathbb{N}$,

$$\left(\sum_{k=0}^{r}\frac{\mathrm{D}^k}{k!}\right)\mathrm{N}' = \mathrm{N}'\left(\sum_{k=0}^{r}\frac{\mathrm{D}^k}{k!}\right),$$

ce qui nous donne la relation voulue en faisant tendre r vers l'infini. Ainsi, $\exp(\mathrm{D})\mathrm{N}'$ est nilpotente (car N' est nilpotente et commute avec $\exp(\mathrm{D})$) et donc l'unicité de la décomposition de Dunford de $\exp(\mathrm{A})$ fournit que

$$\exp(\mathrm{D})\mathrm{N}' = 0 \ \textit{i.e.}\ \mathrm{N}' = 0$$

car $\exp(\mathrm{A})$ est diagonalisable. D'après la proposition VII.5 du chapitre 28 de réduction d'endomorphismes, $(\mathrm{I}_n, \mathrm{N}, \ldots, \mathrm{N}^{d-1})$ est une base de $\mathbb{K}[\mathrm{N}]$ et donc, en particulier, une famille libre. Ainsi, $\mathrm{N}' = 0$ implique que $d = 1$ *i.e.* $\mathrm{N} = 0$ et A est donc diagonalisable car égale à D, qui est diagonalisable.

Remarque. En examinant la démonstration ci-dessus, on peut voir que le résultat suivant est vrai : pour tout $\mathrm{A} \in \mathrm{M}_n(\mathbb{C})$, on a l'équivalence

$$\exp(\mathrm{A}) = \mathrm{I}_n \iff \mathrm{A} \text{ est diagonalisable et } \mathrm{Spec}(\mathrm{A}) \subset 2\pi i\mathbb{Z}.$$

3. La première chose à laquelle on pense ici est la décomposition de Dunford. Soit $\mathrm{A} \in \mathrm{GL}_d(\mathbb{C})$. On veut trouver une matrice Ω telle que $\mathrm{A} = \exp(\Omega)$. En posant $\mathrm{A} = \mathrm{D} + \mathrm{N}$ avec D diagonalisable et N nilpotente avec $\mathrm{DN} = \mathrm{ND}$, on peut voir que

$$\mathrm{A} = \mathrm{D}\,\underbrace{(\mathrm{I} + \underbrace{\mathrm{D}^{-1}\mathrm{N}}_{\text{nilpotente}})}_{\mathrm{H}}.$$

BY-NC-SA

Il semble donc intuitif de chercher à montrer qu'il existe $\Delta, \mathrm{V} \in \mathcal{M}_d(\mathbb{C})$ tels que $\mathrm{D} = e^{\Delta}$ et $\mathrm{H} = e^{\mathrm{V}}$ et de montrer que Δ et V commutent. On va commencer par montrer ce résultat pour H : on montre que pour toute matrice de la forme $\mathrm{I}+\mathrm{N}'$ avec N' nilpotente, il existe $\mathrm{M} \in \mathcal{M}_d(\mathbb{C})$ telle que $\mathrm{I}+\mathrm{N}' = \exp(\mathrm{M})$. Soit donc $\mathrm{B} = \mathrm{I}+\mathrm{N}'$ avec N' nilpotente.

On veut s'inspirer du développement en série entière de $x \longmapsto \ln(1+x)$. On sait que le développement en série entière de cette fonction s'écrit

$$\ln(1+x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k}.$$

Notre intuition est donc d'utiliser un équivalent matriciel de ce développement en série entière, et donc on souhaite montrer que

$$\mathrm{M} = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{\mathrm{N}'^k}{k} \underset{\substack{\uparrow \\ \mathrm{N}'^d=0}}{=} \sum_{k=1}^{d-1}(-1)^{k+1}\frac{\mathrm{N}'^k}{k}$$

est un bon candidat pour la matrice M. On a

$$e^y - 1 = \underbrace{y + \frac{y^2}{2!} + \cdots + \frac{y^{d-1}}{(d-1)!}}_{\mathrm{Q}(y)} + \mathrm{O}(y^d),$$

$$\ln(1+x) = \underbrace{x - \frac{x^2}{2} + \cdots + (-1)^d\frac{x^{d-1}}{d-1}}_{\mathrm{P}(x)} + \mathrm{O}(x^d),$$

et donc

$$x = e^{\ln(x+1)} - 1 = \mathrm{Q}(\mathrm{P}(x)) + \mathrm{O}((\ln(1+x))^d) = \mathrm{Q}(\mathrm{P}(x)) + \mathrm{O}(x^d).$$

On en déduit donc que

$$\mathrm{Q}(\mathrm{P}(x)) - x = \underbrace{\mathrm{O}(x^d).}_{\text{polynôme noté R}}$$

On voit clairement que R ne contient que des termes de degré supérieur ou égal à d. On en déduit donc, en utilisant le fait que $\mathrm{M}^d = 0$, que

$$e^{\mathrm{M}} - \mathrm{I} = \mathrm{Q}(\mathrm{M}) = \mathrm{Q}(\mathrm{P}(\mathrm{N}')) = \underbrace{\mathrm{Q}(\mathrm{P}(\mathrm{N}')) - \mathrm{N}'}_{=0} + \mathrm{N}' = \mathrm{N}'.$$

On rappelle que $\mathrm{R}(\mathrm{N}') = \mathrm{Q}(\mathrm{P}(\mathrm{N}')) - \mathrm{N}'$ est une combinaison de puissances de N' de degré supérieur à d et est donc nulle. On en déduit donc que

$$\mathrm{I} + \mathrm{N}' = \exp\left(\sum_{k=1}^{d-1}(-1)^{k+1}\frac{\mathrm{N}'^k}{k}\right).$$

En appliquant ce résultat pour $\mathrm{N}' = \mathrm{D}^{-1}\mathrm{N}$, on peut écrire

$$\mathrm{A} = \mathrm{D}(\mathrm{I} + \mathrm{D}^{-1}\mathrm{N}) = \mathrm{D}\exp\underbrace{\left(\sum_{k=1}^{d-1}(-1)^{k+1}\frac{(\mathrm{D}^{-1}\mathrm{N})^k}{k}\right)}_{\mathrm{M}}.$$

(cc) BY-NC-SA

De plus, D étant diagonalisable, on en déduit qu'il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ et $\mathrm{U} \in \mathrm{GL}_d(\mathbb{C})$ tels que

$$\mathrm{D} = \mathrm{U}\mathrm{diag}(\lambda_1, \ldots, \lambda_d)\mathrm{U}^{-1}.$$

Soit $\mu_1, \ldots, \mu_d$ tels que pour tout $i \in [\![1; d]\!]$, $e^{\mu_i} = \lambda_i$ et vérifiant $\forall i, j \in [\![1; n]\!],\ \lambda_i = \lambda_j \Longrightarrow \mu_i = \mu_j$. On veut choisir Δ de manière à garder la commutation avec M. Soit F un polynôme complexe tel que pour tout $i \in [\![1; n]\!],\ \mathrm{F}(\lambda_i) = \mu_i$ (existe, pour le voir, on peut utiliser les polynômes interpolateurs de Lagrange). On a alors, posant $\Delta = \mathrm{U}\mathrm{diag}(\mu_1, \ldots, \mu_n)\mathrm{U}^{-1}$,

$$\mathrm{F}(\mathrm{D}) = \mathrm{U}\mathrm{diag}(\mathrm{F}(\lambda_1), \ldots, \mathrm{F}(\lambda_n))\mathrm{U}^{-1} = \mathrm{U}\mathrm{diag}(\mu_1, \ldots, \mu_n)\mathrm{U}^{-1} = \Delta.$$

De plus, il est clair que $e^{\Delta} = \mathrm{D}$. Δ est polynomiale en D et D commute avec M donc Δ aussi, et alors en utilisant la proposition I.4, on en déduit que

$$\mathrm{A} = \mathrm{D}\exp(\mathrm{M}) = \exp(\Delta)\exp(\mathrm{M}) = \exp(\Delta + \mathrm{M}),$$

d'où le résultat.

Remarque. En examinant la démonstration ci-dessus, on peut en déduire facilement le résultat suivant qui est plus puissant : Pour tout $\mathrm{A} \in \mathrm{GL}_n(\mathbb{C})$, il existe $\mathrm{B} \in \mathbb{C}[\mathrm{A}]$ tel que $\mathrm{A} = \exp(\mathrm{B})$.

En effet, en reprenant les notations de la démonstration, on peut montrer que $\Delta + \mathrm{M}$ est un polynôme en A. On écrit encore une fois $\mathrm{A} = \mathrm{D} + \mathrm{N}$ la décomposition de Dunford de A. D'après la proposition VI.9 du chapitre 28 sur la réduction d'endomorphismes, D et N sont des polynômes en A. En particulier, en utilisant la même astuce avec les polynômes interpolateurs de Lagrange, on peut voir que D^{-1} est un polynôme en D et donc un polynôme en A. Par conséquent, on voit que $\mathrm{D}^{-1}\mathrm{N}$ est également un polynôme en A. On en déduit alors finalement que M est un polynôme en A. On a vu que Δ aussi et alors $\mathrm{M} + \Delta$ est également un polynôme en A.

(cc) BY-NC-SA

Chapitre 34

Espaces préhilbertiens réels

Dans tout ce chapitre, E désigne un $\mathbb{R}$-espace vectoriel non réduit à $\{0\}$.

I Géométrie d'un espace préhilbertien

Définition I.1.

Un produit scalaire sur E est une application $\langle\cdot\,,\,\cdot\rangle : \mathrm{E}^2 \longrightarrow \mathbb{R}$ bilinéaire qui vérifie les propriétés suivantes.

→ Elle est symétrique, *i.e.* pour tout $(x,y) \in \mathrm{E}^2$, $\langle x,\ y\rangle = \langle y,\ x\rangle$.

→ Elle est positive, *i.e.* pour tout $x \in \mathrm{E}$, $\langle x,\ x\rangle \geq 0$.

→ Elle est définie positive, *i.e.* pour tout $x \in \mathrm{E} \setminus \{0\}$, $\langle x,\ x\rangle \neq 0$.

Exemples.

→ Pour tout $n \in \mathbb{N}^*$, le produit scalaire canonique sur $\mathbb{R}^n$ est défini par :

$$\forall x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n, \quad \langle x,\ y\rangle = \sum_{i=1}^{n} x_i y_i.$$

→ Pour tout entier $n \geq 2$, le produit scalaire canonique sur $\mathcal{M}_n(\mathbb{R})$ est défini par :

$$\forall \mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{R}), \quad \langle \mathrm{A},\ \mathrm{B}\rangle = \mathrm{Tr}(\mathrm{A}^\top \mathrm{B}).$$

De plus, si $\mathrm{O} \in \mathcal{O}_n(\mathbb{C})$, et A et $\mathrm{B} \in \mathcal{M}_n(\mathbb{R})$, alors $\langle \mathrm{OA},\ \mathrm{OB}\rangle = \mathrm{Tr}(\mathrm{A}^\top \underbrace{\mathrm{O}^\top \mathrm{O}}_{=\mathrm{I}_n} \mathrm{B}) = \langle \mathrm{A},\ \mathrm{B}\rangle$.

→ Pour tout intervalle réel I, le produit scalaire canonique sur l'espace $\mathrm{L}^2_c(\mathrm{I}, \mathbb{R})$ des fonctions réelles continues sur I et de carré intégrable est défini par :

$$\forall f, g \in \mathrm{L}^2_c(\mathrm{I}), \quad \langle f,\ g\rangle = \int_{\mathrm{I}} f(t)g(t)\mathrm{d}t.$$

→ Le produit scalaire canonique sur l'espace $\ell^2(\mathbb{N}, \mathbb{R})$ des suites réelles de carré sommable est défini par :

$$\forall (u_n), (u_n) \in \ell^2(\mathbb{N}, \mathbb{R}), \quad \langle u,\ v\rangle = \sum_{n=0}^{+\infty} u_n v_n.$$

Remarque. L'existence des deux derniers produits scalaires est assurée par l'inégalité valable pour tous réels x et y : $|xy| \leq \dfrac{x^2 + y^2}{2}$.

Dans tout le reste du chapitre, on munit E d'un produit scalaire $\langle\cdot\,,\,\cdot\rangle$, si bien que $(\mathrm{E}, \langle\cdot\,,\,\cdot\rangle)$ est un espace préhilbertien réel.

BY-NC-SA

Théorème (Inégalité de Cauchy-Schwarz) I.2.

Pour tout $(x, y) \in \mathrm{E}$,
$$|\langle x,\, y\rangle| \leq \sqrt{\langle x,\, x\rangle}\sqrt{\langle y,\, y\rangle}$$
et l'inégalité est une égalité si, et seulement si, x et y sont colinéaires.

Démonstration. Supposons que $y \neq 0$. Introduisons la fonction réelle

$$\varphi : t \longmapsto \langle x + ty,\, x + ty\rangle = \langle x,\, x\rangle + 2\langle x,\, y\rangle t + \langle y,\, y\rangle t^2.$$

Cette fonction polynomiale est positive, donc son discriminant, qui vaut $4\langle x,\, y\rangle^2 - 4\langle x,\, x\rangle\langle y,\, y\rangle$, est négatif ou nul, d'où l'inégalité.
Le cas d'égalité est réalisé si, et seulement si, il existe $t_0 \in \mathbb{R}$ tel que $\varphi(t_0) = 0$ donc si, et seulement si, $x + t_0 y = 0$.

Remarque I.3.

Supposons que $\langle\,\cdot\,,\,\cdot\,\rangle$ ne soit pas défini positif. Soit $y \in \mathrm{E} \setminus \{0\}$ tel que $\langle y,\, y\rangle = 0$. Alors, pour tout $x \in \mathrm{E}$, $\langle x,\, y\rangle = 0$.
En effet, φ serait affine et positive, donc constante.

Exemple. Si $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace probabilisé dénombrable, l'application $f : (\mathrm{X}, \mathrm{Y}) \mapsto \mathbb{E}(\mathrm{XY})$ est une forme bilinéaire, symétrique, et positive sur l'espace $\mathrm{L}^2_0(\Omega, \mathbb{R})$ des variables aléatoires réelles sur Ω d'espérances nulles et qui admettent un moment d'ordre 2, mais elle n'est pas définie positive car si $\mathrm{X} \in \mathrm{L}^2(\Omega, \mathbb{R})$, on a l'équivalence $\mathbb{E}(\mathrm{X}^2) = 0 \iff \mathbb{P}(\mathrm{X} = 0) = 1$ mais X n'est pas nécessairement nulle.
Analogie avec le théorème de Pythagore : La variance peut être interprétée comme le carré de la « norme » associée au « produit scalaire » f. En effet, pour tout entier $n \geq 2$, si $\mathrm{X}_1, \ldots, \mathrm{X}_n \in \mathrm{L}^2(\Omega, \mathbb{R})$ sont indépendantes, alors, pour tout $(i, j) \in [\![1; n]\!]^2$ tel que $i \neq j$, $\mathrm{Cov}(\mathrm{X}_i, \mathrm{X}_i) = \mathbb{E}(\mathrm{X}_i\mathrm{X}_j) = 0$, et on a bien
$$\mathbb{V}\left(\sum_{k=1}^{n}\mathrm{X}_k\right) = \sum_{k=1}^{n}\mathbb{V}(\mathrm{X}_k).$$

Théorème (Inégalité de Minkowski) I.4.

Pour tout $x \in \mathrm{E}$, on note $\|x\|$ le réel $\sqrt{\langle x,\, x\rangle}$. Alors, pour tout $(x, y) \in \mathrm{E}^2$,
$$\|x + y\| \leq \|x\| + \|y\|\,.$$
et l'inégalité est une égalité si, et seulement si, il existe $\lambda \geq 0$ tel que $x = \lambda y$ ou $y = \lambda x$.

Démonstration. $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x,\, y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2$ où l'on a utilisé l'inégalité de Cauchy-Schwarz.
$\|x + y\| = \|x\| + \|y\|$ si, et seulement si, $\langle x,\, y\rangle = \|x\|\,\|y\|$, donc, si et seulement si le cas d'égalité de l'inégalité de Cauchy-Schwarz est vérifié (car $\langle x,\, y\rangle \leq |\langle x,\, y\rangle| \leq \|x\|\,\|y\|$). Quitte à échanger x et y, supposons $x \neq 0$. Alors, l'inégalité de Minkowski est une égalité si, et seulement si, il existe $\lambda \in \mathbb{R}$ tel que $\lambda\cancel{\langle x,\, x\rangle} = |\lambda|\,\cancel{\|x\|^2}$, ce qu'il fallait démontrer.
Remarque. Si n est un entier supérieur ou égal à 2, l'inégalité se généralise avec $x_1, \ldots, x_n \in \mathrm{E}$, et, s'ils ne sont pas tous nuls, le cas d'égalité se traduit géométriquement par l'appartenance des x_i à une demi-droite d'origine O, dont le sens est déterminé par l'un des points non confondu avec l'origine (*cf.* chapitre 1).

(cc) BY-NC-SA

Dans tout le reste du chapitre, E est également muni de la norme euclidienne $\|\cdot\|$ associée à $\langle\cdot,\cdot\rangle$.

Exercice I.5.

Soit $(E, \|\cdot\|)$ est un espace vectoriel normé non trivial. Soient C une partie convexe de E, et $x \in E$. x est un point extrémal de C s'il n'est contenu dans aucun segment de C non réduit à un point, donc si et seulement si, on a la propriété

$$\forall y, z \in C,\ (\exists t \in [0,1],\ x = ty + (1-t)z \Longrightarrow x = y \text{ ou } x = z).$$

Déterminer les points extrémaux de $\overline{B}(0,1)$.

Égalité de la médiane : Pour tout $(x,y) \in E^2$, $\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$. **Application.**

Soit $(a,b) \in E^2$ tel que $a \neq b$. Soit $r > 0$. Montrons que $\operatorname{diam}(\overbrace{\overline{B}(a,r) \cap \overline{B}(b,r)}^{\text{notée } \Gamma}) < 2r$.
Si, $\|a-b\| > 2r$, Γ est vide, et il n'y a rien à démontrer.
Si $\|a-b\| \leq 2r$, Γ n'est pas vide et $\dfrac{a+b}{2} \in \Gamma$. Soit $x \in \Gamma$.

$$\begin{aligned}
\left\|x - \frac{a+b}{2}\right\|^2 &= \left\|\frac{1}{2}(\underbrace{x-a}_{\text{noté } u}) + \frac{1}{2}(\underbrace{x-b}_{\text{noté } v})\right\|^2 \\
&= \frac{1}{4}\|u+v\|^2 \\
&\underset{\substack{\uparrow \\ \text{égalité de la médiane}}}{=} \frac{1}{2}\Big(\underbrace{\overbrace{\|u\|^2}^{<r^2} + \overbrace{\|v^2\|}^{<r^2}}_{<2r^2} - \frac{1}{2}\|u-v\|^2\Big) < r^2.
\end{aligned}$$

Or, pour tout $(x,y) \in \Gamma^2$, $\|x-y\| \leq \left\|x - \dfrac{a+b}{2}\right\| + \left\|y - \dfrac{a+b}{2}\right\| < 2r$, ce qu'on voulait.

Exercice I.6.

Soit $(E, \mathcal{N})$ un $\mathbb{R}$-espace vectoriel normé. On suppose que pour tout $(x,y) \in E^2$,

$$\mathcal{N}(x+y)^2 + \mathcal{N}(x-y)^2 = 2\left(\mathcal{N}(x)^2 - \mathcal{N}(y)^2\right).$$

Montrer qu'il existe un produit scalaire sur E tel que $\mathcal{N}$ soit la norme euclidienne qui lui est associée.

II Orthogonalité

1. Généralités

Définition II.1.

Soit I un ensemble non vide. Une famille $(x_i) \in E^I$ est dite orthogonale lorsque pour tout $(i,j) \in I^2$ tel que $i \neq j$, $\langle x_i, x_j\rangle = 0$.
Elle est dite orthonormée si, de plus, pour tout $i \in I$, $\|x_i\| = 1$.

(cc) BY-NC-SA

Proposition II.2.

Soit I un ensemble non vide, et $(x_i)_{i\in \mathrm{I}}$ une famille orthogonale de E.

→ Pour toute partie finie J de I, et pour tout $(\lambda_i) \in \mathbb{R}^{\mathrm{J}}$, $\left\|\sum_{i\in \mathrm{J}} \lambda_i x_i\right\|^2 = \sum_{i\in \mathrm{J}} |\lambda_i| \, \|x_i\|^2$.

→ Si, pour tout $i \in \mathrm{I}$, $x_i \neq 0$, alors $(x_i)_{i\in \mathrm{I}}$ est libre.

Exercice II.3.

On admet que $\ell^2(\mathbb{N},\mathbb{R})$ est de dimension infinie non dénombrable. Montrer que $\ell^2(\mathbb{N},\mathbb{R})$, muni de son produit scalaire canonique et de la norme associée notée $\|\cdot\|$, n'admet pas de base orthonormée.

2. Dimension finie

Dans cette partie, on suppose que E est de dimension finie n, *i.e.* $(\mathrm{E}, \langle\,\cdot\,,\,\cdot\,\rangle)$ est un espace euclidien.

Théorème II.4.

L'application

$$j : \begin{cases} \mathrm{E} & \longrightarrow \mathrm{E}^* \\ u & \longmapsto \langle u,\,\cdot\,\rangle \end{cases}$$

est un isomorphisme.

Démonstration. La bilinéarité du produit scalaire entraîne la linéarité de j.
Si $u \in \mathrm{Ker}(j)$, alors, pour tout $v \in \mathrm{E}$ $\langle u,\, v\rangle = 0$, donc $\langle u,\, u\rangle = 0$, donc $u = 0$ *i.e.* j est injective. L'égalité des dimensions de E et E^* permet de conclure.

Corollaire II.5.

Supposons que $n \geq 2$. Soit H un hyperplan de E. Alors, il existe $u \in \mathrm{E}\setminus\{0\}$ tel que, pour tout $x \in \mathrm{E}$, on ait l'équivalence

$$x \in \mathrm{H} \iff \langle u,\, x\rangle = 0.$$

L'ensemble de tels u est inclus dans une droite d'origine O, et $\mathbb{R}u$ est la normale à H.

Démonstration. H est le noyau d'une forme linéaire sur E non nulle, donc d'après le théorème, il existe un unique $u \in \mathrm{E}\setminus\{0\}$, dépendant du choix de cette forme linéaire, tel que $\mathrm{H} = \mathrm{Ker}(\langle u,\,\cdot\,\rangle)$, ce qui démontre l'équivalence.
Soit φ une forme linéaire sur E non nulle de noyau H, et notons v_φ l'unique vecteur tel que $\langle v_\varphi,\,\cdot\,\rangle = \varphi$. Alors v_φ convient également, et $\mathrm{Ker}(\langle v_\varphi,\,\cdot\,\rangle) = \mathrm{Ker}(\langle u,\,\cdot\,\rangle) = \mathrm{H}$ donc $\langle u,\,\cdot\,\rangle$ et $\langle v_\varphi,\,\cdot\,\rangle$ sont proportionnelles, *i.e.* il existe $\lambda \in \mathbb{R}^*$ tel que $\langle v_\varphi,\,\cdot\,\rangle = \lambda\,\langle u,\,\cdot\,\rangle = \langle \lambda u,\,\cdot\,\rangle$. j étant injective, $v_\varphi = \lambda u$.

(cc) BY-NC-SA

Théorème II.6.

→ E possède des bases orthonormées.

→ Supposons que $n \geq 2$. Soit $p \in \mathbb{N}^*$ tel que $p < n$. Si $(e_1, \dots, e_p)$ est une famille orthonormée de E, alors il existe une famille $(e_{p+1}, \dots, e_n)$ telle que $(e_1, \dots, e_n)$ soit une base de E.

Démonstration.

→ Si $n = 1$, un vecteur $x \in \mathrm{E}$ de norme 1 forme à lui tout seul une base de E. Supposons que $n \geq 2$, et pour tout $k \in [\![1; n-1]\!]$ introduisons par récurrence descendante un hyperplan E_k de E_{k+1}. Pour tout $k \in [\![1; n]\!]$ introduisons l'assertion

$$\mathcal{P}_k : \text{« } \mathrm{E}_k \text{ possède une base orthonormée ».}$$

$\mathcal{P}_1$ est vraie. Supposons qu'il existe $k \in [\![1; n-1]\!]$ tel que $\mathcal{P}_k$ soit vraie. Alors, E_k, possède une base orthonormée $(e_1, \dots, e_k)$. Soit $u \in \mathrm{E}_{k+1} \setminus \{0\}$ tel que $\mathbb{R}u$ soit la normale à E_k dans E_{k+1}, et posons $e_{k+1} = \dfrac{1}{\|u\|} u$ de sorte que $\|e_{k+1}\| = 1$. Alors la famille $(e_1, \dots, e_{k+1})$ est libre car e_{k+1} est orthogonal aux vecteurs de $(e_1, \dots, e_k)$. De plus, elle comporte $k+1$ vecteurs, et $\dim(\mathrm{E}_{k+1}) = k+1$. Donc $(e_1, \dots, e_{k+1})$ est une base orthonormée de E_{k+1}. Donc $\mathcal{P}_{k+1}$ est vraie. Par récurrence finie, E possède une base orthonormée. Chaque vecteur de la base construite peut être changé en son opposé, donc E admet des bases orthonormée.

→ En complétant la famille libre (car orthonormée) $(e_1, \dots, e_p)$ en une base $(e_1, \dots, e_p, u_{p+1}, \dots, u_p)$ de E, puis en posant $\mathrm{E}_p = \mathrm{Vect}(e_1, \dots, e_p)$, et, pour tout $k \in [\![p{+}1; n]\!]$, $\mathrm{E}_k = \mathrm{Vect}(e_1, \dots, e_p, u_{p+1}, \dots, u_k)$, la récurrence ci-dessus permet de compléter $(e_1, \dots, e_p)$ en une base orthonormée de E.

Théorème (Procédé de Gram-Schmidt) II.7.

Soit $(e_1, \dots, e_n)$ une base de E. Il existe une base orthonormée $(\varepsilon_1, \dots, \varepsilon_n)$ de E telle qu'on ait la propriété :

$$\forall k \in [\![1; n]\!], \quad \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_k) = \mathrm{Vect}(e_1, \dots, e_k).$$

De plus, si $(\varepsilon'_1, \dots, \varepsilon'_n)$ est une base orthonormée vérifiant la même propriété, alors, il existe $(\alpha_1, \dots, \alpha_n) \in \{-1, 1\}^n$ tel que pour tout $i \in [\![1; n]\!]$, $\varepsilon'_i = \alpha_i \varepsilon_i$.

Démonstration. Pour tout $k \in [\![1; n]\!]$, introduisons l'assertion

$$\mathcal{A}_k : \text{« Il existe une famille orthonormée } (\varepsilon_1, \dots, \varepsilon_k) \text{ de E telle que pour tout } i \in [\![1; k]\!], \underbrace{\mathrm{Vect}(e_1, \dots, e_i)}_{\text{noté } \mathrm{E}_i} = \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_i) \text{ »}$$

→ En posant $\varepsilon_1 = \pm \dfrac{1}{\|e_1\|} e_1$, on montre que $\mathcal{A}_1$ est vraie. Ces deux choix sont également les seuls possibles pour que $\mathcal{A}_1$ soit vraie.

→ Soit $k \in [\![1; n-1]\!]$ tel que $\mathcal{A}_k$ soit vraie. E_k est un hyperplan de E_{k+1} donc il existe $u \in \mathrm{E}_{k+1} \setminus \{0\}$ tel que $\mathbb{R}u$ soit l'unique normale à E_k dans E_{k+1} passant par l'origine. Ainsi, pour que $(\varepsilon_1, \dots, \varepsilon_{k+1})$ soit une base orthonormée de E_{k+1}, il faut et il suffit que $\varepsilon_{k+1} = \pm \dfrac{1}{\|u\|} u$.

Remarque. Le caractère suffisant de cette condition suffit pour mener à terme cette récurrence. Alors, $\mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_{k+1}) = \mathrm{E}_{k+1} = \mathrm{Vect}(e_1, \dots, e_{k+1})$. L'hypothèse de récurrence assure par ailleurs que pour tout $i \in [\![1; k]\!]$, $\mathrm{Vect}(e_1, \dots, e_i) = \mathrm{Vect}(\varepsilon_1, \dots, \varepsilon_i)$. Donc $\mathcal{A}_{k+1}$ est vraie.

→ Finalement, le principe de récurrence assure que $\mathcal{A}_n$ est vraie.

(cc) BY-NC-SA

Nous avons exhibé, lors de la construction de $(\varepsilon_1, \ldots, \varepsilon_n)$, que chaque ε_i peut être échangé avec son opposé, et seulement son opposé, si bien que l'ensemble des bases orthonormées vérifiant la propriété démontrée par récurrence est $\{(\alpha_1\varepsilon_1, \ldots, \alpha_n\varepsilon_n) \mid (\alpha_1, \ldots, \alpha_n) \in \{-1, 1\}^n\}$

Expression matricielle : La matrice de $(e_1, \ldots, e_n)$ relativement la base $(\varepsilon_1, \ldots, \varepsilon_n)$ est triangulaire supérieure. En effet, pour tout $i \in [\![1; n]\!]$, e_i est une combinaison linéaire des vecteurs de la famille $(\varepsilon_1, \ldots, \varepsilon_i)$ en vertu de l'égalité $\mathrm{Vect}(e_1, \ldots, e_i) = \mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_i)$.

Proposition II.8.

Soient x et $y \in \mathrm{E}$, et introduisons leurs décompositions dans une base orthonormée $(e_1, \ldots, e_n)$ de E :

$$x = \sum_{i=1}^{n} x_i e_i \qquad y = \sum_{i=1}^{n} y_i e_i.$$

Alors,

$$\langle x,\ y\rangle = \sum_{i=1}^{n} x_i y_i \qquad \|x\|^2 = \sum_{i=1}^{n} \langle x,\ e_i\rangle^2 = \sum_{i=1}^{n} x_i^2.$$

Remarque II.9.

La décomposition de x étant unique dans la base $(e_1, \ldots, e_n)$, l'application

$$\begin{aligned} \mathrm{E} &\longrightarrow \mathbb{R}^n \\ x &\longmapsto (\langle x,\ e_i\rangle)_{1 \le i \le n} \end{aligned}$$

est un isomorphisme.

Exercice II.10.

Soit $(\mathrm{E}, \langle\cdot\,,\,\cdot\rangle)$ un espace euclidien de dimension $n \geq 1$, et notons $\|\cdot\|$ la norme euclidienne associée à $\langle\cdot\,,\,\cdot\rangle$. Soit m un entier supérieur ou égal à n. Soit $(e_1, \ldots, e_m) \in \mathrm{E}^m$ tel que pour tout $x \in \mathrm{E}$,

$$\|x\|^2 = \sum_{i=1}^{n} \langle x,\ e_i\rangle^2 .$$

1. Montrer que $\mathrm{Vect}(e_1, \ldots, e_m) = \mathrm{E}$.
2. Donner un exemple d'espace euclidien et des exemples de vecteurs vérifiant les hypothèses de l'exercice lorsque $m = 3$, $n = 2$, et $\|e_1\| = \|e_2\| = \|e_3\|$.
3. Supposons que $n = m$. Montrer que $(e_1, \ldots, e_n)$ est une base orthonormée de E.

3. Orthogonal d'une partie

On rappelle que $(\mathrm{E}, \langle\cdot\,,\,\cdot\rangle)$ est un espace préhilbertien réel, et que $\|\cdot\|$ est la norme associée à $\langle\cdot\,,\,\cdot\rangle$.

Définition II.11.

Soit A une partie de E. On appelle *orthogonal de* A l'ensemble

$$\mathrm{A}^\perp = \{x \in \mathrm{E} \mid \forall y \in \mathrm{A},\ \langle x,\ y\rangle = 0\}.$$

(cc) BY-NC-SA

Proposition II.12.

Soient A et B des parties de E.

→ $\{0\}^{\perp} = \mathrm{E}$ et $\mathrm{E}^{\perp} = \{0\}$.

→ $\mathrm{A}^{\perp} = \bigcap_{x \in \mathrm{A}} \mathrm{Ker}(\langle x,\, \cdot\, \rangle)$ est un sous-espace vectoriel de E.

→ Si $\mathrm{A} \subset \mathrm{B}$, alors $\mathrm{B}^{\perp} \subset \mathrm{A}^{\perp}$.

Conséquence. Pour toute partie A de E, on a $\mathrm{A} \subset \left(\mathrm{A}^{\perp}\right)^{\perp}$, donc $\mathrm{A}^{\perp} = \left(\left(\mathrm{A}^{\perp}\right)^{\perp}\right)^{\perp}$.

Définition II.13.

Soit F un sous-espace vectoriel de E. On dit que F admet un supplémentaire orthogonal dans E s'il existe un sous-espace vectoriel G de $\mathrm{F}^{\perp}$ tel que $\mathrm{E} = \mathrm{F} \oplus \mathrm{G}$.

Conséquence. Dans ce cas, $\mathrm{G} = \mathrm{F}^{\perp}$.

Démonstration de l'unicité sous réserve d'existence. Si, $x \in \mathrm{F}^{\perp}$, il existe $y \in \mathrm{F}$, et $z \in \mathrm{G}$ tels que $x = y + z$, donc $y = x - z \in \mathrm{F}^{\perp}$ donc $y = 0$, donc $x = z \in \mathrm{G}$. Ainsi, $\mathrm{F}^{\perp} \subset \mathrm{G}$.

Observation. Soit F une partie de E. Alors $\mathrm{F}^{\perp}$ est fermé dans E.

Démonstration. Pour tout $y \in \mathrm{F}$, l'application linéaire $x \mapsto \langle x,\, y\rangle$ est continue en vertu de l'inégalité de Cauchy-Schwarz, son noyau est donc fermé en tant qu'image réciproque de $\{0\}$, fermé dans $\mathbb{R}$.
Or $\mathrm{F}^{\perp}$ est l'intersection de tels fermés, d'où l'observation.

4. Orthogonal d'un sous-espace vectoriel de dimension finie

Dans cette partie, on suppose que $(\mathrm{E}, \langle\, \cdot\,,\, \cdot\, \rangle)$ est un espace euclidien de dimension n.

Théorème II.14.

Soit F un sous-espace vectoriel de E. Alors $\mathrm{F} \oplus \mathrm{F}^{\perp} = \mathrm{E}$.

Démonstration. Notons p la dimension de F. Soit $(e_1, \ldots, e_p)$ une base orthonormée de F que l'on complète en une base orthonormée $(e_1, \ldots, e_n)$ de E. Alors, pour tout $x \in \mathrm{E}$, il existe $(x_1, \ldots, x_n) \in \mathbb{R}^n$ tel que

$$x = \underbrace{\sum_{i=1}^{p} x_i e_i}_{\in \mathrm{F}} + \underbrace{\sum_{i=p+1}^{n} x_i e_i}_{\in \mathrm{F}^{\perp}},$$

donc $\mathrm{F} + \mathrm{F}^{\perp} = \mathrm{E}$. De plus, $\mathrm{F} \cap \mathrm{F}^{\perp} = \{0\}$, donc $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}$.

(cc) BY-NC-SA

Proposition II.15.

Soient F et G des sous-espaces vectoriels de E.

→ $\dim(\mathrm{F}^\perp) = n - \dim(\mathrm{F})$.

→ $\left(\mathrm{F}^\perp\right)^\perp = \mathrm{F}$. En effet, $\mathrm{F} \subset \left(\mathrm{F}^\perp\right)^\perp$, et $\dim\left(\left(\mathrm{F}^\perp\right)^\perp\right) = \dim(\mathrm{F})$.

→ $(\mathrm{F}+\mathrm{G})^\perp = \mathrm{F}^\perp \cap \mathrm{G}^\perp$. En effet, par bilinéarité du produit scalaire, $\mathrm{F}^\perp \cap \mathrm{G}^\perp \subset (\mathrm{F}+\mathrm{G})^\perp$, et l'inclusion réciproque est vraie car $\mathrm{F} \subset \mathrm{F}+\mathrm{G}$ et $\mathrm{G} \subset \mathrm{F}+\mathrm{G}$.

→ $(\mathrm{F}\cap\mathrm{G})^\perp = \left(\left(\mathrm{F}^\perp\right)^\perp \cap \left(\mathrm{G}^\perp\right)^\perp\right)^\perp = \left(\left(\mathrm{F}^\perp + \mathrm{G}^\perp\right)^\perp\right)^\perp = \mathrm{F}^\perp + \mathrm{G}^\perp$.

Remarque. L'orthogonal d'une somme est l'intersection des orthogonaux, et l'orthogonal d'une intersection est la somme des orthogonaux.

5. Projecteurs orthogonaux

On rappelle que $(\mathrm{E}, \langle\cdot\,,\cdot\rangle)$ est un espace préhilbertien réel, et que $\|\cdot\|$ est la norme euclidienne associée à $\langle\cdot\,,\cdot\rangle$.

Définition II.16.

Soit $p \in \mathscr{L}(\mathrm{E})$. On dit que p est un projecteur orthogonal lorsque $p = p\circ p$ et $\mathrm{Ker}(p) \perp \mathrm{Im}(p)$.

Proposition II.17.

Soit $p \in \mathscr{L}(\mathrm{E})$ tel que p soit un projecteur orthogonal.

→ $\mathrm{Ker}(p) = \mathrm{Im}(p)^\perp$ et $\mathrm{Im}(p) = \mathrm{Ker}(p)^\perp$.

→ $\mathrm{Ker}(p)$ et $\mathrm{Im}(p)$ sont fermés.

→ Si $p \neq 0$, $|||p||| = 1$.

Démonstration.

→ $\mathrm{Ker}(p) \oplus \mathrm{Im}(p) = \mathrm{E}$ et $\mathrm{Im}(p) \subset \mathrm{Ker}(p)^\perp$ donc $\mathrm{Im}(p) = \mathrm{Ker}(p)^\perp$.

→ La première propriété entraîne immédiatement la deuxième.

→ Soit $x = \underbrace{y}_{\in\mathrm{Ker}(p)} + \underbrace{z}_{\in\mathrm{Im}(p)} \in \mathrm{E}$. Alors $\|x\|^2 = \|y\|^2 + \|z\|^2 \geq \|z\|^2 = \|p(x)\|^2$.

Donc, pour tout $x \in \mathrm{E}\setminus\{0\}$, $\dfrac{\|p(x)\|^2}{\|x\|^2} \leq 1$, ce qui assure que $|||p||| \leq 1$. De plus, ce majorant est atteint pour tout $x \in \mathrm{Im}(p)\setminus\{0\}$, ce qui assure que $|||p||| = 1$.

Théorème II.18.

Soit F un sous-espace vectoriel de E. Alors F possède un supplémentaire orthogonal si, et seulement si, il existe un projecteur orthogonal d'image F.

Démonstration. $(\Leftarrow)$ S'il existe un projecteur orthogonal p de E tel que $\mathrm{F} = \mathrm{Im}(p)$, alors un supplémentaire orthogonal de F est $\mathrm{Ker}(p)$.

(cc) BY-NC-SA

($\Rightarrow$) Le projecteur sur F parallèlement à $\mathrm{F}^\perp$ est bien un projecteur orthogonal d'image F.

Proposition II.19.

Soit F un sous-espace vectoriel de E tel qu'il soit image d'un projecteur orthogonal p. Alors, pour tout $x \in \mathrm{E}$,

$$\mathrm{d}(x, \mathrm{F}) = \|x - p(x)\|$$

et $p(x)$ est l'unique vecteur de F tel que cette égalité est vérifiée.

Démonstration. Soit $y \in \mathrm{F} = \mathrm{Im}(p)$. $x - y = \underbrace{x - p(x)}_{\in \mathrm{F}^\perp} + \underbrace{p(x) - y}_{\in \mathrm{F}}$, donc

$$\|x - y\|^2 = \|x - p(x)\|^2 + \|p(x) - y\|^2 \geq \|x - p(x)\|^2$$

et ce minorant est réalisé si, et seulement si, $y = p(x)$ (car sinon, la norme ne serait pas séparante).

Théorème II.20.

Soit F un sous-espace vectoriel de E de dimension finie $d \geq 1$. Soit $(e_1, \ldots, e_d)$ une base orthonormée de F. Alors l'application

$$\pi : \begin{cases} \mathrm{E} & \longrightarrow \mathrm{F} \\ x & \longmapsto \displaystyle\sum_{i=1}^{d} \langle e_i,\, x\rangle\, e_i \end{cases}$$

est un projecteur orthogonal de E sur F et $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^\perp$.

De plus, pour tout $x \in \mathrm{E}$, $\mathrm{d}(x, \mathrm{F})^2 = \|x\|^2 - \displaystyle\sum_{i=1}^{d} \langle e_i,\, x\rangle^2$.

Démonstration. π est linéaire et $\mathrm{Im}(\pi) = \mathrm{F}$. En effet, pour tout $x \in \mathrm{F}$, $x = \displaystyle\sum_{i=1}^{d} \langle x,\, e_i\rangle\, e_i = \pi(x)$.

Montrons que $\mathrm{Ker}(\pi) \perp \mathrm{F}$. Soit $x \in \mathrm{Ker}(\pi)$. Alors, $0 = \pi(x) = \displaystyle\sum_{i=1}^{d} \langle x,\, e_i\rangle\, e_i$. Or $(e_1, \ldots, e_d)$ est libre, donc pour tout $i \in [\![1; d]\!]$, $\langle e_i,\, x\rangle = 0$. Donc $\mathrm{Ker}(\pi) \subset \mathrm{F}^\perp$. Donc $\mathrm{Ker}(\pi) = \mathrm{F}^\perp$, et $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^\perp$.

De plus, pour tout $x \in \mathrm{E}$, $\|x\|^2 = \|x - \pi(x)\|^2 + \|\pi(x)\|^2 = \mathrm{d}(x, \mathrm{F})^2 + \displaystyle\sum_{i=1}^{d} \langle x,\, e_i\rangle^2$.

Remarque. On prouve au passage l'inégalité de Bessel : pour toute famille orthonormée $(e_1, \ldots, e_d)$ de E, et pour tout $x \in \mathrm{E}$,

$$\sum_{i=1}^{d} \langle x,\, e_i\rangle^2 \leq \|x\|^2 .$$

Caractérisation des projecteurs orthogonaux parmi les projecteurs

Soit $p \in \mathscr{L}(\mathrm{E})$ tel que $p^2 = p$. Montrons que p est orthogonal si, et seulement si, p est continu et $|\!|\!|p|\!|\!| \leq 1$.

Nous avons déjà démontré que si p est un projecteur orthogonal de E, alors $|\!|\!|p|\!|\!| = 1 \leq 1$, ce qui entraîne sa continuité.

Supposons que p soit continu et que $|\!|\!|p|\!|\!| \leq 1$.

(cc) BY-NC-SA

Soit $(x, y) \in \mathrm{Ker}(p) \times \mathrm{Im}(p)$. Alors, pour tout $t \in \mathbb{R}$,

$$\|p(x+ty)\|^2 \leq \|x+ty\|^2 \qquad \text{donc} \qquad \cancel{\|x\|^2} \leq \cancel{\|x\|^2} + 2t\,\langle x,\ y\rangle + \|y\|^2\, t^2,$$

donc pour tout $t > 0$, $2\,\langle x,\ y\rangle + t\,\|y\|^2 \geq 0$, puis par passage à la limite, $t \to 0^+$, $\langle x,\ y\rangle \geq 0$.
De même, pour tout $t < 0$, $2\,\langle x,\ y\rangle + t\,\|y\|^2 \leq 0$, puis par passage à la limite, $t \to 0^-$, $\langle x,\ y\rangle \leq 0$.
Donc $\langle x,\ y\rangle = 0$, donc $\mathrm{Im}(p) \subset \mathrm{Ker}(p)^{\perp}$. Ainsi, p admet un supplémentaire orthogonal dans E, ce qui assure que $\mathrm{Im}(p) = \mathrm{Ker}(p)^{\perp}$. Donc p est orthogonal.

Exercice II.21.

Soit $(\mathrm{E}, \langle\,\cdot\,,\,\cdot\,\rangle)$ un espace préhilbertien réel, et notons $\|\cdot\|$ la norme euclidienne associée à $\langle\,\cdot\,,\,\cdot\,\rangle$.
Soient p et q deux projecteurs orthogonaux de E. Montrer que $p \circ q$ est un projecteur orthogonal de E si $p \circ q = q \circ p$.

(cc) BY-NC-SA

Correction de l'exercice I.5. :

Remarque. On conjecture aisément que les points extrémaux sont ceux de la sphère unité.

→ Soit $x \in \mathrm{B}(0,1)$. Si $x = 0$ alors en fixant $a \in \overline{\mathrm{B}}(0,1) \setminus \{0\}$, $x = \dfrac{1}{2}(a - a)$ et pourtant $x \neq a$ et $x \neq -a$.
Si $x \neq 0$, alors $x = (1 - \|x\|) \cdot 0 + \|x\| \dfrac{x}{\|x\|}$ et pourtant $x \neq 0$ et $x \neq \dfrac{x}{\|x\|}$.
Donc les points de $\mathrm{B}(0,1)$ ne sont pas des points extrémaux de $\overline{\mathrm{B}}(0,1)$.

→ Soit $x \in \mathrm{S}(0,1)$. Soient $y, z \in \mathrm{C}$. Supposons qu'il existe $t \in [0,1]$ tel que $x = ty + (1-t)z$. Si $t = 0$ ou $t = 1$, il n'y a rien à démontrer. Sinon, d'après l'inégalité triangulaire,

$$1 \leq t\,\|y\| + (1-t)\,\|z\| \leq t + (1-t) = 1,$$

donc $t\,\|y\| + (1-t)\,\|z\| = \|ty + (1-t)z\|$. D'après le cas d'égalité de l'inégalité de Minkowski, il existe $\lambda \geq 0$ tel que $y = \lambda z$ ou $\lambda y = z$. Par ailleurs,

$$\underbrace{t(1 - \|y\|)}_{\geq 0} + \underbrace{(1-t)(1 - \|z\|)}_{\geq 0} = 0,$$

donc $\|y\| = \|z\| = 1$, donc $\lambda = 1$. Donc $x = y$ ou $x = z$.

Donc $\mathrm{S}(0,1)$ est l'ensemble des points extrémaux de $\mathrm{B}(0,1)$.

Correction de l'exercice I.6. :

Introduisons l'application

$$\varphi : (x,y) \longmapsto \mathcal{N}(x+y)^2 - \mathcal{N}(x-y)^2$$

définie sur $\mathrm{E} \times \mathrm{E}$. φ est clairement symétrique et définie positive. Il reste à montrer qu'elle est bilinéaire. Soient x, y et $z \in \mathrm{E}$. D'après l'égalité de la médiane,

$$\begin{aligned}
\varphi(x, y+z) &= 2\left(\mathcal{N}(x+y)^2 + \mathcal{N}(z)^2\right) - \mathcal{N}(x+y-z)^2 - \mathcal{N}(x-y-z)^2 \\
&= 2\left(\mathcal{N}(x+y)^2 + \mathcal{N}(z)^2\right) - 2\left(\mathcal{N}(x-z)^2 + \mathcal{N}(y)^2\right), \\
\varphi(-x, y+z) &= 2\left(\mathcal{N}(x-y)^2 + \mathcal{N}(z)^2\right) - \mathcal{N}(x-y+z)^2 - \mathcal{N}(x+y+z)^2 \\
&= 2\left(\mathcal{N}(x-y)^2 + \mathcal{N}(z)^2\right) - 2\left(\mathcal{N}(x+z)^2 + \mathcal{N}(y)^2\right).
\end{aligned}$$

Or, $\varphi(-x, y+z) = -\varphi(x, y+z)$, donc en soustrayant membre à membre,

$$\varphi(x, y+z) = \left(\mathcal{N}(x+y)^2 - \mathcal{N}(x-y)^2\right) + \left(\mathcal{N}(x+z)^2 - \mathcal{N}(x-z)^2\right) = \varphi(x,y) + \varphi(x,z).$$

En remarquant que $\varphi(x, 0) = 0$, une récurrence immédiate permet de montrer que pour tout $n \in \mathbb{N}$,

$$\varphi(x, ny) = n\varphi(x,y).$$

Or, $\varphi(x, -y) = -\varphi(x,y)$, donc pour tout $m \in \mathbb{Z}$, $\varphi(x, my) = m\varphi(x,y)$.
Soit $r = \dfrac{p}{q} \in \mathbb{Q}$ tel que $(p,q) \in \mathbb{Z} \times \mathbb{N}^*$. Alors

$$p\varphi(x,y) = \varphi(x, py) = \varphi(x, qry) = q\varphi(x, ry),$$

donc $\varphi(x, ry) = \dfrac{p}{q}\varphi(x,y) = r\varphi(x,y)$ et ceci est valable pour tout rationnel r.

Enfin, soit $\lambda \in \mathbb{R}$. Par densité de $\mathbb{Q}$ dans $\mathbb{R}$, il existe $(\lambda_n) \in \mathbb{Q}^{\mathbb{N}}$ tel que $\lambda_n \xrightarrow[n\to+\infty]{} \lambda$.
Or $u \mapsto \varphi(x,u)$ est continue, donc $\underbrace{\lambda_n\varphi(x,y)}_{\xrightarrow[n\to+\infty]{}\lambda\varphi(x,y)} = \varphi(x,\lambda_n y) \xrightarrow[n\to+\infty]{} \varphi(x,\lambda y)$, donc par unicité de la limite,

$$\varphi(x,\lambda y) = \lambda\varphi(x,y),$$

et ceci est valable pour tout $\lambda \in \mathbb{R}$. Donc φ est linéaire selon la deuxième variable. Étant de plus symétrique, elle est bilinéaire. Donc φ est un produit scalaire sur E.

Correction de l'exercice II.3. :

Soit I un ensemble infini non dénombrable.
Raisonnons par l'absurde en supposant que $\ell^2(\mathbb{N},\mathbb{R})$ admet une base orthonormée $(e^i) \in \ell^2(\mathbb{N},\mathbb{R})^{\mathrm{I}}$.

→ Montrons que $\ell^2(\mathbb{N},\mathbb{R})$ est séparable, *i.e.* $\ell^2(\mathbb{N},\mathbb{R})$ contient une partie dénombrable dense.
Soit $(x_n) \in \ell^2(\mathbb{N},\mathbb{R})$. Soit $\varepsilon > 0$. Il existe $n_\varepsilon \in \mathbb{N}$ tel que $\sum\limits_{n=n_\varepsilon+1}^{+\infty} x_n^2 \leq \frac{\varepsilon^2}{2}$.
Par ailleurs, pour tout $n \in [\![0;n_\varepsilon]\!]$, il existe $a_n \in \mathbb{Q}$ tel que $|a_n - x_n| \leq \frac{\varepsilon}{\sqrt{2(n_\varepsilon+1)}}$.
Prolongeons $(a_0,\dots,a_{n_\varepsilon})$ en une suite presque nulle $(a_n) \in \mathbb{Q}[\mathrm{X}]$ en posant, pour tout entier $n > n_\varepsilon$, $a_n = 0$.
Rappel : Rappelons qu'on se donne une telle suite si, et seulement si, on se donne un polynôme à coefficients rationnels, d'où la notation commune des deux ensembles.
Ainsi,

$$\|(x_n)_{n\in\mathbb{N}} - (a_n)_{n\in\mathbb{N}}\|^2 = \sum_{n=1}^{n_\varepsilon}(x_n-a_n)^2 + \sum_{n=n_\varepsilon+1}^{+\infty} x_n^2 \leq \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2,$$

ce qui prouve que $\mathbb{Q}[\mathrm{X}]$ est dense dans $\ell^2(\mathbb{N},\mathbb{R})$.

→ $\mathbb{Q}[\mathrm{X}]$ est dénombrable. En effet, $\mathbb{Q}[\mathrm{X}] = \bigcup\limits_{n=0}^{+\infty} \mathbb{Q}_n[\mathrm{X}]$, et il est clair que pour tout $n \in \mathbb{N}$, $\mathbb{Q}_n[\mathrm{X}]$ est en bijection avec $\mathbb{Q}^{n+1}$ qui est dénombrable en tant que produit cartésien d'ensembles dénombrables.

→ Ainsi, pour tout $i \in \mathrm{I}$, il existe $(a_n^i)_{n\in\mathbb{N}} \in \mathbb{Q}[\mathrm{X}]$ telle que $\|(e_n^i)_{n\in\mathbb{N}} - (a_n^i)_{n\in\mathbb{N}}\| < \frac{1}{2}$. Soit $(i,j) \in \mathbb{N}^2$ tel que $a^i = a^j$. Alors,

$$\underbrace{\left\|(e_n^i)_{n\in\mathbb{N}} - (e_n^j)_{n\in\mathbb{N}}\right\|}_{=2\delta_{i,j}} \leq \left\|(e_n^i)_{n\in\mathbb{N}} - (a_n^i)_{n\in\mathbb{N}}\right\| + \left\|(e_n^j)_{n\in\mathbb{N}} - (a_n^j)_{n\in\mathbb{N}}\right\| < 1,$$

donc $i = j$. Donc $i \longmapsto (a_n^i)$ est une injection de I dans $\mathbb{Q}[\mathrm{X}]$, donc I est dénombrable, ce qui est faux.

En conclusion, $\ell^2(\mathbb{N},\mathbb{R})$ n'admet pas de base orthonormée.

Correction de l'exercice II.10. :

1. Supposons qu'il existe un hyperplan H de E tel que $\mathrm{Vect}(e_1,\dots,e_m) \subset \mathrm{H}$. Soit $u \in \mathrm{E}\setminus\{0\}$ tel que $\mathrm{H}^\perp = \mathbb{R}u$. Alors

$$\|u\|^2 = \sum_{i=1}^{n}\langle u,\ e_i\rangle = 0,$$

ce qui est faux. Donc $\mathrm{E} = \mathrm{Vect}(e_1,\dots,e_m)$.

2. Considérons $\mathbb{R}^2$ muni de produit scalaire canonique, et posons $u_1 = (1,0)$, $u_2 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ et

(cc) BY-NC-SA

$u_3 = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$. Alors, $\|u_1\| = \|u_2\| = \|u_3\|$. De plus, pour tout $x = (x_1, x_2) \in \mathbb{R}^2$,

$$\sum_{i=1}^{3} \langle x,\ u_i\rangle^2 = x_1^2 + \left(-\frac{1}{2}x_1 + \frac{\sqrt{3}}{2}x_2\right)^2 + \left(-\frac{1}{2}x_1 - \frac{\sqrt{3}}{2}x_2\right)^2 = \frac{3}{2}\|x\|^2,$$

donc en posant $(e_1, e_2, e_3) = \sqrt{\frac{2}{3}}(u_1, u_2, u_3)$, on a bien $\|e_1\| = \|e_2\| = \|e_3\|$, et

$$\sum_{i=1}^{3} \langle x,\ e_i\rangle^2 = \|x\|^2$$

3. Soit $i \in [\![1;n]\!]$. $\|e_i\|^2 = \sum_{j=1}^{n} \langle e_i,\ e_j\rangle^2 = \|e_i\|^4 + \sum_{\substack{j=1\\ j\neq i}}^{n} \underbrace{\langle e_i,\ e_j\rangle^2}_{\geq 0}$, donc $\|e_i\|^2 \geq \|e_i\|^4$ donc $\|e_i\| \leq 1$.

   Soit $u \in \operatorname{Vect}\left((e_j)_{1\leq j\neq i\leq n}\right)^{\perp}$ tel que $\|u\| = 1$. Alors,

$$1 = \|u\|^2 = \langle e_i,\ u\rangle^2 \underset{\substack{\uparrow\\ \text{Cauchy-Schwarz}}}{\leq} \|u\|^2\|e_i\|^2 \leq \|e_i\|^2,$$

   donc $\|e_i\| = 1$. Ainsi, $\cancel{1} = \|e_i\|^2 = \underbrace{\cancel{\|e_i\|^4}}_{=1} + \sum_{\substack{j=1\\ j\neq i}}^{n} \underbrace{\langle e_i,\ e_j\rangle^2}_{\geq 0}$, donc $\sum_{\substack{j=1\\ j\neq i}}^{n} \langle e_i,\ e_j\rangle^2 = 0$, donc pour tout $j \in [\![1;n]\!]$ tel que $j \neq i$, $\langle e_i,\ e_j\rangle = 0$. Donc $(e_1, \ldots, e_n)$ est une base orthonormée de E.

Correction de l'exercice II.21. :

Si $p\circ q = q\circ p$, alors $(p\circ q)^2 = p\circ q\circ p\circ q = p\circ q\circ q\circ p = p\circ q\circ p = p\circ p\circ q = p\circ q$, donc $p\circ q$ est un projecteur. Or, $|\!|\!|p\circ q|\!|\!| \leq |\!|\!|p|\!|\!|\,|\!|\!|q|\!|\!| \leq 1$ donc $p\circ q$ est orthogonal.

(cc) BY-NC-SA

COMPLÉMENTS

Espaces préhilbertiens réels, compléments

Dans tout le document, $(\mathrm{E}, \langle \cdot\,, \ \cdot \rangle)$ désigne espace préhilbertien réel et on note $\|\cdot\|$ la norme associée au produit scalaire $\langle \cdot\,, \ \cdot \rangle$. De plus, n désigne un entier supérieur ou égal à 1.

I Décomposition de Cartan (Iwasawa)

Dans cette partie, on suppose que E est de dimension finie n.

1. Changements de base, orientation

Rappel I.1.

Soient $\mathscr{E}$, $\mathscr{F}$, et $\mathscr{G}$ des bases de E.

→ La matrice de passage de $\mathscr{E}$ à $\mathscr{F}$ est la matrice $[\mathscr{F}]_{\mathscr{E}}$ de $\mathcal{M}_n(\mathbb{R})$ telle que, pour tout $k \in [\![1;n]\!]$, la $k-$ième colonne de $[\mathscr{F}]_{\mathscr{E}}$ est la représentation du $k-$ième vecteur de $\mathscr{F}$ dans la base $\mathscr{E}$.

→ $[\mathscr{F}]_{\mathscr{E}}[\mathscr{G}]_{\mathscr{F}}$ est la matrice de passage de $\mathscr{E}$ à $\mathscr{G}$: elle vaut $[\mathscr{G}]_{\mathscr{E}}$.

Définition I.2.

Soient $\mathscr{E}$ et $\mathscr{F}$ deux bases de E. On dit que $\mathscr{E}$ et $\mathscr{F}$ ont la même orientation, et on note $\mathscr{E} \sim \mathscr{F}$, lorsque

$$\det_{\mathscr{E}}(\mathscr{F}) := \det([\mathscr{F}]_{\mathscr{E}}) > 0.$$

Proposition I.3.

$\sim$ est une relation d'équivalence possédant exactement deux classes.

Démonstration. Soit $\mathscr{E} = (e_1, \dots, e_n)$ une base de E. On a

$$\det_{\mathscr{E}}(-e_1, e_2, \dots, e_n) = \begin{vmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{vmatrix} = -1 < 0,$$

donc $\mathscr{E} \nsim \underbrace{(-e_1, e_2, \dots, e_n)}_{\text{notée } \tilde{\mathscr{E}}}$. Donc $\sim$ admet au moins deux classes d'équivalence.

Soit $\mathscr{F}$ une base de E.

→ Si $\det_{\mathscr{E}}(\mathscr{F}) > 0$, alors on a $\mathscr{F} \in \overline{\mathscr{E}}$.

→ Sinon, $\det_{\mathscr{E}}(\mathscr{F}) < 0$. Or $[\mathscr{F}]_{\tilde{\mathscr{E}}} = [\mathscr{E}]_{\tilde{\mathscr{E}}}[\mathscr{F}]_{\mathscr{E}} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} [\mathscr{F}]_{\mathscr{E}}$ donc $\det_{\tilde{\mathscr{E}}}(\mathscr{F}) = -\det_{\mathscr{E}}(\mathscr{F}) > 0$,

(cc) BY-NC-SA

i.e. $\mathscr{F} \in \overline{\tilde{\mathscr{E}}}$.

On en déduit que les deux seules classes d'équivalences sont celles de $\mathscr{E}$ et $\tilde{\mathscr{E}}$.
Dans la suite de cette partie, E est associé à l'une de ses bases orthonormées $\mathscr{E}$. On dit alors que E est orienté.

Définition I.4.

Soit $\mathscr{F}$ une base de E. On dit que $\mathscr{F}$ est positive ou directe dans E lorsque $\mathscr{F} \in \overline{\mathscr{E}}$.
Dans le cas contraire, $\mathscr{F}$ est négative ou indirecte.

Proposition I.5.

Soient $\mathscr{F}$ une base orthonormée de E et $(x_1, \dots, x_n) \in \mathrm{E}^n$

→ Si $\mathscr{F}$ est directe, alors $\det_{\mathscr{F}}(x_1, \dots, x_n) = \det_{\mathscr{E}}(x_1, \dots, x_n)$.

→ Si $\mathscr{F}$ est indirecte, alors $\det_{\mathscr{F}}(x_1, \dots, x_n) = -\det_{\mathscr{E}}(x_1, \dots, x_n)$

Démonstration.

→ On pose $\mathrm{P} = [\mathscr{F}]_{\mathscr{E}}$ et, pour tout $i \in [\![1; n]\!]$, $\mathrm{X}_i = [x_i]_{\mathscr{E}}$ et $\mathrm{X}'_i = [x_i]_{\mathscr{F}}$. Ainsi, pour tout $i \in [\![1; n]\!]$, $\mathrm{X}_i = \mathrm{P}\mathrm{X}'_i$. Donc

$$\det_{\mathscr{E}}(x_1, \dots, x_n) = \det(\mathrm{X}_1, \dots, \mathrm{X}_n) = \det(\mathrm{P}\mathrm{X}'_1, \dots, \mathrm{P}\mathrm{X}'_n) = \underbrace{\det(\mathrm{P})}_{=1} \det(\mathrm{X}'_1, \dots, \mathrm{X}'_n) = \det_{\mathscr{F}}(x_1, \dots, x_n),$$

car P est une matrice orthogonale et $\mathscr{F}$ est orientée positivement (relativement à $\mathscr{E}$).

→ Le calcul est identique, seulement $\det(\mathrm{P}) = -1$ car P est orthogonale et orientée négativement (relativement à $\mathscr{E}$).

Proposition I.6.

On suppose que $n = 3$. Soit $(x, y) \in \mathrm{E}^2$. Il existe un unique $w \in \mathrm{E}$ vérifiant la propriété

$$\forall z \in \mathrm{E}, \ \det_{\mathscr{E}}(x, y, z) = \langle w, \ z \rangle .$$

w est le produit vectoriel de x et y et est noté $x \wedge y$.

Démonstration. On considère la forme linéaire $\varphi : z \longmapsto \det_{\mathscr{E}}(x, y, z)$. D'après le théorème à la page 4 du chapitre sur les espaces préhilbertien réels, il existe un unique vecteur w tel que $\varphi = \langle w, \ \cdot \rangle$, ce qu'on voulait.

Remarque. Comme en physique, les coordonnées de w s'obtiennent à partir des coordonnées de x et y :

$$w_1 = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} \qquad w_2 = \begin{vmatrix} x_3 & y_3 \\ x_1 & y_1 \end{vmatrix} \qquad w_3 = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

Exercice I.7.

Soit $(\mathrm{E}, \langle \cdot, \ \cdot \rangle)$ un espace euclidien de dimension 3.
Montrer que pour tous $a, b, c \in \mathrm{E}$

$$a \wedge (b \wedge c) = \langle a, \ c \rangle \, b - \langle a, \ b \rangle \, c.$$

(cc) BY-NC-SA

2. Décomposition de Cartan

On note $\mathcal{T}_n^+(\mathbb{R})$ l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ qui sont triangulaires supérieures et dont les coefficients diagonaux sont strictement positifs.

Proposition I.8.

Soit $A \in GL_n(\mathbb{R})$. Il existe un unique couple $(O, T) \in \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R})$ tel que $A = OT$.

Démonstration.

→ **Unicité**
S'il existe deux tels couples (O_1, T_1) et (O_2, T_2) alors $O_1T_1 = O_2T_2$, donc $O_1^{-1}O_2 = T_1T_2^{-1}$.
$U := O_1^{-1}O_2$ est orthogonale, triangulaire supérieure à coefficients diagonaux strictement positifs, car l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ inversibles et triangulaires supérieures est un sous-groupe de $(GL_n(\mathbb{R}), \times)$.
Or, $U^\top = U^{-1} = (T_1T_2^{-1})^{-1} = T_2T_1^{-1} \in \mathcal{T}_n^+(\mathbb{R})$. Donc $U^\top$ est aussi triangulaire supérieure à coefficients diagonaux strictement positifs.
Donc U est diagonale et à coefficients strictement positifs. On en déduit que $U = I_n$, donc $O_1 = O_2$ et $T_1 = T_2$.

→ **Existence**
Notons $\mathscr{A} = (A_1, \dots, A_n)$ la base de $\mathcal{M}_{n,1}(\mathbb{R})$ donnée par les colonnes de A, et $\mathscr{F}$ une base de $\mathcal{M}_{n,1}(\mathbb{R})$ obtenue à l'aide du procédé de Gram-Schmidt appliqué à $\mathscr{A}$.
Quitte à échanger certains vecteurs de $\mathscr{F}$ par leurs opposés, le théorème de Gram-Schmidt assure que $T := [\mathscr{F}]_{\mathscr{A}}$ est triangulaire supérieure et ses coefficients diagonaux sont strictement positifs.
Posons $O = [\mathscr{F}]_{\mathrm{Can}(\mathcal{M}_{n,1}(\mathbb{R}))}$. Alors, $AT = O$ car $A = [\mathscr{A}]_{\mathrm{Can}(\mathcal{M}_{n,1}(\mathbb{R}))}$. Ainsi $A = OT^{-1}$.

Exercice I.9.

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

$$f : \begin{cases} \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R}) & \longrightarrow GL_n(\mathbb{R}) \\ (O, T) & \longmapsto OT \end{cases}$$

est un homéomorphisme.

Exercice I.10.

Soit $n \in \mathbb{N}^*$. On note $GL_n^+(\mathbb{R})$ l'ensemble des matrices inversibles de déterminant strictement positif.
Montrer que $GL_n^+(\mathbb{R})$ est connexe.

BY-NC-SA

3. Inégalité de Hadamard

Exercice I.11.

Soit $(\mathrm{E}, \langle \cdot\,,\ \cdot \rangle)$ un espace euclidien de dimension n orienté par l'une de ses bases orthonormées $\mathscr{E}$. Soit $(x_1, \ldots, x_n) \in \mathrm{E}^n$. Montrer que

$$\left|\det_{\mathscr{E}}(x_1, \ldots, x_n)\right| \leq \prod_{k=1}^{n} \|x_k\|.$$

Montrer de plus qu'il y a égalité si, et seulement si, $(x_1, \ldots, x_n)$ est orthogonale.

II Matrices de Gram

On rappelle que $(\mathrm{E}, \langle \cdot\,,\ \cdot \rangle)$ un espace préhilbertien réel.

Définition II.1.

Soit $(x_1, \ldots, x_n) \in \mathrm{E}^n$. On appelle matrice de Gram associée à $(x_1, \ldots, x_n)$ la matrice

$$\mathrm{G}(x_1, \ldots, x_n) = (\langle x_i,\ x_j \rangle)_{1 \leq i,j \leq n}.$$

On note aussi $|\mathrm{G}|\,(x_1, \ldots, x_n)$ le déterminant de cette matrice de Gram.

Proposition II.2.

Soit $(x_1, \ldots, x_n) \in \mathrm{E}^n$.

→ Soit $(\varepsilon_1, \ldots, \varepsilon_n)$ un système orthonormé tel que $x_1, \ldots, x_p \in \mathrm{Vect}(\varepsilon_1, \ldots, \varepsilon_n)$.
Notons $\mathrm{A} = [(x_1, \ldots, x_n)]_{(\varepsilon_1, \ldots, \varepsilon_n)}$. Alors,

$$\mathrm{G}(x_1, \ldots, x_n) = \mathrm{A}^\top \mathrm{A}.$$

→ $|\mathrm{G}|\,(x_1, \ldots, x_n) > 0$ si, et seulement si, $(x_1, \ldots, x_n)$ est libre.

→ Notons M la matrice $\mathrm{G}(x_1, \ldots, x_n)$. Pour tout $\Lambda = (\lambda_k)_{1 \leq k \leq n} \in \mathcal{M}_{n,1}(\mathbb{R})$,

$$\|\lambda_1 x_1 + \cdots + \lambda_n x_n\|^2 = \Lambda^\top \mathrm{M} \Lambda.$$

→ Pour tout $\lambda \in \mathrm{Sp}(\mathrm{M})$, $\lambda \geq 0$.

→ On suppose E de dimension finie. Soient $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n) \in \mathrm{E}^n$. On a l'équivalence

$$\mathrm{G}(x_1, \ldots, x_n) = \mathrm{G}(y_1, \ldots, y_n) \iff \exists u \in \mathcal{O}(\mathrm{E}), \quad \forall i \in [\![1; n]\!], \quad u(x_i) = y_i.$$

Démonstration.

→ On note $\mathrm{A}_1, \ldots, \mathrm{A}_n$ les colonnes de A. Alors

$$\mathrm{A}^\top \mathrm{A} = [\mathrm{A}_i^\top \mathrm{A}_j]_{1 \leq i,j \leq n} = (\langle x_i,\ x_j \rangle)_{1 \leq i,j \leq n} = \mathrm{G}(x_1, \ldots, x_n).$$

→ D'après la proposition précédente, $|\mathrm{G}|\,(x_1, \ldots, x_p) = \det(\mathrm{A}^\top \mathrm{A}) = (\det(\mathrm{A}))^2 > 0$ donc $\det(\mathrm{A}) = 0$, d'où l'équivalence.

CC BY-NC-SA

→ On rappelle une identité classique :

$$\forall \mathrm{M} = (m_{i,j})_{1\le i,j\le n} \in \mathcal{M}_n(\mathbb{C}), \forall x = (x_k)_{1\le k\le n} \in \mathcal{M}_{n,1}(\mathbb{C}), \qquad \mathrm{X}^\top \mathrm{M}\mathrm{X} = \sum_{1\le i,j\le n} x_i x_j m_{i,j}.$$

Ainsi $\|\lambda_1 x_1 + \cdots + \lambda_n x_n\|^2 = \sum_{1\le i,j\le n} \lambda_i\lambda_j \langle x_i,\ x_j\rangle = {}^t\Lambda \mathrm{M}\Lambda$.

→ Il existe $\Lambda \in \mathcal{M}_{n,1}(\mathbb{R})$ tel que $\mathrm{M}\Lambda = \lambda\Lambda$, donc $\Lambda^\top \mathrm{M}\Lambda = \lambda\Lambda^\top\Lambda$ et d'après la proposition précédente

$$\|\lambda_1 x_1 + \cdots + \lambda_n x_n\|^2 = \Lambda^\top \mathrm{M}\Lambda = \mu(\lambda_1^2 + \cdots + \lambda_p^2),$$

Donc $\lambda \ge 0$.

→ ($\Leftarrow$) Supposons qu'un tel u existe. Alors on a pour tout $i, j \in [\![1; n]\!]$,

$$\langle y_i,\ y_j\rangle = \langle u(x_i),\ u(x_j)\rangle \underset{\substack{\uparrow\\ u \text{ conserve le produit scalaire}}}{=} \langle x_i,\ x_j\rangle\,,$$

donc les deux matrices de Gram sont égales.
($\Rightarrow$) Remarquons que l'implication est vraie si tous les x_i (et donc tous les y_i) sont nuls. Supposons que les x_i ne sont pas tous nuls.
À l'aide d'une permutation σ de $[\![1; n]\!]$, permutons les x_i de sorte à considérer l'entier naturel non nul $r \le n$ tel que $(x_{\sigma(1)}, \ldots, x_{\sigma(r)})$ soit libre et $x_{\sigma(r+1)}, \ldots, x_{\sigma(n)} \in \underbrace{\mathrm{Vect}(x_{\sigma(1)}, \ldots, x_{\sigma(r)})}_{\text{noté A}}$.

Posons, pour tout $i \in [\![1; r]\!]$, $a_i = x_{\sigma(i)}$ et $b_i = y_{\sigma(i)}$. Alors

(i) D'après la deuxième proposition, $(b_1, \ldots, b_r)$ est libre car $|\mathrm{G}|\,(b_1, \ldots, b_r) = |\mathrm{G}|\,(a_1, \ldots, a_r) > 0$.

(ii) Pour tout entier $k \ge r+1$,

$$|\mathrm{G}|\,(b_1, \ldots, b_r, y_{\sigma(k)}) \underset{\substack{\uparrow\\ \text{égalité des matrices de Gram}}}{=} |\mathrm{G}|\,(a_1, \ldots, a_r, x_{\sigma(k)}) \underset{\substack{\uparrow\\ \text{deuxième proposition}}}{=} 0,$$

donc $y_{\sigma(k)} \in \underbrace{\mathrm{Vect}(b_1, \ldots, b_r)}_{\text{noté B}}$.

Notons p la dimension de E. Si $r < p$, introduisons les systèmes orthonormés $(a_{r+1}, \ldots, a_p)$ et $(b_{r+1}, \ldots, b_p)$ de sorte que les familles $(a_1, \ldots, a_p)$ et $(b_1, \ldots, b_p)$ soient des bases adaptées aux décompositions $\mathrm{E} = \mathrm{A} \oplus \mathrm{A}^\perp$ et $\mathrm{E} = \mathrm{B} \oplus \mathrm{B}^\perp$.
On considère $u \in \mathscr{L}(\mathrm{E})$ tel que pour tout $i \in [\![1; n]\!]$, $u(a_i) = b_i$.
u est une isométrie sur E. En effet, pour tout $x \in \mathrm{E}$, il existe $\Lambda = (\lambda_k)_{1\le k\le n} \in \mathcal{M}_{r,1}(\mathbb{R})$ et $(\lambda_{r+1}, \ldots, \lambda_p) \in \mathbb{R}^{p-r}$ tel que $x = \lambda_1 a_1 + \cdots + \lambda_p a_p$, et

$$\begin{aligned}
\|x\|^2 &= \|\lambda_1 a_1 + \cdots + \lambda_r a_r\|^2 + \sum_{i=r+1}^{n} \lambda_i^2 \|a_i\|^2 && \text{(Pythagore)}\\
&= \Lambda^\top \mathrm{G}(a_1, \ldots, a_r)\Lambda + \sum_{i=r+1}^{n} \lambda_i^2 \|b_i\|^2 && \text{(troisième proposition)}\\
&= \Lambda^\top \mathrm{G}(b_1, \ldots, b_r)\Lambda + \|\lambda_{r+1}u(a_{r+1}) + \cdots + \lambda_p u(a_p)\|^2 &&\\
&= \|\lambda_1 b_1 + \cdots + \lambda_r b_r\|^2 + \|\lambda_{r+1}u(a_{r+1}) + \cdots + \lambda_p u(a_p)\|^2 && \text{(troisième proposition)}\\
&= \|u(x)\|^2. && \text{(Pythagore)}
\end{aligned}$$

BY-NC-SA

Par ailleurs, pour tout $k \in [\![r+1;n]\!]$, et pour tout $i \in [\![1;r]\!]$

$$\left\langle u(x_{\sigma(k)}) - y_{\sigma(k)},\ b_i\right\rangle = \left\langle u(x_{\sigma(k)}),\ b_i\right\rangle - \left\langle y_{\sigma(k)},\ b_i\right\rangle \underset{\substack{\uparrow \\ u \text{ conserve le produit scalaire}}}{=} \left\langle x_{\sigma(k)},\ a_i\right\rangle - \left\langle y_{\sigma(k)},\ b_i\right\rangle \underset{\substack{\uparrow \\ \text{égalité des matrices de Gram}}}{=} 0.$$

Ainsi, pour tout $k \in [\![r+1;n]\!]$, $u(x_{\sigma(k)}) - y_{\sigma(k)} \perp \mathrm{B}$ et $u(x_{\sigma(k)}) - y_{\sigma(k)} \in \mathrm{B}$, donc $u(x_{\sigma(k)}) - y_{\sigma(k)} = 0$.
Donc, pour tout $i \in [\![1;n]\!]$ $u(x_i) = y_i$.

Exercice II.3.

Soit $(\mathrm{E}, \langle \cdot\,,\ \cdot \rangle)$ un espace préhilbertien réel. On suppose que $(x_1, \ldots, x_n)$ est libre dans E et on pose $\mathrm{F} = \mathrm{Vect}(x_1, \ldots, x_n)$. Soit $a \in \mathrm{E}$.
Montrer que

$$d(a, \mathrm{F})^2 = \frac{|\mathrm{G}|\,(a, x_1, \ldots, x_p)}{|\mathrm{G}|\,(x_1, \ldots, x_p)}.$$

(cc) BY-NC-SA

Correction de l'exercice I.7. :

Soit $\mathscr{E}$ une base orthonormée de E qui détermine le sens direct. On note (a_1, a_2, a_3), (b_1, b_2, b_3), et (c_1, c_2, c_3) les coordonnées respectives de a, b et c dans cette base.

Calculons : $[b \wedge c]_{\mathscr{E}} = \begin{pmatrix} b_2c_3 - b_3c_2 \\ b_3c_1 - b_1c_3 \\ b_1c_2 - b_2c_1 \end{pmatrix}$, puis $[a \wedge (b \wedge c)]_{\mathscr{E}} = \begin{pmatrix} a_2(b_1c_2 - b_2c_1) - a_3(b_3c_1 - b_1c_3) \\ a_3(b_2c_3 - b_3c_2) - a_1(b_1c_2 - b_2c_1) \\ a_1(b_3c_1 - b_1c_3) - a_2(b_2c_3 - b_3c_2) \end{pmatrix}$.

Par ailleurs : $[\langle a,\, c\rangle\, b - \langle a,\, b\rangle\, c]_{\mathscr{E}} = \begin{pmatrix} (\cancel{a_1c_1} + a_2c_2 + a_3c_3)b_1 - (\cancel{a_1b_1} + a_2b_2 + a_3b_3)c_1 \\ (a_1c_1 + \cancel{a_2c_2} + a_3c_3)b_2 - (a_1b_1 + \cancel{a_2b_2} + a_3b_3)c_2 \\ (a_1c_1 + a_2c_2 + \cancel{a_3c_3})b_3 - (a_1b_1 + a_2b_2 + \cancel{a_3b_3})c_3 \end{pmatrix}$.

En factorisant par le couple de coordonnées de a qui apparaît dans chaque ligne, ou en développant chaque ligne des deux résultats, on obtient que les deux colonnes sont égales.

Correction de l'exercice I.9. :

Il s'agit de montrer que :

- $\to$ f est continue : c'est vrai par continuité du produit matriciel.
- $\to$ f est bijective : c'est vrai d'après la décomposition de Cartan.
- $\to$ f^{-1} est continue : c'est le plat de résistance de l'exercice.

Soit $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$.

Considérons une suite $(\mathrm{M}_m)_{m\in\mathbb{N}}$ à valeurs dans $\mathrm{GL}_n(\mathbb{R})$ convergeant vers $\mathrm{M} \in \mathrm{GL}_n(\mathbb{R})$, et montrons que $f^{-1}(\mathrm{M}_m) \xrightarrow[m\to+\infty]{} f^{-1}(\mathrm{M})$.

Introduisons la décomposition de Cartan $(\mathrm{O}, \mathrm{T}) \in \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R})$ de M, et pour tout $m \in \mathbb{N}$, la décomposition de Cartan $(\mathrm{O}_m, \mathrm{T}_m) \in \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R})$ de M_m.

Rappelons que $\mathcal{O}_n(\mathbb{R})$ est compact : il est fermé (c'est l'image réciproque du fermé $\{\mathrm{I}_n\}$ par l'application continue $\mathrm{X} \longmapsto \mathrm{X}^\top\mathrm{X}$) et borné car tous les coefficients d'une matrice orthogonale sont majorés par 1 en valeur absolue.

Ainsi, de $(\mathrm{O}_m)_{m\in\mathbb{N}}$ on peut extraire une sous-suite $\left(\mathrm{O}_{\varphi(m)}\right)_{m\in\mathbb{N}}$ convergeant vers $\tilde{\mathrm{O}} \in \mathcal{O}_n(\mathbb{R})$.

Pour tout $m \in \mathbb{N}$, $\mathrm{T}_{\varphi(m)} = \left(\mathrm{O}_{\varphi(m)}\right)^{-1} \mathrm{M}_{\varphi(m)} \xrightarrow[m\to+\infty]{} \tilde{\mathrm{O}}^{-1}\mathrm{M}$.

Or $\tilde{\mathrm{T}} := \tilde{\mathrm{O}}^{-1}\mathrm{M}$ est aussi triangulaire supérieure à coefficients diagonaux positifs, car l'ensemble des matrices triangulaires supérieures à coefficients diagonaux positifs est fermé.

Ainsi, $\mathrm{M} = \tilde{\mathrm{O}}\tilde{\mathrm{T}}$ et $(\tilde{\mathrm{O}}, \tilde{\mathrm{M}}) \in \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R})$. Par unicité de la décomposition de Cartan de M, $\mathrm{O} = \tilde{\mathrm{O}}$ et $\mathrm{T} = \tilde{\mathrm{T}}$.

Remarquons alors que toute extraction convergente de la suite $(\mathrm{O}_m)_{m\in\mathbb{N}}$ exhibe une unique valeur d'adhérence pour cette suite, et cette valeur est O.

Cette suite est à valeurs dans un compact et admet une unique valeur d'adhérence, c'est donc une suite convergente de limite O.

Ainsi, $\mathrm{T}_m = \mathrm{O}_m^{-1}\mathrm{M}_m \xrightarrow[m\to+\infty]{} \mathrm{O}^{-1}\mathrm{M} = \mathrm{T}$. Donc,

$$f^{-1}(\mathrm{M}_m) = (\mathrm{O}_m, \mathrm{T}_m) \xrightarrow[m\to+\infty]{} (\mathrm{O}, \mathrm{T}) = f^{-1}(\mathrm{M}),$$

ce qui prouve la continuité de f^{-1}.

Correction de l'exercice I.10. :

L'application

$$f : \begin{cases} \mathcal{SO}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R}) & \longrightarrow \mathrm{GL}_n^+(\mathbb{R}) \\ (\mathrm{O}, \mathrm{T}) & \longmapsto \mathrm{OT} \end{cases}$$

est continue (car le produit matriciel l'est) et surjective : pour toute $\mathrm{M} \in \mathrm{GL}_n^+(\mathbb{R})$, la décomposition de

(cc) BY-NC-SA

Cartan donne un couple $(\mathrm{O}, \mathrm{T}) \in \mathcal{O}_n(\mathbb{R}) \times \mathcal{T}_n^+(\mathbb{R})$ tel que $\mathrm{OT} = \mathrm{M}$, mais en fait, $\mathrm{O} \in \mathcal{SO}_n(\mathbb{R})$ car sinon, le déterminant de M serait négatif, celui de T étant positif.
Par ailleurs, $\mathcal{SO}_n(\mathbb{R})$ est connexe (*cf.* chapitre 35) et $\mathcal{T}_n^+(\mathbb{R})$ est connexe car convexe, donc leur produit cartésien est connexe.
$\mathrm{GL}_n^+(\mathbb{R})$ est donc connexe en tant qu'image d'un connexe par une fonction continue.

Correction de l'exercice I.11. :

On pose $\mathrm{M} = [x_1, \ldots, x_n]_{\mathscr{E}}$. Remarquons que les colonnes de M sont les colonnes X_i représentatives des vecteurs x_i.

- → Si $(\mathrm{X}_1, \ldots, \mathrm{X}_n)$ est liée, $\det \mathrm{M} = 0$ donc l'inégalité est vraie.
- → Sinon, remarquons que M est la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ à $\mathscr{X} := (\mathrm{X}_1, \ldots, \mathrm{X}_n)$.
  Notons $\mathscr{U} = (\mathrm{U}_1, \ldots, \mathrm{U}_n)$ une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ donnée par le procédé de Gram-Schmidt, puis T la matrice de passage de $\mathscr{U}$ à $\mathscr{X}$. D'après le théorème de Gram-Schmidt, T est triangulaire supérieure. Ses colonnes étant les représentations des vecteurs de $\mathscr{X}$ dans $\mathscr{U}$ qui est orthonormée, pour tout $i \in [\![1; n]\!]$, le i-ème coefficient diagonal de T est $\mathrm{U}_i^\top \mathrm{X}_i$.
  Ainsi, en notant U la matrice de passage de la base canonique de $\mathcal{M}_{n,1}(\mathbb{R})$ à $\mathscr{U}$,

$$|\det(\mathrm{M})| = |\det(\mathrm{UT})| = |\det(\mathrm{T})| = \prod_{k=1}^n \left|\mathrm{U}_k^\mathrm{T} \mathrm{X}_k\right| \underset{\substack{\uparrow \\ \text{Cauchy-Schwarz}}}{\leq} \prod_{k=1}^n \sqrt{\mathrm{X}_k^\top \mathrm{X}_k}\, \underbrace{\sqrt{\mathrm{U}_k^\top \mathrm{U}_k}}_{=1} = \prod_{k=1}^n \sqrt{\mathrm{X}_k^\top \mathrm{X}_k}.$$

Correction de l'exercice II.3. :

Introduisons la décomposition $a = u + v$ adaptée à $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^\perp$.

$$\begin{aligned}
|\mathrm{G}|\,(a, x_1, \ldots, x_p) &= \begin{vmatrix} \langle a,\ a\rangle & \overbrace{\langle a,\ x_1\rangle \quad \ldots \quad \langle a,\ x_p\rangle}^{\text{ligne notée L}} \\ \langle x_1,\ a\rangle & \langle x_1,\ x_1\rangle \quad \ldots \quad \langle x_1,\ x_p\rangle \\ \vdots & \vdots \qquad\quad \ddots \qquad\quad \vdots \\ \langle x_p,\ a\rangle & \langle x_p,\ x_1\rangle \quad \ldots \quad \langle x_p,\ x_p\rangle \end{vmatrix} \\
&= \begin{vmatrix} \|u\|^2 + \|v\|^2 & \mathrm{L} \\ \mathrm{L}^\top & \mathrm{G}(x_1, \ldots, x_n) \end{vmatrix} \\
&= \begin{vmatrix} \|v\|^2 & \mathrm{L} \\ 0 & \mathrm{G}(x_1, \ldots, x_n) \end{vmatrix} + \begin{vmatrix} \|u\|^2 & \mathrm{L} \\ \mathrm{L}^\top & \mathrm{G}(x_1, \ldots, x_n) \end{vmatrix} \\
&= \|v\|^2\, |\mathrm{G}|\,(x_1, \ldots, x_p) + \underbrace{|\mathrm{G}|\,(u, x_1, \ldots, x_p)}_{=0 \text{ car } (u, x_1, \ldots, x_p) \text{ est liée}} \\
&= d(a, \mathrm{F})^2\, |\mathrm{G}|\,(x_1, \ldots, x_p).
\end{aligned}$$

(cc) BY-NC-SA

Chapitre 36

Espaces préhilbertiens complexes

Dans tout ce chapitre, E désigne un $\mathbb{C}$-espace vectoriel non trivial et n désigne un entier supérieur ou égal à 1.

I Espaces hermitiens

Définition I.1.

Une application $\langle \cdot\,,\ \cdot \rangle : \mathrm{E}^2 \longrightarrow \mathbb{C}$ est un produit scalaire hermitien si

- → Pour tout $x \in \mathrm{E}$, $y \longmapsto \langle x,\ y\rangle$ est $\mathbb{C}$-linéaire.
- → Pour tout $y \in \mathrm{E}$, $x \longmapsto \langle x,\ y\rangle$ est semi-linéaire, *i.e.* pour tout $\lambda \in \mathbb{C}$ et pour tout $x \in \mathrm{E}$,

$$\langle \lambda x,\ y\rangle = \overline{\lambda}\,\langle x,\ y\rangle\,.$$

- → Pour tout $(x, y) \in \mathrm{E}^2$, $\langle x,\ y\rangle = \overline{\langle y,\ x\rangle}$
- → Pour tout $x \in \mathrm{E}\setminus\{0\}$, $\langle x,\ x\rangle \in \mathbb{R}^{+*}$

Dorénavant, E est muni d'un produit scalaire hermitien $\langle \cdot\,,\ \cdot \rangle$. On dit que $(\mathrm{E}, \langle \cdot\,,\ \cdot \rangle)$ est un espace préhilbertien complexe ou encore un espace hermitien. On notera également $\|\cdot\|$ la norme associée à ce produit scalaire : pour tout $x \in \mathrm{E}$, $\|x\| = \sqrt{\langle x,\ x\rangle}$.

Identités remarquables I.2.

Les identités remarquables sur un espace préhilbertien complexe sont différentes de celles sur un espace préhilbertien réel. Si la norme est toujours réelle, le produit scalaire peut ne pas l'être. Ainsi, pour tous $x, y \in \mathrm{E}$,

$$\|x+y\|^2 = \|x\|^2 + 2\,\mathrm{Re}(\langle x,\ y\rangle) + \|y\|^2 \qquad \|x-iy\|^2 = \|x\|^2 + 2\,\mathrm{Im}(\langle x,\ y\rangle) + \|y\|^2\,,$$

$$\|x-y\|^2 = \|x\|^2 - 2\,\mathrm{Re}(\langle x,\ y\rangle) + \|y\|^2 \qquad \|x+iy\|^2 = \|x\|^2 - 2\,\mathrm{Im}(\langle x,\ y\rangle) + \|y\|^2\,,$$

$$\langle x,\ y\rangle \underset{\substack{\uparrow \\ \text{identité de polarisation}}}{=} \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i\left(\|x-iy\|^2 - \|x+iy\|^2\right)\right).$$

Quelques produits scalaires usuels

- → $\mathrm{E} = \mathbb{C}$, et $\langle \cdot\,,\ \cdot \rangle : (z, w) \longmapsto \overline{z}w$.
- → $\mathrm{E} = \mathbb{C}^n$, $\langle \cdot\,,\ \cdot \rangle : (z, w) \longmapsto \displaystyle\sum_{k=0}^{n} \overline{x_k} y_k$.
- → $\mathrm{E} = \mathcal{M}_n(\mathbb{C})$, $\langle \cdot\,,\ \cdot \rangle : (\mathrm{A}, \mathrm{B}) \longmapsto \mathrm{Tr}({}^t\overline{\mathrm{A}}\mathrm{B})$.
- → $\mathrm{E} = \mathscr{C}_{2\pi}(\mathbb{R}, \mathbb{C})$, $\langle \cdot\,,\ \cdot \rangle : (f, g) \mapsto \displaystyle\int_0^{2\pi} \overline{f} g$ où $\mathscr{C}_{2\pi}(\mathbb{R}, \mathbb{C})$ est le sous-espace de $\mathscr{C}^0(\mathbb{R}, \mathbb{C})$ ne contenant que des fonctions 2π-périodiques.

(cc) BY-NC-SA

Proposition I.3.

L'inégalité de Cauchy-Schwarz est valide sur l'espace préhilbertien complexe E :

$$\forall (x, y) \in \mathrm{E}, \quad |\langle x,\ y\rangle| \leq \|x\|^2 \|y\|^2 .$$

Démonstration. On suppose que $y \neq 0$. Soit $\theta \in \mathbb{R}$ vérifiant $e^{i\theta} \langle x,\ y\rangle = |\langle x,\ y\rangle|$. Pour tout $t \in \mathbb{R}$, posons

$$\varphi(t) = \left\| x + te^{i\theta} y \right\|^2 .$$

Alors, pour tout $t \in \mathbb{R}$,

$$\varphi(t) = \|x\|^2 + 2\,\mathrm{Re}\left\langle x,\ te^{i\theta} y\right\rangle + t^2 \|y\|^2 = \|x\|^2 + 2t\,|\langle x,\ y\rangle| + t^2 \|y\|^2 .$$

Ainsi, φ est une fonction polynomiale de degré 2 et positive, donc son discriminant est négatif ou nul. Or, celui-ci vaut $4\,|\langle x,\ y\rangle|^2 - 4\,\|x\|^2 \|y\|^2$, ce qu'on voulait.

Nombre de notions relatives aux espaces préhilbertiens réels définies en première année et au chapitre 34 sont toujours valables sur un espace préhilbertien complexe. Plus précisément :

→ La notion d'orthogonalité se définit de la même manière sur E que sur un espace préhilbertien réel : produit scalaire nul. On peut donc considérer des familles orthogonales ou orthonormées de E, ainsi que l'orthogonal d'une partie de E ou deux parties orthogonales de E, de même qu'un vecteur normal à un hyperplan de E.

→ L'inégalité de Minkowski, le théorème de Pythagore et l'égalité de la médiane sont toujours valides sur E.

→ Si F est un sous-espace vectoriel de dimension finie de E, on a toujours $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^{\perp}$.

→ E possède des bases orthonormées.

→ Les formules de projections orthogonales sont toujours valables.

→ L'inégalité de Bessel est toujours valable.

→ Le procédé de Gram-Schmidt appliqué à une base de E donne une base orthonormée de E.

Avertissement I.4.

Soit F un sous-espace vectoriel de E de dimension finie n et $(e_1, \ldots, e_n)$ une base orthonormée de F. Soit $x \in \mathrm{E}$. Le projeté de x sur F s'écrit toujours

$$\sum_{k=1}^{n} \langle e_k,\ x\rangle\, e_k .$$

Le **vecteur à projeter** est dans la **case de droite** du produit scalaire hermitien. Tout fonctionne identiquement au cas réel pour les projections, le procédé de Gram-Schmidt et l'inégalité de Bessel à condition de ne pas négliger ce petit détail.

Remarque I.5.

On suppose que E est de dimension finie n.
Soit $(e_1, \ldots, e_n)$ une base orthonormée de E. Soient $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Considérons l'hyperplan de E défini par $\mathrm{H} = \left\{ x := \sum_{k=1}^{n} x_k e_k \;\middle|\; \lambda_1 x_1 + \cdots + \lambda_n x_n = 0 \right\}$. Alors $y := \sum_{k=1}^{n} \overline{\lambda_k} e_k$ est un vecteur normal à H.

(cc) BY-NC-SA

Exercice I.6.

Soit G un groupe fini d'ordre n. On pose $\mathrm{E} = \mathscr{F}(\mathrm{G}, \mathbb{C})$.
On considère le produit scalaire hermitien sur E vérifiant, pour tous x et y dans E,

$$\langle f,\ g\rangle = \frac{1}{|\mathrm{G}|}\sum_{a\in \mathrm{G}} \overline{f(a)}g(a).$$

On dit que χ est un caractère de G si $\chi \in \mathrm{Hom}(\mathrm{G}, \mathbb{C}^*)$.

1. Montrer que tout caractère de G est à valeurs dans $\mathbb{U}_n$.
2. Montrer que si Φ est un caractère non constant, alors $\displaystyle\sum_{a\in \mathrm{G}} \Phi(a) = 0$.
3. Montrer que si χ et ψ sont deux caractères distincts, alors ils sont orthogonaux et que $\|\chi\|^2 = 1$.

II Opérations unitaires

Définition II.1.

Soit $u \in \mathscr{L}(\mathrm{E})$. On dit que u est unitaire lorsque pour tout $x \in \mathrm{E}$,

$$\|u(x)\| = \|x\|\,.$$

On note $\mathcal{U}(\mathrm{E})$ l'ensemble de tels endomorphismes.

On dit qu'un endomorphisme unitaire u de E préserve la norme. De plus, en vertu de l'identité de polarisation, il préserve aussi le produit scalaire *i.e.* pour tous $x, y \in \mathrm{E}$,

$$\langle u(x),\ u(y)\rangle = \langle x,\ y\rangle\,.$$

Proposition II.2.

Soit $u \in \mathcal{U}(\mathrm{E})$ et $v \in \mathscr{L}(\mathrm{E})$.

→ $(\mathcal{U}(\mathrm{E}), \circ)$ est un sous-groupe de $\mathrm{GL}(\mathrm{E})$.

→ Si E est de dimension finie n et $(e_1, ..., e_n)$ est une base orthonormée de E, alors v est unitaire si, et seulement si, $(v(e_1), \dots, v(e_n))$ est une base orthonormée de E.

→ F est stable par u si, et seulement si, $\mathrm{F}^\perp$ est stable par u

→ v est unitaire si, et seulement si, v est diagonalisable en base orthonormée et $\mathrm{Sp}(u) \subset \mathbb{U}$.

Notation. Soit $\mathrm{U} \in \mathcal{M}_n(\mathbb{C})$. On note U^*, et on appelle matrice adjointe, la matrice $\overline{\mathrm{U}}^\top$.

Définition II.3.

Soit $\mathrm{U} \in \mathcal{M}_n(\mathbb{C})$. On dit que U est unitaire si $\mathrm{UU}^* = \mathrm{I}_n$.

Remarque. Remarquons alors que U est inversible et que l'on a aussi $\mathrm{U}^*\mathrm{U} = \mathrm{I}_n$. Ces deux égalités sont en fait équivalentes.

(cc) BY-NC-SA

Proposition II.4.

On suppose que E est de dimension finie n. Soit $\mathrm{U} \in \mathcal{M}_n(\mathbb{C})$. Les propositions suivantes sont équivalentes.

(1) U est unitaire.

(2) Les colonnes de U forment une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$.

(3) U est la matrice d'un endomorphisme $u \in \mathcal{U}(\mathrm{E})$ dans une base orthonormée.

Démonstration.

→ **(1) ⇒ (2)**
Pour tout $k \in [\![1;n]\!]$, on note C_k la k-ième colonne de U.
Pour tous $k, j \in [\![1;n]\!]$, le coefficient à la ligne i et à la colonne j dans $\mathrm{U}^*\mathrm{U}$ est égal à $\mathrm{C}_k^*\mathrm{C}_j$.
Or, $\mathrm{C}_k^*\mathrm{C}_j = \delta_{k,j}$, donc les colonnes de U forment bien une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$.

→ **(2) ⇒ (3)**
Considérons une base orthonormée $(e_1, ..., e_n)$ de E et définissons $u \in \mathscr{L}(\mathrm{E})$ tel que sa matrice relativement à cette base soit U.
Soit $x \in \mathrm{E}$ et X sa colonne représentative relativement à cette même base. Alors

$$\|u(x)\|^2 = \mathrm{X}^* \underbrace{\mathrm{U}^*\mathrm{U}}_{=\mathrm{I}_n} \mathrm{X} = \mathrm{X}^*\mathrm{X} = \|x\|^2 ,$$

donc $u \in \mathcal{U}(\mathrm{E})$.

→ **(3) ⇒ (1)**
Rappelons que u préserve aussi le produit scalaire. Ainsi, pour tous $\mathrm{X}, \mathrm{Y} \in \mathcal{M}_{n,1}(\mathbb{C})$,

$$\mathrm{X}^*\mathrm{U}^*\mathrm{U}\mathrm{Y} = \mathrm{X}^*\mathrm{Y}. \qquad (\star)$$

Rappelons que pour tous $k, j \in [\![1;n]\!]$ et pour toute $\mathrm{M} \in \mathcal{M}_n(\mathbb{C})$,

$$\mathrm{E}_k^\top \mathrm{M} \mathrm{E}_j = \mathrm{M}[k,j] \qquad \text{et} \qquad \mathrm{E}_k^\top \mathrm{E}_j = \delta_{k,j}.$$

Ainsi, en évaluant l'identité $(\star)$ en chaque E_k, il vient que $\mathrm{U}^*\mathrm{U} = \mathrm{I}_n$. Donc U est unitaire.

(cc) BY-NC-SA

Correction de l'exercice I.6. :

1. Soit χ un caractère de G. Soit $a \in$ G. Le théorème de Lagrange assure que l'ordre de a divise n, donc $a^n = e$, donc $\chi(a)^n = \chi(a^n) = 1$. Donc χ est à valeurs dans $\mathbb{U}_n$.
2. Soit Φ un caractère non constant. Il existe $b \in$ G tel que $\Phi(b) \neq 1$.
   L'application $a \longmapsto ab$ est une bijection de G dans G, donc
   $$\sum_{a\in \mathrm{G}} \Phi(a) = \sum_{a\in \mathrm{G}} \Phi(ab) = \Phi(b) \sum_{a\in \mathrm{G}} \Phi(a),$$
   donc $\displaystyle\sum_{a\in \mathrm{G}} \Phi(a) = 0$.
3. Soient χ et ψ deux caractères distincts. Calculons leur produit scalaire :
   $$\langle \psi,\ \chi\rangle = \frac{1}{|\mathrm{G}|}\sum_{a\in \mathrm{G}} \overline{\psi(a)}\chi(a) \underset{\substack{\uparrow \\ \psi(a)\,\in\,\mathbb{U}_n}}{=} \frac{1}{|\mathrm{G}|}\sum_{a\in \mathrm{G}} \frac{\chi}{\psi}(a) \underset{\substack{\uparrow \\ \frac{\chi}{\psi} \text{ est un caractère}}}{=} 0.$$
   De plus,
   $$\langle \chi,\ \chi\rangle = \frac{1}{|\mathrm{G}|}\sum_{a\in \mathrm{G}} |\chi(a)|^2 \underset{\substack{\uparrow \\ \chi(a)\,\in\,\mathbb{U}_n}}{=} \frac{1}{|\mathrm{G}|}\sum_{a\in \mathrm{G}} 1 = 1.$$

(cc) BY-NC-SA

Chapitre 38

Calcul différentiel

Soit $\mathrm{E}, \mathrm{F}, \mathrm{G}$ des $\mathbb{R}$-espaces vectoriels normés de **dimension finie** munis de produits scalaires et normes respectives $\langle\ ,\ \rangle_{\mathrm{E}}$, $\langle\ ,\ \rangle_{\mathrm{F}}$, $\langle\ ,\ \rangle_{\mathrm{G}}$, $\|.\|_{\mathrm{E}}$, $\|.\|_{\mathrm{F}}$, $\|.\|_{\mathrm{G}}$. Lorsqu'il n'y a pas ambiguïté sur l'espace où l'on travaille, on notera la norme et le produit scalaire tout simplement $\|.\|$ et $\langle\ ,\ \rangle$. Soit Ω un ouvert de E.

I Applications linéaires et hyperplans

On commence par voir quelques éléments qui nous seront utiles par la suite.

Proposition I.1.

Soit $a \in \mathrm{E}$ et soit f l'application

$$f : \begin{cases} \mathrm{E} & \longrightarrow \mathbb{R} \\ x & \longmapsto \langle a,\ x\rangle\,. \end{cases}$$

L'ensemble $\mathrm{H}(f) = \{(x, f(x)),\ x \in \mathrm{E}\}$ est un hyperplan de l'espace vectoriel $\mathrm{E} \times \mathbb{R}$.

Démonstration. On a pour tout $(u, v) \in \mathrm{E} \times \mathbb{R}$,

$$(u, v) \in \mathrm{H}(f) \iff v = f(u) \iff v - \langle a,\ u\rangle = 0.$$

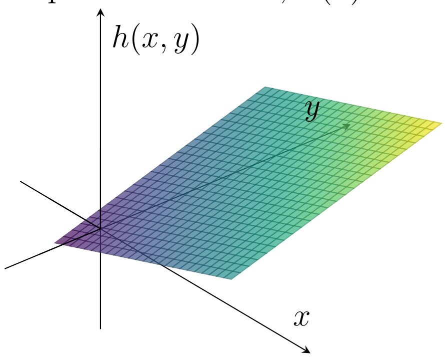
L'application $\psi : (u, v) \longmapsto v - \langle a,\ u\rangle$ est une forme linéaire de $\mathbb{E} \times \mathbb{R}$. L'équivalence précédente nous donne que $\mathrm{H} = \operatorname{Ker}\psi$, *i.e.* H est le noyau d'une forme linéaire, ce qui signifie que H est un hyperplan de $\mathrm{E} \times \mathbb{R}$. □

Exemple. Observons les deux exemples suivants.

Considérons le cas où $\mathrm{E} = \mathbb{R}^2$, et l'application

$$h : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x, y) & \longmapsto \langle (x, y),\ (1, 1)\rangle = x + y. \end{cases}$$

La surface représentative de h, $\mathrm{H}(h)$ est la suivante.

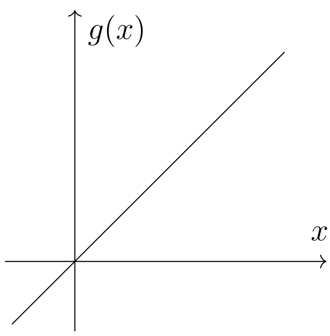


On voit bien ici que l'ensemble $\mathrm{H}(h)$ est un plan (hyperplan en dimension 3) de $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$.

Considérons le cas où $\mathrm{E} = \mathbb{R}$, et l'application

$$g : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto 2x. \end{cases}$$

La courbe représentative de g, $\mathrm{H}(g)$ est la suivante.



On voit bien ici que l'ensemble $\mathrm{H}(g)$ est une droite (hyperplan en dimension 2) de $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$.

Remarque. Bien entendu, dans la proposition I.1, si f était de la forme $x \longmapsto b + \langle a,\ x\rangle$ avec $b \in \mathbb{R}$, $\mathrm{H}(f)$ serait un hyperplan affine de $\mathrm{E} \times \mathbb{R}$.

BY-NC-SA

II Différentiabilité

Dans cette partie, on suppose que f est une application de Ω dans E.

Définition II.1.

Soit $g : \Omega \longrightarrow \mathrm{F}$ et $h : \Omega \longmapsto \mathrm{F}$ deux applications de Ω dans F et $a \in \Omega$. On dit que h est dominée par g au voisinage de a lorsque

$$\exists \eta > 0,\ \exists \mathrm{C} \geq 0,\ \forall x \in \Omega,\ x \in \mathrm{B}(a, \eta) \Longrightarrow \|h(x)\|_{\mathrm{F}} \leq \mathrm{C}\,\|g(x)\|_{\mathrm{F}}\,.$$

Dans ce cas, on note $h(x) \underset{x \to a}{=} \mathrm{O}(g(x))$. De même, on dit que h est négligeable devant g au voisinage de a lorsque

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in \Omega,\ x \in \mathrm{B}(a, \eta) \Longrightarrow \|h(x)\|_{\mathrm{F}} \leq \varepsilon\,\|g(x)\|_{\mathrm{F}}\,.$$

Dans ce cas, on note $h(x) \underset{x \to a}{=} o(g(x))$. Lorsqu'il n'y a pas d'ambiguïté sur le point a, on note tout simplement dans les deux cas respectivement $h(x) = \mathrm{O}(g(x))$ et $h(x) = o(g(x))$.

Remarques.

→ Bien entendu, on a

$$\begin{aligned} h(x) \underset{x \to a}{=} o(g(x)) &\Longleftrightarrow \frac{\|h(x)\|_{\mathrm{F}}}{\|g(x)\|_{\mathrm{F}}} \underset{x \to a}{\longrightarrow} 0 \text{ lorsque } g \text{ est non nulle sur un voisinage de } a \\ &\Longleftrightarrow \exists \varepsilon \in \mathrm{F}^{\mathrm{E}},\ \exists \eta > 0, \forall x \in \mathrm{B_E}(a, \eta),\ h(x) = \varepsilon(x)\,\|g(x)\|_{\mathrm{F}}\,,\ \text{et } \varepsilon(x) \underset{x \to a}{\longrightarrow} 0, \end{aligned}$$

où, $\mathrm{B_E}(a, \eta)$ est la boule de rayon η centrée en a pour la distance $(x, y) \longmapsto \|x - y\|_{\mathrm{E}}$. Remarquons que lorsque $g(x) = x - a$ pour tout x, on n'a pas besoin que x soit dans un voisinage de a. En effet, la condition imposant que x soit dans un voisinage de a nous permet d'éviter le cas où $h(x) \neq 0$ et $g(x) = 0$, ce qui donne $h(x) = 0$. Dans le cas où $g(x) = x - a$, g est toujours non nulle (excepté en a où par définition $h(a) = 0$ et donc on définit $\varepsilon(a) = 0$), et donc on peut définir $\varepsilon(x)$ pour tout $x \neq a$ comme $\dfrac{h(x)}{\|g(x)\|_{\mathrm{F}}}$.

→ En dimension finie, on peut remplacer les normes dans la définition par n'importe quelles normes sur E et sur F, car toutes les normes sont équivalentes en dimension finie (voir chapitre 11.4).

→ Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction réelle et $a \in \mathbb{R}$. On sait que g est dérivable en a si et seulement s'il existe $\ell \in \mathbb{R}$ tel que

$$\frac{g(x) - g(a)}{x - a} \underset{x \to a}{\longrightarrow} \ell \Longleftrightarrow g(x) \underset{x \to a}{=} g(a) + \ell \times (x - a) + o(x - a).$$

On veut étendre cette définition de dérivée aux fonctions dont l'espace de départ est un espace vectoriel de dimension supérieure à 1. La définition suivant cette remarque nous donne un moyen de le faire.

(cc) BY-NC-SA

Définition II.2.

Soit $a \in \Omega$. On dit que f est différentiable en a si et seulement s'il existe une application linéaire $\varphi \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ telle que

$$f(x) \underset{x \to a}{=} f(a) + \varphi(x-a) + o(x-a),$$

ou d'une manière équivalente

$$f(a+h) \underset{h \to 0}{=} f(a) + \varphi(h) + o(h).$$

Lorsque φ existe, il s'agit de la seule application linéaire qui vérifie la propriété ci-dessus. on dit que φ est la différentielle de f en a, et on la note df_a.

Démonstration. Montrons l'unicité de φ. Supposons qu'il existe φ et ψ vérifiant la même propriété. Dans ce cas, on a

$$f(a) + \varphi(h) + o(h) = f(a+h) = f(a) + \psi(h) + o(h),$$

et donc $(\varphi - \psi)(h) = o(h)$. Posons $v = \varphi - \psi$. v est clairement une application linéaire de E dans F, et on a $v\left(\dfrac{h}{\|h\|_{\mathrm{F}}}\right) \xrightarrow[h \to 0]{} 0$. Soit $x \in \mathrm{E} \setminus \{0\}$. Posons $x_n = \dfrac{x}{n+1}$. On a, étant donné que $x_n \xrightarrow[n \to +\infty]{} 0$,

$$v\left(\frac{x}{\|x\|}\right) = v\left(\frac{x_n}{\|x_n\|}\right) \xrightarrow[n \to +\infty]{} 0$$

et donc $v(x) = 0$. On en déduit donc que v et nulle sur E et que finalement $\varphi = \psi$.

□

Remarque.

→ En dimension 1, i.e. lorsque $\Omega \subset \mathbb{R}$, la définition II.2 est cohérente avec celle de la dérivée. En effet, si $f : \Omega \longrightarrow \mathbb{R}$ est différentiable en a, on écrit

$$f(x) \underset{x \to a}{=} f(a) + \varphi(x-a) + o(x-a). \tag{19}$$

φ est une forme linéaire de $\mathbb{R}$, donc il existe $u \in \mathbb{R}$ tel que pour tout $x \in \mathbb{R}$, $\varphi(x) = ux$ (le lecteur non convaincu de cette affirmation peut la montrer lui-même). On en déduit donc que l'égalité 19 s'écrit

$$f(x) \underset{x \to a}{=} f(a) + u \times (x-a) + o(x-a) \quad i.e. \quad \frac{f(x) - f(a)}{x - a} \xrightarrow[x \to a]{} u.$$

On en déduit que la différentiabilité est équivalente à la dérivabilité sur $\mathbb{R}$, et donc qu'il s'agit bien d'une généralisation de la dérivabilité aux $\mathbb{R}$-espaces vectoriels normés quelconques.

→ En réexaminant la définition II.2, on voit qu'écrire $f(x) \underset{x \to a}{=} f(a) + \varphi(x-a) + o(x-a)$ signifie qu'au voisinage de a, on peut dire que $f(x) \simeq f(a) + \varphi(x-a)$, c'est-à-dire qu'on peut approximer f par la somme d'une constante et une application linéaire appliquée à $x - a$ (*i.e.* une application linéaire affine dont la courbe représentative est une hyperplan affine dans le cas $\mathrm{F} = \mathbb{R}$, comme vu dans la première section de ce chapitre). Autrement dit, à une constante près, l'application linéaire φ est l'application linéaire qui approxime f le mieux au voisinage de a. En particulier, écrire $g(x) \underset{x \to a}{=} g(a) + \ell \times (x-a) + o(x)$ signifie qu'au voisinage de a, on peut dire que $g(x) \simeq g(a) + \ell \times (x-a)$, c'est-à-dire qu'on peut approximer g au voisinage de a par une fonction affine, *i.e.* une droite. Une

(cc) BY-NC-SA

autre manière de le voir est de dire que lorsque x se déplace légèrement à partir de a, la variation de $f(x)$ (égale à $f(x)-f(a)$) est approximativement linéaire en la variation de x (égale à $x-a$) à l'ordre 1 près (c'est-à-dire en négligeant les termes négligeables devant $x-a$ au voisinage de a).

Définition II.3.

Lorsque $\mathrm{F}=\mathbb{R}$, φ est une forme linéaire, et donc il existe $u \in \mathrm{E}$ tel que

$$f(a+h) \underset{h\to 0}{=} f(a)+\langle u,\ h\rangle_{\mathrm{E}}+o(h).$$

Dans ce cas, le vecteur u est noté ∇f_a et est appelé gradient de f en a.

Remarques.

→ Attention, le gradient dépend du produit scalaire considéré. Lorsqu'on mentionnera le gradient d'une fonction sans préciser pour quel produit scalaire, il s'agira du produit scalaire duquel E est muni, $\langle .,\ .\rangle$.

→ En fait, le gradient de f en a ∇f_a, lorsqu'il est non nul, est la direction dans laquelle f croît le plus rapidement au voisinage de a. En effet, pour une petite variation $\delta \in \mathrm{E}$, de norme fixée $\|\delta\|=r>0$, la variation de f, $f(a+\delta)-f(a)$ est approximativement égale à $\langle \nabla f_a,\ \delta\rangle$. Cette variation est maximale lorsque δ est positivement proportionnel à ∇f_a, *i.e.* $\delta = r\dfrac{\nabla f_a}{\|\nabla f_a\|}$.

Exemples.

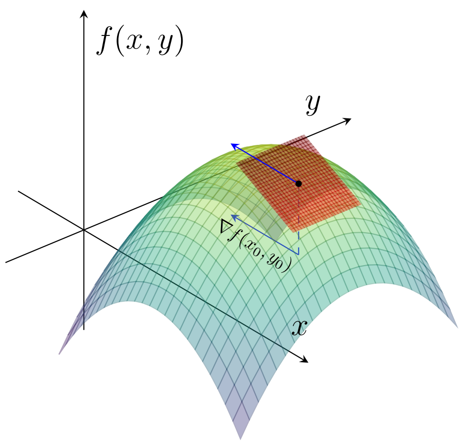
→ Observons l'exemple suivant. On considère l'application

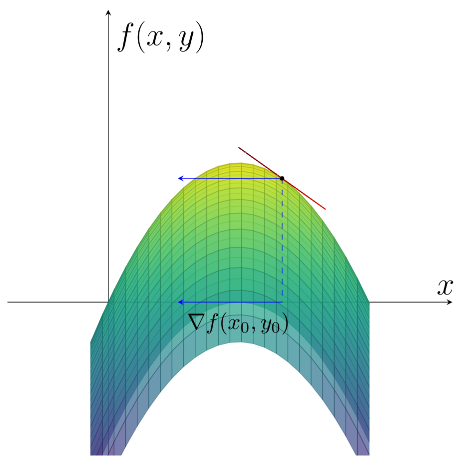
$$h : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto 1-(x-1)^2-(y-1)^2. \end{cases}$$

On a pour tout $(\delta_1,\delta_2)\in\mathbb{R}^2$ et $(x_0,y_0)\in\mathbb{R}^2$, $dh_{(x_0,y_0)}(\delta_1,\delta_2)=2(1-x_0)\delta_1+2(1-y_0)\delta_2$ i.e. $\nabla h_{(x_0,y_0)}=(2(1-x_0),2(1-y_0))$. Nous montrerons un peu plus tard un moyen pratique de trouver cette différentielle. On en déduit par exemple lorsque $(x_0,y_0)=(3,1)$, et donc $df_{(x_0,y_0)}(\delta_1,\delta_2)=-4\delta_1$. On peut donc écrire

$$\begin{aligned} f(x,y) &\underset{(x,y)\to(x_0,y_0)}{=} f(x_0,y_0)+df_{(x_0,y_0)}(x-x_0,y-y_0)+o((x-x_0,y-y_0)) \\ &\underset{(x,y)\to(3,1)}{=} -3-4(x-3)+o((x-3,y-1)). \end{aligned}$$

On en déduit qu'on peut écrire au voisinage de $(3,1)$, $f(x,y)\simeq -3-4(x-3)=9+\langle (x,y),\ (-4,0)\rangle$. Observons ces éléments sur les deux figures ci-dessus : la courbe représentative de f (en vert), la courbe de la fonction à droite de l'égalité précédente (en rouge) en (x_0,y_0) et le gradient de f en (x_0,y_0) (en bleu).





CPGE paradise
BY-NC-SA

→ Si $u \in \mathcal{L}(\mathrm{E}, \mathrm{F})$, alors pour tout $a \in \mathrm{E}$, $du_a = u$. En effet, on a bien $u(a+h) = u(a) + u(h) = u(a) + u(h) + o(h)$. Ce résultat est très naturel, car la meilleure application linéaire qui approxime u au voisinage de tout point est u elle-même.

→ Soit $u \in \mathrm{E}$. Considérons l'application de E dans $\mathbb{R}$, $h : x \longmapsto \langle u,\, x\rangle_{\mathrm{E}}$. Il s'agit d'une application linéaire, et en utilisant le point précédent, on peut facilement voir que pour tout $a \in \mathrm{E}$, $\nabla h_a = u$.

→ Si φ est une application linéaire et $\mathrm{C} \in \mathrm{F}$ un vecteur constant, alors pour tout $a \in \mathrm{E}$, $d(\varphi + \mathrm{C})_a = d\varphi_a$.

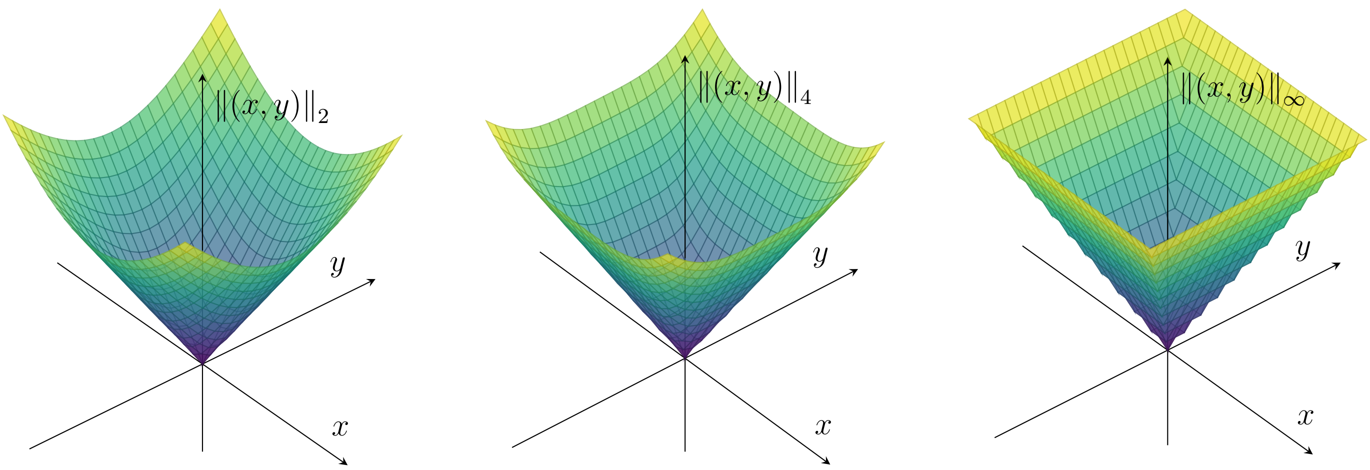
Remarque. Toute norme $\mathrm{N} : \mathrm{E} \longrightarrow \mathbb{R}$ sur E n'est pas différentiable en 0. Montrons ce résultat par l'absurde. Supposons que N est différentiable en 0. On pose $\varphi = d\mathrm{N}_0$. Soit $\varepsilon : \mathrm{E} \longrightarrow \mathbb{R}$ telle que $\varepsilon(x) \xrightarrow[x \to 0]{} 0$ et

$$\mathrm{N}(x) = \mathrm{N}(0) + \varphi(x) + o(x) = \varphi(x) + \|x\|_{\mathrm{E}}\, \varepsilon(x).$$

On en déduit donc que pour tout $t \in \mathbb{R}^*$,

$$\mathrm{N}(tx) = \varphi(tx) + \varepsilon(tx)\, \|tx\|_{\mathrm{E}} \;\; i.e. \;\; \frac{|t|}{t}\mathrm{N}(x) = \varphi(x) + \frac{|t|}{t}\, \|x\|_{\mathrm{E}}\, \varepsilon(tx). \tag{20}$$

L'égalité 20 devient, lorsqu'on fait tendre t vers 0^+ et 0^- respectivement $\mathrm{N}(x) = \varphi(x)$ et $\mathrm{N}(x) = -\varphi(x)$. Ceci signifie que pour tout x, $\mathrm{N}(x) = 0$, ce qui est absurde. On retrouve en particulier le fait que la valeur absolue n'est pas dérivale en 0, car il s'agit d'une norme sur $\mathbb{R}$.



Les trois figures ci-dessus nous donnent une idée lorsque $\mathrm{E} = \mathbb{R}^2$. En effet, pour ces 3 normes, on remarque que leurs courbes forment un point anguleux en 0, ce qui reflète le fait que ces fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ ne sont pas différentiables en 0. En général, si l'on trace la courbe d'une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ et qu'on observe un point anguleux en $a \in \mathbb{R}^2$, alors nécessairement cette fonction n'est pas différentiable en a. L'inverse n'est pas nécessairement vrai, car en général, lorsqu'une fonction n'est pas différentiable en $a \in \mathbb{R}^2$, on peut observer quelque chose de plus compliqué qu'un point anguleux, comme des oscillations.

De plus, les courbes des normes ci-dessus sont symétriques par rapport à l'axe vertical (l'axe des z). En effet, ceci est cohérent, car pour tout $u \in \mathbb{R}^2$ et toute norme $\|.\|$ sur $\mathbb{R}^2$, $\|-u\| = \|u\|$.

(cc) BY-NC-SA

Exercice II.4.

Montrer que chaque application ci-dessous est différentiable et calculer sa différentielle (et son gradient lorsque l'espace d'arrivée est une partie de $\mathbb{R}$) en tout point.

1. $\pi_i : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x_1, \ldots, x_n) & \longmapsto x_i, \end{cases}$ pour tout $i \in [\![1; n]\!]$ où $n \in \mathbb{N}^*$.
2. $\psi_1 : \begin{cases} \mathrm{E} & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto \mathrm{N}(x)^2, \end{cases}$ où N est une norme euclidienne dans E.
3. $\psi_2 : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow \mathcal{M}_n(\mathbb{R}) \\ \mathrm{A} & \longmapsto \mathrm{A}^2, \end{cases}$ où $n \in \mathbb{N}^*$.

On énonce maintenant quelques propriétés importantes de la différentielle et du gradient avant de passer à la suite.

Proposition II.5.

Si $f : \mathrm{E} \longrightarrow \mathrm{F}$ et $g : \mathrm{E} \longrightarrow \mathrm{F}$ sont différentiables en $a \in \mathrm{E}$, alors pour tout $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ est également différentiable en a, et

$$d(\lambda f + \mu g)_a = \lambda df_a + \mu dg_a.$$

De même, lorsque $\mathrm{G} = \mathbb{R}$,

$$\nabla(\lambda f + \mu g)_a = \lambda \nabla f_a + \mu \nabla g_a.$$

Démonstration. Cette preuve ne présente pas de difficulté et est laissée comme exercice au lecteur. □

Proposition II.6.

Soit $p \in \mathbb{N}^*$, et $f_1, \ldots, f_p$ des applications de E dans F. Supposons que $f : \mathrm{E} \longmapsto \mathrm{F}^p$ est définie sur l'espace produit de la manière suivante

$$\forall x \in \mathrm{E},\ f(x) = (f_1(x), \ldots, f_p(x)).$$

Pour tout $a \in \mathrm{E}$, f est différentiable si et seulement si $f_1, \ldots, f_p$ le sont. De plus, on a pour tout $h \in \mathrm{E}$,

$$df_a(h) = (df_{1,a}(h), \ldots, df_{p,a}(h)).$$

Démonstration.

→ (⇒) Supposons que f est différentiable en $a \in \mathrm{E}$ et posons $df_a = \varphi = (\varphi_1, \ldots, \varphi_p)$. On a

$$f(a+h) - f(a) - \varphi(h) = o(h),$$

ce qui donne, en substituant f et φ par leurs expressions,

$$(f_1(a+h) - f_1(a) - \varphi_1(h), \ldots, f_p(a+h) - f_p(a) - \varphi_p(h)) = o(h) = \|h\|_{\mathrm{E}}\, \varepsilon(h).$$

BY-NC-SA

où ε est une fonction de E dans F^p qui tend vers 0 en 0. On note dans ce cas pour tout $h \in \mathrm{E}$, $\varepsilon(h) = (\varepsilon_1(h), \ldots, \varepsilon_p(h))$. On en déduit donc que pour tout $i \in [\![1;p]\!]$,

$$f_i(a+h) - f_i(a) - \varphi_i(a) = \|h\|_{\mathrm{E}}\, \varepsilon_i(h) = o(h).$$

Ceci nous permet d'affirmer que pour tout $i \in [\![1;p]\!]$, f_i est différentiable et sa différentielle est égale à φ_i.

→ ($\Leftarrow$) Supposons que pour tout $i \in [\![1;p]\!]$, f_i est différentiable en $a \in \mathrm{E}$ et posons $\varphi_i = df_{i,a}$. Posons $\varphi : \mathrm{E} \longrightarrow \mathrm{F}^p$ l'application définie par

$$\forall h \in \mathrm{E},\ \varphi(h) = (\varphi_1(h), \ldots, \varphi_p(h)).$$

On a

$$\begin{aligned} f(a+h) - f(h) - \varphi(h) &= (f_1(a+h) - f_1(a) - \varphi_1(h), \ldots, f_p(a+h) - f_p(a) - \varphi_p(h)) \\ &= (\|h\|_{\mathrm{E}}\, \varepsilon_1(h), \ldots, \|h\|_{\mathrm{E}}\, \varepsilon_p(h)), \end{aligned}$$

où $\varepsilon_1, \ldots, \varepsilon_p$ sont des applications de E dans F qui tendent vers 0 en 0. On en déduit finalement que

$$f(a+h) - f(h) - \varphi(h) = \|h\|_{\mathrm{E}} \underbrace{(\varepsilon_1(h), \ldots, \varepsilon_p(h))}_{\xrightarrow[h\to 0]{} 0} = o(h).$$

f est donc bien différentiable et sa différentielle est égale à φ.

□

Proposition II.7.

Si $f : \mathrm{E} \longrightarrow \mathrm{F}$ est différentiable en $a \in \mathrm{E}$, alors f est continue en a.

Démonstration. On a

$$f(x) = f(a) + \underbrace{df_a(x-a)}_{\xrightarrow[x\to a]{} 0} + \underbrace{o(x-a)}_{\xrightarrow[x\to a]{} 0}.$$

Bien entendu, $df_a(x-a) \xrightarrow[x\to a]{} 0$ car df_a est une application linéaire en dimension finie (voir le chapitre 11.6 sur les applications linéaires continues). □

Proposition II.8.

Soit $\Omega_1 \subset \mathrm{E}$, $\Omega_2 \subset \mathrm{F}$, et $\Omega_3 \subset \mathrm{G}$ des ouverts. Si $f : \Omega_1 \longrightarrow \Omega_2$ est différentiable en $a \in \Omega_1$, et $g : \Omega_2 \longrightarrow \Omega_3$ est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et

$$d(g \circ f)_a = dg_{f(a)} \circ df_a.$$

Démonstration. On note $|\!|\!|.|\!|\!|$ la norme d'opérateur associée à $\|.\|_{\mathrm{E}}$ (pour l'espace de départ) et à $\|.\|_{\mathrm{F}}$ (pour l'espace d'arrivée). Soit $a \in \Omega_1$. On a

$$g(f(x)) - g(f(a)) \underset{(*)}{=} g(f(a)) + dg_{f(a)}(f(x) - f(a)) + o(f(x) - f(a)).$$

(cc) BY-NC-SA

L'égalité $(*)$ est vraie, car g est différentiable en $f(a)$ et $f(x) \xrightarrow[x\to a]{} f(a)$. On a de plus

$$f(x) - f(a) = df_a(x-a) + o(x-a),$$

et pour tout $h \in \mathrm{E}$, $\|df_a(h)\|_{\mathrm{F}} \leq \|\!|df_a|\!\|\, \|h\|_{\mathrm{E}}$, et alors $df_a(h) \underset{h\to 0}{=} \mathrm{O}(h)$, et de même $dg_{f(a)}(h) \underset{h\to 0}{=} \mathrm{O}(h)$. On en déduit donc que

$$\begin{aligned} g(f(x)) - g(f(a)) &= g(f(a)) + dg_{f(a)}(df_a(x-a)) + dg_{f(a)}(o(x-a)) + o(df_a(x-a)) + o(o(x-a)) \\ &= g(f(a)) + dg_{f(a)} \circ df_a(x-a) + \underbrace{\mathrm{O}(o(x-a)) + o(\mathrm{O}(x-a)) + o(x-a)}_{o(x-a)} \\ &= g(f(a)) + dg_{f(a)} \circ df_a(x-a) + o(x-a). \end{aligned}$$

On a donc bien montré le résultat voulu.

□

Exemple. Dans le cas où $\mathrm{E} = \mathrm{F} = \mathrm{G} = \mathbb{R}$, regardons ce que devient la relation démontrée ci-dessus. On a pour tout $(a,b) \in \Omega_1 \times \Omega_2$ et $h \in \mathbb{R}$, $df_a(h) = f'(a)h$ et $dg_b(h) = g'(b)h$. On en déduit donc que pour tout $h \in \mathbb{R}$,

$$d(g\circ f)_a(h) = dg_{f(a)} \circ df_a(h) = g'(f(a))f'(a)h, \ \textit{i.e.}\ (g\circ f)'(a) = g'(f(a))f'(a).$$

On retrouve bien donc la formule de la dérivée d'une composée de deux fonctions.

La propriété suivante est un cas particulier de la précédente, qui est assez utile en pratique, en particulier en physique.

Proposition (Règle de la chaine) II.9.

Soit I un intervalle ouvert de $\mathbb{R}$, $\gamma : \mathrm{I} \longrightarrow \mathrm{E}$ dérivable en $t_0 \in \mathrm{I}$, et $f : \Omega \longrightarrow \mathrm{F}$ différentiable en $\gamma(t_0)$ qu'on suppose dans Ω. $f\circ\gamma$ est dérivable en t_0, et

$$(f\circ\gamma)'(t_0) = df_{\gamma(t_0)}(\gamma'(t_0)).$$

Lorsque $\mathrm{F} = \mathbb{R}$, cette égalité devient

$$(f\circ\gamma)'(t_0) = \left\langle \nabla f_{\gamma(t_0)},\, \gamma'(t_0) \right\rangle.$$

Remarque. La deuxième égalité ci-dessus est très similaire à la formule de la dérivée d'une composée de deux fonctions, sauf qu'on remplace le produit sur $\mathbb{R}$ par le produit scalaire sur E.

Démonstration. En utilisant la formule de la proposition II.8, on obtient pour tout $h \in \mathbb{R}$,

$$d(f\circ\gamma)_{t_0}(h) = df_{\gamma(t_0)}(d\gamma_{t_0}(h)) = df_{\gamma(t_0)}(\gamma'(t_0))h.$$

Ceci nous donne bien la formule voulue. La deuxième égalité est une conséquence directe de la première.

□

Exemples.

→ Lorsque $a, b \in \mathbb{R}$, $a \leq b$, et γ est définie sur $\mathrm{I} =]0,1[$ par

$$\forall t \in]0,1[,\ \gamma(t) = tb + (1-t)a,$$

alors on a pour tout $t_0 \in]0,1[$

$$(f\circ\gamma)'(t_0) = df_{t_0 b + (1-t_0)a}(b-a).$$

(cc) BY-NC-SA

→ On peut également retrouver une formule assez connue en physique de deuxième année. Soit Γ une courbe de $\mathbb{R}^3$ et soit $\gamma : [0,1] \longmapsto \Gamma$ un arc continu bijectif de classe $\mathcal{C}^1$ (il s'agit d'un arc continu parcourant Γ). Soit $\overrightarrow{\mathrm{E}} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ un champ vectoriel. On définit, comme en physique, mais cette fois-ci formellement, l'intégrale de E sur l'arc Γ de la manière suivante.

$$\int_\Gamma \overrightarrow{\mathrm{E}}(\mathrm{M})\overrightarrow{d\mathrm{M}} = \int_0^1 \left\langle \overrightarrow{\mathrm{E}}(\gamma(t)),\, \gamma'(t)\right\rangle dt.$$

Lorsqu'il existe un potentiel $\mathrm{V} : \mathbb{R}^3 \longrightarrow \mathbb{R}$ tel que $\overrightarrow{\mathrm{E}} = (x \longmapsto \nabla \mathrm{V}_x)$, alors on a

$$\int_\Gamma \overrightarrow{\mathrm{E}}(\mathrm{M})\overrightarrow{d\mathrm{M}} = \int_0^1 \left\langle \nabla \mathrm{V}_{\gamma(t)},\, \gamma'(t)\right\rangle dt \underset{\substack{\uparrow \\ \text{règle de la chaine}}}{=} \int_0^1 \frac{d}{dt}\mathrm{V}(\gamma(t))dt = \mathrm{V}(\gamma(1)) - \mathrm{V}(\gamma(0)).$$

On retrouve alors l'égalité ci-dessus bien connue en physique.

Proposition II.10.

Lorsque $\mathrm{F} = \mathbb{R}$, et f, g sont des applications de Ω dans $\mathbb{R}$, alors si f et g sont différentiables en $a \in \Omega$, $f \times g$ l'est aussi.

Démonstration. L'application

$$\Phi : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto xy \end{cases}$$

est différentiable en tout point de $\mathbb{R}^2$. De plus, si f, g sont différentiables en $a \in \Omega$, alors d'après la proposition II.6, $h : x \longmapsto (f(x), g(x))$ est aussi différentiable en a. Enfin, Φ est différentiable en $(f(a), g(a))$, donc d'après la proposition II.8, on peut affirmer que $\Phi \circ h = f \times g$ est différentiable en a. □

Exercice II.11.

Soit $f : \Omega \longrightarrow \mathbb{R}$ une application différentiable sur Ω.

1. Calculer la différentielle de f^2.
2. Supposons que $f(\Omega) \subset]0, +\infty[$. Calculer la différentielle de $\sqrt{f}$.

Exercice II.12.

Supposons que E est un espace euclidien, et que $\|.\|_\mathrm{E}$ est la norme euclidienne associée à $\langle , \rangle$, et soit $b \in \mathrm{E}$. Calculer la différentielle en tout point lorsqu'elle existe de l'application

$$d : \begin{cases} \mathrm{E} & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto d(a,x) = \|x - b\|_\mathrm{E}. \end{cases}$$

III Dérivées partielles

1. Généralités

On commence par introduire une notion assez utile.

(cc) BY-NC-SA

Définition III.1.

Soit $u \in \mathrm{E}$ et $f : \Omega \longrightarrow \mathrm{F}$ une application de E dans F. On appelle dérivée de f en $x \in \Omega$ selon la direction u, lorsqu'elle existe, la quantité définie par

$$\mathrm{D}_u(f)(x) = \lim_{t \to 0} \frac{f(x+tu) - f(x)}{t}.$$

Une fois cette notion définie, on peut caractériser d'une manière assez élégante la différentielle d'une fonction.

Proposition III.2.

Soit $n = \dim \mathrm{E} \in \mathbb{N}^*$, $(e_1, \dots, e_n)$ une base de E, et $f : \Omega \longrightarrow \mathrm{F}$ différentiable en $a \in \Omega$. Pour tout u, la dérivée de f en a selon u existe et pour tout $h = h_1 e_1 + \cdots + h_n e_n \in \mathrm{E}$,

$$df_a(h) = \sum_{k=1}^{n} \mathrm{D}_{e_k}(f)(a) h_k.$$

Lorsque $\mathrm{F} = \mathbb{R}$ et $(e_1, \dots, e_n)$ est une base orthonormée, l'égalité ci-dessus devient,

$$\nabla f_a = \mathrm{D}_{e_1}(f)(a) e_1 + \cdots + \mathrm{D}_{e_n}(f)(a) e_n. \tag{21}$$

Démonstration. Pour tout $u \in \mathrm{E}$ et $a \in \Omega$, on a

$$f(a+tu) - f(a) = df_a(tu) + o(tu) = t \cdot df_a(u) + o(t).$$

Ceci implique que $\dfrac{f(a+tu) - f(a)}{t} \underset{t \to 0}{\longrightarrow} df_a(u)$, *i.e.* la dérivée directionnelle en a selon u de f existe et $\mathrm{D}_u(f)(a) = df_a(u)$. Montrons maintenant la formule de la proposition. On a pour tout $h = h_1 e_1 + \cdots + h_n e_n \in \mathrm{E}$,

$$df_a(h) = df_a\left(\sum_{k=1}^{n} h_k e_k\right) = \sum_{k=1}^{n} df_a(e_k) h_k = \sum_{k=1}^{n} \mathrm{D}_{e_k}(f)(a) h_k.$$

Ceci est bien la formule voulue. La seconde inégalité est une conséquence directe de la première (le lecteur ayant un doute peut essayer de la montrer lui-même en utilisant les définitions). □

Remarques.

→ Lorsque $\mathrm{E} = \mathbb{R}^n$ muni du **produit scalaire canonique** $((x_1, \dots, x_n), (y_1, \dots, y_n)) \longmapsto \sum_{k=1}^{n} x_i y_i$, et $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$, alors en notant pour tout $i \in [\![1; n]\!]$, $x = (x_1, \dots, x_n) \in \mathrm{E}$,

$$\frac{\partial f}{\partial x_i}(x) = \mathrm{D}_{e_i}(f)(x) = \lim_{t \to 0} \frac{f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{t},$$

alors, on peut réécrire l'égalité 21 pour tout $a \in \Omega$ où a est différentiable de la manière suivante

$$\nabla f_a = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}. \tag{22}$$

(cc) BY-NC-SA

Pour tout $i \in [\![1;n]\!]$, on appelle $\dfrac{\partial f}{\partial x_i}(a)$ la dérivée partielle de f par rapport à la $i-$ème variable en a. En fait, l'égalité 22 nous permet de calculer la différentielle (ou le gradient) de fonctions à valeurs dans $\mathbb{R}$ assez facilement en pratique. En pratique, pour $i \in [\![1;n]\!]$, pour calculer $\dfrac{\partial f}{\partial x_i}$, il suffit de considérer toutes les variables autres que x_i comme constantes et de dériver uniquement par rapport à x_i.

→ Attention, la proposition III.2 ferait croire que la réciproque est vraie. La réciproque est en fait fausse : admettre une dérivée directionnelle, même selon toutes les directions d'une base, n'implique pas être différentiable. En effet, on peut le voir via le contrexemple suivant. Considérons l'application

$$g : \begin{cases} \mathbb{R}^2 & \longrightarrow \mathbb{R} \\ (x,y) & \longmapsto \begin{cases} \frac{x^2y}{x^4+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{sinon.} \end{cases} \end{cases}$$

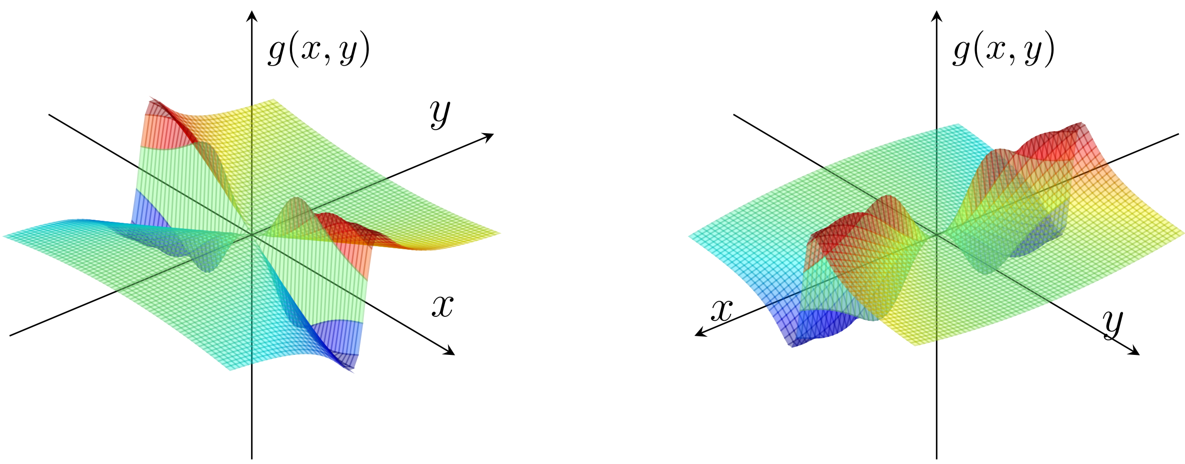
On remarque qu'en $(0,0)$, g admet des dérivées directionnelles selon $(1,0)$ et $(0,1)$. En effet, on a

$$\frac{g(t,0)-g(0,0)}{t} = 0 \underset{t\to 0}{\longrightarrow} 0 \text{ et } \frac{g(0,t)-g(0,0)}{t} = 0 \underset{t\to 0}{\longrightarrow} 0.$$

Cependant, g n'est pas différentiable en 0. En effet, en supposant que g est différentiable en 0 et en utilisant ce qui précède et la proposition 21, on peut affirmer que $dg_{(0,0)} = 0$. Ceci implique que la dérivée en $(0,0)$ selon n'importe quelle direction doit être nulle, mais pour une direction $u = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{R}^2$ telle que $\beta \neq 0$, on a

$$\frac{g(tu)-g(0,0)}{t} = \frac{t^3\alpha^2\beta/(t^4\alpha^4+t^2\beta^2)}{t} = \frac{\alpha^2\beta}{t^2\alpha^4+\beta^2} \underset{t\to 0}{\longrightarrow} \underbrace{\frac{\alpha^2}{\beta}}_{\neq 0 \text{ si } \alpha\neq 0},$$

ce qui est absurde, donc g n'est pas différentiable en $(0,0)$.



On observe sur le dessin ci-dessus qu'en 0, il y a des oscillations qui font que g n'est pas "lisse" et donc ne peut pas être différentiable en 0. Il s'agit donc d'un autre cas de figure que le point anguleux en 0 qu'on a observé auparavant pour les normes. Remarquons que dans des directions particulières au voisinage de 0, ces oscillations divergent vers l'infini, mais on ne peut pas l'observer sur le dessin, car on aurait besoin d'une précision beaucoup plus grande.

→ En réexaminant la définition II.2 et la proposition III.2, on a jusqu'à présent deux manières de montrer qu'une application $f : \Omega \longmapsto \mathrm{F}$ est différentiable et deux moyens pratiques de calculer la différentielle d'une application en un point $a \in \Omega$.

(cc) BY-NC-SA

- Si f est "simple", on peut potentiellement montrer grâce aux propositions II.6, II.8, II.5, II.10 que f est différentiable. Par exemple, la fonction $\mathrm{M} \longmapsto \det \mathrm{M}$ est somme et produit d'applications identiques à celles de la question 1 de l'exercice II.11, et est donc différentiable. On pourra calculer la différentielle de cette fonction dans un exercice ultérieur (exercice III.9)

- Si on sait déjà que f est différentiable en a (ou on l'a montré via le point précédent), pour calculer la différentielle de f, il suffit de calculer les dérivées selon les vecteurs d'une base convenable de E (en général, on calculera plutôt les dérivées partielles dans le cas où $\mathrm{E} = \mathbb{R}^n$ avec $n \in \mathbb{N}^*$ muni de son produit scalaire canonique ou est un espace euclidien). Ensuite, il suffit d'écrire pour tout $h = h_1 e_1 + \cdots + h_n e_n \in \mathrm{E}$,
$$df_a(h) = \sum_{k=1}^{n} \mathrm{D}_{e_k}(f)(a) h_k,$$
et dans le cas plus fréquent où $\mathrm{E} = \mathbb{R}^n$, pour tout $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$,
$$df_a(h) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) h_k.$$
Si de plus $\mathrm{F} = \mathbb{R}$, alors $df_a(h) = \langle \nabla f_a,\ h \rangle$ où ∇f_a est donné par l'égalité 22.

- Si on ne sait pas si f est différentiable à priori, on calcule $f(a+h) - f(a)$, et on essaye d'écrire cette différence conformément à la définition II.2. Une fois cela fait, on aura montré que f est différentiable en a, et on aura potentiellement aussi calculé la différentielle de f en a en même temps.

→ Le premier point de cette remarque nous montre également qu'un moyen possible pour démontrer que f n'est pas différentiable est de démontrer que f n'admet pas une dérivée selon une certaine direction $u \in \mathrm{E}$, ou alors montrer par l'absurde que si c'était le cas, les valeurs de la différentielle et des dérivées directionnelles ne sont pas cohérentes entre elles.

Exercice III.3.

Soit $n \in \mathbb{N}$. On munit $\mathcal{M}_n(\mathbb{R})$ du produit scalaire $\langle .,\ . \rangle : (\mathrm{A}, \mathrm{B}) \longmapsto \mathrm{Tr}(\mathrm{AB}^{\mathrm{T}})$ et de la norme $\|.\| : \mathrm{A} \longmapsto \mathrm{Tr}(\mathrm{AA}^{\mathrm{T}})$. Montrer que les applications suivantes sont différentiables (sauf potentiellement en un point) et calculer leurs différentielles.

1. $h_1 : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R}^+ \\ (x_1, \ldots, x_n) & \longmapsto \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2}, \end{cases}$ où $a_1, \ldots, a_n \in \mathbb{R}$.
2. $h_2 : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow \mathbb{R} \\ \mathrm{M} & \longmapsto \det \mathrm{M}. \end{cases}$
3. $h_3 : \begin{cases} \mathcal{M}_n(\mathbb{R}) & \longrightarrow \mathcal{M}_n(\mathbb{R}) \\ \mathrm{M} & \longmapsto \mathrm{MAM} \end{cases}$ où $\mathrm{A} \in \mathcal{M}_n(\mathbb{R})$.

BY-NC-SA

Exercice III.4.

Soit $k \in \mathbb{N}^*$ et f une fonction de E^k dans F. On dit que f est $k-$linéaire lorsque pour tout $x_1, \ldots, x_k \in \mathrm{E}$ et tout $i \in [\![1; k]\!]$, l'application $x \longmapsto f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k)$ est linéaire.

1. Montrer que f est continue sur E.
2. Montrer qu'il existe $\mathrm{C} \geq 0$ tel que pour tout $x_1, \ldots, x_k \in \mathrm{E}$,
$$\|f(x_1, \ldots, x_n)\|_{\mathrm{F}} \leq \mathrm{C} \|x_1\|_{\mathrm{E}} \times \cdots \times \|x_k\|_{\mathrm{E}}.$$
3. Calculer la différentielle de l'application
$$g : \begin{cases} \mathrm{E} & \longrightarrow \mathrm{F} \\ x & \longmapsto f(x, \ldots, x). \end{cases}$$
On dit que f est symétrique lorsque pour tout $x_1, \ldots, x_k \in \mathrm{E}$ et toute permutation $\sigma \in \mathcal{S}_k$, $f(x_1, \ldots, x_k) = f(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$. Que devient cette différentielle de g lorsque f est symétrique ?

2. Fonctions de classe $\mathcal{C}^1$

Dans cette sous-partie, on suppose que $\mathrm{E} = \mathbb{R}^n$ avec $n \in \mathbb{N}^*$ muni de son produit scalaire canonique et Ω est un ouvert de E.

Définition III.5.

Soit f une application de $\mathbb{R}^n$ dans F. On dit que f est de classe $\mathcal{C}^1$ lorsque f admet des dérivées partielles $\dfrac{\partial f}{\partial x_i}$ pour tout $i \in [\![1; n]\!]$ (par rapport à toutes ses variables), et que pour tout $i \in [\![1; n]\!]$, $x \longmapsto \dfrac{\partial f}{\partial x_i}(x)$ est continue sur Ω.

Exemple. Les fonctions polynomiales, i.e. les fonctions de l'ensemble

$$\mathbb{R}[x_1, \ldots, x_n] = \left\{ (x_1, \ldots, x_n) \longmapsto \sum_{\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathrm{A}} a_\alpha x_1^{\alpha_1} \ldots x_n^{\alpha_n},\ \mathrm{A} \subset \mathbb{N}^n \text{ fini}, (a_\alpha)_{\alpha \in \mathrm{A}} \in \mathbb{R}^{\mathrm{A}} \right\}.$$

sont de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$.

Proposition III.6.

Si $f : \Omega \longrightarrow \mathrm{F}$ est de classe $\mathcal{C}^1$, alors elle est différentiable en tout point de Ω.

Démonstration. Soit $a \in \Omega$. Si f est différentiable en a, alors d'après la proposition III.2, sa différentielle en a doit être égale à la fonction φ définie par

$$\forall h = (h_1, \ldots, h_n) \in \mathbb{R}^n,\ \varphi(h) = \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k}(a).$$

On va donc essayer de montrer que $f(a+h) - f(a) - \varphi(h) \underset{h \to 0}{=} o(h)$. On sait que les dérivées partielles sont bien définies et continues, on a donc des informations utiles sur les variations du type $f(a + te_i) - f(a)$

CC BY-NC-SA

(approximativement égales, à l'ordre 1, à $t\dfrac{\partial f}{\partial x_i}(a)$) où e_i est le $i-$ème vecteur de la base canonique de $\mathbb{R}^n$ et t un nombre réel assez petit en valeur absolue. On va donc essayer de décomposer la variation $f(a+h)-f(a)$ en somme de variations égales à $f(b+te_i)-f(b)$ où b est proche de a.
Soit $\varepsilon>0$. Par continuité des dérivées partielles et le fait que Ω est ouvert, on peut affirmer qu'il existe $\eta>0$ tel que pour tout $y=(y_1,\dots,y_n)\in\mathbb{R}^n$,

$$\max_{i\in[\![1;n]\!]}|a_i-y_i|<\eta \implies y\in\Omega \text{ et } \max_{i\in[\![1;n]\!]}\left(\left\|\frac{\partial f}{\partial x_i}(a)-\frac{\partial f}{\partial x_i}(y)\right\|_{\mathrm{F}}\right)<\varepsilon. \tag{23}$$

Posons pour tout $h\in\mathbb{R}^n$ tel que $\|h\|_\infty<\eta$ (cette condition sur h nous donne en particulier que $a+h\in\Omega$),

$$\Delta(h)=f(a+h)-f(a)-\varphi(h).$$

On cherche à montrer que $\Delta(h)=o(h)$. On écrit

$$\begin{aligned}\Delta(h)&=\overbrace{f(a_1+h_1,a_2+h_2,\dots,a_n+h_n)-f(a_1,a_2+h_2,\dots,a_n+h_n)-h_1\frac{\partial f}{\partial x_1}(a)}^{\Delta_1(h)}\\&\quad+\underbrace{f(a_1,a_2+h_2,\dots,a_n+h_n)-f(a_1,a_2,a_3+h_3,\dots,a_n+h_n)-h_2\frac{\partial f}{\partial x_2}(a)}_{\Delta_2(h)}+\dots\\&\quad+\underbrace{f(a_1,a_2,\dots,a_{n-1},a_n+h_n)-f(a_1,a_2,\dots,a_n)-h_n\frac{\partial f}{\partial x_n}(a)}_{\Delta_n(h)}\\&=\sum_{k=1}^{n}\Delta_k(h).\end{aligned}$$

Notre but maintenant est de trouver une bonne majoration pour $\Delta_k(h)$ pour tout $k\in[\![1;n]\!]$. On écrit pour tout $k\in[\![1;n]\!]$,

$$\Delta_k(h)=\underbrace{f(a_1,\dots,a_{k-1},a_k+h_k,\dots,a_n+h_n)-f(a_1,\dots,a_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)}_{\delta_k(h)}-h_k\frac{\partial f}{\partial x_k}(a).$$

On voit que $\delta_k(h)$ ressemble fortement à une variation du type $f(b+te_i)-f(b)$ où b est proche de a, on veut donc trouver une relation entre cette quantité et la dérivée partielle. Un outil assez pratique pour le faire (plutôt de le faire directement, ce qui peut être assez pénible) est l'inégalité des accroissements finis. En effet, en posant pour tout $t\in[0,1]$

$$\Phi(t)=f(a_1,\dots,a_{k-1},a_k+th_k,\dots,a_n+h_n)-f(a_1,\dots,a_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)-th_k\frac{\partial f}{\partial x_k}(a),$$

qui est une fonction continue sur $[0,1]$ et dérivable sur $]0,1[$, on peut dire que

$$\begin{aligned}\forall c\in]0,1[,\ \|\Phi'(c)\|_{\mathrm{F}}&\underset{\substack{\uparrow\\ \text{règle de la chaine}}}{=}\left\|h_k\left(\frac{\partial f}{\partial x_k}(a_1,\dots,a_k+ch_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)-\frac{\partial f}{\partial x_k}(a)\right)\right\|_{\mathrm{F}}\\&=|h_k|\left\|\frac{\partial f}{\partial x_k}\underbrace{(a_1,\dots,a_k+ch_k,a_{k+1}+h_{k+1},\dots,a_n+h_n)}_{\alpha_{k,c}}-\frac{\partial f}{\partial x_k}(a)\right\|_{\mathrm{F}}\\&\underset{(*)}{\le}|h_k|\,\varepsilon.\end{aligned}$$

(cc) BY-NC-SA

L'inégalité $(*)$ est vraie car $\|a-\alpha_{k,c}\|_\infty < \eta$ (on peut facilement le voir en utilisant le fait que $\|h\|_\infty < \eta$, $c \in]0,1[$ et en utilisant l'implication 23).
Maintenant, on peut appliquer l'inégalité des accroissements finis (voir le chapitre sur les fonctions à valeurs vectorielles) sur l'intervalle $[0,1]$, ce qui nous permet de dire que

$$\|\Delta_k(h)\|_{\mathrm{F}} = \|\Phi(1)-\Phi(0)\|_{\mathrm{F}} \leq |h_k|\,\varepsilon.$$

On en déduit donc que

$$\|\Delta(h)\|_{\mathrm{F}} \leq \sum_{k=1}^{n} \|\Delta_k(h)\|_{\mathrm{F}} \leq \varepsilon \sum_{k=1}^{n} |h_k| \leq n\varepsilon \|h\|_\infty .$$

Ceci nous permet de dire finalement que $\Delta(h) = o(h)$, ce qui est bien le résultat voulu : f est différentiable en a et sa différentielle est égale à φ.

□

Notation. On note $\mathcal{C}^1(\Omega, \mathrm{F})$ l'ensemble des applications de Ω dans F de classe $\mathcal{C}^1$.

Proposition III.7.

Soit f une application de Ω dans F. Les propositions suivantes sont équivalentes.

1. f est de classe $\mathcal{C}^1$ sur Ω.
2. f est différentiable sur Ω et l'application $\psi : \begin{cases} \Omega & \longrightarrow \mathcal{L}(\mathrm{E},\mathrm{F}) \\ a & \longmapsto df_a \end{cases}$ est continue.

Démonstration. On note $|||.|||$ la norme d'opérateur associée à $\|.\|_\infty$ (pour l'espace de départ) et $\|.\|_{\mathrm{F}}$ (pour l'espace d'arrivée).

→ (1) ⇒ (2) D'après la proposition III.6, f est différentiable, car elle est de classe $\mathcal{C}^1$. De plus, pour tout $a, b \in \Omega$ et $h \in \mathrm{E}$, on a

$$\|(df_a - df_b)(h)\|_{\mathrm{F}} = \left\| \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(a) - \frac{\partial f}{\partial x_k}(b) \right) h_k \right\|_{\mathrm{F}} \leq \left(\sum_{k=1}^{n} \left\| \frac{\partial f}{\partial x_k}(a) - \frac{\partial f}{\partial x_k}(b) \right\|_{\mathrm{F}} \right) \|h\|_\infty .$$

On en déduit donc par définition de la norme d'opérateur que

$$|||df_a - df_b||| \leq \sum_{k=1}^{n} \left\| \frac{\partial f}{\partial x_k}(a) - \frac{\partial f}{\partial x_k}(b) \right\|_{\mathrm{F}} \underset{b \to a}{\longrightarrow} 0.$$

On a donc bien montré la continuité de ψ.

→ (2) ⇒ (1) f est différentiable en tout point de Ω et donc admet des dérivées partielles en tout point de Ω. On a alors pour tout $a, b \in \Omega$, en notant pour $i \in [\![1;n]\!]$, e_i le $i-$ème vecteur de la base canonique de $\mathbb{R}^n$,

$$\left\| \frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right\|_{\mathrm{F}} = \|df_a(e_i) - df_b(e_i)\|_{\mathrm{F}} \leq |||df_a - df_b|||\, \|e_i\|_\infty = |||df_a - df_b||| \underset{b \to a}{\longrightarrow} 0.$$

f est donc bien de classe $\mathcal{C}^1$ sur Ω.

□

On introduit ensuite quelques propriétés générales utiles des fonctions de classe $\mathcal{C}^1$.

(cc) BY-NC-SA

Proposition III.8.

Les propositions suivantes sont vraies.

1. $\mathcal{C}^1(\Omega, \mathrm{F})$ est un $\mathbb{R}-$espace vectoriel.
2. $(\mathcal{C}^1(\Omega, \mathrm{F}), +, \times)$ est une $\mathbb{R}-$algèbre.

Démonstration. La preuve de ces deux propositions ne présente pas de difficulté et est laissée comme exercice au lecteur. $\square$

Exemples.

→ Si $f : \Omega \longrightarrow \mathbb{R}$ ne s'annule pas sur Ω et est de classe $\mathcal{C}^1$, alors $\dfrac{1}{f}$ est de classe $\mathcal{C}^1$.

→ Si $f : \Omega \longrightarrow \mathbb{R}$ est strictement positive sur Ω et est de classe $\mathcal{C}^1$, alors $\sqrt{f}$ est de classe $\mathcal{C}^1$.

Exercice III.9.

Soit $\sum a_n z^n$ une série entière complexe de rayon de convergence égal à $\mathrm{R} > 0$. Montrer que $(x, y) \longmapsto \sum_{k=0}^{\infty} a_k (x + iy)^k$ est de classe $\mathcal{C}^1$ sur

$$\mathrm{B}_{\mathbb{R}^2}(0, \mathrm{R}) = \{(x, y) \in \mathbb{R}^2,\ \|(x, y)\|_2 = \sqrt{x^2 + y^2} < \mathrm{R}\}.$$

3. Matrices jacobiennes

Dans cette sous-partie, on suppose que $\mathrm{E} = \mathbb{R}^n$ et $\mathrm{F} = \mathbb{R}^p$ où $n, p \in \mathbb{N}^*$ et sont munis de leurs produits scalaires canoniques. On suppose aussi encore une fois que Ω est un ouvert de E.

Définition III.10.

Soit $a \in \Omega$ et $f : \Omega \longrightarrow \mathbb{R}^p$ différentiable en a. On appelle matrice jacobienne de f en a, la matrice de l'application linéaire df_a dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$. On note cette matrice $\mathrm{J}_f(a)$.

Proposition III.11.

Si $f : \Omega \longrightarrow \mathbb{R}^p$ est différentiable en $a \in \Omega$, et pour tout $x \in \Omega$, $f(x) = (f_1(x), \dots, f_p(x))$, alors on a

$$\mathrm{J}_f(a) = \begin{pmatrix} \nabla f_{1,a}^{\mathrm{T}} \\ \vdots \\ \nabla f_{p,a}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \dots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \in \mathcal{M}_{p,n}(\mathbb{R}).$$

BY-NC-SA

Démonstration. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. On a pour tout $i \in [\![1; n]\!]$,

$$df_a(e_i) = \begin{pmatrix} df_{1,a}(e_i) \\ \vdots \\ df_{p,a}(e_i) \end{pmatrix} = \begin{pmatrix} \langle \nabla f_{1,a},\ e_i \rangle \\ \vdots \\ \langle \nabla f_{p,a},\ e_i \rangle \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_i}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_i}(a) \end{pmatrix}.$$

On a donc bien retrouvé l'image de la base canonique de $\mathbb{R}^n$ dans la base canonique de $\mathbb{R}^p$ par df_a, ce qui nous permet d'écrire la matrice complète dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.

$$\mathrm{J}_f(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \dots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \nabla f_{1,a}^{\mathrm{T}} \\ \vdots \\ \nabla f_{p,a}^{\mathrm{T}} \end{pmatrix}.$$

$\square$

Remarques.

→ Bien évidemment, on déduit de ce qui précède que si $f : \Omega \longrightarrow \mathbb{R}^p$ est différentiable en $a \in \Omega \subset \mathbb{R}^n$, alors pour toute matrice colonne $h \in \mathbb{R}^n$,

$$f(a+h) - f(a) = \mathrm{J}_f(a) \times h + o(h).$$

→ On déduit également de ce qui précède que pour tout $f : \Omega \longrightarrow \mathbb{R}^p$ telle que pour tout $x \in \Omega$, $f(x) = (f_1(x), \dots, f_p(x))$ où $f_1, \dots, f_p$ sont des applications de Ω dans $\mathbb{R}$, et $a \in \Omega$ où f est différentiable, on a pour tous $(i,j) \in [\![1; p]\!] \times [\![1; n]\!]$,

$$[\mathrm{J}_f(a)]_{i,j} = \frac{\partial f_i}{\partial x_j}(a),$$

et par conséquent, si $g = (g_1, \dots, g_m)$ est une application de $\mathbb{R}^p$ dans $\mathbb{R}^m$ différentiable en $f(a)$ avec $m \in \mathbb{N}^*$, on a en utilisant la proposition II.8,

$$\mathrm{J}_{g \circ f}(a) = \mathrm{J}_g(f(a)) \times \mathrm{J}_f(a)$$

et donc en posant pour tous $x \in \mathbb{R}^n$,

$$g \circ f(x) = (h_1(x), \dots, h_m(x)) = (g_1(f(x)), \dots, g_m(f(x))),$$

on a pour tous $(i,j) \in [\![1; m]\!] \times [\![1; n]\!]$

$$\begin{aligned} \frac{\partial h_i}{\partial x_j}(a) &= \frac{\partial (g_i \circ f)}{\partial x_j}(a) = [\mathrm{J}_g(f(a)) \times \mathrm{J}_f(a)]_{i,j} \\ &= \left[\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}(f(a)) & \dots & \dfrac{\partial g_1}{\partial x_p}(f(a)) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_1}(f(a)) & \dots & \dfrac{\partial g_m}{\partial x_p}(f(a)) \end{pmatrix} \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(f(a)) & \dots & \dfrac{\partial f_1}{\partial x_n}(f(a)) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(f(a)) & \dots & \dfrac{\partial f_p}{\partial x_n}(f(a)) \end{pmatrix} \right]_{i,j} \\ &= \sum_{k=1}^{p} \frac{\partial g_i}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a). \end{aligned}$$

On aurait aussi pu montrer cette égalité en utilisant la règle de la chaine. En effet, l'application

(cc) BY-NC-SA

$\varphi : x_k \longmapsto f(a_1, \dots, a_{k-1}, x_k, a_k, \dots, a_n)$ est dérivable en a_k, et

$$\varphi'(a_k) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_j}(a) \end{pmatrix}.$$

On a alors en utilisant la règle de la chaine,

$$\frac{\partial h_i}{\partial x_j}(a) = (g_i \circ \varphi)'(a_k) = \left\langle \nabla g_{i,\varphi(a_k)},\ \varphi'(a_k) \right\rangle = \left\langle \nabla g_{i,f(a)},\ \varphi'(a_k) \right\rangle = \left\langle \begin{pmatrix} \dfrac{\partial g_i}{\partial x_1}(f(a)) \\ \vdots \\ \dfrac{\partial g_i}{\partial x_p}(f(a)) \end{pmatrix}, \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_j}(a) \end{pmatrix} \right\rangle$$

$$= \sum_{k=1}^{p} \frac{\partial g_i}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

→ En physique, on utilise quelques opérateurs différentiels dans quelques situations, par exemple en électromagnétique, en particulier la divergence et le laplacien. En fait, on peut définir ces opérateurs avec ce que l'on vient de voir. En effet, lorsque $p = n$, si f est une application de Ω dans $\mathbb{R}^n$ différentiable en $a \in \Omega$, alors

$$(\mathrm{div} f)(a) = \mathrm{Tr}(\mathrm{J}_f(a)) = \sum_{k=1}^{n} \frac{\partial f_k}{\partial x_k}(a),$$

et lorsque f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ admet des dérivées partielles secondes en a (dérivées deux fois, notées $\dfrac{\partial^2 f}{\partial x_i^2} = \left(\dfrac{\partial}{\partial x_i}\dfrac{\partial f}{\partial x_i}\right)(a)$ pour tout $i \in [\![1; n]\!]$), on définit le laplacien Δf de la manière suivante

$$\Delta f(a) = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k^2}(a),$$

et en notant

$$\psi : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R}^n \\ x & \longmapsto \nabla f_x, \end{cases}$$

on obtient la formule connue suivante

$$(\mathrm{div}\ \nabla f)(a) = (\mathrm{div}\ \psi)(a) = \Delta f.$$

IV Accroissements finis

Dans cette partie, on ne suppose plus que $\mathrm{E} = \mathbb{R}^n$, mais juste que E est un espace vectoriel normé de dimension finie, et on considère Ω un ouvert de E.

1. Théorème des accroissements finis

On énonce maintenant un équivalent du théorème des accroissements finis pour les fonctions d'un $\mathbb{R}-$espace vectoriel normé dans un autre, assez utile en exercices.

(cc) BY-NC-SA

Théorème (Accroissements finis) IV.1.

Soit $f \in \mathcal{C}^1(\Omega, \mathrm{F})$ et $a, b \in \Omega$. Posons $[a,b] = \{(1-t)a + tb,\ t \in [0,1]\}$, et supposons que $[a,b] \subset \Omega$. Les propositions suivantes sont vraies.

1. $f(b) - f(a) = \displaystyle\int_0^1 df_{(1-t)a+tb}(b-a)dt$.
2. Lorsque $\mathrm{F} = \mathbb{R}$, $f(b) - f(a) = \displaystyle\int_0^1 \left\langle \nabla f_{(1-t)a+tb},\ b-a \right\rangle dt$.
3. $\|f(b) - f(a)\|_{\mathrm{F}} \leq \displaystyle\sup_{t\in[0,1]} \left\|\left| df_{(1-t)a+tb} \right|\right\| \times \|b-a\|_{\mathrm{E}}$.

Démonstration.

1. Considérons l'application

$$\Phi : \begin{cases} [0,1] & \longrightarrow \mathrm{F} \\ t & \longmapsto f((1-t)a+tb). \end{cases}$$

Φ est une composition de deux fonctions différentiables et est donc différentiable sur $]0,1[$, et est donc dérivable. De plus, étant donné que $\Phi \in \mathcal{C}^1([0,1], \mathrm{F})$ et donc que Φ' est continue, on a pour tous $a, b \in \Omega$,

$$f(b) - f(a) = \Phi(1) - \Phi(0) = \int_0^1 \Phi'(t)dt \underset{\substack{\uparrow \\ \text{règle de la chaine}}}{=} \int_0^1 df_{(1-t)a+tb}(b-a)dt.$$

2. Cette égalité est une conséquence directe de l'égalité précédente. En effet, lorsque $\mathrm{F} = \mathbb{R}$, on a pour tous $a, b \in \Omega$

$$df_{(1-t)a+tb}(b-a) = \left\langle \nabla f_{(1-t)a+bt},\ b-a \right\rangle_{\mathrm{E}},$$

ce qui donne bien l'égalité voulue.

3. Cette inégalité est une conséquence directe de l'égalité (1). En effet, on a pour tous $a, b \in \Omega$

$$\begin{aligned} \|f(b) - f(a)\|_{\mathrm{F}} = \left\| \int_0^1 df_{(1-t)a+tb}(b-a)dt \right\|_{\mathrm{F}} &\leq \sup_{t\in[0,1]} \left\| df_{(1-t)a+tb}(b-a) \right\|_{\mathrm{F}} \\ &\leq \sup_{t\in[0,1]} \left\|\left| df_{(1-t)a+tb} \right|\right\| \times \|b-a\|_{\mathrm{E}}. \end{aligned}$$

$\square$

Conséquences.

→ Si $f \in \mathcal{C}^1$, alors f est localement lipschitzienne. En effet, considérons $a \in \Omega$ et $\eta > 0$ tel que $\mathrm{B}_f(a,\eta) \subset \Omega$. $f \in \mathcal{C}^1$, donc d'après la proposition III.7, l'application

$$\varphi : \begin{cases} \mathrm{B}_f(a,\eta) & \longrightarrow \mathcal{L}(\mathrm{E}, \mathrm{F}) \\ x & \longmapsto df_x \end{cases}$$

est continue sur $\mathrm{B}_f(a,\eta)$. De plus, E est de dimension finie et $\mathrm{B}_f(a,\eta)$ est fermé borné, et est donc compact. On en déduit que φ est bornée sur $\mathrm{B}_f(a,\eta)$ par un réel qu'on note K. On a alors pour

(cc) BY-NC-SA

tous $x, y \in \mathrm{B}(a,\eta)$, en utilisant la proposition précédente,

$$\|f(x)-f(y)\|_{\mathrm{F}} \leq \sup_{t\in[0,1]} \left\|df_{(1-t)x+ty}\right\| \times \|x-y\|_{\mathrm{E}} \leq \mathrm{K}\,\|x-y\|_{\mathrm{E}}\,.$$

→ Si pour tout $x \in \Omega$, $df_x = 0$ et Ω est connexe, alors f est constante sur Ω. Montrons ce résultat. Considérons $a \in \Omega$ et $\eta > 0$ tel que $\mathrm{B}(a,\eta) \subset \Omega$. D'après L'égalité (1) de la proposition précédente, on a pour tout $x \in \mathrm{B}(a,\eta)$,

$$f(x)-f(a) = \int_0^1 df_{(1-t)a+tx}(x-a)dt = 0.$$

On en déduit donc que pour tout $x \in \mathrm{B}(a,\eta)$, $f(x) = f(a)$. Posons maintenant

$$\mathrm{A} = \{x \in \Omega,\ f(x) = f(a)\} = f^{-1}(\{f(a)\}).$$

Cet ensemble est fermé, car il s'agit de l'image réciproque d'un ensemble fermé de Ω par une fonction continue. De plus, le raisonnement précédent nous permet aussi d'affirmer que A est ouvert (car tout point de cet ensemble peut être entouré d'une boule ouverte incluse dans A). finalement $a \in \mathrm{A}$, donc A est non vide, et alors par connexité de Ω, $\mathrm{A} = \Omega$. Au passage, on aurait également pu montrer ce résultat si on avait supposé que Ω est connexe par arcs (qui est une propriété plus forte que la connexité) sans passer par la connexité. Pour ce faire, il suffit de considérer deux points x, y de Ω et considérer un arc γ de classe $\mathcal{C}^1$ par morceaux (c'est-à-dire qu'on peut subdiviser $[0,1]$ en un nombre fini d'intervalles disjoints où γ est de classe $\mathcal{C}^1$) les liant (en fait, la définition de connexité par arcs donne uniquement que γ est continu, mais on peut montrer que dans tout $\mathbb{R}-$espace vectoriel, la connexité pour un ouvert implique la connexité par arcs de classe $\mathcal{C}^1$. Nous montrons ce résultat en annexe pour le lecteur curieux). On peut montrer facilement que $f \circ \gamma$ est constante sur son support, car sa dérivée est nulle et que finalement $f(b) = f(a)$.

→ Si Ω est connexe et qu'il existe $u \in \mathcal{L}(\mathrm{E},\mathrm{F})$ telle que pour tout $x \in \Omega$, $df_x = u$, alors il existe une constante $\mathrm{C} \in \mathrm{F}$ telle que pour tout $x \in \Omega$, $f(x) = u(x) + \mathrm{C}$. En effet, pour voir que c'est vrai, il suffit d'appliquer le raisonnement précédent pour $f - u$ qui a une différentielle nulle sur tout Ω (pour tout x, $d(f-u)_x = 0$). On en déduit donc qu'il existe une constante $\mathrm{C} \in \mathrm{F}$ telle que $f(x) - u(x) = \mathrm{C}$ pour tout x, ce qui donne bien l'égalité voulue.

Exercice IV.2.

On suppose que Ω est convexe et on considère $f \in \mathcal{C}^1(\Omega,\mathbb{R})$. Montrer que pour tous $a, b \in \Omega$

$$\exists \theta \in [0,1],\ f(b)-f(a) = \left\langle \nabla f_{(1-\theta)a+\theta b},\ b-a \right\rangle_{\mathrm{E}}.$$

Exercice IV.3.

Soient $m \in \mathbb{N}^*$ et $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}_+^*$. Montrer que

$$\left(\prod_{k=1}^m a_k\right)^{\frac{1}{m}} + \left(\prod_{k=1}^m b_k\right)^{\frac{1}{m}} \leq \left(\prod_{k=1}^m (a_k+b_k)\right)^{\frac{1}{m}}.$$

2. Inversion locale

Cette partie est hors programme, mais potentiellement utile aux oraux X-ENS. Par souci de simplicité, on suppose dans cette sous-partie que E et F sont égaux à $\mathbb{R}^n$ et munis d'une même norme $\|.\|$.

(cc) BY-NC-SA

Définition IV.4.

Soit Ω, Ω' deux ouverts de $\mathrm{E} = \mathbb{R}^n$. Une fonction $f : \Omega \longrightarrow \Omega'$ est un $\mathcal{C}^1-$difféomorphisme si f est de classe $\mathcal{C}^1$, bijective, et f^{-1} est de classe $\mathcal{C}^1$ sur Ω'.

Remarque. Lorsque f est un $\mathcal{C}^1-$difféomorphisme comme définit dans la définition ci-dessus, on a $f^{-1} \circ f = \mathrm{Id}$ et donc en appliquant la formule de la différentielle pour une composée de deux fonctions (proposition II.8), on obtient pour tout $a \in \Omega$

$$df^{-1}_{f(a)} \circ df_a = \mathrm{Id}, \ \ i.e. \ \mathrm{J}_{f^{-1}}(f(a)) \times \mathrm{J}_f(a) = \mathrm{I}$$

On remarquera en particulier que la différentielle en a (resp. la matrice jacobienne) de f en tout point $a \in \Omega$ est inversible, d'inverse $df^{-1}_{f(a)}$ (resp. $\mathrm{J}_{f^{-1}}(f(a))$).

Exemple. Posons $\Omega = \mathbb{R}_+^* \times]-\pi, \pi[$ et $\Omega' = \mathbb{R}^2 \setminus (\mathbb{R}^- \times \{0\})$ et considérons l'application

$$\varphi : \begin{cases} \Omega & \longrightarrow \Omega' \\ (r, \theta) & \longmapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \end{cases}. \tag{24}$$

φ est bijective de classe $\mathcal{C}^1$ sur Ω et $\forall (x,y) \in \Omega'$, $\varphi^{-1}(x,y) = \left(\sqrt{x^2+y^2}, 2\arctan\left(\dfrac{y}{x+1}\right)\right)$. φ^{-1} est également de classe $\mathcal{C}^1$ sur Ω', donc φ est bien in difféomorphisme de Ω dans Ω'.

Le théorème hors programme suivant peut être également vu comme un exercice. Il arrive que quelque chose de similaire tombe aux oraux X-ENS.

Théorème (Théorème d'inversion locale) IV.5.

Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R}^n)$. Soit $a \in \Omega$ tel que $\det \mathrm{J}_f(a) \neq 0$. Il existe $\mathrm{U} \subset \Omega$ un voisinage ouvert de a et $\mathrm{V} \subset \mathbb{R}^n$ un voisinage ouvert de $f(a)$ tels que

$$\tilde{f} : \begin{cases} \mathrm{U} & \longrightarrow \mathrm{V} \\ x & \longmapsto f(x) \end{cases}$$

est un $\mathcal{C}^1-$difféomorphisme.

Démonstration. On commence par faire les réductions suivantes du problème.

→ Quitte à remplacer f par la fonction $x \longmapsto f(x+a)$, on peut supposer que $a = 0$.

→ Quitte à remplacer f par $x \longmapsto f(x) - f(0)$, on suppose que $f(0) = 0$.

→ Enfin, en posant $u = df_0$, quitte à remplacer f par $x \longmapsto u^{-1} \circ f$, on peut supposer que $df_0 = \mathrm{Id}$.

Le lecteur ayant un doute peut montrer lui-même que ces simplifications ne changent pas le résultat, en particulier, l'image d'un voisinage ouvert d'un point $b \in \mathbb{R}^n$ par une application linéaire inversible $v \in \mathcal{L}(\mathbb{R}^n)$ est un voisinage ouvert de $v(b)$.

Citons les étapes clés de la preuve.

1. Dans un premier temps, on montre qu'il existe un voisinage U de 0 dans lequel f est injective.
2. Ensuite, en posant $\mathrm{V} = f(\mathrm{U})$, il est clair que $\tilde{f} : \mathrm{U} \longrightarrow \mathrm{V}$ est bijective. Il faut alors montrer que V est un voisinage de $f(0) = 0$, et qu'on peut supposer que V est ouvert, quitte à changer U par un sous-ensemble ouvert de U.
3. Enfin, il reste à montrer que $\tilde{f}^{-1} \in \mathcal{C}^1(\mathrm{V}, \mathrm{U})$.

(cc) BY-NC-SA

On commence par la première étape. Inspirons-nous du cas où $n = 1$. Dans ce cas, $f'(0) = 1$. Si on se restreint à un voisinage convexe de 0 où $f'(x) > 0$ (par exemple, voisinage assez petit pour qu'on ait $|f'(x) - 1| < \frac{1}{2}$) pour tout x dans ce voisinage, alors f y est strictement croissante et par conséquent injective.

Revenons au cas général $n \in \mathbb{N}^*$. f étant de classe $\mathcal{C}^1$, d'après le théorème III.7, on peut considérer $r > 0$ tel que pour tout $x \in \mathrm{B}(0, r)$, $|||df_x - \mathrm{Id}||| < \frac{1}{2}$. Pour montrer que f est nécessairement injective, il faut étudier sa variation entre deux points $x, y \in \mathrm{B}(0, r)$ en exploitant la relation entre les variations de f sur le segment $[x, y]$ et sa différentielle. Ceci nous fait penser directement au théorème des accroissements finis (théorème IV.1). Posons $\mathrm{U} = \mathrm{B}(0, r)$. On a pour tous $x, y \in \mathrm{U}$ tels que $f(x) = f(y)$, d'après le théorème IV.1,

$$0 = f(y) - f(x) = \int_0^1 df_{(1-t)x+ty}(y - x)dt.$$

Dans le segment $[x, y] \subset \mathrm{B}(0, r)$, df est proche de Id et alors pour tout $t \in [0, 1]$, $df_{(1-t)x+ty}$ est également proche de Id. Par conséquent, $df_{(1-t)x+ty}(y - x)$ est proche de $y - x$, et enfin son intégrale (qui est égale à 0) est assez proche de $y - x$, ce qui n'est pas possible sauf si $x = y$. Formellement, ceci s'écrit

$$\underbrace{\left\| \int_0^1 df_{(1-t)x+ty}(y - x)dt - (y - x) \right\|}_{=\|y-x\|} = \left\| \int_0^1 (df_{(1-t)x+ty} - \mathrm{Id})(y - x)dt \right\| \tag{25}$$

$$\leq \int_0^1 \left\| (df_{(1-t)x+ty} - \mathrm{Id})(y - x) \right\| dt \tag{26}$$

$$\leq \int_0^1 |||df_{(1-t)x+ty} - \mathrm{Id}|||\ \|y - x)\| \, dt \tag{27}$$

$$\leq \int_0^1 \frac{1}{2} \|y - x\| \, dt = \frac{1}{2} \|y - x\| . \tag{28}$$

Ceci implique donc que $\|y - x\| = 0$, c'est-à-dire que $y = x$. On a donc bien montré que F est injective sur U.

Passons maintenant à la deuxième étape. Posons $\mathrm{V} = f(\mathrm{U})$. On veut montrer que quitte à remplacer U à un certain sous-ensemble de U, on peut supposer que V est ouvert. On va montrer que $\mathrm{B}\left(0, \frac{r}{4}\right) \subset \mathrm{V}$, puis remplacer U par $f^{-1}\left(\mathrm{B}\left(0, \frac{r}{4}\right)\right)$. Soit $x \in \mathrm{U}$ tel que $\|x\| < \frac{r}{2}$ et $y \in \mathbb{R}^n$ tel que $\|y\| < \frac{r}{4}$. Considérons la suite $(x_m)_{m\in\mathbb{N}}$ définie par

$$\begin{cases} x_0 = x \\ \forall m \in \mathbb{N},\ x_{m+1} = x_m + y - f(x_m). \end{cases}$$

On va montrer que (x_n) converge vers un point $x' \in \mathrm{U}$ tel que $f(x') = y$. Pour cela, on va montrer que (x_m) est à valeurs dans U et qu'elle converge. Montrons le premier point par récurrence. La propriété

(cc) BY-NC-SA

$\|x_0\| \leq \dfrac{r}{2}$ est vérifiée pour $m = 0$. Soit $m \in \mathbb{N}$. Supposons que $\|x_m\| \leq \dfrac{r}{2}$. On a

$$\begin{aligned}
\|x_{m+1}\| = \|x_m + y - f(x_m)\| &\leq \|y\| + \left\| x_m - \int_0^1 df_{tx_m}(x_m)dt \right\| \\
&\leq \frac{r}{4} + \left\| \int_0^1 \left(df_{tx_m} - \mathrm{Id}\right)(x_m)dt \right\| \\
&\leq \frac{r}{4} + \int_0^1 \left|\!\left|\!\left| df_{tx_m} - \mathrm{Id} \right|\!\right|\!\right| \|x_m\| \, dt \\
&\leq \frac{r}{4} + \frac{1}{2}\|x_m\| \leq \frac{r}{2}.
\end{aligned}$$

On a donc bien montré par récurrence que pour tout $m \in \mathbb{N}$, $\|x_m\| \leq \dfrac{r}{2}$. Montrons maintenant qu'elle converge. On a pour tout $m \in \mathbb{N}^*$,

$$\begin{aligned}
\|x_{m+1} - x_m\| &= \|x_m - x_{m-1} - (f(x_m) - f(x_{m-1}))\| \\
&= \left\| \int_0^1 \left(df_{(1-t)x_{m-1}+tx_m} - \mathrm{Id}\right)(x_m - x_{m-1})dt \right\| \\
&\leq \int_0^1 \left|\!\left|\!\left| df_{(1-t)x_{m-1}+tx_m} - \mathrm{Id} \right|\!\right|\!\right| \|x_m - x_{m-1}\| \, dt \\
&\underset{x_m, x_{m-1} \in \mathrm{U}}{\leq} \frac{1}{2}\|x_m - x_{m-1}\|.
\end{aligned}$$

On en déduit donc par une récurrence immédiate que pour tout $m \in \mathbb{N}$,

$$\|x_{m+1} - x_m\| \leq \frac{1}{2^m}\|x_1 - x_0\| \leq \frac{1}{2^m}\left(\|x_1\| + \|x_0\|\right) \leq \frac{r}{2^m},$$

et enfin pour tout $\varepsilon > 0$ et tous $k, m \in \mathbb{N}$ tels que $m \geq k$ et $k \geq -\log_2\left(\dfrac{\varepsilon}{r}\right) + 1$,

$$\|x_m - x_k\| = \left\| \sum_{i=m}^{k-1} x_{i+1} - x_i \right\| \leq \sum_{i=m}^{k-1} \|x_{i+1} - x_i\| \leq \sum_{i=k}^{m} \frac{r}{2^i} \leq \frac{r}{2^{k-1}} \leq \varepsilon.$$

(x_m) est donc une suite de Cauchy dans un espace normé de dimension finie (donc complet), ce qui nous permet d'affirmer qu'elle converge. Notons sa limite $x' \in \mathrm{E}$. On a montré précédemment que pour tout $m \in \mathbb{N}$, $\|x_m\| \leq \dfrac{r}{2}$, et donc $\|x'\| \leq \dfrac{r}{2}$, ce qui implique $x' \in \mathrm{U}$. Enfin, en passant à la limite dans la relation de récurrence de (x_m), par continuité de f, on obtient $y - f(x') = 0$, et alors $y \in f(\mathrm{U}) \subset \mathrm{V}$. En résumé, on a considéré $y \in \mathrm{B}\left(0, \dfrac{r}{4}\right)$ et on a montré que $y \in \mathrm{V}$, c'est-à-dire que $\mathrm{B}\left(0, \dfrac{r}{4}\right) \subset \mathrm{V}$. On en déduit donc que V est bien un voisinage de 0. Quitte à restreindre f à $f^{-1}\left(\mathrm{B}\left(0, \dfrac{r}{4}\right)\right)$ (qui est ouvert, car il s'agit de l'image réciproque d'un ouvert par une application continue) et remplacer V par $\mathrm{B}\left(0, \dfrac{r}{4}\right)$, on peut supposer que V est ouvert.

Il reste maintenant à montrer que $\tilde{f}^{-1} \in \mathcal{C}^1(\mathrm{V}, \mathrm{U})$. On commence par montrer que $\tilde{f}^{-1}$ est différentiable sur V. Soit $k \in \mathbb{R}^n$ et $b = f(c) \in \mathrm{V}$ avec $c \in \mathrm{U}$. On suppose que la norme de k est assez petite pour que $b + k \in \mathrm{V}$, et on pose

$$h = \tilde{f}^{-1}(b+k) - \tilde{f}^{-1}(b) \text{ ce qui implique } f(c+h) - f(c) = k.$$

Si $\tilde{f}^{-1}$ est différentiable en b, la formule de la différentielle de la composée de deux fonctions nous permet de dire que $d\tilde{f}_b^{-1} = \left(df_{\tilde{f}^{-1}(b)}\right)^{-1} = (df_c)^{-1}$. Cette propriété suggère de procéder comme ce qui suit.

(cc) BY-NC-SA

$$\begin{aligned}\left\|\tilde{f}^{-1}(b+k)-\tilde{f}^{-1}(b)-(df_c)^{-1}(k)\right\| &= \left\|h-(df_c)^{-1}(f(c+h)-f(c))\right\| \\ &= \left\|(df_c)^{-1}(f(c+h)-f(c)-df_c(h))\right\| \\ &\leq \left\|\left|(df_c)^{-1}\right|\right\| \left\|f(c+h)-f(c)-df_c(h)\right\| = o(h).\end{aligned}$$

Une suite naturelle à ce qu'on a écrit ci-dessus est de comparer h et k. On a

$$\|h-k\| = \|f(c+h)-f(c)-(c+h-c)\| \underset{(*)}{\leq} \frac{1}{2}\|c+h-c\| = \frac{1}{2}\|h\|,$$

L'inégalité $(*)$ ci-dessus peut être obtenue en utilisant l'inégalité 28. L'inégalité ci-dessus implique que $\frac{\|h\|}{2} \leq \|k\| \leq 2\|h\|$ et alors $h = \mathrm{O}(k)$. Enfin, on a

$$\left\|\tilde{f}^{-1}(b+k)-\tilde{f}^{-1}(b)-(df_c)^{-1}(k)\right\| = o(h) = o(k),$$

ce qui signifie que $\tilde{f}^{-1}$ est différentiable (et donc continue) sur V de différentielle égale en tout point $b \in \mathrm{V}$ à $\left(df_{f^{-\tilde{1}}(b)}\right)^{-1}$. Enfin, il reste à prouver que $\tilde{f}^{-1}$ est de classe $\mathcal{C}^1$, ce qui est clair, car l'application $b \longmapsto \left(df_{\tilde{f}^{-1}(b)}\right)^{-1}$ est continue sur V, étant donné qu'il s'agit d'une composition d'applications continues. $\square$

Remarque. Nous citons sans démonstration une formule hors programme très souvent utilisée en physique pour le changement de variables.

Soit φ un $\mathcal{C}^1-$difféomorphisme d'un ouvert $\mathrm{U} \subset \mathbb{R}^n$ vers un ouvert $\mathrm{V} \subset \mathbb{R}^n$. Cette formule est souvent utilisée en physique pour des ensembles fermés d'intérieur non vides $\mathrm{K} \subset \mathrm{U}$ et $\mathrm{L} = \varphi(\mathrm{K})$. En fait, on peut formaliser ceci et garder la véracité de la formule ci-haut, pourvu que lesdits fermés vérifient certaines conditions qu'on ne précisera pas (mais qui sont, en général en physique de spé, toujours vérifiées).

Soit $\mathrm{L} = \varphi(\mathrm{K}) \subset \mathbb{R}^n$, et g une fonction intégrable de L dans $\mathbb{R}$. Lorsque K et L vérifient ces conditions (qu'on ne précisera pas par souci de simplicité), on admet la formule suivante.

$$\int_{\mathrm{L}} g(x)dx = \int_{\mathrm{K}} g(\varphi(y))\left|\det \mathrm{J}_\varphi(y)\right| dy.$$

Un exemple d'application de cette formule en physique est le passage en coordonnées polaires. En effet, en reconsidérant l'application φ définie en (24), on a pour toute fonction g d'un ouvert $\mathrm{L} \subset \mathbb{R}^* \times]-\pi,\pi[$ dans $\mathbb{R}$, on a

$$\begin{aligned}\int_{\mathrm{L}} g(x,y)dxdy &= \int_{\varphi^{-1}(\mathrm{L})} g(r\cos\theta, r\sin\theta)\left|\det \mathrm{J}_\varphi(r,\theta)\right| drd\theta \\ &= \int_{\varphi^{-1}(\mathrm{L})} g(r\cos\theta, r\sin\theta)\left|\det\begin{pmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{pmatrix}\right| drd\theta \\ &= \int_{\varphi^{-1}(\mathrm{L})} g(r\cos\theta, r\sin\theta) r drd\theta.\end{aligned}$$

Exercice IV.6.

Calculer $\int_0^{+\infty} e^{-x^2}$. *Indice : calculer* $\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy$.

(cc) BY-NC-SA

3. Conditions d'optimalité

On généralise les conditions d'optimalité vues pour les fonctions définies sur une partie de $\mathbb{R}$. Il est naturel de déduire de ce qui précède et par analogie avec le cas de dimension 1 que la différentielle en un point est nulle si ce point est un extremum local. On formalise cette propriété via la proposition suivante.

Proposition IV.7.

Soit $f : \Omega \longrightarrow \mathbb{R}$ et $a \in \Omega$. On suppose que f est différentiable en a. Les propositions suivantes sont vraies.

1. Si f admet un extremum local selon la direction $u \in \mathbb{R}^n$ en a, *i.e.* la fonction $\Phi : t \longmapsto f(a+tu)$ admet extremum local en 0, alors $\mathrm{D}_u(f)(a) = 0$.
2. Si f admet un extremum local en a, alors $df_a = 0$ (*i.e.* $\nabla f_a = 0$).

Démonstration.

1. Il s'agit d'un résultat classique d'analyse sur $\mathbb{R}$, étant donné que $\Phi'(0) = \mathrm{D}_u(f)(a)$ et donc le fait que Φ admette un extremum local en 0 implique directement que $\mathrm{D}_u(f)(a) = 0$.

2. L'intuition derrière cette preuve est simple : si ∇f_a est non nul, alors pour h assez petit, on peut dire que $f(a+h) - f(a) \simeq \langle \nabla f_a,\ h\rangle$, et alors si la variation h a une direction opposée à ∇f_a (égal à $-\alpha \nabla f_a / \|\nabla f_a\|$ avec $\alpha > 0$ assez petit), la variation $f(a+h) - f(a)$ est nécessairement strictement négative, ce qui est en contradiction avec le fait que a est un minimum local car cela implique que pour tout $h \in \mathbb{R}^n$ assez petit, la variation $f(a+h) - f(a)$ est positive. Maintenant, formalisons ce raisonnement via la preuve ci-dessous.

   Supposons sans perte de généralité que f admet un minimum local en a. On peut alors considérer $\varepsilon > 0$ tel que pour tout $h \in \mathrm{B}(0, \varepsilon)$, $f(a+h) - f(a) \geq 0$. Supposons par l'absurde que $\nabla f_a \neq 0$. On a

$$\forall h \in \mathrm{B}(0, \varepsilon),\ 0 \leq f(a+h) - f(a) = \langle \nabla f_a,\ h\rangle + o(h).$$

   On a alors, pour $\alpha \in]0, \varepsilon[$,

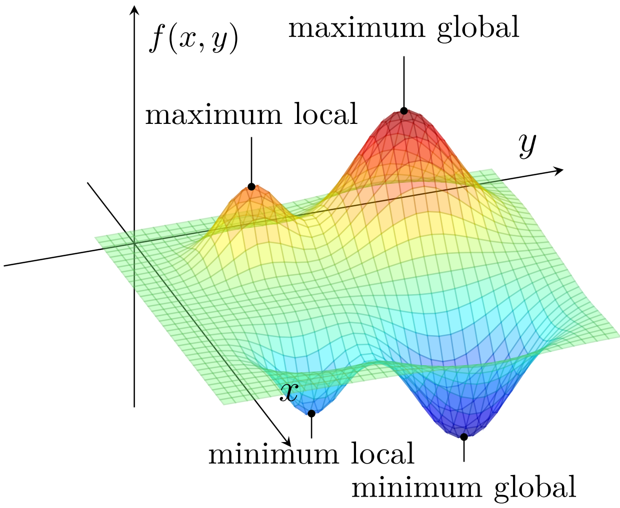
$$0 \leq \left\langle \nabla f_a,\ -\alpha \frac{\nabla f_a}{\|\nabla f_a\|} \right\rangle + o(\alpha) = -\alpha \|\nabla f_a\| + o(\alpha).$$

   En divisant des deux côtés par α et en faisant tendre α vers 0, on obtient que $-\|\nabla f_a\| \geq 0$, *i.e.* $\nabla f_a = 0$, ce qui est absurde. On a donc bien montré le résultat voulu.

   Remarquons qu'on aurait pu prouver ce second point plus rapidement en utilisant le premier point : a étant un extremum local, sa dérivée selon toutes les directions est nulle, ce qui signifie que pour tout vecteur $u \in \mathbb{R}^n$, $df_a(u) = 0$ et donc $\nabla f_a = 0$.

□

(cc) BY-NC-SA



Remarque. Bien entendu, comme sur $\mathbb{R}$, un minimum (resp. maximum) local d'une fonction convexe (resp. concave) est un minimum (resp. maximum) global.

Exercice (Régression linéaire) IV.8.

Soit $A \in \mathcal{M}_{r,n}(\mathbb{R})$, $r, n \in \mathbb{N}^*$ tels que $r \leq n$, et $B \in \mathbb{R}^r$.

1. Montrer que si A est de rang maximal (*i.e.* $\operatorname{Ker} A = \{0\}$), alors $A^T A$ est inversible.
2. On suppose que A est de rang maximal. Trouver le minimum global sur $\mathbb{R}^n$ de la fonction

$$g : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ X & \longmapsto \|AX - B\|_2^2 \,. \end{cases}$$

V Dérivées d'ordre supérieur à 2

Dans cette partie, on considère $n, p \in \mathbb{N}^*$ et on suppose que $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ sont respectivement munis de leurs normes euclidiennes canoniques.

Définition V.1.

Soit Ω un ouvert de $E = \mathbb{R}^n$, $a \in \Omega$ et f une fonction de Ω dans F. Lorsque cela a un sens, on définit récursivement la quantité suivante.

$$\forall k \in \mathbb{N}^*,\ \forall (i_1, \dots, i_k) \in [\![1; n]\!]^k,\ \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} \right).$$

Toutes les dérivées de cette forme sont appelées dérivées d'ordre k. L'ensemble des fonctions admettant des dérivées à l'ordre $m \in \mathbb{N}^*$ continues en tout point est appelé ensemble des fonctions de classe $\mathcal{C}^m$, et est noté, lorsque les ensembles de départ et d'arrivée sont respectivement Ω et F, $\mathcal{C}^m(\Omega, F)$.

En fait, on peut montrer qu'on peut librement permuter les dérivées d'ordre inférieur à $k \in \mathbb{N}^*$ pour les fonctions de classe $\mathcal{C}^k$.

BY-NC-SA

Proposition V.2.

Soit $k \in \mathbb{N}^*$ et $f \in \mathcal{C}^k(\Omega, \mathrm{F})$. Pour tout $(i_1, \ldots, i_k) \in \mathbb{N}^k$ et $\sigma \in \mathcal{S}_k$ permutation de $[\![1; k]\!]$, on a

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \ldots \partial x_{i_{\sigma(k)}}}.$$

Démonstration. Étant donné que pour tout $k \geq 2$, l'ensemble de permutations de $\mathcal{S}_k$ est généré par l'ensemble des transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ avec $i \in [\![1; k-1]\!]$, il suffit de montrer le lemme suivant.

Lemme V.3.

Pour tout $g \in \mathcal{C}^2(\Omega', \mathbb{R})$ avec $\Omega' \subset \mathbb{R}^2$ ouvert, on a

$$\frac{\partial^2 g}{\partial x_1 \partial x_2} = \frac{\partial^2 g}{\partial x_2 \partial x_1}.$$

Démonstration du lemme. Sans perte de généralité, on suppose que $(0,0) \in \Omega'$. Pour montrer le résultat voulu, on peut remplacer tout point quelconque de Ω par $(0,0)$, car la preuve est la même, et alors il suffit de montrer l'égalité

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

On va utiliser le théorème de Fubini, qui nous permet de permuter les intégrales, pour montrer qu'on peut permuter les dérivées. On pose pour tout $(x,y) \in \Omega'$,

$$\mathrm{F}(x,y) = \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) dt \right) ds \text{ et } \mathrm{G}(x,y) = \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) dt \right) ds.$$

On a alors pour tous $(x,y) \in \Omega'$,

$$\mathrm{F}(x,y) = \int_0^x \left(\frac{\partial g}{\partial x_1}(s,y) - \frac{\partial g}{\partial x_1}(s,0) \right) ds = g(x,y) - g(0,y) - g(x,0) + g(0,0).$$

$$\begin{aligned} \mathrm{G}(x,y) = \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) dt \right) ds &\underset{\text{Fubini}}{=} \int_0^y \left(\int_0^x \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) ds \right) dt \\ &= \int_0^y \left(\frac{\partial g}{\partial x_2}(x,t) - \frac{\partial g}{\partial x_2}(0,t) \right) dt \\ &= g(x,y) - g(x,0) - g(0,y) + g(0,0). \end{aligned}$$

Ces deux égalités entrainent

$$\forall (x,y) \in \Omega',\ \mathrm{F}(x,y) = \mathrm{G}(x,y) \ \textit{i.e.} \int_0^x \left(\int_0^y \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) ds \right) dt = 0.$$

Supposons maintenant par l'absurde, sans perte de généralité que

$$\frac{\partial^2 g}{\partial x_1 \partial x_2}(0,0) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(0,0) = \lambda > 0.$$

(cc) BY-NC-SA

f est de classe $\mathcal{C}^2$, donc la fonction ci-dessus est continue. Il existe donc $\eta > 0$ tel que pour tout $(s,t) \in \Omega'$, on ait

$$\|(s,t)\|_\infty < \eta \implies \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) > \frac{\lambda}{2}.$$

On en déduit donc que

$$\int_0^\eta \left(\int_0^\eta \frac{\partial^2 g}{\partial x_1 \partial x_2}(s,t) - \frac{\partial^2 g}{\partial x_2 \partial x_1}(s,t) ds \right) dt \geq \frac{\eta^2 \lambda}{2} > 0,$$

ce qui est absurde. On a donc bien démontré lemme voulu.

Revenons maintenant à la preuve du théorème. Soit $k \in \mathbb{N}^*$ et $\sigma \in \mathcal{S}_k$. On sait que l'ensemble des permutations est généré par l'ensemble des transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ avec $i \in [\![1; k-1]\!]$, il existe donc $m \in \mathbb{N}$ et $\tau_1, \ldots, \tau_m \in \mathcal{S}_k$ des transpositions de la forme $\begin{pmatrix} i & i+1 \end{pmatrix}$ telles que $\sigma = \tau_1 \circ \cdots \circ \tau_m$. Étant donné que toute dérivée d'ordre au plus $k-2$ de f est de classe $\mathcal{C}^2$ par définition, on peut affirmer d'après le lemme V.3 qu'on peut permuter les dérivées voisines c'est-à-dire pour toute transposition $\tau \in \mathcal{S}_k$ de la forme $\begin{pmatrix} d & d+1 \end{pmatrix}$ avec $d \in [\![1; k-1]\!]$,

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_{d-1}} \partial x_{i_{d+1}} \partial x_{i_d} \partial x_{i_{d+2}} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\tau(1)}} \ldots \partial x_{i_{\tau(k)}}}.$$

Enfin, en appliquant successivement les permutations, on peut affirmer que pour tous $i_1, \ldots, i_k \in [\![1; n]\!]^k$,

$$\frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{i_{\tau_1 \circ \cdots \circ \tau_m(1)}} \ldots \partial x_{i_{\tau_1 \circ \cdots \circ \tau_m(k)}}} = \frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \ldots \partial x_{i_{\sigma(k)}}}.$$

□

Conséquence. Cette propriété de permutation des dérivées a plusieurs applications intéressantes. Citons-en une qui est courante en physique de deuxième année.
Soit $g_1, g_2, g_3 \in \mathcal{C}^1(\Omega_3, \mathbb{R})$ et $g : x \longmapsto (g_1(x), g_2(x), g_3(x))$ où Ω_3 est un ouvert de $\mathbb{R}^3$. On définit l'opérateur rotationnel, comme définit en physique, de la manière suivante.

$$\text{Rot } g \underset{\substack{\uparrow \\ \text{abus de notation}}}{=} \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \wedge \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial g_3}{\partial x_2} - \frac{\partial g_2}{\partial x_3} \\ \frac{\partial g_1}{\partial x_3} - \frac{\partial g_3}{\partial x_1} \\ \frac{\partial g_2}{\partial x_1} - \frac{\partial g_1}{\partial x_2} \end{pmatrix}$$

Le fait de pouvoir permuter les variables nous permet d'avoir l'égalité suivante pour toute fonction $f \in \mathcal{C}^2(\Omega_3, \mathbb{R})$, connue au programme de physique de deuxième année.

$$\text{Rot } \nabla f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_2 \partial x_3} - \frac{\partial^2 f}{\partial x_3 \partial x_2} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} - \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} - \frac{\partial^2 f}{\partial x_2 \partial x_1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

(cc) BY-NC-SA

Exercice V.4.

Considérons $u : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}$ définie par

$$\forall (t,x) \in \mathbb{R}^+ \times \mathbb{R},\ u(t,x) = \sum_{k=0}^{\infty} \alpha_k e^{-k^2 t} \cos(kx),$$

où (α_k) est le terme général d'une série absolument convergente. Montrer que u est continue sur $[0,+\infty[\times\mathbb{R}$ et que $u \in \mathcal{C}^2(]0,+\infty[\times\mathbb{R},\mathbb{R})$.

Exercice V.5.

Déterminer toutes les fonctions $u \in \mathcal{C}^1(\mathbb{R}^n,\mathbb{R})$ telles qu'il existe $(a_1,\ldots,a_n) \in \mathbb{R}^n \setminus \{0\}$ tel que

$$\forall x \in \mathbb{R}^n,\ \sum_{i=1}^{n} a_i \frac{\partial u}{\partial x_i}(x) = 0.$$

Exercice V.6.

Soit O un ouvert de $\mathbb{R}^n$ tel que pour tout $x \in \mathrm{O}$ et $\lambda > 0$, $\lambda x \in \mathrm{O}$. On dit qu'une fonction $f : \mathrm{O} \longrightarrow \mathbb{R}$ est homogène de degré $r > 0$ si pour tout $x \in \mathbb{R}^n$, et $\lambda > 0$, $f(\lambda x) = \lambda^r f(x)$. Montrer que $f \in \mathcal{C}^1(\mathrm{O},\mathbb{R})$ est homogène de degré $r > 0$ si et seulement si

$$\forall x = (x_1,\ldots,x_n) \in \mathrm{O},\ \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(x) = r f(x).$$

(cc) BY-NC-SA

Correction de l'exercice II.4 :

1. π_i est une application linéaire en dimension finie, elle est donc différentiable. De plus, pour tout $a \in \mathbb{R}^n$, $d\pi_{i,a} = \pi_i$ et

$$\nabla\pi_{i,a} = \begin{pmatrix} \dfrac{\partial \pi_i}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial \pi_i}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial \pi_i}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial \pi_i}{\partial x_n}(a) \end{pmatrix} = \begin{pmatrix} \mathbb{1}_{i=1} \\ \vdots \\ \mathbb{1}_{i=n} \end{pmatrix} = e_i,$$

où e_i est le $i-$ème élément de la base canonique de $\mathbb{R}^n$.

2. On a pour tout $x, h \in \mathrm{E}$, en notant $\langle ., .\rangle_{\mathrm{N}}$ le produit scalaire associé à la norme euclidienne N,

$$\mathrm{N}(x+h)^2 - \mathrm{N}(x)^2 = \langle x+h,\ x+h\rangle_{\mathrm{N}} - \langle x,\ x\rangle_{\mathrm{N}} = 2\langle x,\ h\rangle_{\mathrm{N}} + \langle h,\ h\rangle_{\mathrm{N}} = 2\langle x,\ h\rangle_{\mathrm{N}} + \mathrm{N}(h)^2.$$

Toutes les normes sont équivalentes en dimension finie, il existe donc $\mathrm{C} > 0$ tel que pour tout $h \in \mathrm{E}$, $\mathrm{N}(h) \leq \mathrm{C}\,\|h\|_{\mathrm{E}}$, et alors on a $\mathrm{N}(h)^2 = o(h)$, ce qui donne

$$\mathrm{N}(x+h)^2 - \mathrm{N}(x)^2 = 2\langle x,\ h\rangle_{\mathrm{N}} + o(h),$$

et finalement pour tout $x, h \in \mathrm{E}$, $d\psi_{1,x}(h) = 2\langle x,\ h\rangle$, et $\nabla\psi_{1,x} = 2x$.

3. On travaille avec une norme d'opérateur $|||.|||$ quelconque, associée à la même norme pour l'espace de départ et l'espace d'arrivée (en dimension finie le choix de la norme importe peu). On a pour tout $\mathrm{A}, \mathrm{H} \in \mathcal{M}_n(\mathbb{R})$,

$$(\mathrm{A}+\mathrm{H})^2 - \mathrm{A}^2 = \mathrm{AH} + \mathrm{HA} + \mathrm{H}^2.$$

De plus, on a $|||\mathrm{H}^2||| \leq |||\mathrm{H}|||^2 = o(\mathrm{H})$, et alors

$$(\mathrm{A}+\mathrm{H})^2 - \mathrm{A}^2 = \mathrm{AH} + \mathrm{HA} + o(\mathrm{H}),$$

ce qui donne $d\psi_{2,\mathrm{A}}(\mathrm{H}) = \mathrm{AH} + \mathrm{HA}$.

Correction de l'exercice II.11 :

1. f^2 est la composition de $\psi : x \longmapsto x^2$ et de f. On a alors, en appliquant la proposition II.8, pour tout $a \in \Omega$, ψ est différentiable en $f(a)$ et f est différentiable en a, donc $\psi \circ f$ est différentiable en a et

$$df^2_a = d\psi \circ f_a = d\psi_{f(a)} \circ df_a = 2f(a)df_a.$$

2. $\sqrt{f}$ est la composition de $\omega : x \longmapsto \sqrt{x}$ et de f. On a alors, en appliquant la proposition II.8, pour tout $a \in \Omega$, f est différentiable en a et $f(a) > 0$ donc ω est différentiable en $f(a)$ et

$$d\omega \circ f_a = d\omega_{f(a)} \circ df_a = \frac{1}{2\sqrt{f(a)}}df_a.$$

Correction de l'exercice II.12 :

CPGE paradise

(cc) BY-NC-SA

f n'est pas différentiable en b car $x \longmapsto \|x\|_{\mathrm{E}}$ n'est pas différentiable en 0. Soit $a \in \mathrm{E} \setminus \{b\}$. On a

$$\begin{aligned}
\|a+h-b\|_{\mathrm{E}} &= \sqrt{\langle a+h-b,\ a+h-b\rangle} \\
&= \sqrt{\|a-b\|_{\mathrm{E}}^2 + 2\langle a-b,\ h\rangle + \|h\|_{\mathrm{E}}^2} \\
&= \|a-b\|_{\mathrm{E}} \sqrt{1 + \frac{2\langle a-b,\ h\rangle}{\|a-b\|_{\mathrm{E}}^2} + \frac{\|h\|_{\mathrm{E}}^2}{\|a-b\|_{\mathrm{E}}^2}} \\
&= \|a-b\|_{\mathrm{E}} \left(1 + \frac{\langle a,\ h\rangle}{\|a-b\|_{\mathrm{E}}^2} + \underbrace{\frac{\|h\|_{\mathrm{E}}^2}{2\|a-b\|_{\mathrm{E}}^2} + o\left(\frac{2\langle a,\ h\rangle}{\|a-b\|_{\mathrm{E}}^2} + \frac{\|h\|_{\mathrm{E}}^2}{\|a-b\|_{\mathrm{E}}^2}\right)}_{=o(h)}\right) \\
&= \|a-b\|_{\mathrm{E}} + \left\langle \frac{a-b}{\|a-b\|_{\mathrm{E}}},\ h\right\rangle + o(h).
\end{aligned}$$

On en déduit donc que pour tout $h \in \mathrm{E}$, $df_a(h) = \left\langle \dfrac{a-b}{\|a-b\|_{\mathrm{E}}},\ h\right\rangle$ et $\nabla f_a = \dfrac{a-b}{\|a-b\|_{\mathrm{E}}}$.

Correction de l'exercice III.3 :

1. h_1 est composition de la fonction racine et de la fonction polynomiale $(x_1, \ldots, x_n) \longmapsto (x_1-a_1)^2 + \cdots + (x_n-a_n)^2$, qui est strictement positive sur $\mathbb{R}^n \setminus \{a\}$. On en déduit que h_1 est différentiable sur $\mathbb{R}^n \setminus \{a\}$, et en utilisant la proposition 21, on peut affirmer que

$$\forall u = (u_1, \ldots, u_n) \in \mathbb{R}^n \setminus \{a\},\ \forall h \in \mathbb{R}^n,\ dh_{1,u}(h) = \langle \nabla h_{1,u},\ h\rangle,$$

où

$$\nabla h_{1,u} = \begin{pmatrix} \dfrac{\partial h_1}{\partial x_1}(u) \\ \vdots \\ \dfrac{\partial h_1}{\partial x_n}(u) \end{pmatrix} = \begin{pmatrix} \dfrac{u_1-a_1}{\sqrt{(u_1-a_1)^2+\cdots+(u_n-a_n)^2}} \\ \vdots \\ \dfrac{u_n-a_n}{\sqrt{(u_1-a_1)^2+\cdots+(u_n-a_n)^2}} \end{pmatrix} = \frac{u-a}{\|u-a\|_2}.$$

Remarquons que ce résultat est cohérent avec celui de l'exercice précédent.

2. det est différentiable sur $\mathcal{M}_n(\mathbb{R})$ car il s'agit d'une fonction polynomiale. On veut calculer la différentielle de det. Pour toute matrice $\mathrm{M} = (m_{i,j})_{i,j\in[\![1;n]\!]}$ et $i_0, j_0 \in [\![1;n]\!]$, on note $\Delta_{i_0,j_0}(\mathrm{M})$ le mineur de la matrice M d'indices (i_0, j_0), *i.e.* le déterminant de la matrice M à laquelle on a enlevé le i_0-ème ligne et la j_0-ème colonne, qui s'écrit

$$\Delta_{i_0,j_0}(\mathrm{M}) = \begin{vmatrix} m_{1,1} & \ldots & m_{1,j_0-1} & m_{1,j_0+1} & \ldots & m_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ m_{i_0-1,1} & \ldots & m_{i_0-1,j_0-1} & m_{i_0-1,j_0+1} & \ldots & m_{i_0-1,n} \\ m_{i_0+1,1} & \ldots & m_{i_0+1,j_0-1} & m_{i_0+1,j_0+1} & \ldots & m_{i_0+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ m_{n,1} & \ldots & m_{n,j_0-1} & m_{n,j_0+1} & \ldots & m_{n,n} \end{vmatrix}.$$

Soit $i_0, j_0 \in [\![1;n]\!]$. On développant par rapport à colonne j_0, on peut écrire pour toute matrice

BY-NC-SA

$\mathrm{A} = (a_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{R})$

$$\det \mathrm{A} = \sum_{i=1}^{n} \underbrace{(-1)^{i+j_0}\, a_{i,j_0} \quad \underbrace{\Delta_{i,j_0}(\mathrm{A})}_{\text{independant de } a_{i_0,j_0}}}_{\mathrm{H}_{i,j_0}(\mathrm{A})}.$$

On a alors, en notant $\mathrm{E}_{i_0,j_0} = (\delta_{i,i_0}\delta_{j,j_0})_{i,j\in[\![1;n]\!]}$ la matrice avec un 1 dans la coordonnée i_0, j_0 et 0 partout ailleurs, et

$$\frac{\partial \det}{\partial x_{i_0,j_0}}(\mathrm{A}) = \lim_{t\to 0} \frac{\det(\mathrm{A} + t\mathrm{E}_{i_0,j_0}) - \det \mathrm{A}}{t}$$

la dérivée de det par rapport à la variable de coordonnées i_0, j_0, on peut écrire

$$\frac{\partial \det}{\partial x_{i_0,j_0}}(\mathrm{A}) = \sum_{i=1}^{n}(-1)^{i+j_0}\frac{\partial \mathrm{H}_{i,j_0}}{\partial x_{i_0,j_0}}(\mathrm{A}) = (-1)^{i_0+j_0}\Delta_{i_0,j_0}(\mathrm{A}).$$

On en déduit donc que pour tout $\mathrm{H} = (h_{i,j})_{i,j\in[\![1;n]\!]} \in \mathcal{M}_n(\mathbb{R})$,

$$d\det{}_{\mathrm{A}}(\mathrm{H}) = \sum_{i,j\in[\![1;n]\!]} h_{i,j}\frac{\partial \det}{\partial x_{i,j}}(\mathrm{A}) = \sum_{i,j\in[\![1;n]\!]}(-1)^{i+j}h_{i,j}\Delta_{i,j}(\mathrm{A}) = \mathrm{Tr}(\mathrm{H}(\mathrm{Com}\ \mathrm{A})^{\mathrm{T}}),$$

et alors, on peut écrire $\nabla \det_{\mathrm{A}} = \mathrm{ComA}$.

3. Soit $\mathrm{M} \in \mathcal{M}_n(\mathbb{R})$. On a

$$(\mathrm{M}+\mathrm{H})\mathrm{A}(\mathrm{M}+\mathrm{H}) = \mathrm{MAM} + \mathrm{HAM} + \mathrm{MAH} + \mathrm{HAH}.$$

$\mathrm{HAH} = o(\mathrm{H})$ car on a $\|\mathrm{HAH}\| \leq \|\mathrm{A}\|\,\|\mathrm{H}\|^2$, et $\mathrm{H} \longmapsto \mathrm{HAM} + \mathrm{MAH}$ est linéaire, on en déduit alors que

$$h_3(\mathrm{M}+\mathrm{H}) - h_3(\mathrm{M}) = \mathrm{HAM} + \mathrm{MAH} + o(\mathrm{H}),$$

i.e. h_3 est différentiable et pour tout $\mathrm{H} \in \mathcal{M}_n(\mathbb{R})$, $dh_{3,\mathrm{M}}(\mathrm{H}) = \mathrm{MAH} + \mathrm{HAM}$.

Correction de l'exercice III.4 :

1. Soit $e_1, \ldots, e_n$ une base de E. On a pour tout $x_1, \ldots, x_k \in \mathrm{E}$, en posant pour tout $i \in [\![1;k]\!]$, $x_i = x_{i,1}e_1 + \ldots, x_{i,n}e_n$,

$$\begin{aligned} f(x_1,\ldots,x_n) &= \sum_{r_1=1}^{n} x_{1,r_1} f(e_{r_1}, x_2, \ldots, x_k) \\ &= \sum_{r_1=1}^{n}\sum_{r_2=1}^{n} x_{1,r_1}x_{2,r_2} f(e_{r_1}, e_{r_2}, x_3, \ldots, x_k) \\ &\ \vdots \\ &= \sum_{r_1=1}^{n}\sum_{r_2=1}^{n}\ldots\sum_{r_k=1}^{n} x_{1,r_1}x_{2,r_2}\ldots x_{k,r_k}\ \underbrace{f(e_{r_1}, e_{r_2}, \ldots, e_{r_k})}_{\text{indépendant de } x_1,\ldots,x_n} \end{aligned}$$

On en déduit alors que f est une fonction polynomiale (somme et produit de fonctions de la forme $(x_1, \ldots, x_n) \longmapsto x_{i,r_i}$, $i \in [\![1;n]\!]$) en les coordonnées de ses arguments en dimension finie, ce qui nous permet d'affirmer que f est continue sur E^k.

(cc) BY-NC-SA

2. Soit $x_1, \dots, x_k \in \mathrm{E}$ non nuls. On écrit

$$\|f(x_1, \dots, x_n)\|_{\mathrm{F}} = \|x_1\|_{\mathrm{E}} \dots \|x_k\|_{\mathrm{E}} \underbrace{\left\| f\left(\frac{x_1}{\|x_1\|_{\mathrm{E}}}, \dots, \frac{x_k}{\|x_k\|_{\mathrm{E}}}\right) \right\|_{\mathrm{F}}}_{\mathrm{A}}.$$

On veut montrer que A est bornée. Pour ce faire, on remarque que A est égale à f appliquée à des éléments de l'ensemble $\mathrm{S}(0,1) = \{x \in \mathrm{E},\ \|x\|_{\mathrm{E}} = 1\}$. D'après la question précédente, f est continue sur E^k et alors, elle est bornée sur le compact $\mathrm{S}(0,1)^k$ (car le produit de compacts est compact d'après le chapitre sur les compacts). En posant alors C un majorant de $(y_1, \dots, y_k) \longmapsto \|f(y_1, \dots, y_k)\|_{\mathrm{F}}$ sur $\mathrm{S}(0,1)^k$, on obtient

$$\|f(x_1, \dots, x_n)\|_{\mathrm{F}} = \|x_1\|_{\mathrm{E}} \dots \|x_k\|_{\mathrm{E}} \left\| f\left(\frac{x_1}{\|x_1\|_{\mathrm{E}}}, \dots, \frac{x_k}{\|x_k\|_{\mathrm{E}}}\right) \right\|_{\mathrm{F}} \leq \mathrm{C}\, \|x_1\|_{\mathrm{E}} \times \dots \times \|x_k\|_{\mathrm{E}}.$$

L'inégalité reste vraie lorsque $x_1, \dots, x_k$ sont potentiellement nuls.

3. Soit $a, h \in \mathrm{E}$. On veut écrire $f(a+h, \dots, a+h) - f(a, \dots, a)$ comme somme d'une application linéaire appliquée à h et un terme négligeable devant h. Lorsqu'on développe $f(a+h, \dots, a+h)$ comme un "produit" non commutatif, on obtient trois types de termes : le terme $f(a, \dots, a)$, les termes de la forme $f(a, \dots, a, h, a, \dots, a)$ où h apparaît une fois, et les termes où h apparait plus d'une fois. Les termes où h apparait une fois sont linéaires en h, il est donc naturel d'essayer de montrer que les termes où h apparait plus d'une fois sont négligeables devant h. Soit $r \in [\![2; k]\!]$. On va montrer que $f(a, \dots, a, \underbrace{h, \dots, h}_{r \text{ fois}}) = o(h)$. Une fois cela fait, la preuve pour tous les autres termes où h apparait plus d'une fois se fait exactement de la même manière.
D'après la question précédente, il existe $\mathrm{C} \geq 0$ tel que pour tout $h \in \mathrm{E}$,

$$\|f(a, \dots, a, h \dots, h)\|_{\mathrm{F}} \leq \mathrm{C}\, \|h\|_{\mathrm{E}}^r \, \|a\|_{\mathrm{E}}^{k-r} = o(h).$$

On en déduit donc que

$$f(a+h, \dots, a+h) = f(a, \dots, a) + f(h, a, \dots, a) + f(a, h, a, \dots, a) + \dots + f(a, \dots, a, h) + o(h),$$

et alors

$$dg_a(h) = f(h, a, \dots, a) + f(a, h, a, \dots, a) + \dots + f(a, \dots, a, h).$$

Enfin, lorsque f est symétrique, on a $dg_a(h) = kf(a, \dots, a, h)$.

Remarques.

→ En fait, on aurait pu faire la question 3 de l'exercice précédent en utilisant le résultat ci-dessus. En effet, pour une matrice A de taille n, considérons l'application $g : \mathrm{M} \longmapsto f(\mathrm{M}, \mathrm{M})$ où pour toutes matrices U, V de taille n, $f(\mathrm{U}, \mathrm{V}) = \mathrm{UAV}$. f est 2−linéaire, donc la différentielle de g en une matrice $\mathrm{M} \in \mathcal{M}_n(\mathbb{R})$ est égale à $\mathrm{H} \longmapsto f(\mathrm{M}, \mathrm{H}) + f(\mathrm{H}, \mathrm{M}) = \mathrm{MAH} + \mathrm{HAM}$.

→ Dans le cas où f est symétrique, la formule du binôme de Newton s'applique, *i.e.* on peut écrire pour tout $a, b \in \mathrm{E}$,

$$f(a+b, \dots, a+b) = \sum_{i=0}^{k} \binom{k}{i} f(\underbrace{a, \dots, a}_{i \text{ fois}}, \underbrace{b, \dots, b}_{k-i \text{ fois}}),$$

et dans le cas général où f n'est pas forcément symétrique,

$$f(a+b, \dots, a+b) = \sum_{\mathrm{U} \subset [\![1;k]\!]} f(\mathbb{1}_{1 \in \mathrm{U}} a + \mathbb{1}_{1 \notin \mathrm{U}} b, \dots, \mathbb{1}_{k \in \mathrm{U}} a + \mathbb{1}_{k \notin \mathrm{U}} b).$$

BY-NC-SA

→ On aurait aussi pu faire la question 2 de la manière suivante. Pour tout $x_1, \dots, x_n \in \mathrm{E}$, on a

$$\begin{aligned}
\|f(x_1, \dots, x_n)\|_{\mathrm{F}} &= \left\| \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} \dots \sum_{r_k=1}^{n} x_{1,r_1} x_{2,r_2} \dots x_{k,r_k} f(e_{r_1}, e_{r_2}, \dots, e_{r_k}) \right\|_{\mathrm{F}} \\
&\leq \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} \dots \sum_{r_k=1}^{n} |x_{1,r_1} x_{2,r_2} \dots x_{k,r_k}| \, \|f(e_{r_1}, e_{r_2}, \dots, e_{r_k})\|_{\mathrm{F}} \\
&\leq \sum_{r_1=1}^{n} \sum_{r_2=1}^{n} \dots \sum_{r_k=1}^{n} \left(\max_{r \in [\![1;n]\!]} |x_{1,r}| \right) \times \dots \times \left(\max_{r \in [\![1;n]\!]} |x_{k,r}| \right) \times \max_{m_1, \dots, m_k \in [\![1;n]\!]} \|f(e_{m_1}, e_{m_2}, \dots, e_{m_k})\|_{\mathrm{F}} \\
&= n^k \times \left(\max_{r \in [\![1;n]\!]} |x_{1,r}| \right) \times \dots \times \left(\max_{r \in [\![1;n]\!]} |x_{k,r}| \right) \times \max_{m_1, \dots, m_k \in [\![1;n]\!]} \|f(e_{m_1}, e_{m_2}, \dots, e_{m_k})\|_{\mathrm{F}} \\
&\underset{(*)}{\leq} \underbrace{\mathrm{M}^k n^k \times \left(\max_{m_1, \dots, m_k \in [\![1;n]\!]} \|f(e_{m_1}, e_{m_2}, \dots, e_{m_k})\|_{\mathrm{F}} \right)}_{\mathrm{C}} \times \|x_1\|_{\mathrm{E}} \times \dots \times \|x_k\|_{\mathrm{E}} \\
&= \mathrm{C} \, \|x_1\|_{\mathrm{E}} \times \dots \times \|x_k\|_{\mathrm{E}} \, .
\end{aligned}$$

où M est une constante positive telle que pour tout $y = y_1 e_1 + \dots + y_n e_n \in \mathrm{E}$, $\max_{r \in [\![1;n]\!]} |y_r| \leq \mathrm{M} \, \|y\|_{\mathrm{F}}$ (qui existe par équivalence des normes en dimension finie). L'inégalité $(*)$ est vraie d'après l'inégalité $\max_{r \in [\![1;n]\!]} |y_r| \leq \mathrm{M} \, \|y\|_{\mathrm{F}}$.

→ Nous encourageons le lecteur à montrer ce résultat plus général à titre d'exercice. La différentielle de la fonction $(x_1, \dots, x_k) \longmapsto f(x_1, \dots, x_n)$ en tout point $a = (a_1, \dots, a_k) \in \mathrm{E}^k$ est définie par

$$\forall (h_1, \dots, h_k) \in \mathrm{E}^k, \; df_a(h_1, \dots, h_k) = f(h_1, a_2, \dots, a_k) + f(a_1, h_2, \dots, a_k) + \dots + f(a_1, \dots, a_{k-1}, h_k).$$

Correction de l'exercice III.9 :

L'application $g : (x, y) \longmapsto x + iy$ de $\mathbb{R}^2$ dans $\mathbb{C}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ (où $\mathbb{C}$ est vu comme un $\mathbb{R}-$espace vectoriel de dimension 2). De plus, d'après le cours des séries entières, $f : z \longmapsto \sum_{k=0}^{\infty} a_k z^k$ est de classe $\mathcal{C}^1$ sur $\mathrm{B}_{\mathbb{C}}(0, \mathrm{R}) = \{z \in \mathbb{C}, \; |z| < \mathrm{R}\}$ et en posant $\mathrm{B}_{\mathbb{R}^2}(0, \mathrm{R}) = \{(x, y) \in \mathbb{R}^2, \; \sqrt{x^2 + y^2} < \mathrm{R}\}$, on a clairement $g(\mathrm{B}_{\mathbb{R}^2}(0, \mathrm{R})) = \mathrm{B}_{\mathbb{C}}(0, \mathbb{R})$ et donc $f \circ g$ est de classe $\mathcal{C}^1$ sur $\mathrm{B}_{\mathbb{R}^2}(0, \mathrm{R})$.

Correction de l'exercice IV.2 :

On pose $\Phi : t \longmapsto f((1-t)a + tb)$, qu'on définit sur l'intervalle $[0, 1]$. Φ est dérivable sur $]0, 1[$ et continue sur $[0, 1]$. On a donc, d'après l'égalité des accroissements finis,

$$\exists \theta \in]0, 1[, \; \Phi(1) - \Phi(0) = \Phi'(\theta) \underset{\substack{\uparrow \\ \text{règle de la chaine}}}{=} \left\langle \nabla f_{(1-\theta)a + \theta b}, \; b - a \right\rangle,$$

et finalement

$$f(b) - f(a) = \left\langle \nabla f_{(1-\theta)a + \theta b}, \; b - a \right\rangle.$$

(cc) BY-NC-SA

Correction de l'exercice IV.3 :

On considère la fonction

$$f : \begin{cases}]0,+\infty[^n & \longrightarrow \mathbb{R} \\ (x_1,\dots,x_n) & \longmapsto \left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}} \end{cases}.$$

En rejetant un coup d'oeil à l'inégalité qu'on veut montrer, on voit qu'elle est équivalente à

$$\forall a_1,\dots,a_n,b_1,\dots,b_n \in \mathbb{R}^+,\ f(a_1+b_1,\dots,a_n+b_n) - f(a_1,\dots,a_n) \geq f(b_1,\dots,b_n).$$

Cette forme nous faire penser au théorème des accroissements finis sur le segment $[(a_1,\dots,a_n),(a_1+b_1,\dots,a_n+b_n)]$. f est clairement différentiable sur $]0,+\infty[^n$, car il s'agit de la composition de $x \longmapsto x^{\frac{1}{n}}$ et $(x_1\dots,x_n) \longmapsto \prod_{k=1}^{n} x_k$ qui sont toutes les deux des fonctions différentiables en respectivement tout point de $]0,+\infty[$ et $]0,+\infty[^n$. On a de plus, pour tout $i \in [\![1;n]\!]$ et $x_1,\dots,x_n \in]0,+\infty[^n$,

$$\frac{\partial f}{\partial x_i}(x_1,\dots,x_n) = \frac{1}{n}\left(\prod_{\substack{k=1\\k\neq i}}^{n} x_k\right)\left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}-1} = \frac{1}{n}\frac{1}{x_i}\left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}}.$$

En appliquant l'égalité des accroissements finis montrée dans l'exercice précédent, on peut dire qu'il existe $\theta \in]0,1[$ tel qu'en posant pour tout $k \in [\![1;n]\!]$, $x_k = (1-\theta)a_k + \theta(a_k+b_k) = a_k + \theta b_k$,

$$\begin{aligned} f(a_1+b_1,\dots,a_n+b_n) - f(a_1,\dots,a_n) &= \left\langle \nabla f_{(x_1,\dots,x_n)},\ (b_1,\dots,b_n)\right\rangle \\ &= \frac{1}{n}\sum_{k=1}^{n} b_k \frac{1}{x_k}\left(\prod_{h=1}^{n} x_h\right)^{\frac{1}{n}} \underset{\substack{\uparrow\\ \text{AM-GM}}}{\geq} \left(\prod_{k=1}^{n} b_k \frac{1}{x_k}\left(\prod_{h=1}^{n} x_h\right)^{\frac{1}{n}}\right)^{\frac{1}{n}} = \left(\prod_{k=1}^{n} b_k\right)^{\frac{1}{n}} = f(b_1,\dots,b_k), \end{aligned}$$

ce qui est bien l'inégalité voulue.

Correction de l'exercice IV.6 :

L'indice fourni dans cet exercice est de calculer $\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy$. Essayons de comprendre pourquoi. Il est clair que cette intégrale est bien définie, car la fonction sous le signe intégral est intégrable. De plus, on a

$$\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy = \left(\int_0^{+\infty} e^{-x^2}dx\right)\left(\int_0^{+\infty} e^{-y^2}dx\right) = \left(\int_0^{+\infty} e^{-x^2}dx\right)^2.$$

Il est donc clair que la valeur de cette intégrale double nous donne directement la valeur de l'intégrale qu'on souhaite calculer. On va donc maintenant trouver la valeur de cette intégrale double. En la regardant de plus près, la quantité x^2+y^2 nous fait penser à un changement de variable polaire. En effet, on a, en utilisant l'application φ définie en 24,

$$\begin{aligned} \int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)}dxdy &= \int_0^{\frac{\pi}{2}}\int_0^{+\infty} e^{-r^2}\left|\det \mathrm{J}_\phi(r,\theta)\right| drd\theta \\ &= \int_0^{\frac{\pi}{2}}\int_0^{+\infty} e^{-r^2} rdrd\theta = \frac{\pi}{4}. \end{aligned}$$

(cc) BY-NC-SA

On en déduit alors finalement que

$$\int_0^{+\infty} e^{-x^2}dx = \frac{\sqrt{\pi}}{2}.$$

Correction de l'exercice IV.8 :

1. On a pour tout $X \in \mathbb{R}^n$,

$$A^T AX = 0 \implies X^T A^T AX = 0 \implies \|AX\|_2 = 0 \implies AX = 0 \implies X = 0.$$

   On en déduit que $\operatorname{Ker} A^T A = \{0\}$ et donc $A^T A$ étant une matrice carrée (de taille n), elle est inversible.

2. g est la composition d'une application affine $g_1 : X \longmapsto AX - B$ et de $g_2 : X \longmapsto \|X\|_2^2$ qui est clairement convexe (le lecteur ayant un doute peut essayer de le montrer lui-même). On en déduit donc que g est convexe. Il suffit alors de trouver où son gradient s'annule pour trouver son minimum. On a pour tout $U, H \in \mathbb{R}^n$,

$$\begin{aligned} dg_U(H) &= dg_{2,g_1(U)} \circ dg_{1,U}(H) = \left\langle \nabla g_{2,g_1(U)},\ dg_{2,U}(H) \right\rangle \\ &= 2\left\langle AU - B,\ AH \right\rangle = (AU - B)^T AH = (A^T AU - A^T B)^T H. \end{aligned}$$

   On en déduit donc que

$$dg_U = 0 \iff A^T AU - A^T B = 0 \iff U = (A^T A)^{-1} A^T B.$$

Correction de l'exercice V.4 :

Dans cet exercice, on utilise des propriétés du cours des suites et séries de fonctions. Posons pour tout $k \in \mathbb{N}$ et $(x,t) \in \mathbb{R} \times \mathbb{R}^+$, $u_k(t,x) = \alpha_k e^{-k^2 t}\cos(kx)$. On a pour tout $k \in \mathbb{N}$, $\sup\limits_{(t,x)\in\mathbb{R}^+\times\mathbb{R}} |u_k(t,x)| \leq |\alpha_k|$, donc la série de fonctions (continues) converge normalement, car $\sum |\alpha_k|$ converge. On en déduit donc que u est continue sur $\mathbb{R}^+ \times \mathbb{R}$. On a de même pour tout $(t,x) \in \mathbb{R}_+^* \times \mathbb{R}$,

$$\forall k \in \mathbb{N},\ \frac{\partial u_k}{\partial x}(t,x) = -k\alpha_k e^{-k^2 t}\sin(kt) \text{ et } \frac{\partial u}{\partial t}(t,x) = -k^2 \alpha_k e^{-k^2 t}\cos(kx).$$

On a alors pour tout $k \in \mathbb{N}$ et $\tau > 0$,

$$\begin{aligned} \sup_{(t,x)\in[\tau,+\infty[\times\mathbb{R}} \left|\frac{\partial u}{\partial x}(t,x)\right| &\leq |\alpha_k|\, k^2 e^{-k^2\tau} = o(\alpha_k), \\ \sup_{(t,x)\in[\tau,+\infty[\times\mathbb{R}} \left|\frac{\partial u}{\partial t}(t,x)\right| &\leq |\alpha_k|\, k e^{-k^2\tau} = o(\alpha_k), \end{aligned}$$

et donc les deux séries de fonctions $(t,x) \longmapsto \sum \frac{\partial u}{\partial x}(t,x)$ et $(t,x) \longmapsto \sum \frac{\partial u}{\partial t}(t,x)$ convergent normalement sur $[\tau, +\infty[$ pour tout $\tau > 0$. On en déduit donc que u est dérivable en x et en t et ses dérivées sont continues sur $[\tau, +\infty[\times\mathbb{R}$ pour tout $\tau > 0$ et par conséquent sur $]0, +\infty[\times\mathbb{R}$. On a alors bien montré que $u \in \mathcal{C}^1(]0,+\infty[\times\mathbb{R}, \mathbb{R})$. Pour montrer que u est de classe $\mathcal{C}^2$, il suffit de réappliquer le même argument que précédemment.

(cc) BY-NC-SA

Correction de l'exercice V.5 :

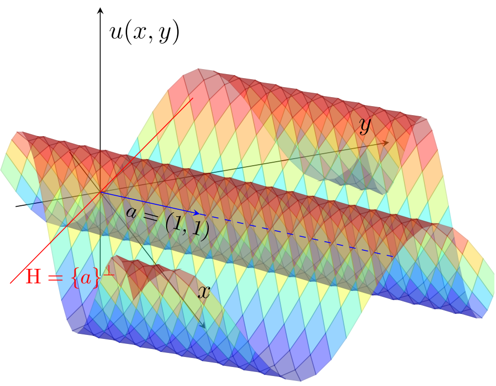
En fait, l'égalité de cet exercice peut être réécrite de la manière suivante, en posant $a = (a_1, \ldots, a_n)$, $\langle \nabla u_x,\, a\rangle = 0$. On déduit donc que localement, au voisinage de tout $x \in \mathbb{R}^n$ et pour tout $t \in \mathbb{R}$, $u(x+ta) - u(x) \simeq \langle \nabla u_x,\, ta\rangle = 0$, c'est-à-dire que u est approximativement constante dans la direction a. On va montrer qu'en fait u est constante dans la direction a, et que donc toutes ses valeurs dans $\mathbb{R}^n$ peuvent être entièrement caractérisées par ses valeurs dans l'hyperplan $\mathrm{H} = \{a\}^\perp$. Pour cela, on va considérer un élément $x \in \mathbb{R}^n$ et montrer qu'ajouter un vecteur proportionnel à a ne change pas son image par u. Soit $x \in \mathbb{R}^n$. On pose pour tout $t \in \mathbb{R}$, $v(t) = u(x + ta)$. On a, en utilisant la règle de la chaine

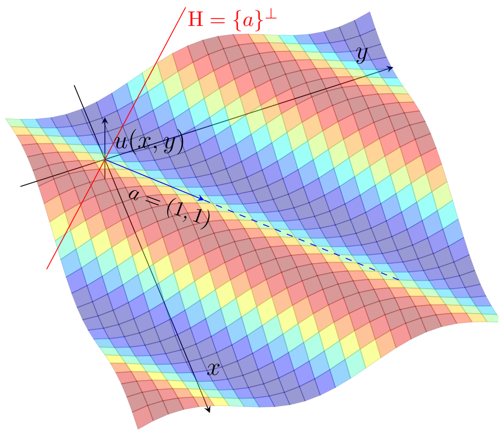
$$\forall t \in \mathbb{R},\ u'(t) = \langle \nabla u_{x+ta},\, a\rangle = 0.$$

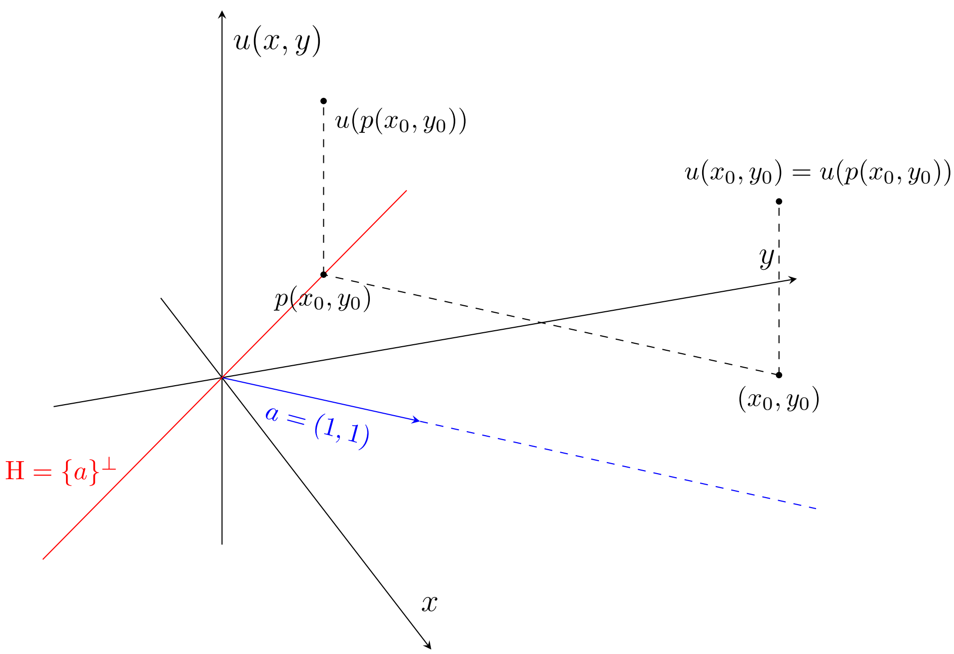
On en déduit donc que v est constante, et que u ne varie pas dans la direction de a et qu'alors si on considère $p : x \longmapsto x - \left\langle x,\, \dfrac{a}{\|a\|}\right\rangle \dfrac{a}{\|a\|}$ la projection sur $\mathrm{H} = \{a\}^\perp$ parallèlement à $\mathrm{Vect}\{a\}$, on a pour tout $x \in \mathbb{R}^n$, $u(p(x)) = u(x)$.

Montrons à présent la réciproque. Supposons que u est constante sur la direction de a, *i.e.* pour tout $x \in \mathbb{R}^n$, $v_x : t \longmapsto u(x + ta)$ est constante. On a alors pour tout $x \in \mathbb{R}^n$, $v_x'(0) = 0$, *i.e.* $\langle \nabla u_x,\, a\rangle = 0$, ce qui bien l'égalité voulue. En conclusion, l'ensemble des fonctions u qui vérifient l'égalité sont celles qui sont constantes selon la direction de a.

Pour avoir une meilleure intuition géométrique de ce qui ce passe, observons la figure ci-dessous qui illustre l'exemple en dimension 2 où $u = (x \longmapsto \sin(x - y))$ et $a = (1, 1)$.







CC BY-NC-SA

Correction de l'exercice V.6 :

Procédons par double implication.

→ (⇒) On dérive des deux côtés l'égalité $f(\lambda x) = \lambda^r f(x)$ pour tout $\lambda > 0$, et on obtient en appliquant la règle de la chaine, pour tout $x \in \mathrm{O}$,

$$\langle \nabla f_{\lambda x},\, x\rangle = r\lambda^{r-1} f(x) \ \textit{i.e.} \ \sum_{k=1}^{n} x_i \frac{\partial f}{\partial x_i}(\lambda x_i) = r\lambda^{r-1} f(x).$$

Ensuite, en évaluant en $\lambda = 1$, on obtient l'égalité désirée.

→ (⇐) Soit $x = (x_1, \dots, x_k) \in \mathrm{O}$. Posons pour tout $t > 0$, $\gamma(t) = f(tx)$. Le domaine de définition de γ est ouvert car égal à $\mathbb{R}_+^*$. γ est dérivable, car f est différentiable et pour tout $t > 0$,

$$\gamma'(t) = \langle \nabla f_{tx},\, x\rangle = \sum_{k=1}^{n} x_k \frac{\partial f}{\partial x_k}(tx) = \frac{r}{t} f(tx) = \frac{r}{t}\gamma(t).$$

On a de plus

$$\begin{aligned}\gamma'(t) = \frac{r}{t}\gamma(t) &\implies e^{-r\log t}\Big(\gamma'(t) - \frac{r}{t}\gamma(t)\Big) = 0\\ &\implies \Big(\gamma(t)e^{-r\log t}\Big)' = 0\\ &\implies \exists \mathrm{C} \in \mathbb{R},\ \gamma(t) = \mathrm{C}e^{r\log t} = \mathrm{C}t^r.\end{aligned}$$

En évaluant en $t = 1$, on obtient que $\mathrm{C} = f(x)$, et finalement, on a $f(tx) = t^r f(x)$ pour tout $t > 0$ tel que $tx \in \mathrm{O}$.

Annexe

On présente dans cette annexe un résultat intéressant de connexité pour le lecteur curieux. On rappelle que les annexes ne sont pas nécessaires pour se préparer aux concours et sont uniquement à but culturel.

Définition V.7.

Un arc (ou chemin) $\mathcal{C}^1$ par morceaux liant (ou entre) $x, y \in \mathrm{E}$ est une application $\gamma : [0,1] \to \mathrm{E}$ vérifiant les propriétés suivantes.

→ γ est continue.

→ $\gamma(0) = x$ et $\gamma(1) = y$ (ou $\gamma(0) = y$ et $\gamma(1) = x$).

→ Il existe $m \geq 2$ et $\sigma_1, \dots, \sigma_m \in [0,1]$ tels que $\sigma_0 = 0 < \sigma_1 < \dots < \sigma_n = 1$, et pour tout $i \in [\![1; m-1]\!]$ la restriction de γ à $]\sigma_i, \sigma_{i+1}[$ est de classe $\mathcal{C}^1$.

Si de plus, pour tout $i \in [\![1; m-1]\!]$, la restriction de γ à $]\sigma_i, \sigma_{i+1}[$ est affine, *i.e.* la restriction de γ à $]\sigma_i, \sigma_{i+1}[$ est une application de la forme $t \longmapsto u + tv$ où $u, v \in \mathrm{E}$, alors γ est dit être un arc (ou chemin) affine par morceaux (ou polygonal).

Vocabulaire. On dit qu'un arc γ est contenu dans un sous-ensemble S de E si $\gamma([0,1]) \subset \mathrm{S}$.

Définition V.8.

Soit S un sous-ensemble non vide de E. On dit que S est connexe par arcs $\mathcal{C}^1$ par morceaux (resp. connexe par arcs affines par morceaux) si pour tout $x, y \in \mathrm{S}$, il existe un arc de classe $\mathcal{C}^1$ par morceaux (resp. affine par morceaux) liant x et y et contenu dans S.

CC BY-NC-SA

Proposition V.9.

Soit U un ouvert non vide de E. Les propriétés suivantes sont équivalentes.

1. U est connexe.
2. U est connexe par arcs C^1 par morceaux.
3. U est connexe par arcs affines par morceaux.

Démonstration.

→ (3) ⇒ (2) ⇒ (1) Ces deux implications sont évidentes.

→ (1) ⇒ (3) On rappelle qu'ici, la connexité est définie en utilisant la topologie induite sur U *i.e.* en considérant comme espace métrique $(\mathrm{U}, d_{\mathrm{U}})$ où d_{U} désigne la restriction de d, la distance induite par la norme de $\|.\|_{\mathrm{E}}$ à U^2.

Notons que même si un fermé de U ne l'est pas forcément dans E, les ouverts de U sont bien des ouverts de E, étant donné que U est ouvert dans E. Pour $a \in \mathrm{U}$ et $r > 0$, On note

$$\mathrm{B}_{\mathrm{E}}(a, r) = \{x \in \mathrm{E},\ d(x, a) < r\}$$

la boule ouverte de centre a et de rayon r dans E, et par $\mathrm{B}_{\mathrm{U}}(a, r) = \mathrm{B}_{\mathrm{E}}(a, r) \cap \mathrm{U}$ la boule centrée en a de rayon r dans U.

Supposons désormais que U connexe. Soit $x \in \mathrm{U}$. Remarquons d'abord que pour montrer que U est connexe par arcs polygonaux, il suffit de montrer qu'on peut lier x à tout autre élément $y \in \mathrm{U}$ par un arc polygonal contenu dans U. En effet, si y et z peuvent être liés à x par des arcs γ et ψ ($\gamma(0) = y, \gamma(1) = x, \psi(0) = x, \psi(1) = z$), alors l'arc polygonal

$$\gamma \bullet \psi : \begin{cases} [0, 1] & \longrightarrow \mathrm{A} \\ t & \longmapsto \begin{cases} \gamma(2t) & \text{si } t \in [0, 1/2] \\ \psi(1 - 2t) & \text{sinon} \end{cases} \end{cases}$$

lie y et z dans u. Posons F l'ensemble des $z \in \mathrm{U}$ tel qu'on puisse lier x à z par un arc polygonal contenu dans U. On a alors,

- F est non vide. En effet, F contient x, étant donné que l'arc polygonal

$$\begin{cases} [0, 1] & \longrightarrow \mathrm{U} \\ t & \longmapsto x \end{cases}$$

  lie x à x et donc $x \in \mathrm{F}$.

- F est un ouvert de U. Soit $y \in \mathrm{F}$ et $r > 0$ assez petit pour que la boule $\mathrm{B}_{\mathrm{E}}(y, r)$ soit incluse dans U (possible vu que U est un ouvert de E). Soit $z \in \mathrm{B}_{\mathrm{E}}(y, r)$, l'arc polygonal

$$\begin{cases} [0, 1] & \longrightarrow \mathrm{U} \\ t & \longmapsto y + t(z - y) \end{cases}$$

  lie y à z et est inclus dans $\mathrm{B}(y, r) \subset \mathrm{U}$. En liant cet arc à un autre arc polygonal liant x et y (qui existe par définition de F), on obtient un arc polygonal entre x et z et alors $\mathrm{B}(y, r) \subset \mathrm{F}$. F est alors bien ouvert dans Ω.

- F est un fermé de U. Soit y un élément de l'adhérence de F dans U et $r > 0$ assez petit tel que $\mathrm{B}_{\mathrm{E}}(y, r)$ soit incluse dans U (possible, car U est un ouvert de E). Par définition de l'adhérence,

(cc) BY-NC-SA

l'intersection $\mathrm{F} \cap \mathrm{B}_{\mathrm{E}}(y, r) \neq \varnothing$. Soit $z \in \mathrm{F} \cap \mathrm{B}_{\mathrm{E}}(y, r)$. L'arc polygonal

$$\begin{cases} [0,1] & \longrightarrow \mathrm{U} \\ t & \longmapsto z + t(y-z) \end{cases}$$

lie y à z dans U, et alors en liant cet arc à un autre arc polygonal liant x et y (qui existe par définition de F), on obtient un arc polygonal entre x et z ce qui donne que $y \in \mathrm{F}$, et finalement l'adhérence de F dans Ω est bien incluse dans F.

On conclut donc par connexité de U, que $\mathrm{F} = \mathrm{U}$, ce qui est bien le résultat voulu.

□

(cc) BY-NC-SA

Chapitre 39

Réduction d'endomorphismes symétriques

Notations, rappels et remarques

→ Dans ce chapitre, on considère E un espace vectoriel euclidien de dimension $n \geq 1$ et de produit scalaire $\langle\ ,\ \rangle$ et sa norme associée $\|.\|$.

→ On pose $\mathcal{S}_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}),\ A^T = A\}$. $\mathcal{S}_n(\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ et admet la famille suivante comme base

$$\{E_{i,i}\}_{i\in[\![1;n]\!]} \cup \left\{\frac{E_{i,j}+E_{j,i}}{2}\right\}_{1\leq i<j\leq n}.$$

Étant donné le cardinal de cette base, on peut directement en déduire que $\dim \mathcal{S}_n(\mathbb{R}) = \dfrac{n(n+1)}{2}$.

→ On note $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$, *i.e.* les matrices $O \in \mathcal{M}_n(\mathbb{R})$ telles que $O^TO = I_n$.

→ Si $A, B \in \mathcal{S}_n(\mathbb{R})$, alors $AB \in \mathcal{S}_n(\mathbb{R}) \iff AB = BA$.

→ Pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on note $\mathrm{VP}(A)$ l'ensemble des valeurs propres de A.

→ Soit $\beta = (e_1, \ldots, e_n) \in E$ une base de E. On dit que β est une base orthonormée de E si pour tous $i, j \in [\![1;n]\!]$, $\langle e_i,\ e_j\rangle = \mathbb{1}_{i=j}$.

→ Une matrice $O \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si ses colonnes forment une base orthonormée de $\mathbb{R}^n$.

→ Si $\beta = (e_1, \ldots, e_n)$ est une base orthonormée quelconque de E, alors on a pour tous $x = x_1e_1 + \cdots + x_ne_n \in E$ et $y = y_1e_1 + \cdots + y_ne_n \in E$,

$$\langle x,\ y\rangle = x_1y_1 + \cdots + x_ny_n \text{ et en particulier } \|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Autrement dit, la forme du produit scalaire $\langle\ ,\ \rangle$ ne dépend pas de la base orthonormée où on regarde les vecteurs x et y. Cette propriété s'écrit également, en posant $X = [x]_\beta$ et $Y = [y]_\beta$, $\langle x,\ y\rangle = X^TY$.

→ Finalement, on reprend les notations du chapitre 29 sur la réduction d'endomorphismes.

I Généralités

Définition I.1.

Soit $u \in \mathcal{L}(E)$. On dit que u est symétrique si

$$\forall x, y \in E,\ \langle u(x),\ y\rangle = \langle x,\ u(y)\rangle.$$

On note $\mathcal{S}(E) = \{u \in \mathcal{L}(E),\ u \text{ symétrique}\}$.

Exemples.

→ Soit $p \in \mathcal{L}(E)$ un projecteur sur F parallèlement à G ($\operatorname{Ker} p = G$, $\operatorname{Im} p = F$). Alors $p \in \mathcal{S}(E)$ si et seulement si $F \perp G$. Un projecteur (qui est d'ailleurs unique) vérifiant cette propriété est appelé projecteur orthogonal sur F (parallèlement à G). La preuve de cette équivalence est laissée comme exercice au lecteur.

BY-NC-SA

→ Soit $s \in \mathcal{L}(\mathrm{E})$ une symétrie par rapport à F parallèlement à G. s est symétrique si et seulement si $\frac{1}{2}(\mathrm{Id}-s)$ est une projection orthogonale. En effet, on a

$$s \in \mathcal{S}(\mathrm{E}) \iff \frac{1}{2}(\mathrm{Id}-s) \in \mathcal{S}(\mathrm{E}) \iff \frac{1}{2}(\mathrm{Id}-s) \text{ est un projecteur orthogonal.}$$

Lorsque cette propriété est vérifiée, on dit que s est une symétrie orthogonale.

→ On peut facilement montrer que $u \in \mathcal{O}(\mathrm{E}) \cap \mathcal{S}(\mathrm{E})$ si et seulement si u est une symétrie orthogonale.

Proposition I.2.

Soit $u \in \mathcal{L}(\mathrm{E})$. Alors les propositions suivantes sont équivalentes.

1. u est symétrique.
2. Il existe une base orthonormée $\beta = (e_1, \ldots, e_n)$ de E telle que $[u]_\beta \in \mathcal{S}_n(\mathbb{R})$.
3. pour toute base orthonormée $\beta = (e_1, \ldots, e_n)$ de E, on a $[u]_\beta \in \mathcal{S}_n(\mathbb{R})$.

Démonstration.

→ (1) $\Longrightarrow$ (3) Soit $(e_1, \ldots, e_n)$ une base orthonormée de E. Posons $[u]_\beta = (a_{i,j})_{i,j \in [\![1;n]\!]}$. On a, en utilisant le fait que la base choisie est orthonormée, pour tout $i, j \in [\![1;n]\!]$,

$$a_{i,j} = \langle e_i,\, u(e_j)\rangle = \langle e_j,\, u(e_i)\rangle = a_{j,i}.$$

Ceci donne bien que $[u]_\beta \in \mathcal{S}_n(\mathbb{R})$.

→ (3) $\Longrightarrow$ (2) Cette implication est évidente.

→ (2) $\Longrightarrow$ (1) Soit $\beta = (e_1, \ldots, e_n)$ une base orthonormée de E telle que $[u]_\beta \in \mathcal{S}_n(\mathbb{R})$. Posons $[u]_\beta = (a_{i,j})_{i,j \in [\![1;n]\!]}$. On a pour tous $x = x_1e_1 + \cdots + x_ne_n \in \mathrm{E}$ et $y = y_1e_1 + \cdots + y_ne_n \in \mathrm{E}$, en posant $\mathrm{M} = [u]_\beta$, $\mathrm{X} = [x]_\beta$ et $\mathrm{Y} = [y]_\beta$

$$\langle x,\, u(y)\rangle = \mathrm{X}^\mathrm{T}\mathrm{M}\mathrm{Y} = (\mathrm{M}^\mathrm{T}\mathrm{X})^\mathrm{T}\mathrm{Y} = (\mathrm{M}\mathrm{X})^\mathrm{T}\mathrm{Y} = \langle u(x),\, y\rangle$$

Ceci donne bien que $u \in \mathcal{S}(\mathrm{E})$.

□

Corollaire. En fixant β une base orthonormée de E, on peut voir que $\mathrm{S}(\mathrm{E}) \simeq \mathcal{S}_n(\mathbb{R})$ et donc en particulier $\dim \mathcal{S}(\mathrm{E}) = \dim \mathcal{S}_n(\mathbb{R}) = \dfrac{n(n+1)}{2}$.

Remarque. Dans le cas où $u \in \mathcal{L}(\mathrm{E})$ et $\beta = (e_1, \ldots, e_n)$ une base orthonormée, il est intéressant de mémoriser les identités suivantes. En posant $[u]_\beta = (a_{i,j})_{i,j \in [\![1;n]\!]}$, et en considérant $x = x_1e_1 + \cdots + x_ne_n \in \mathrm{E}$ et $y = y_1e_1 + \cdots + y_ne_n \in \mathrm{E}$, on a

$$\langle x,\, u(y)\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_ja_{i,j} \text{ et } \langle u(x),\, y\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_ja_{j,i},$$

et de même, en posant $\mathrm{A} = [u]_\beta$, $\mathrm{X} = [x]_\beta$ et $\mathrm{Y} = [y]_\beta$, les égalités ci-dessus deviennent

$$\mathrm{X}^\mathrm{T}\mathrm{A}\mathrm{Y} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_ja_{i,j} \text{ et } (\mathrm{A}\mathrm{X})^\mathrm{T}\mathrm{Y} = \mathrm{X}^\mathrm{T}\mathrm{A}^\mathrm{T}\mathrm{Y} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_ja_{j,i}.$$

Remarquons également qu'on n'a pas supposé que u est symétrique.

(cc) BY-NC-SA

II Réduction

Théorème (Théorème spectral) II.1.

Soit $u \in \mathcal{S}(\mathrm{E})$ et soit F un sous-espace vectoriel de E.

1. Si F est stable par u alors, $\mathrm{F}^{\perp}$ aussi.
2. Si $\lambda, \mu \in \mathbb{R}$ sont deux valeurs propres distinctes de u, alors $\mathrm{E}_\lambda \perp \mathrm{E}_\mu$.
3. $\mathrm{E} = \bigoplus\limits_{\lambda \in \mathrm{Spec}_{\mathbb{R}}(u)} \mathrm{E}_\lambda$.
4. u est diagonalisable dans une base orthonormée.

Démonstration.

1. Supposons que F est stable par u. On a

$$\begin{aligned} u(\mathrm{F}) \subset \mathrm{F} &\Longrightarrow \forall (x,y) \in \mathrm{F} \times \mathrm{F}^{\perp},\ \langle u(x),\ y\rangle = 0 \\ &\Longrightarrow \forall (x,y) \in \mathrm{F} \times \mathrm{F}^{\perp},\ \langle x,\ u(y)\rangle = 0 \Longrightarrow u(\mathrm{F}^{\perp}) \subset \mathrm{F}^{\perp}. \end{aligned}$$

2. Soit λ, μ deux valeurs propres de u distinctes. Soit $(x,y) \in \mathrm{E}_\lambda \times \mathrm{E}_\mu$. On a

$$\lambda \langle x,\ y\rangle = \langle u(x),\ y\rangle = \langle x,\ u(y)\rangle = \mu \langle x,\ y\rangle .$$

Ceci nous donne directement que $\langle x,\ y\rangle = 0$ ce qui est bien le résultat voulu.

3. On envisage plusieurs méthodes, mais on expose ici que la plus classique. Deux méthodes supplémentaires sont disponibles en annexe.

Procédons par récurrence forte sur n. Le cas $n = 1$ est évident. Intéressons-nous au cas $n = 2$. Supposons que $n = 2$ et soit β une base orthonormée de E. On a, en posant $[u]_\beta = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$,

$$\chi_u = \det(\mathrm{XI}_2 - [u]_\beta) = \det \begin{pmatrix} \mathrm{X} - a & -b \\ -b & \mathrm{X} - d \end{pmatrix} = \mathrm{X}^2 - (a+d)\mathrm{X} + (ad - b^2).$$

Calculons le discriminant de ce polynôme. On a

$$\Delta = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geq 0.$$

Deux cas alors se présentent. Si $\Delta > 0$, alors χ_u possède deux racines distinctes et donc deux espaces propres de dimension 1 en somme directe, ce qui est bien le résultat voulu. Sinon, si $\Delta = 0$, alors $a = d$ et $b = 0$ ce qui donne $u = a\,\mathrm{Id}$, et donc il n'y a qu'un seul espace propre égal à l'espace tout entier. Ceci correspond également à ce qu'on cherche à démontrer.

Passons au cas général. Fixons $n \geq 2$ et supposons que pour tout $k \leq n$, la propriété est vérifiée. Supposons que $\dim \mathrm{E} = n+1$. Soit P un facteur unitaire irréductible de μ_u, $x \in \operatorname{Ker} \mathrm{P}(u) \setminus \{0\}$ et $\mathrm{F} = \mathrm{Vect}(x, u(x))$. Posons $\mathrm{P} = \mathrm{X}^2 + c_1\mathrm{X} + c_2$. On a $u^2(x) + c_1 u(x) + c_2 x = 0$ donc $u^2(x) \in \mathrm{F}$. On a également pour tout $\theta_1, \theta_2 \in \mathbb{R}$,

$$u(\theta_1 x + \theta_2 u(x)) = \theta_1 u(x) + \theta_2 u^2(x) \in \mathrm{F},$$

donc F est stable par u. D'après le point 1, on sait que $\mathrm{F}^{\perp}$ est également stable par u. Il suffit donc de considérer $w = u\Big|_{\mathrm{F}^{\perp}}^{\mathrm{F}^{\perp}} \in \mathcal{S}(\mathrm{F}^{\perp})$ et $v = u\Big|_{\mathrm{F}}^{\mathrm{F}} \in \mathcal{S}(\mathrm{F})$ et appliquer l'hypothèse de récurrence à ces deux endomorphismes pour avoir le résultat voulu.

CC BY-NC-SA

4. Il suffit de considérer pour chaque $\lambda \in \mathrm{Spec}_{\mathbb{R}}(u)$, β_λ une base orthonormée de E_λ, et de considérer $\beta = \bigcup_{\lambda \in \mathrm{Spec}_{\mathbb{R}}(u)} \beta_\lambda$. Le point 2 permet de montrer facilement que cette base est orthonormée et la matrice $[u]_\beta$ est diagonale étant donné que β est constituée de vecteurs de propres de u.

$\square$

Corollaire. Pour tout $A \in \mathcal{M}_n(\mathbb{R})$, on a

$$A \in \mathcal{S}_n(\mathbb{R}) \iff \exists O \in \mathcal{O}_n(\mathbb{R}),\ \exists \Delta \in \mathcal{D}_n(\mathbb{R}),\ A = O\Delta O^{-1} = O\Delta O^{T}.$$

Démonstration. Le sens réciproque étant évident, on se propose de vérifier le sens direct. Considérons

$$\phi : \begin{cases} \mathbb{R}^n & \to \mathbb{R}^n \\ X & \mapsto AX. \end{cases}$$

$\phi \in \mathcal{S}(\mathbb{R}^n)$ car est symétrique dans la base canonique qui est orthonormale, ce qui nous donne le résultat voulu en se rappelant qu'un changement d'une base orthonormale vers une autre est fait par une matrice orthogonale. $\square$

Exercice II.2.

Soit $A \in \mathcal{S}_n(\mathbb{R})$. Notons $\lambda_1 < \cdots < \lambda_n$ ses valeurs propres distinctes et soit $\lambda \in \mathbb{R}$ quelconque.

1. Montrer que $\mathrm{Ker}(A - \lambda I_n)^\perp = \mathrm{Im}(A - \lambda I_n) = \bigoplus_{\substack{\mu \in VP(A) \\ \mu \neq \lambda}} \mathrm{Ker}(A - \mu I)$.
2. En déduire que si $n = 2$, alors

$$\mathrm{Ker}(A - \lambda_1 I_n)^\perp = \mathrm{Im}(A - \lambda_1 I_n) = \mathrm{Ker}(A - \lambda_2 I_n) = \mathrm{Im}(A - \lambda_2 I_n)^\perp.$$

Exercice II.3.

Diagonaliser la matrice $A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$. En particulier, exhiber une base orthonormée où A est diagonale.

Exercice II.4.

Soit $A \in \mathcal{S}_n(\mathbb{R})$. Montrer que $(\mathrm{Tr}\, A)^2 \leq \mathrm{rg}(A)\, \mathrm{Tr}(A^2)$.

Exercice II.5.

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Montrer que $A^T A$ et AA^T sont équivalentes.

(cc) BY-NC-SA

III Estimations et valeurs propres pour $u \in \mathcal{S}(\mathrm{E})$

Proposition III.1.

Soit $u \in \mathcal{S}(\mathrm{E})$. Soit $\beta = (e_1, \dots, e_n)$ une base orthonormée de vecteurs propres de u telle que $[u]_\beta = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ avec $\lambda_1 \leq \dots \leq \lambda_n$. Les propriétés suivantes sont vérifiées.

1. Pour tous $s = s_1 e_1 + \dots + s_n e_n \in \mathbb{R}^n$ et $t = t_1 e_1 + \dots + t_n e_n \in \mathbb{R}^n$, on a

$$\langle u(s),\, t\rangle = \sum_{i=1}^{n} \lambda_i s_i t_i, \text{ et en particulier } \langle u(s),\, s\rangle = \sum_{i=1}^{n} \lambda_i s_i^2.$$

2. $\min\limits_{||x||=1} \langle u(x),\, x\rangle = \lambda_1$ et $\max\limits_{||x||=1} \langle u(x),\, x\rangle = \lambda_n$.

Démonstration.

1. On a

$$\langle u(s),\, t\rangle = \left\langle u\left(\sum_{i=1}^{n} s_i e_i\right),\, \sum_{i=1}^{n} t_i e_i \right\rangle = \left\langle \sum_{i=1}^{n} \lambda_i s_i e_i,\, \sum_{i=1}^{n} t_i e_i \right\rangle = \sum_{i=1}^{n} \lambda_i s_i t_i.$$

La seconde inégalité découle directement du résultat ci-dessus.

2. Soit $x = x_1 e_1 + \dots + x_n e_n \in \mathrm{E}$ tel que $\|x\| = 1$, *i.e.* $x_1^2 + \dots + x_n^2 = 1$. On a

$$\langle u(x),\, x\rangle = \sum_{i=1}^{n} \lambda_i x_i^2.$$

De plus,

$$\lambda_1 = \lambda_1 \sum_{i=1}^{n} x_i^2 \leq \sum_{i=1}^{n} \lambda_i x_i^2 \leq \lambda_n \sum_{i=1}^{n} x_i^2 = \lambda_n$$

La borne supérieure est atteinte lorsque $x = e_n$ et la borne inférieure est atteinte lorsque $x = e_1$. On en déduit donc qu'on a bien $\min\limits_{||x||=1} \langle u(x),\, x\rangle = \lambda_1$ et $\max\limits_{||x||=1} \langle u(x),\, x\rangle = \lambda_n$.

□

Exercice III.2.

On reprend les mêmes notations que ci-haut. Soit $\mu \in \mathbb{R}$. Posons

$$\mathrm{F}_\mu = \{x \in \mathrm{E},\ \langle u(x),\, x\rangle = \mu \langle x,\, x\rangle\}.$$

Montrer que F_μ est un sous-espace vectoriel de E non trivial si et seulement si $\mu = \lambda_1$ ou $\mu = \lambda_n$.

(cc) BY-NC-SA

Correction de l'exercice II.2 :

1. Soit $X \in \mathrm{Im}(A - \lambda I)$ et soit $U \in \mathbb{R}^n$ tel que $X = AU - \lambda U$. On a pour tout $Y \in \mathrm{Ker}(A - \lambda I)$
$$X^T Y = (AU - \lambda U)^T Y = U^T AY - \lambda U^T Y = U^T(AY - \lambda Y) = 0.$$
Ceci nous donne l'inclusion $\mathrm{Im}(A - \lambda I) \subset \mathrm{Ker}(A - \lambda I)^\perp$. On a de plus, d'après la formule du rang, $\dim \mathrm{Im}(A - \lambda I_n) = n - \dim \mathrm{Ker}(A - \lambda I)$ et étant donné que $\mathrm{Ker}(A - \lambda I)^\perp$ est un supplémentaire (en dimension finie) de $\mathrm{Ker}(A - \lambda I)$ dans $\mathbb{R}^n$, on a également
$$\dim \mathrm{Ker}(A - \lambda I)^\perp = n - \dim \mathrm{Ker}(A - \lambda I) = \dim \mathrm{Im}(A - \lambda I),$$
et donc finalement $\mathrm{Im}(A - \lambda I) = \mathrm{Ker}(A - \lambda I)^\perp$. Pour finir, l'égalité
$$\mathrm{Ker}(A - \lambda I)^\perp = \bigoplus_{\substack{\mu \in \mathrm{VP}(A) \\ \mu \neq \lambda}} \mathrm{Ker}(A - \mu I)$$
est une conséquence directe du point 1 du théorème II.1.
2. L'égalité de cette question est une conséquence immédiate de la question précédente.

Correction de l'exercice II.3 :

On commence d'abord pour calculer le polynôme caractéristique de A. On a

$$\begin{aligned}
\det(XI_3 - A) &= \begin{vmatrix} X-2 & -1 & 1 \\ -1 & X-2 & 1 \\ 1 & 1 & X-2 \end{vmatrix} = \begin{vmatrix} X-2 & -1 & 0 \\ -1 & X-2 & X-1 \\ 1 & 1 & X-1 \end{vmatrix} \\
&= (X-1) \begin{vmatrix} X-2 & -1 & 0 \\ -1 & X-2 & 1 \\ 1 & 1 & 1 \end{vmatrix} \\
&= (X-1) \left(- \begin{vmatrix} X-2 & -1 \\ 1 & 1 \end{vmatrix} + \begin{vmatrix} X-2 & -1 \\ -1 & X-2 \end{vmatrix} \right) \\
&= (X-1)^2 (X-4).
\end{aligned}$$

Étant donné que A est diagonalisable (car réelle symétrique), on peut affirmer que $\dim \mathrm{Ker}(I_3 - A) = 2$ (voir proposition V.6 du chapitre 29 sur la réduction d'endomorphismes).
$$(A - I_3)X = 0 \text{ où } X = \begin{pmatrix} x & y & z \end{pmatrix}^T \in \mathbb{R}^3.$$
On a
$$(A - I_3)X = 0 \iff \begin{cases} x + y - z &= 0 \\ x + y - z &= 0 \\ -x - y + z &= 0 \end{cases} \iff x + y - z = 0.$$
Ceci signifie que $\mathrm{Ker}(A - I_3)$ est l'hyperplan caractérisé par l'équation $x + y - z = 0$. Cet hyperplan est de vecteur normal $U = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}^T$. On en déduit que pour trouver une base de diagonalisation de A, il suffit de trouver une base orthonormée de $\mathrm{Ker}(A - I_3)$ et de la concaténer avec $\tilde{U} = \dfrac{U}{\|U\|}$ (car $\mathrm{Ker}(A - I_3)$ est caractérisé par l'équation $\langle X,\ U \rangle = 0$). Pour ce faire, deux approches sont possibles.

→ **Approche 1 :** Si on peut facilement trouver une famille libre de l'hyperplan, on peut l'orthonormaliser avec le procédé de Schmidt (ici pour deux vecteurs, qui est donc assez simple). On considère

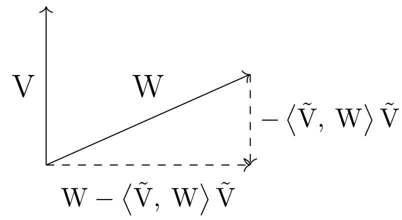
(cc) BY-NC-SA

les deux vecteurs

$$\mathrm{V} = \begin{pmatrix}1 & 0 & 1\end{pmatrix}^{\mathrm{T}} \text{ et } \mathrm{W} = \begin{pmatrix}1 & 1 & 2\end{pmatrix}^{\mathrm{T}}.$$

Il est facile de montrer que $(\mathrm{V}, \mathrm{W}) \in \mathrm{Ker}(\mathrm{A} - \mathrm{I}_3)$ et qu'il s'agit d'une famille libre. Il reste donc à orthonormaliser cette famille. Pour ce faire, on considère les transformations de V, W suivantes

$$\tilde{\mathrm{V}} = \frac{\mathrm{V}}{\|\mathrm{V}\|} \text{ et } \tilde{\mathrm{W}} = \frac{\mathrm{W} - \left\langle \tilde{\mathrm{V}},\, \mathrm{W}\right\rangle \tilde{\mathrm{V}}}{\left\|\mathrm{W} - \left\langle \tilde{\mathrm{V}},\, \mathrm{W}\right\rangle \tilde{\mathrm{V}}\right\|}.$$



Le calcul de ces quantités donne

$$\tilde{\mathrm{V}} = \begin{pmatrix}1/\sqrt{2}\\ 0\\ 1/\sqrt{2}\end{pmatrix},\ \tilde{\mathrm{W}} = \begin{pmatrix}-1/\sqrt{6}\\ \sqrt{2/3}\\ 1/\sqrt{6}\end{pmatrix} \text{ et } \tilde{\mathrm{U}} = \begin{pmatrix}1/\sqrt{3}\\ 1/\sqrt{3}\\ -1/\sqrt{3}\end{pmatrix}$$

On en déduit donc qu'en posant

$$\mathrm{O} = \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 0 & \sqrt{2/3} & 1/\sqrt{3}\\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3}\end{pmatrix},$$

$\mathrm{O}^{\mathrm{T}}\mathrm{AO}$ devrait être diagonale, et en effet, après calcul, on trouve

$$\mathrm{O}^{\mathrm{T}}\mathrm{AO} = \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 0 & \sqrt{2/3} & 1/\sqrt{3}\\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3}\end{pmatrix}^{\mathrm{T}} \begin{pmatrix}2 & 1 & -1\\ 1 & 2 & -1\\ -1 & -1 & 2\end{pmatrix} \begin{pmatrix}1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 0 & \sqrt{2/3} & 1/\sqrt{3}\\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3}\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4\end{pmatrix}.$$

→ **Approche 2 :** Lorsque trouver une base n'est pas évident, on peut procéder de la manière suivante. On commence par considérer un vecteur quelconque V de $\mathrm{Ker}(\mathrm{A}-\mathrm{I}_3)$, par exemple le même qu'avant, $\mathrm{V} = \begin{pmatrix}1 & 0 & 1\end{pmatrix}^{\mathrm{T}}$. Ensuite, on cherche un vecteur dans l'orthogonal de $\mathrm{Vect}(\mathrm{U}, \mathrm{V})$ en résolvant les équations suivantes

$$\begin{cases}\langle \mathrm{U},\, \mathrm{X}\rangle & = 0\\ \langle \mathrm{V},\, \mathrm{X}\rangle & = 0\end{cases} \quad i.e. \quad \begin{cases}x + z & = 0\\ x + y - z & = 0,\end{cases}$$

ce qui donne $\mathrm{X} = (x, y, z)^{\mathrm{T}} = (1, 2, -1)^{\mathrm{T}}$. Il suffit donc de considérer $\mathrm{W} = (1, 2, -1)$ et de normaliser la base orthonormée $(\mathrm{U}, \mathrm{V}, \mathrm{W})$, ce qui donne bien une base de diagonalisation de A.

Remarque. En utilisant l'exercice II.2, on peut remarquer qu'il y a un moyen plus rapide de trouver une base de diagonalisation de A.

$$\mathrm{Ker}(\mathrm{A} - 4\mathrm{I}) = \mathrm{Im}(\mathrm{A} - \mathrm{I}) \text{ et } \mathrm{Ker}(\mathrm{A} - \mathrm{I}) = \mathrm{Im}(\mathrm{A} - 4\mathrm{I}).$$

$\mathrm{Ker}(\mathrm{A} - 4\mathrm{I}) = \mathrm{Im}(\mathrm{A} - \mathrm{I})$ étant de dimension 1, il suffit de prendre une colonne non nulle de $\mathrm{A} - \mathrm{I}$ comme base de $\mathrm{Ker}(\mathrm{A} - 4\mathrm{I})$. De même, $\mathrm{Ker}(\mathrm{A} - \mathrm{I}) = \mathrm{Im}(\mathrm{A} - 4\mathrm{I})$ étant de dimension 2, il suffit de prendre deux vecteurs formant une famille libre des colonnes de $\mathrm{A} - 4\mathrm{I}$ pour voir une base de $\mathrm{Ker}(\mathrm{A} - 4\mathrm{I})$. Enfin, il suffit d'orthonormaliser la base qu'on a trouvé de $\mathrm{Ker}(\mathrm{A} - \mathrm{I})$, puis normaliser le vecteur formant une base de $\mathrm{Ker}(\mathrm{A} - 4\mathrm{I})$, et on obtient une base orthonormée de diagonalisation de A.

BY-NC-SA

Correction de l'exercice II.4 :

A est symétrique et est donc diagonalisable en base orthonormée. Posons donc $A = O\Delta O^T$ avec $O \in \mathcal{O}_n(\mathbb{R})$ et $\Delta = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ telle que les coefficients diagonaux nuls soient exactement $\lambda_{r+1}, \ldots, \lambda_n$. En appliquant l'inégalité Cauchy-Schwarz à $(\lambda_1, \ldots, \lambda_r)$ et $(1, \ldots, 1)$. On obtient que $\left(\sum_{i=1}^{r}\lambda_i\right)^2 \leq r\sum_{1=1}^{r}\lambda_i^2$.

Pour conclure, il suffit de remarquer les points suivants.

→ $r = \operatorname{rg} A$.

→ $\operatorname{Tr} A = \operatorname{Tr}(O\Delta O^T) = \operatorname{Tr}\Delta = \sum_{i=1}^{r}\lambda_i$.

→ $\operatorname{Tr}(A^2) = \operatorname{Tr}(O\Delta^2 O^T) = \operatorname{Tr}(\Delta^2) = \sum_{i=11}^{r}\lambda_i^2$.

Correction de l'exercice II.5 :

On a $(A^TA)^T = A^TA$ et $(AA^T)^T = AA^T$, donc ces deux matrices sont symétriques. De plus, il est facile de voir que pour toutes matrices $M_1, M_2 \in \mathcal{M}_n(\mathbb{R})$, $\chi_{M_1M_2} = \chi_{M_2M_1}$. On en déduit donc que $\chi_{A^TA} = \chi_{AA^T}$. Ces deux matrices sont symétriques, donc diagonalisables, et alors l'égalité des polynômes caractéristiques nous permet d'affirmer qu'elles ont les mêmes valeurs propres et leurs espaces propres sont de même dimension, et donc finalement que ces matrices sont équivalentes (car équivalentes à une même matrice diagonale).

Correction de l'exercice III.2 :

Soit $(e_1, \ldots, e_n)$ une base de diagonalisation de u telle que pour tout $i \in [\![1;n]\!]$, $u(e_i) = \lambda_i e_i$. Soit $x \in F_\mu \setminus \{0\}$ tel que $\|x\| = 1$ (quitte à diviser par $\|x\|$, par stabilité de F_μ par multiplication par un scalaire). On a d'après la proposition III.1,

$$\lambda_1 \leq \underbrace{\langle u(x),\ x\rangle}_{=\mu\langle x,\, x\rangle = \mu} \leq \lambda_n.$$

On a donc nécessairement $\mu \in [\lambda_1, \lambda_n]$, sinon $F_\mu = \{0\}$ (il est facile de vérifier que $0 \in F_\mu$). De plus, si $\lambda_1 < \mu < \lambda_n$, F_μ n'est pas un espace vectoriel. Supposons que F_μ est un espace vectoriel et considérons l'espace vectoriel $G = F_\mu \cap \operatorname{Vect}(e_1, e_n)$. Soit $s = s_1e_1 + s_ne_n \in G$. On a

$$\begin{aligned} s \in G &\iff \mu s_1^2 + \mu s_2^2 = \mu\,\langle s,\ s\rangle = \langle u(x),\ x\rangle = \lambda_1 s_1^2 + \lambda_2 s_2^2 \\ &\iff s_2 = \sqrt{\frac{\lambda_n - \mu}{\mu - \lambda_1}}s_1 \text{ ou } s_2 = -\sqrt{\frac{\lambda_n - \mu}{\mu - \lambda_1}}s_1. \end{aligned}$$

On en déduit que

$$G = \left\{s = s_1e_1 + s_2e_2 \in E,\ s_2 = \sqrt{\frac{\lambda_n - \mu}{\mu - \lambda_1}}s_1\right\} \cup \left\{s = s_1e_1 + s_2e_2 \in E,\ s_2 = -\sqrt{\frac{\lambda_n - \mu}{\mu - \lambda_1}}s_1\right\}.$$

G est donc l'union de deux espaces vectoriels distincts et non inclus l'un dans l'autre, donc G n'est pas un espace vectoriel, ce qui est absurde.

Finalement, regardons le cas où $\mu = \lambda_1$ (le cas où $\mu = \lambda_n$ se traite de la même manière). Supposons que $\mu = \lambda_1$. On a pour tout $x = x_1e_1 + \cdots + x_ne_n \in F_\mu$,

$$\lambda_1\sum_{i=1}^{n}x_i^2 = \lambda_1\,\langle x,\ x\rangle = \langle u(x),\ x\rangle = \sum_{i=1}^{n}\lambda_i x_i^2 \ \text{ i.e. } \sum_{i=1}^{n}\underbrace{(\lambda_i - \lambda_1)x_i^2}_{\geq 0} = 0.$$

On en déduit donc finalement que pour tout $i \in [\![1;n]\!]$ tel que $\lambda_i \neq \lambda_1$, $x_i = 0$, i.e. $x \in E_{\lambda_1,u}$. Ceci implique

(cc) BY-NC-SA

que $F_\mu \subset E_{\lambda_1,u}$. L'inclusion réciproque est évidente, ce qui nous donne que F_μ est un espace vectoriel non trivial. On a donc bien montré le résultat voulu.

Annexe

Nous exhibons ici deux méthodes alternatives pour montrer le point 3 du théorème spectral (Théorème II.1). Cette annexe est à but purement culturel. Nous conseillons donc au lecteur de l'aborder uniquement s'il est à l'aise avec les notions du cours et s'il a le temps de le faire.

→ **Méthode alternative 1 (espaces hermitiens) :** En fait, en examinant la preuve de ce résultat fournie, on peut facilement voir que si on exhibe un sous-espace vectoriel F de E de dimension 1 qui est stable par u, alors on pouvait se passer de montrer la propriété pour $n = 2$. Pour ce faire, on va montrer que u admet une valeur propre réelle. On va passer aux matrices. Soit β une base orthonormée de E. On pose $A = [u]_\beta$. On sait que (voir le chapitre 29 sur la réduction) A admet une valeur propre complexe, qu'on note λ. Fixons $X \in \mathbb{C}^n \setminus \{0\}$ quelconque tel que $AX = \lambda X$. On a

$$\overline{\lambda}||X||^2 = \overline{\lambda \overline{X^T} X} = \overline{(\overline{X^T} A X)^T} = \overline{X^T}\ \overline{A^T} X = \overline{X^T} A X = \lambda \overline{X^T} X = \lambda ||X||^2.$$

On a donc $\lambda \in \mathbb{R}$. L'idée qui vient à l'esprit consiste à considérer l'espace

$$F = \mathrm{Vect}_{\mathbb{R}}(X) = \{\lambda X,\ \lambda \in \mathbb{R}\}.$$

Cependant, on a à priori $X \in \mathbb{C}^n$, ce qui fait que cet espace n'est pas valable, car F doit être inclus dans $\mathbb{R}^n$. On propose deux manières pour éviter ce problème.

- **Possibilité 1 :** Posons $X = \mathrm{Re}(X) + i\,\mathrm{Im}(X)$ (où Re(X) et Im(X) sont respectivement les vecteurs dont les coordonnées sont les parties réelles et imaginaires de celles de X). On a $AX = \lambda X$ et donc

$$\begin{aligned} A\,\mathrm{Re}(X) + iA\,\mathrm{Im}(X) &= \lambda\,\mathrm{Re}(X) + i\lambda\,\mathrm{Im}(X) \\ \textit{i.e. } A\,\mathrm{Re}(X) = \lambda\,\mathrm{Re}(X) &\text{ et } A\,\mathrm{Im}(X) = \lambda\,\mathrm{Im}(X). \end{aligned}$$

  L'un des deux vecteurs Re(X) ou Im(X) est non nul, et est donc un vecteur propre de A associé à λ. Il suffit alors de considérer $\mathrm{Vect}_{\mathbb{R}}(\mathrm{Re}(X))$ ou $\mathrm{Vect}_{\mathbb{R}}(\mathrm{Im}(X))$ selon les cas.
- **Possibilité 2 :** Rappelons que le rang d'une matrice réelle est le même sur $\mathbb{R}$ et $\mathbb{C}$. Plus exactement, on a le lemme suivant.

Lemme III.3.

Si $\mathbb{K}, \mathbb{L}$ sont deux corps tels que $\mathbb{K} \subset \mathbb{L}$ et $A \in \mathcal{M}_n(\mathbb{K})$, alors $\mathrm{rg}_{\mathbb{K}} A = \mathrm{rg}_{\mathbb{L}} A$.

Démonstration. Posons $r = \mathrm{rg}_{\mathbb{K}} A$. A est équivalente à la matrice

$$J_{r,n} = \mathrm{diag}(\underbrace{1,\dots,1}_{r \text{ fois}}, \underbrace{0,\dots,0}_{n-r \text{ fois}}).$$

On peut donc écrire $A = PJ_{r,n}Q$ avec $P, Q \in GL_n(\mathbb{K})$. Cette écriture reste valable sur $\mathbb{L}$ et donc, sur $\mathbb{L}$, A est toujours équivalente à la matrice $J_{r,n}$. Par conséquent, $rg_{\mathbb{L}} A = r$ □

Grâce à ce lemme, on a $rg_{\mathbb{R}}(A - \lambda I_n) = rg_{\mathbb{C}}(A - \lambda I_n) < n$ et donc λ est bien valeur propre de A, vue comme matrice réelle. On dispose, en particulier, d'un vecteur propre $X \in \mathbb{R}^n \setminus \{0\}$ de A associé à λ et $\mathrm{Vect}_{\mathbb{R}}(X)$ convient donc.

CC BY-NC-SA

→ **Méthode alternative 2 (calcul différentiel et optimisation) :** Cette partie utilise quelques outils du chapitre du calcul différentiel. Le lecteur n'ayant pas encore abordé ce chapitre est encouragé à n'aborder cette méthode seulement après avoir maitrisé le chapitre en question.

Toujours dans le même esprit que la méthode alternative précédente, on expose un autre moyen de montrer que u admet une valeur propre réelle. On se propose de le montrer de manière analytique cette fois. On considère l'application

$$\phi : \begin{cases} \mathrm{E} & \to \mathbb{R} \\ x & \mapsto \langle x,\ u(x)\rangle \,. \end{cases}$$

$\phi \in \mathcal{C}^1(\mathrm{E}, \mathbb{R})$ (car c'est une application polynomiale) sur le compact $\mathrm{S}(0,1) = \{x \in \mathrm{E},\ \|x\| = 1\}$ et donc atteint son maximum en $x_0 \in \mathrm{S}(0,1)$.

$\mathrm{S}(0,1)$ admettant $x_0 + \mathrm{Vect}(x_0)^\perp$ comme espace tangent (affine) en x_0, on a alors que $\nabla\phi_{x_0}$ est perpendiculaire à $\mathrm{Vect}(x_0)^\perp$ et est donc de la forme $\nabla\phi_{x_0} = \alpha x_0$. Pour finir, remarquons que

$$\phi(x_0 + h) - \phi(x_0) = \langle h,\ u(x_0)\rangle + \langle x_0,\ u(h)\rangle + \underbrace{\langle h,\ u(h)\rangle}_{\mathrm{O}(||h||^2)} \underset{u \in \mathrm{S(E)}}{=} \langle 2u(x_0),\ h\rangle + \underbrace{\langle h,\ u(h)\rangle}_{\mathrm{O}(||h||^2)}$$

et que donc $\nabla\phi_{x_0} = 2u(x_0) = \alpha x_0$ d'où x_0 est un vecteur propre de u de valeur propre (réelle) associée $\dfrac{\alpha}{2}$.

(cc) BY-NC-SA

CHAPITRE 42

Endomorphismes d'un espace hermitien

Dans tout ce chapitre, $(\mathrm{E}, \langle \cdot\,, \ \cdot \rangle)$ désigne un espace hermitien de dimension $n \geq 1$.
Exemples.

→ Sur $\mathbb{C}^n$: pour tous $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n$, $\langle x,\ y \rangle = \sum_{k=1}^{n} \overline{x_k} y_k$.

→ Sur $\mathscr{T}_n$: pour tous $\mathrm{P}, \mathrm{Q} \in \mathscr{T}_n$, $\langle \mathrm{P},\ \mathrm{Q} \rangle = \dfrac{1}{2\pi} \displaystyle\int_0^{2\pi} \overline{\mathrm{P}(t)} \mathrm{Q}(t) \mathrm{d}t$.

Si, pour tout $k \in [\![-n; n]\!]$, l'application $x \mapsto e^{ikx}$ est notée e_k, rappelons que $\mathscr{T}_n = \mathrm{Vect}(e_{-n}, \dots, e_n)$.

I Adjonction matricielle

Définition I.1.

Soit $\mathrm{A} \in \mathcal{M}_n(\mathbb{C})$. On appelle *matrice adjointe de* A, et on note A^*, la matrice $\overline{\mathrm{A}}^\top$.

Proposition I.2.

Pour toutes $\mathrm{A}, \mathrm{B} \in \mathcal{M}_n(\mathbb{C})$, et pour tout $\lambda \in \mathbb{C}$,

→ $(\mathrm{A} + \mathrm{B})^* = \mathrm{A}^* + \mathrm{B}^*$, $(\lambda \mathrm{A})^* = \overline{\lambda} \mathrm{A}^*$, et $(\mathrm{AB})^* = \mathrm{B}^* \mathrm{A}^*$.

→ $\det(\mathrm{A}^*) = \overline{\det(\mathrm{A})}$, et $\chi_{\mathrm{A}^*} = \overline{\chi_{\mathrm{A}}}$.

→ $\mathrm{rg}(\mathrm{A}^*) = \mathrm{rg}(\mathrm{A})$. En effet, $\mathrm{rg}(\mathrm{A}^*) = \mathrm{rg}(\overline{\mathrm{A}}) = \mathrm{rg}(\mathrm{A})$.

Si un polynôme P s'écrit $\sum_{k=0}^{n} a_k \mathrm{X}^k$, alors $\overline{\mathrm{P}} = \sum_{k=0}^{n} \overline{a_k} \mathrm{X}^k$

Matrices hermitiennes : Une matrice carrée est hermitienne si elle est égale à sa matrice adjointe.

Proposition I.3.

→ L'ensemble des matrices hermitiennes d'ordre n

$$\mathcal{H}_n(\mathbb{C}) = \{\mathrm{A} \in \mathcal{M}_n(\mathbb{C}) \mid \mathrm{A}^* = \mathrm{A}\}$$

est un $\mathbb{R}$-espace vectoriel de dimension n^2.

→ Si $\mathrm{A} \in \mathcal{H}_n(\mathbb{C})$, alors $\chi_{\mathrm{A}} \in \mathbb{R}[\mathrm{X}]$.

Démonstration.

→ Soit $\mathrm{A} = \mathrm{B} + i\mathrm{C} \in \mathcal{M}_n(\mathbb{C})$ telle que B et $\mathrm{C} \in \mathcal{M}_n(\mathbb{R})$. On a les équivalences suivantes :

$$\mathrm{A} \in \mathcal{H}_n(\mathbb{C}) \iff \mathrm{A}^* = \mathrm{A} \iff \mathrm{B}^\top = \mathrm{B} \quad \text{et} \quad \mathrm{C}^\top = -\mathrm{C},$$

donc $\mathcal{H}_n(\mathbb{C})$ est isomorphe $\mathcal{S}_n(\mathbb{R}) \times \mathcal{A}_n(\mathbb{R})$. Donc $\dim(\mathcal{H}_n(\mathbb{C})) = \dfrac{n(n+1)}{2} + \dfrac{n(n-1)}{2} = n^2$.

→ Si $\mathrm{A} \in \mathcal{H}_n(\mathbb{C})$, alors $\chi_{\mathrm{A}} = \chi_{\mathrm{A}^*} = \overline{\chi_{\mathrm{A}}}$.

CC BY-NC-SA

Matrices unitaires : Les matrices unitaires sont les matrices de l'ensemble

$$\mathcal{U}_n(\mathbb{C}) = \{\mathrm{U} \in \mathcal{M}_n(\mathbb{C}) \mid \mathrm{U}^*\mathrm{U} = \mathrm{I}_n\}.$$

Proposition I.4.

→ $\mathcal{U}_n(\mathbb{C})$ est un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$.

→ $\mathrm{U} \in \mathcal{U}_n(\mathbb{C}) \iff$ les colonnes de U forment une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{C})$
$\iff$ les lignes de U forment une base orthonormée de $\mathcal{M}_{1,n}(\mathbb{C})$.

→ $\mathrm{U} \in \mathcal{U}_n(\mathbb{C})$ si, et seulement si, il existe deux bases orthonormée $\mathscr{B}$ et $\mathscr{B}'$ de E telles que U soit la matrice de la famille $\mathscr{B}'$ relativement à la base $\mathscr{B}$.

Démonstration de la deuxième proposition. Soit $\mathrm{U} = (u_{k,\ell})_{1\leq k,\ell\leq n} \in \mathcal{U}_n(\mathbb{C})$. Alors pour tous $p, q \in [\![1; n]\!]$, $(\mathrm{U}^*\mathrm{U})(p,q) = \sum_{k=1}^{n} \overline{u_{k,p}} u_{k,q}$, donc pour $p \neq q$, $\sum_{k=1}^{n} \overline{u_{k,p}} u_{k,q} = 0$ et pour $p = q$, $\sum_{k=1}^{n} |u_{k,p}|^2 = 1$.
Donc, les colonnes de U forment une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{C})$.
De même, en examinant les coefficients de UU^*, il vient que les lignes de U forment une base orthonormée de $\mathcal{M}_{1,n}(\mathbb{C})$. Les réciproques sont immédiates.

Matrices normales Une matrice M de $\mathcal{M}_n(\mathbb{C})$ est dite normale si $\mathrm{AA}^* = \mathrm{A}^*\mathrm{A}$.

II Adjonction dans $\mathscr{L}(\mathrm{E})$

Théorème - définition II.1.

Soit $u \in \mathscr{L}(\mathrm{E})$. Alors, il existe un unique endomorphisme v de E tel que :

$$\forall (x,y) \in \mathrm{E}^2, \qquad \langle u(x),\ y\rangle = \langle x,\ v(y)\rangle\,.$$

De plus, la matrice de v relativement à toute base orthonormée $\mathscr{B}$ de E est la matrice adjointe de la matrice de u relativement à la même base $\mathscr{B}$, et v est appelé *endomorphisme adjoint* de u et est noté u^*.

Démonstration. Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base orthonormée de E, et $\mathrm{A} = (a_{k,\ell})_{1\leq k,\ell\leq n}$ la matrice de u relativement à $\mathscr{B}$. Posons v tel que sa matrice relativement à $\mathscr{B}$ soit A^*.
Alors, pour tous $x = x_1e_1 + \cdots + x_ne_n$, $y = y_1e_1 + \cdots + y_ne_n \in \mathrm{E}$,

$$\langle u(x),\ y\rangle = \left\langle \sum_{\ell=1}^{n} x_\ell u(e_\ell),\ \sum_{\ell=1}^{n} y_\ell e_\ell \right\rangle = \left\langle \sum_{\ell=1}^{n} x_\ell \left(\sum_{k=1}^{n} a_{k,\ell} e_k\right),\ \sum_{j=1}^{n} y_j e_j \right\rangle \underset{\substack{\uparrow \\ \text{on conserve les termes où } k=j}}{=} \sum_{1\leq \ell,k\leq n} \overline{x_\ell a_{k,\ell}} y_k,$$

et

$$\langle x,\ v(y)\rangle = \left\langle \sum_{j=1}^{n} x_j e_j,\ \sum_{\ell=1}^{n} y_\ell \left(\sum_{k=1}^{n} \overline{a_{\ell,k}} e_k\right)\right\rangle = \sum_{1\leq \ell,k\leq n} \overline{x_k} y_\ell \overline{a_{\ell,k}} = \langle u(x),\ y\rangle\,.$$

De plus, si un endomorphisme w, de matrice notée $(w_{j,k})_{1\leq j,k\leq n}$ relativement à la base $\mathscr{B}$, vérifie la même identité, alors pour tous $j, k \in [\![1; n]\!]$, $w_{j,k} = \langle e_j,\ w(e_k)\rangle = \langle u(e_j),\ e_k\rangle = \overline{a_{k,j}}$, donc $w = v$.

BY-NC-SA

Proposition II.2.

Soit F un sous-espace vectoriel de E, et u un endomorphisme de E.
Si F est stable par u, alors $F^\perp$ est stable par u^*.

Démonstration. Si $y \in F^\perp$, alors, pour tout $x \in F$, alors $0 = \langle u(x),\ y\rangle = \langle x,\ u^*(y)\rangle$, donc $u^*(y) \in F^\perp$.

Proposition II.3.

Soit u un endomorphisme de E.

1. Les trois assertions suivantes sont équivalentes :
   - → u est hermitien *i.e.* $u = u^*$.
   - → Il existe une base orthonormée de E telle que la matrice de u relativement à cette base soit hermitienne.
   - → Pour toute base orthonormée $\mathscr{B}$ de E, la matrice de u relativement à $\mathscr{B}$ est hermitienne.
2. Les trois assertions suivantes sont également équivalentes :
   - → u est unitaire *i.e.* pour tous $x, y \in E$, $\langle u(x),\ u(y)\rangle = \langle x,\ y\rangle$.
   - → Il existe une base orthonormée de E telle que la matrice de u relativement à cette base soit unitaire.
   - → Pour toute base orthonormée $\mathscr{B}$ de E, la matrice de u relativement à $\mathscr{B}$ est unitaire.

Démonstration. La première chaîne d'équivalences est immédiate car la matrice de u relativement à une base orthonormée $\mathscr{B}$ de E (et il en existe) est la matrice adjointe de la matrice de u^* relativement à $\mathscr{B}$.
Pour la deuxième chaîne d'équivalences, il convient de remarquer que u est unitaire si, et seulement si, pour tout $(x, y) \in E^2$, $\langle x,\ (u^* \circ u)(y)\rangle = \langle x,\ y\rangle$, donc si, et seulement si, $u^* \circ u = \mathrm{Id}_E$ (car pour tout $x \in E$, $\langle x,\ (u^* \circ u)(x)\rangle = \langle x,\ x\rangle$).
Notation. On note $\mathcal{U}(E)$ les endomorphismes unitaires de E.

Exercice II.4.

Soit $(E, \langle\cdot\,,\ \cdot\rangle)$ un espace hermitien de dimension $n \geq 1$ et on note $\|\cdot\|$ la norme associée au produit scalaire $\langle\cdot\,,\ \cdot\rangle$. Montrer que $\mathcal{U}(E)$ est compact.

Théorème II.5.

Soit $u \in \mathscr{L}(E)$.

- → u est normal si, et seulement si, u est diagonalisable dans une base orthonormée de E.
- → u est hermitien si, et seulement si, u est diagonalisable dans une base orthonormée de E et $\mathrm{Sp}(u) \subset \mathbb{R}$.
- → u est unitaire si, et seulement si, u est diagonalisable dans une base orthonormée de E et $\mathrm{Sp}(u) \subset \mathbb{U}$.

(cc) BY-NC-SA

Correction de l'exercice II.4. :

Montrons que $\mathcal{U}_n(\mathbb{C})$ est un compact de $\mathcal{M}_n(\mathbb{C})$. Notons $\|\cdot\|$ la norme associée au produit scalaire canonique sur $\mathcal{M}_n(\mathbb{C})$. Pour toute $\mathrm{U} \in \mathcal{U}_n(\mathbb{C})$, $\|\mathrm{U}\| = 1$, donc $\mathcal{U}_n(\mathbb{C})$ est borné.
De plus, $\mathcal{U}_n(\mathbb{C})$ est fermé dans $\mathcal{M}_n(\mathbb{C})$ en tant qu'image réciproque de $\{0\}$, fermé dans $\mathbb{R}$, par l'application continue $\mathrm{U} \mapsto \mathrm{U}^*\mathrm{U} - \mathrm{I}_n$. Donc $\mathcal{U}_n(\mathbb{C})$ est compact.
Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base orthonormée de E. Introduisons l'application de $\mathcal{M}_n(\mathbb{C})$ à valeurs dans $\mathscr{L}(\mathrm{E})$

$$\varphi : \mathrm{M} = (m_{j,k})_{1 \leq j,k \leq n} \longmapsto \left(x = \sum_{j=1}^{n} x_j e_j \longmapsto \sum_{j=1}^{n} x_i \left(\sum_{k=1}^{n} m_{k,j} e_j \right) \right).$$

Cette application est linéaire donc continue car $\mathcal{M}_n(\mathbb{C})$ est de dimension finie. Il est clair que $\mathcal{U}(\mathrm{E})$ est l'image par φ de $\mathcal{U}_n(\mathbb{C})$, qui est compact, donc $\mathcal{U}(\mathrm{E})$ est compact.

* *
*

(cc) BY-NC-SA