Mercredi 10 juin

Référence	Origine	Thèmes
130-319	X MP	Équivalent de la somme d'une série entière
130-469	Mines MP	Décomposition d'une permutation
130-470	"	Morphismes de groupes
130-471	"	Groupe cyclique
130-483	"	Vers l'application det
130-498	"	Réduction d'un endomorphisme
130-511	"	Vers la co-trigonalisation
130-516	"	Sous-espaces stables par un endomorphisme
130-529	"	Matrices symétriques et polynômes
130-532	"	Similitude de A et de $A^\top$
130-541	"	Classe de similitude d'une matrice
130-544	"	Compacité et point fixe
130-566	"	Zéros des solutions d'une équation différentielle
130-574	"	Calcul d'une intégrale (archi-classique)
130-600	"	Inégalité de Gronwall
130-601	"	Étude qualitative d'une équation différentielle
130-624	"	Variables aléatoires indépendantes
130-630	"	Cardinal d'un ensemble aléatoire
135-375	X-MP	Espace euclidien et probabilités

rms130-319

Soit $(u_n)_{n\in\mathbb{N}}$, une suite complexe de limite nulle. On pose

$$f(t)=\sum_{n=0}^{+\infty}u_nt^n.$$

[**1.**] *Démontrer que f est définie sur $]-1,1[$.*
[**2.**] *Démontrer que*

$$\lim_{t\to1^-}(1-t)f(t)=0.$$

[**1.**] Comme la suite $(u_n)_{n\in\mathbb{N}}$ tend vers 0, alors

$$\forall\, t\in\,]-1,1[,\qquad u_nt^n\underset{n\to+\infty}{=}o(t^n)$$

et comme la série géométrique $\sum t^n$ est absolument convergente pour tout $t\in\,]-1,1[$, la série $\sum u_nt^n$ est (absolument) convergente. Donc la somme f est définie (au moins) sur $]-1,1[$.
[**2.**] Si le rayon de convergence R de la série entière $\sum u_nt^n$ est strictement supérieur à 1, alors la somme f est continue sur $]-R,R[$ et en particulier continue en 1. Le résultat est alors évident.

✍ *La démonstration ci-dessous suit les grandes lignes de la démonstration du Théorème de Cesaro.*

❧ Soit $\varepsilon>0$. Il existe un rang $N_\varepsilon\in\mathbb{N}$ tel que

$$\forall\, n\geqslant N_\varepsilon,\qquad |u_n|\leqslant\varepsilon.$$

Par conséquent, pour tout $0<t<1$,

$$\big|f(t)\big|\leqslant\Big|\sum_{k=0}^{N_\varepsilon}u_kt^k\Big|+\varepsilon\sum_{k=N_\varepsilon+1}^{+\infty}t^n\leqslant\Big|\sum_{k=0}^{N_\varepsilon}u_kt^k\Big|+\frac{\varepsilon}{1-t}$$

donc

$$\forall\, 0<t<1,\qquad \big|(1-t)f(t)\big|\leqslant(1-t)\Big|\sum_{k=0}^{N_\varepsilon}u_kt^k\Big|+\varepsilon.$$

❧ L'expression étant une fonction continue de $t\in\mathbb{R}$, il est clair que

$$\lim_{t\to1}(1-t)\Big|\sum_{k=0}^{N_\varepsilon}u_kt^k\Big|=0.$$

Par conséquent, il existe un réel $0<\alpha<1$ tel que

$$\forall\, 1-\alpha<t<1,\qquad 0\leqslant(1-t)\Big|\sum_{k=0}^{N_\varepsilon}u_kt^k\Big|\leqslant\varepsilon.$$

❧ On a ainsi démontré que, pour tout $\varepsilon>0$, il existait un réel $0<\alpha<1$ tel que

$$\forall\, t\in\,]1-\alpha,1[,\qquad \big|(1-t)f(t)\big|\leqslant2\varepsilon.$$

Autrement dit :

$$\lim_{t\to1^-}(1-t)f(t)=0.$$

✍ ***Variante.*** Merci VO-D!
Développons le produit :

$$\forall\, 0\leqslant t<1,\qquad (1-t)f(t)=\sum_{n=0}^{+\infty}u_nt^n-\sum_{n=0}^{+\infty}u_nt^{n+1}=\sum_{n=0}^{+\infty}u_nt^n-\sum_{n=1}^{+\infty}u_{n-1}t^n$$
$$=u_0-\sum_{n=1}^{+\infty}(u_n-u_{n-1})t^n.$$

Comme la suite $(u_n)_{n\in\mathbb{N}}$ est convergente, la série télescopique $\sum(u_n - u_{n-1})$ est convergente. D'après le Théorème d'Abel radial, la somme de la série entière $\sum(u_n - u_{n-1})t^n$ est continue sur $[0,1]$ et en particulier

$$\lim_{t\to 1^-} \sum_{n=1}^{+\infty} (u_n - u_{n-1})t^n = \sum_{n=1}^{+\infty} (u_n - u_{n-1}) = -u_0.$$

Par conséquent,

$$\lim_{t\to 1^-} (1-t)f(t) = u_0 - u_0 = 0.$$

❧ *Cette démonstration est beaucoup plus courte que la précédente car la démonstration du Théorème d'Abel radial contient* beaucoup *d'informations.*

rms130-469

Soient $n \in \mathbb{N}^*$ et $\gamma \in \mathfrak{S}_n$, un cycle. Écrire γ comme un produit de transpositions.

✍ *Une chose est de savoir comment le groupe symétrique $\mathfrak{S}_n$ est engendré (par les transpositions, par les cycles...), autre chose est de savoir expliciter une factorisation d'une permutation σ !*

Soit $\sigma \in \mathfrak{S}_n$.

❧ Si $\sigma(n) = n$, alors on pose $\sigma_1 = \sigma$. Sinon, on considère la transposition

$$\tau_n = \begin{pmatrix} n & \sigma(n) \end{pmatrix}$$

et on pose $\sigma_1 = \tau_n \circ \sigma$. Dans les deux cas, on observe que

$$\sigma_1(n) = n.$$

❧ Il ne reste plus qu'à continuer le processus, en augmentant à chaque étape le nombre de points fixés par la permutation.

❧ (HR) En supposant connue une permutation σ_k telle que

$$\forall\, n-k+1 \leqslant i \leqslant n, \quad \sigma_k(i) = i,$$

— ou bien $\sigma_k(n-k) = n-k$, c'est-à-dire que σ_k fixe $k+1$ points (tous les entiers compris entre n et $(n-k)$ inclus) et on pose directement $\sigma_{k+1} = \sigma_k$;

— ou bien $\sigma_k(n-k) \neq n-k$, donc $\sigma_k(n-k) < n-k$ (HR et injectivité de σ_k) et on pose alors $\sigma_{k+1} = \tau_{n-k} \circ \sigma_k$ où τ_{n-k} est la transposition définie par :

$$\tau_{n-k} = \begin{pmatrix} n-k & \sigma_k(n-k) \end{pmatrix}.$$

Dans les deux cas, on dispose d'une permutation σ_{k+1} qui fixe $(k+1)$ points :

$$\forall\, n-(k+1)+1 \leqslant i \leqslant n, \quad \sigma_{k+1}(i) = i.$$

❧ Le processus s'arrête avec la permutation σ_{n-1} qui fixe $(n-1)$ points : les entiers $2, \ldots, n$ et qui fixe par conséquent aussi 1 (puisque $1 \leqslant \sigma_{n-1}(1) < 2$).

En notant $\tau_{i_1}, \ldots, \tau_{i_r}$, les transpositions introduites dans ce processus (avec $0 \leqslant r < n$), on a donc

$$\sigma_{n-1} = \mathrm{Id} = (\tau_{i_r} \circ \cdots \circ \tau_{i_1}) \circ \sigma$$

c'est-à-dire

$$\sigma^{-1} = \tau_{i_r} \circ \cdots \circ \tau_{i_1}$$

et par conséquent

$$\sigma = \tau_{i_1}^{-1} \circ \cdots \circ \tau_{i_r}^{-1} = \tau_{i_1} \circ \cdots \circ \tau_{i_r}$$

puisqu'une transposition est son propre inverse (élément d'ordre deux).

❧ **Variante facile à retenir**

On peut vérifier qu'on peut décomposer un cycle en produit de transpositions rien que par la puissance de la relation de Chasles :

Merci au Chasles fan club !

$$(a_1, a_2, a_3, \ldots, a_{n-2}, a_{n-1}, a_n) = (a_1, a_2) \circ (a_2, a_3) \circ \cdots \circ (a_{n-2}, a_{n-1}) \circ (a_{n-1}, a_n).$$

(Démonstration par récurrence sur n, bien entendu.)

✍ *L'exercice* **[130–470]** *montre que toutes les transpositions de $\mathfrak{S}_n$ sont conjuguées. En particulier, pour quels que soient les entiers $i \neq j$, les transpositions*

$$\begin{pmatrix} i & j \end{pmatrix} \quad \textit{et} \quad \begin{pmatrix} 1 & j \end{pmatrix}$$

sont conjuguées. En suivant au plus près les explications du **[130–470]**, *on constate que*

$$\forall\, i \neq j, \quad \begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} 1 & i \end{pmatrix} \circ \begin{pmatrix} 1 & j \end{pmatrix} \circ \begin{pmatrix} 1 & i \end{pmatrix},$$

ce qui prouve que le groupe symétrique est aussi engendré par les transpositions de la forme $\begin{pmatrix} 1 & i \end{pmatrix}$.

✍ *On rappelle la méthode pour factoriser en produit de cycles de supports deux à deux disjoints.*

Les supports des différents cycles sont les orbites de la permutation σ : chaque entier $1 \leqslant k \leqslant n$ appartient à une, et à une seule, orbite sous l'action de σ et la restriction de σ à chacune de ces orbites est un cycle.

On commence par identifier l'orbite de 1, ce qui définit un premier cycle. On continue en identifier l'orbite du plus petit entier qui n'appartient pas à l'orbite de 1, ce qui définit un second cycle. Etc.

rms130-470

Soit $n \in \mathbb{N}^*$. Déterminer les morphismes de $(\mathfrak{S}_n, \circ)$ dans $(\mathbb{C}^*, \times)$.

✍ *On connaît un morphisme trivial de $(\mathfrak{S}_n, \circ)$ dans $(\mathbb{C}^*, \times)$: l'identité ! Et aussi un morphisme beaucoup moins trivial : la signature. Nous allons démontrer qu'il n'en existe pas d'autre.*

On considère un morphisme de groupes

$$\varphi : (\mathfrak{S}_n, \circ) \to (\mathbb{C}^*, \times).$$

❧ Soit τ, une transposition (ou 2-cycle). Comme $\tau^2 = \mathrm{Id}$ et que φ est un morphisme de groupes, on a donc

$$\big[\varphi(\tau)\big]^2 = \varphi(\mathrm{Id}) = 1$$

et donc $\varphi(\tau) = \pm 1$.

❧ Nous allons maintenant vérifier que la valeur de $\varphi(\tau)$ est la même pour toutes les transpositions.

Considérons quatre entiers $i \neq j$ et $k \neq \ell$ compris entre 1 et n et les transpositions

$$\tau_1 = (i \quad j) \quad \text{et} \quad \tau_2 = (k \quad \ell).$$

Comme $i \neq j$ et $k \neq \ell$, il existe une permutation $\sigma \in \mathfrak{S}_n$ telle que

$$\sigma(k) = i \quad \text{et} \quad \sigma(\ell) = j$$

et nous allons vérifier que

$$\tau_2 = \sigma^{-1} \circ \tau_1 \circ \sigma.$$

✍ *La transformation*

$$\tau \mapsto \sigma^{-1} \circ \tau \circ \sigma$$

est la conjugaison par σ et est l'analogue exact de la relation de similitude sur les matrices carrées

$$M \mapsto P^{-1}MP.$$

Il est en particulier intéressant de comparer les points fixes par τ et les points fixes par $\sigma^{-1} \circ \tau \circ \sigma$ — de même qu'il est intéressant de relier les vecteurs propres de M aux vecteurs propres de $P^{-1}MP$.

Tout d'abord,

$$\begin{array}{ccccccc} k & \overset{\sigma}{\mapsto} & i & \overset{\tau_1}{\mapsto} & j & \overset{\sigma^{-1}}{\mapsto} & \ell \\ \ell & \overset{\sigma}{\mapsto} & j & \overset{\tau_1}{\mapsto} & i & \overset{\sigma^{-1}}{\mapsto} & k \end{array}$$

et d'autre part, pour tout $x \notin \{k, \ell\}$, comme σ est injective,

$$x \mapsto \sigma(x) \notin \{i, j\}$$

puis

$$\tau_1 \circ \sigma(x) = \sigma(x)$$

car $\sigma(x) \notin \{i, j\}$ et donc finalement

$$\sigma^{-1} \circ \tau_1 \circ \sigma(x) = \sigma^{-1} \circ \sigma(x) = x.$$

Tout cela montre bien que $\tau_2 = \sigma^{-1} \circ \tau_1 \circ \sigma$.

❧ Revenons au morphisme φ : puisque c'est un morphisme,

$$\varphi(\tau_2) = \big[\varphi(\sigma)\big]^{-1} \times \varphi(\tau_1) \times \big[\varphi(\sigma)\big] = \varphi(\tau_1).$$

Ainsi, il n'y a que deux possibilités :

— ou bien $\varphi(\tau) = 1$ pour *toute* transposition τ ;

— ou bien $\varphi(\tau) = -1$ pour *toute* transposition τ.

❧ Or le groupe symétrique $\mathfrak{S}_n$ est engendré par les transpositions : pour toute permutation $\sigma \in \mathfrak{S}_n$, il existe un certain nombre p de transpositions $\tau_1, \ldots, \tau_p$ telles que

$$\sigma = \tau_p \circ \cdots \circ \tau_1.$$

Comme φ est un morphisme, on en déduit que

$$\varphi(\sigma) = \prod_{k=1}^{p} \varphi(\tau_k).$$

Par conséquent,
— si $\varphi(\tau) = 1$ pour toute transposition τ, alors $\varphi = \mathrm{Id}$;
— si $\varphi(\tau) = -1$ pour toute transposition τ, alors φ est la signature.

rms130-471

‖ *Soit* G*, un sous-groupe fini de* $\mathrm{GL}_2(\mathbb{C})$ *tel que* $G \cap \mathrm{SL}_2(\mathbb{C}) = \{I_2\}$*. Démontrer que* G *est cyclique.*

L'application $\det : \mathrm{GL}_2(\mathbb{C}) \to \mathbb{C}^*$ est un morphisme de groupes (multiplicatifs) et son noyau est, par définition, le sous-groupe $\mathrm{SL}_2(\mathbb{C})$:

$$\mathrm{SL}_2(\mathbb{C}) = \big\{M \in \mathrm{GL}_2(\mathbb{C}) \ : \ \det M = 1\big\}.$$

❧ Par conséquent, l'image par det du sous-groupe fini G :

$$D = \big\{\det g, \ g \in G\big\}$$

est un sous-groupe fini de $(\mathbb{C}^*, \times)$.

Comme $G \cap \mathrm{SL}_2(\mathbb{C}) = \{I_2\}$, la restriction du morphisme det à G est injective :

$$\begin{cases} \det g = 1 \\ g \in G \end{cases} \iff g \in G \cap \mathrm{SL}_2(\mathbb{C})$$

donc $\#(D) = \#(G)$.

❧ Soit $n \in \mathbb{N}^*$, l'ordre de G (et donc celui de D). Comme l'ordre d'un élément divise l'ordre du groupe,

$$\forall\, g \in G, \quad (\det g)^n = 1$$

ce qui prouve que $\det g$ est une racine n-ième de l'unité.

On a donc

$$D \subset \mathbb{U}_n \qquad \text{et} \qquad \#(D) = \#(\mathbb{U}_n) = n$$

ce qui prouve que $D = \mathbb{U}_n$.

❧ En particulier, il existe $g_0 \in G$ tel que

$$\det g_0 = \zeta_n = e^{2i\pi/n}$$

et par conséquent

$$\mathbb{U}_n = \big\{\det(g_0^k), \ 0 \leqslant k < n\big\}.$$

On a justifié que le morphisme $\det : G \to D = \mathbb{U}_n$ était injectif et comme il est surjectif par construction, on en déduit que

$$G = \{g_0^k, 0 \leqslant k < n\} = \langle g_0 \rangle.$$

✍ *On aurait pu conclure plus vite : ayant établi que* det *était un isomorphisme de groupes de* $(G, \times)$ *sur* $(\mathbb{U}_n, \times)$*, on savait que* G *était cyclique (tout groupe isomorphe à un groupe cyclique est lui-même cyclique).*

On aurait également pu invoquer le premier résultat du ***[468]*** *: comme* det *est un isomorphisme de* G *sur* $\mathbb{U}_n$*, alors* g_0 *est un générateur de* G *si, et seulement si,* $\det g_0$ *est un générateur de* $\mathbb{U}_n$*.*

rms130-483

Soient $n \in \mathbb{N}^*$ et $\mathbb{K}$, un corps. On considère une application non constante

$$f : \mathfrak{M}_n(\mathbb{K}) \to \mathbb{K}$$

telle que

$$\forall A, B \in \mathfrak{M}_n(\mathbb{K}), \qquad f(AB) = f(A).f(B).$$

Démontrer que $M \in \mathfrak{M}_n(\mathbb{K})$ est inversible si, et seulement si, $f(M) \neq 0$.

✍ *L'application f vérifie une propriété de type* morphisme de groupes *mais ce n'est pas un morphisme de groupes ! L'ensemble de départ : $\mathfrak{M}_n(\mathbb{K})$ est un groupe additif, certes ! mais ce n'est pas un groupe multiplicatif...*

L'idée générale pour résoudre un tel exercice consiste à choisir les matrices A et B de manière variée pour exploiter au mieux les propriétés de $\mathfrak{M}_n(\mathbb{K})$.

❧ Avec $A = I_n$, on a

$$\forall B \in \mathfrak{M}_n(\mathbb{K}), \quad f(B) = f(I_n)f(B).$$

Comme f n'est pas constante, il existe au moins une matrice B telle que $f(B) \neq 0$, donc

$$f(I_n) = 1.$$

❧ Si la matrice M est inversible, alors

$$f(M)f(M^{-1}) = f(MM^{-1}) = f(I_n) = 1,$$

ce qui prouve d'une part que $f(M) \neq 0$ et d'autre part que

$$\forall M \in GL_n(\mathbb{K}), \quad f(M^{-1}) = \big[f(M)\big]^{-1}.$$

❧ Avec $A = 0_n$, on a

$$\forall B \in \mathfrak{M}_n(\mathbb{K}), \quad f(0_n)f(B) = f(0_n).$$

Comme la fonction f n'est pas constante, il existe au moins une matrice B telle que $f(B) \neq 1 = f(I_n)$, donc

$$f(0_n) = 0.$$

❧ Considérons la matrice de rang r de référence :

$$J_r = \mathrm{Diag}(I_r, 0_{n-r}) \in \mathfrak{M}_n(\mathbb{K}).$$

Il s'agit d'un projecteur : $J_r^2 = J_r$ donc

$$[f(J_r)]^2 = f(J_r)$$

et par conséquent

$$\forall\, 1 \leqslant r \leqslant n, \quad f(J_r) = 0 \quad \text{ou} \quad f(J_r) = 1.$$

❧ Toute matrice M_r de rang r est équivalente à cette matrice J_r : il existe deux matrices inversibles P et Q telles que

$$J_r = Q^{-1}M_rP$$

et par conséquent ($f(P) \neq 0$ et $f(Q) \neq 0$ car P et Q sont inversibles)

$$f(J_r) = \frac{f(P)}{f(Q)} \cdot f(M_r).$$

Cela prouve que $f(J_r) = 0$ si, et seulement si, $f(M_r) = 0$ pour *toute* matrice M_r de rang r.

❧ On sait donc que $f(J_n) = f(I_n) = 1$ et que

$$\forall\, 1 \leqslant k < n, \quad f(J_r) \in \{0; 1\}.$$

Considérons donc

$$\rho = \min\big\{1 \leqslant r \leqslant n \ : \ f(J_r) = 1\big\}$$

(Il s'agit d'une partie non vide de $\mathbb{N}$ — elle contient au moins $r = n$ — donc l'existence du minimum est donc assurée.)

Par définition, $f(J_\rho) = 1$ et, comme la matrice

$$J'_\rho = \mathrm{Diag}(0_{n-\rho}, I_\rho)$$

est aussi une matrice de rang ρ, alors $f(J'_\rho) \neq 0$. On en déduit que

$$f(J_\rho J'_\rho) = f(J_\rho) f(J'_\rho) \neq 0.$$

Or le rang de la matrice

$$J_\rho J'_\rho = \mathrm{Diag}(0_{n-\rho}, I_{2\rho-n}, 0_{n-\rho})$$

est égal à $2\rho - n$ et si $\rho < n$, alors

$$2\rho - n = \rho - (n - \rho) < \rho.$$

Par définition de ρ, on a $f(M_r) = 0$ pour toute matrice M_r de rang $r < \rho$ et donc en particulier

$$f(J_\rho J'_\rho) = 0.$$

Il faut donc que $\rho = n$, ce qui signifie que

$$\forall\, 1 \leqslant r < n, \qquad f(J_r) = 0$$

et par conséquent que

$$f(M) = 0$$

pour toute matrice non inversible M.

rms130-498

Soient E*, un espace vectoriel de dimension finie sur le corps* $\mathbb{K}$ *et* f*, un endomorphisme de* E *tel que* f^2 *soit un projecteur.*
[**1.**] *Démontrer que* f *est trigonalisable.*
[**2.**] *Démontrer que* f *est diagonalisable si, et seulement si,* $\operatorname{rg}(f) = \operatorname{rg}(f^2)$.

[**1.**] Par hypothèse, $f^4 = f^2$, donc

$$X^4 - X^2 = X^2(X-1)(X+1)$$

est un polynôme annulateur de f.

Comme l'endomorphisme f admet un polynôme annulateur scindé, il est trigonalisable.

[**2.**] On sait que $X^2(X-1)(X+1)$ est un polynôme annulateur, produit de trois facteurs deux à deux premiers entre eux. D'après le théorème de décomposition des noyaux,

$$E = \operatorname{Ker} f^2 \oplus \operatorname{Ker}(f - \mathrm{Id}) \oplus \operatorname{Ker}(f + \mathrm{Id}). \tag{1}$$

► Pour tout endomorphisme f, on sait que $\operatorname{Ker} f \subset \operatorname{Ker} f^2$.

Si $\operatorname{rg} f = \operatorname{rg} f^2$, alors $\dim \operatorname{Ker} f = \dim \operatorname{Ker} f^2$ (Théorème du rang) et par conséquent $\operatorname{Ker} f = \operatorname{Ker} f^2$ (inclusion et égalité des dimensions).

On en déduit que

$$E = \operatorname{Ker} f \oplus \operatorname{Ker}(f - \mathrm{Id}) \oplus \operatorname{Ker}(f + \mathrm{Id})$$

et donc que f est diagonalisable.

► Supposons que f soit diagonalisable. Le polynôme minimal de f est alors scindé à racines simples et il divise $X^2(X-1)(X+1)$, donc il divise aussi $X(X-1)(X+1)$.

Par conséquent, $X(X-1)(X+1)$ est annulateur. Comme ce polynôme est le produit de trois facteurs deux à deux premiers entre eux, on peut appliquer le théorème de décomposition des noyaux et en déduire que

$$E = \operatorname{Ker} f \oplus \operatorname{Ker}(f - \mathrm{Id}) \oplus \operatorname{Ker}(f + \mathrm{Id}). \tag{2}$$

Comme E est de dimension finie, on en déduit de (1) et de (2) que

$$\begin{aligned} \dim E &= \dim \operatorname{Ker} f + \dim \operatorname{Ker}(f - \mathrm{Id}) + \dim \operatorname{Ker}(f + \mathrm{Id}) \\ &= \dim \operatorname{Ker} f^2 + \dim \operatorname{Ker}(f - \mathrm{Id}) + \dim \operatorname{Ker}(f + \mathrm{Id}) \end{aligned}$$

et donc que $\dim \operatorname{Ker} f = \dim \operatorname{Ker} f^2$. D'après le Théorème du rang,

$$\operatorname{rg} f = \operatorname{rg} f^2.$$

✍ *L'endomorphisme* f *est diagonalisable si, et seulement si, il admet un polynôme annulateur scindé à racines simples.*

Comme $X^2(X-1)(X+1)$ *est annulateur, on en déduit que* f *est diagonalisable si, et seulement si, le polynôme*

$$X(X-1)(X+1)$$

est annulateur de f.

rms130-511

Soient E, *un espace vectoriel complexe de dimension finie (non nulle) et* u, v, *deux endomorphismes de* E.
Démontrer que u *et* v *admettent un vecteur propre commun dans chacun des trois cas suivants.*
[**1.**] $u \circ v = 0$
[**2.**] $u \circ v \in \mathbb{C} \cdot u$
[**3.**] $u \circ v \in \mathrm{Vect}(u, v)$

✍ *On considère ici un espace vectoriel complexe de dimension finie non nulle. Sur un tel espace, un endomorphisme admet toujours un polynôme annulateur scindé (polynôme minimal ou caractéristique par exemple) et par conséquent, il admet nécessairement au moins un vecteur propre.*
Cette remarque est fondamentale, elle sert de nombreuses fois dans cet exercice.

[**1.**] Considérons un vecteur propre $x_0 \in E$ pour l'endomorphisme v : il existe un scalaire $\mu \in \mathbb{R}$ tel que

$$v(x_0) = \mu \cdot x_0.$$

On sait alors que

$$u\big(v(x_0)\big) = 0.$$

❧ Si $\mu \neq 0$, le problème est réglé ! En effet, dans ce cas, le vecteur $v(x_0)$ n'est pas nul (ni le scalaire μ, ni le vecteur x_0 ne sont nuls), donc $v(x_0)$ est un vecteur propre de u associé à la valeur propre $\lambda = 0$ et, par linéarité de v,

$$v\big(v(x_0)\big) = v(\mu \cdot x_0) = \mu \cdot v(x_0)$$

donc $v(x_0)$ est aussi un vecteur propre de v associé à la valeur propre μ.
❧ Fort bien, mais si $\mu = 0$? Dans ce cas, $v(x_0) = 0$ n'est pas un vecteur propre !
▶ Si v est l'endomorphisme nul, alors on considère un vecteur propre x_1 de u et comme

$$v(x_1) = 0 = 0 \cdot x_1$$

ce vecteur x_1 est bien un vecteur propre de v aussi.
▶ Supposons que v ne soit pas l'endomorphisme nul. Dans ce cas, $\mathrm{Im}\, v$ est un espace vectoriel complexe de dimension finie non nulle qui est stable par v.
Il existe donc un vecteur propre $y_0 = v(z_0) \in \mathrm{Im}\, v$ pour l'endomorphisme de $\mathrm{Im}\, v$ induit par restriction de v. C'est bien entendu un vecteur propre pour v !
Et comme $u(y_0) = (u \circ v)(z_0) = 0$, on en déduit que y_0 est un vecteur non nul du noyau de u, c'est donc un vecteur propre de u (associé à la valeur propre 0).
Le vecteur y_0 est donc un vecteur propre commun à u et à v.
[**2.**] On suppose ici qu'il existe un scalaire $a \in \mathbb{C}$ tel que

$$u \circ v = a \cdot u.$$

Autrement dit, par linéarité de u,

$$u \circ (v - a \cdot \mathrm{Id}) = 0.$$

D'après la question précédente, les endomorphismes u et $v - a\,\mathrm{Id}$ admettent un vecteur propre commun. Comme les vecteurs propres de $v - a\,\mathrm{Id}$ sont aussi des vecteurs propres de v :

$$(v - a\,\mathrm{Id})(x_0) = \lambda \cdot x_0 \iff v(x_0) = (\lambda + a) \cdot x_0$$

on en déduit que u et v admettent un vecteur propre commun.
[**3.**] On suppose enfin qu'il existe deux scalaires a et b dans $\mathbb{C}$ tels que

$$u \circ v = a \cdot u + b \cdot v.$$

Par linéarité de u et de v, on en déduit que

$$\underbrace{(u - b\,\mathrm{Id})}_{u_b} \circ \underbrace{(v - a\,\mathrm{Id})}_{v_a} = ab\,\mathrm{Id}\,.$$

On distingue alors trois cas.

▶ Si $ab \neq 0$, alors les endomorphismes u_b et v_a sont inversibles et, à un facteur près, réciproques l'un de l'autre. En particulier, ces deux endomorphismes commutent et il en va donc de même pour u et v.

Les valeurs propres de u_b (et donc aussi celles de v_a) sont différentes de 0. Si x_0 est un vecteur propre de u_b associé à λ, alors x_0 est un vecteur propre de v_a associé à ${}^{ab}/_{\lambda}$:

$$\begin{aligned} u_b(x_0) = \lambda \cdot x_0 &\implies (ab) \cdot x_0 = v_a(\lambda \cdot x_0) = \lambda \cdot v(x_0) \\ &\implies v_a(x_0) = \frac{ab}{\lambda} \cdot x_0. \end{aligned}$$

Cela prouve (par symétrie) qu'un vecteur x_0 est un vecteur propre de u_b si, et seulement si, c'est un vecteur propre de v_a.

On en déduit comme plus haut que x_0 est un vecteur propre de u si, et seulement si, x_0 est un vecteur propre de v.

Et comme E est un espace vectoriel complexe de dimension finie non nulle, u et v admettent au moins un vecteur propre en commun.

▶ Si $ab = 0$, on doit distinguer deux sous-cas :

▷ Si $b = 0$, alors $u \circ v = a \cdot u$ et on est ramené au cas précédent.

▷ Si $a = 0$, alors $u_b \circ v = 0$ et on est cette fois ramené au premier cas : il existe au moins un vecteur propre commun à u_b et à v et comme u et u_b ont les mêmes vecteurs propres, on en déduit qu'il existe au moins un vecteur propre commun à u et à v.

rms130-516

Soient E, *un espace vectoriel réel de dimension finie* $n \geqslant 2$ *et* u, *un endomorphisme de* E. *On suppose que les seuls sous-espaces vectoriels stables par* u *sont* $\{0_E\}$ *et* E. *Démontrer que* $n = 2$.

Supposons que u admette une valeur propre (réelle). Dans ce cas, il existerait aussi un vecteur propre x (non nul...) et la droite $\mathbb{R} \cdot x$ serait stable par u.

D'après l'énoncé, seuls $\{0\}$ et E sont stables par u avec $\dim E \geqslant 2$.

Par conséquent, le spectre (réel) de u est vide et u n'admet aucun vecteur propre.

❧ Considérons le polynôme minimal de u et factorisons-le en produit de polynômes irréductibles unitaires :

$$P = \prod_{k=1}^{r} P_k^{m_k}.$$

Les facteurs P_k étant des polynômes irréductibles deux à deux distincts, les facteurs $P_k^{m_k}$ sont deux à deux premiers entre eux. D'après le Théorème de décomposition des noyaux,

$$E = \bigoplus_{k=1}^{r} \operatorname{Ker} P_k^{m_k}(u).$$

Pour tout indice $1 \leqslant k \leqslant r$, le sous-espace $\operatorname{Ker} P_k(u)$ est stable par u, distinct de $\{0\}$ (puisque P_k est un diviseur du polynôme minimal) et

$$\operatorname{Ker} P_k(u) \subset \operatorname{Ker} P_k^{m_k}(u)$$

puisque $m_k \geqslant 1$.

❧ D'après l'hypothèse de l'énoncé,

$$\forall\, 1 \leqslant k \leqslant r, \quad \operatorname{Ker} P_k(u) = E.$$

Mais comme les sous-espaces vectoriels $\operatorname{Ker} P_k^{m_k}(u)$ sont en somme directe, on doit en conclure que $r = 1$ et que $\operatorname{Ker} P_1(u) = E$.

Le polynôme minimal de u est donc irréductible : $P = P_1$.

❧ Dans $\mathbb{R}[X]$, les polynômes irréductibles sont d'une part les polynômes de degré 1 et d'autre part les polynômes de degré 2 dont le discriminant est strictement négatif.

Les racines (réelles) du polynôme minimal sont les valeurs propres de u et on a constaté pour commencer que u n'avait pas de valeurs propres. Par conséquent, son polynôme minimal est un irréductible de degré 2.

❧ Notons $X^2 + aX + b$, le polynôme minimal de u et considérons un vecteur $x_0 \neq 0_E$.

Si les vecteurs x_0 et $u(x_0)$ étaient colinéaires, alors x_0 serait un vecteur propre de u : il n'en existe pas, on l'a déjà vu ! Par conséquent, la famille $\big(x_0, u(x_0)\big)$ est libre et le sous-espace

$$F = \operatorname{Vect}(x_0, u(x_0))$$

est un plan.

Comme $u^2 + au + b\,I = \omega_E$, alors

$$u\big(u(x_0)\big) = -b \cdot x_0 - a \cdot u(x_0) \in F$$

et par suite, le plan F est stable par u.

Or on a supposé que E était le seul sous-espace distinct de $\{0\}$ qui soit stable par u. Donc $E = F$ est bien un plan vectoriel.

rms130-529

Soient $A_1, \dots, A_p$ dans $\mathcal{S}_n(\mathbb{R})$. Déterminer une condition nécessaire et suffisante pour qu'il existe une matrice $A \in \mathcal{S}_n(\mathbb{R})$ telle que

$$\forall\, 1 \leqslant i \leqslant p, \quad A_i \in \mathbb{R}[A].$$

Quelle que soit la matrice $A \in \mathfrak{M}_n(\mathbb{K})$, la sous-algèbre $\mathbb{K}[A]$ est commutative.

Par conséquent, s'il existe une matrice $A \in \mathcal{S}_n(\mathbb{R})$ telle que

$$\forall\, 1 \leqslant i \leqslant p, \qquad A_i \in \mathbb{K}[A],$$

alors les matrices A_i commutent deux à deux :

$$\forall\, 1 \leqslant i, j \leqslant p, \qquad A_i A_j = A_j A_i.$$

❧ Réciproquement, chaque A_i est une matrice symétrique réelle, donc (Théorème spectral) chaque matrice A_i est diagonalisable :

$$\forall\, 1 \leqslant i \leqslant p,\ \exists\, P_i \in O_n(\mathbb{R}), \quad P_i^{-1} A_i P_i = D_i \in D_n(\mathbb{R}).$$

(On note ici $D_n(\mathbb{R})$ le sous-espace des matrices diagonales réelles.)

Si, de plus, les matrices A_i commutent deux à deux, alors les matrices A_i sont co-diagonalisables :

$$\exists\, P \in O_n(\mathbb{R}), \quad \forall\, 1 \leqslant i \leqslant p, \quad P^{-1} A_i P = \Delta_i \in D_n(\mathbb{R}).$$

✍ *Autrement dit, la même matrice de passage P convient pour toutes les matrices A_i !*

Nous allons démontrer cette propriété par récurrence sur le nombre de matrices et, pour des raisons de commodité, nous allons raisonner sur des endomorphismes symétriques plutôt que sur des matrices symétriques réelles.

Le point clé de la démonstration qui suit est le suivant : si deux endomorphismes f et g commutent, alors tout sous-espace propre de f est stable par g.

INITIALISATION.— Comme f_1 est un endomorphisme symétrique, il existe une base orthonormée $\mathscr{B}_1$ de E telle que $\mathfrak{Mat}_{\mathscr{B}_1}(f_1)$ soit diagonale (Théorème spectral).

HYPOTHÈSE DE RÉCURRENCE.— On suppose que, pour un certain entier $p \geqslant 1$, pour tout espace euclidien E, pour toute famille $(f_1, \dots, f_p)$ d'endomorphismes symétriques de E qui commutent deux à deux, il existe une base orthonormée $\mathscr{B}_p(E)$ de E telle que la matrice

$$\mathfrak{Mat}_{\mathscr{B}_p(E)}(f_k)$$

soit diagonale pour tout $1 \leqslant k \leqslant p$.

Ce n'est pas une hypothèse de récurrence qui s'improvise...

HÉRÉDITÉ.— On considère une famille de $(p+1)$ endomorphismes symétriques

$$(f_1, \dots, f_p, f_{p+1})$$

d'un espace euclidien E et on suppose que ces endomorphismes commutent deux à deux.

❧ D'après le Théorème spectral, l'endomorphisme f_{p+1} est diagonalisable et ses sous-espaces propres sont deux à deux orthogonaux :

$$E = \bigoplus_{\lambda \in \mathrm{Sp}(f_{p+1})}^{\perp} E_{p+1}^{\lambda}.$$

Comme f_{p+1} commute à chaque endomorphisme f_k, chaque sous-espace propre E_{p+1}^{λ} est stable par f_k et l'endomorphisme de E_{p+1}^{λ} induit par restriction de f_k est un endomorphisme symétrique :

$$\forall\, \lambda \in \mathrm{Sp}(f_{p+1}),\ \forall\, 1 \leqslant k \leqslant p, \quad f_k^{\lambda} \in \mathcal{S}(E_{p+1}^{\lambda}).$$

✍ *Comme E_{p+1}^{λ} est le sous-espace propre de f_{p+1} associé à λ, il est aussi stable par f_{p+1} et l'endomorphisme de E_{p+1}^{λ} induit par restriction de f_{p+1} est un endomorphisme symétrique particulièrement simple : c'est une homothétie !*

$$f_{p+1}^{\lambda} = [x \mapsto \lambda x]$$

Quels que soient $1 \leqslant j, k \leqslant p$, comme f_j et f_k commutent, alors

$$\forall\, x \in E, \quad (f_j \circ f_k)(x) = (f_k \circ f_j)(x)$$

et comme E_{p+1}^{λ} est stable par f_j et par f_k, alors

$$\forall\, x \in E_{p+1}^{\lambda}, \quad (f_j^{\lambda} \circ f_k^{\lambda})(x) = (f_k^{\lambda} \circ f_j^{\lambda})(x).$$

❧ On peut donc appliquer l'hypothèse de récurrence à l'espace euclidien E_{p+1}^{λ} et aux endomorphismes $f_1^{\lambda}, \ldots, f_p^{\lambda}$. Il existe donc une base orthonormée $\mathscr{B}_{p+1}^{\lambda}$ de E_{p+1}^{λ} telle que la matrice

$$\mathfrak{Mat}_{\mathscr{B}_{p+1}^{\lambda}}(f_k^{\lambda})$$

soit diagonale pour tout $1 \leqslant k \leqslant p$.

Comme f_{p+1}^{λ} est l'homothétie de rapport λ, on en déduit que la matrice

$$\mathfrak{Mat}_{\mathscr{B}_{p+1}^{\lambda}}(f_{p+1}^{\lambda}) = \lambda I$$

est aussi une matrice diagonale.

❧ Comme les sous-espaces propres E_{p+1}^{λ} sont deux à deux orthogonaux et que leur somme est égale à E, en concaténant les familles orthonormées $\mathscr{B}_{p+1}^{\lambda}$ on obtient une base orthonormée $\mathscr{B}_{p+1}$ de E.

Pour tout $1 \leqslant k \leqslant p+1$, la matrice de f_k relative à la base $\mathscr{B}_{p+1}$ est alors diagonale par blocs :

$$\mathfrak{Mat}_{\mathscr{B}_{p+1}}(f_k) = \mathrm{Diag}\big(\mathfrak{Mat}_{\mathscr{B}_{p+1}^{\lambda}}(f_k^{\lambda}),\ \lambda \in \mathrm{Sp}(f_{p+1})\big)$$

puisque chaque sous-espace E_{p+1}^{λ} est stable par f_k.

Et comme chaque bloc diagonal est lui-même une matrice diagonale, on en déduit que chaque matrice $\mathfrak{Mat}_{\mathscr{B}_{p+1}}(f_k)$ est diagonale.

CQFD !

❧ D'après ce qui précède, il existe donc une matrice $P \in O_n(\mathbb{R})$ telle que

$$\forall\, 1 \leqslant k \leqslant p, P^{-1}A_kP = \mathrm{Diag}(\lambda_{k,1}, \ldots, \lambda_{k,n}).$$

On pose alors

$$A = P\,\mathrm{Diag}(1, 2, \ldots, n)P^{-1}$$

de telle sorte que

$$P^{-1}AP = \mathrm{Diag}(1, 2, \ldots, n).$$

Comme les réels 1, 2, …, n sont deux à deux distincts, on déduit de la théorie de Lagrange que, pour tout $1 \leqslant k \leqslant p$, il existe un polynôme interpolateur L_k tel que

$$\forall\, 1 \leqslant i \leqslant n, \quad L_k(i) = \lambda_{k,i}.$$

On a alors

$$P^{-1}A_kP = L_k(P^{-1}AP) = P^{-1}L_k(A)P$$

et par conséquent

$$A_k = L_k(A) \in \mathbb{R}[A].$$

► En conclusion : si A_1, …, A_p sont des matrices symétriques réelles, alors il existe une matrice symétrique réelle A telle que

$$\forall\, 1 \leqslant k \leqslant p, \quad A_k \in \mathbb{R}[A]$$

si, et seulement si, les matrices A_k commutent deux à deux.

✍ ***Complément culturel***

Si E est un espace de dimension finie, on peut généraliser le résultat précédent.
On considère une famille quelconque $(f_i)_{i\in I}$ d'endomorphismes qui commutent deux à deux :

$$\forall\, i, j \in I, \quad f_i \circ f_j = f_j \circ f_i.$$

Comme E est un espace de dimension finie, alors $L(E)$ est un espace de dimension finie et

$$V = \mathrm{Vect}(f_i,\ i \in I)$$

est aussi un espace de dimension finie (en tant que sous-espace de $\mathrm{L}(\mathrm{E})$*).*

D'après le Théorème de la base incomplète (version base extraite*), il existe une famille finie*

$$(f_{i_1}, \dots, f_{i_r})$$

d'endomorphismes qui soit une base de V*.*

On peut alors appliquer le résultat précédent à cette famille finie d'endomorphismes : il existe une base

$$\mathscr{B} = (e_1, \dots, e_n)$$

de E *qui est constituée de vecteurs propres communs aux endomorphismes* $f_{i_1}, \dots, f_{i_r}$ *:*

$$\forall\, 1 \leqslant j \leqslant n,\ \forall\, 1 \leqslant k \leqslant r, \quad f_{i_k}(e_j) = \lambda_{i_k,j} \cdot e_j$$

Or tout endomorphisme f_i *est une combinaison linéaire des endomorphismes* $f_{i_1}, \dots, f_{i_r}$ *:*

$$f_i = \sum_{k=1}^{r} \alpha_k f_{i_k}$$

et les vecteurs de $\mathscr{B}$ *sont donc des vecteurs propres de* f_i *:*

$$\forall\, 1 \leqslant j \leqslant n, \quad f_i(e_j) = \sum_{k=1}^{r} \alpha_k f_{i_k}(e_j) = \left(\sum_{k=1}^{r} \alpha_k \lambda_{i_k,j} \right) \cdot e_j.$$

Il existe donc une base de vecteurs propres commune à tous les endomorphismes f_i*.*

rms130-532

[1.] *La matrice* S *est symétrique et définie positive si, et seulement si, il existe une matrice inversible* P *telle que* $S = P.P^\top$.
[2.] *Quelles que soient* $S \in \mathcal{S}_n^{++}(\mathbb{R})$ *et* $T \in \mathcal{S}_n(\mathbb{R})$, *la matrice* ST *est diagonalisable.*
[3.] *Si* A *est diagonalisable, alors il existe* $S \in \mathcal{S}_n^{++}(\mathbb{R})$ *telle que*

$$A^\top = S^{-1}.A.S.$$

Étudier la réciproque.

[1.] Soit $S \in \mathcal{S}_n^{++}(\mathbb{R})$. D'après le Théorème spectral, il existe une matrice orthogonale Q telle que $Q^\top.S.Q$ soit diagonale et les valeurs propres de S sont strictement positives. Il existe donc des réels $a_1, \dots, a_n$ strictement positifs tels que

$$Q^\top.S.Q = \mathrm{Diag}(a_1, \dots, a_n)^2$$

c'est-à-dire

$$S = [Q.\mathrm{Diag}(a_1, \dots, a_n)].[\mathrm{Diag}(a_1, \dots, a_n)^\top.Q^\top] = P.P^\top$$

en posant $P = Q.\mathrm{Diag}(a_1, \dots, a_n)$, qui est inversible en tant que produit de deux matrices inversibles.

❧ Réciproquement, si P est inversible, alors la matrice $S = P.P^\top$ est symétrique :

$$S^\top = (P.P^\top)^\top = (P^\top)^\top.P^\top = P.P^\top = S$$

et pour toute colonne X,

$$X^\top.S.X = (P.X)^\top.(P.X) = \|PX\|^2 \geqslant 0.$$

En outre, si $X^\top.S.X = 0$, alors $\|PX\| = 0$, donc $PX = 0$ et comme la matrice P est inversible, alors $X = 0$.

Ainsi, la matrice $P.P^\top$ est symétrique définie positive, quelle que soit la matrice inversible P.

[2.] D'après la question précédente,

$$S.T = P.P^\top.T = P.P^\top.T.(P.P^{-1}) = P.(P^\top.T.P).P^{-1}$$

donc S.T est semblable à la matrice $P^\top.T.P$, qui est une matrice symétrique réelle :

$$(P^\top.T.P)^\top = P^\top.T^\top.(P^\top)^\top = P^\top.T.P$$

puisque $T \in \mathcal{S}_n(\mathbb{R})$.

D'après le Théorème spectral, toute matrice symétrique réelle est diagonalisable. Comme S.T est semblable à une matrice diagonalisable, elle est elle-même diagonalisable.

✍ *À moins que* S *et* T *ne commutent, la matrice* S.T *n'est pas symétrique ! Les matrices symétriques réelles ne sont pas les seules matrices diagonalisables...*

[3.] Si la matrice A est diagonalisable, alors il existe une matrice inversible Q et une matrice diagonale Δ telles que

$$A = Q.\Delta.Q^{-1}, \quad \text{c'est-à-dire} \quad \Delta = Q^{-1}.A.Q.$$

Par conséquent, comme Δ est symétrique (elle est diagonale !),

$$A^\top = (Q.\Delta.Q^{-1})^\top = (Q^{-1})^\top.\Delta.Q^\top = (Q^{-1})^\top.(Q^{-1}.A.Q).Q^\top.$$

En posant $S = Q.Q^\top$, on définit une matrice symétrique définie positive, dont l'inverse est égal à

$$(Q.Q^\top)^{-1} = (Q^\top)^{-1}.Q^{-1} = (Q^{-1})^\top.Q^{-1},$$

ce qui prouve que $A^\top = S^{-1}.A.S$.

✍ *Il faut savoir qu'une matrice* Q *est inversible si, et seulement si, sa transposée est inversible et que, le cas échéant,*

$$(Q^\top)^{-1} = (Q^{-1})^\top$$

puisque

$$Q^\top.(Q^{-1})^\top = (Q^{-1}.Q)^\top = I_n = (Q.Q^{-1})^\top = (Q^{-1})^\top.Q^\top.$$

On ne doit pas perdre de temps à vérifier cette propriété à chaque fois qu'on doit s'en servir !

❧ Réciproquement, s'il existe une matrice $S \in \mathcal{S}_n^{++}(\mathbb{R})$ telle que

$$A^\top = S^{-1}.A.S,$$

alors il existe une matrice inversible P telle que $S = P.P^\top$ et

$$A^\top = (P.P^\top)^{-1}.A.(P.P^\top) = (P^\top)^{-1}.(P^{-1}.A.P).P^\top,$$

donc

$$P^{-1}.A.P = P^\top.A^\top.(P^\top)^{-1} = (P^{-1}.A.P)^\top.$$

Cela prouve que la matrice $P^{-1}.A.P$ est symétrique réelle. D'après le Théorème spectral, cette matrice est semblable à une matrice diagonale et par transitivité, la matrice A est elle aussi semblable à une matrice diagonale.

✍ *On a ainsi caractérisé les matrices diagonalisables : la matrice A est diagonalisable si, et seulement si, il existe une matrice $S \in \mathcal{S}_n^{++}(\mathbb{R})$ telle que $A^\top = S^{-1}.A.S$.*
Cela dit, ce critère ne me paraît pas d'une utilité pratique particulière !

rms130-541

On considère deux suites $(A_k)_{k\in\mathbb{N}}$ et $(B_k)_{k\in\mathbb{N}}$ de matrices de $\mathfrak{M}_n(\mathbb{R})$ qui convergent respectivement vers les matrices A et B.
[1.] On suppose que, pour tout $k\in\mathbb{N}$, les matrices A_k et B_k sont semblables. Les matrices A et B sont-elles semblables ?
[2.] Même question en supposant cette fois que les matrices A et B sont orthosemblables ?

[1.] Par hypothèse, pour tout entier $k\in\mathbb{N}$, il existe une matrice inversible P_k telle que

$$A_kP_k = P_kB_k.$$

✍ *Cette manière, peu habituelle, d'écrire la propriété de similitude va nous affranchir des questions relatives à la continuité de la fonction $\left[P\mapsto P^{-1}\right]$.*

On sait que $(A_k)_{k\in\mathbb{N}}$ et $(B_k)_{k\in\mathbb{N}}$ convergent respectivement vers A et B. Si la suite $(P_k)_{k\in\mathbb{N}}$ converge elle aussi vers une matrice P et que cette matrice P est encore inversible, alors on a $AP = PB$ (par continuité de la multiplication matricielle, opération bilinéaire sur un espace de dimension finie), donc les matrices A et B sont bien semblables.

✍ *Mais le groupe $\mathrm{GL}_n(\mathbb{K})$ des matrices inversibles est un ouvert, pas un fermé, donc il se pourrait que la suite $(P_k)_{k\in\mathbb{N}}$ convergeât vers une matrice P non inversible. Et il se pourrait aussi que la suite $(P_k)_{k\in\mathbb{N}}$ ne fût même pas convergente...*

❧ Pour tout $k\in\mathbb{N}^*$, on pose

$$A_k = \begin{pmatrix} 0 & 1 \\ 0 & 1/k \end{pmatrix} \quad\text{et}\quad B_k = \begin{pmatrix} 0 & 0 \\ 0 & 1/k \end{pmatrix}.$$

Pour tout $k\in\mathbb{N}^*$, ces deux matrices sont semblables : la matrice A_k est triangulaire, elle admet $n=2$ valeurs propres distinctes, donc elle est semblable à la matrice diagonale B_k.

Les deux suites $(A_k)_{k\in\mathbb{N}}$ et $(B_k)_{k\in\mathbb{N}}$ convergent pour la norme $\|\cdot\|_\infty$ (et donc pour toutes les normes sur $\mathfrak{M}_2(\mathbb{K})$), respectivement vers

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad\text{et}\quad B = 0_2.$$

La matrice A est nilpotente d'indice 2 et la matrice B est diagonale, donc A et B ne sont pas semblables.

✍ *Une matrice de passage qui mène de A_k à B_k est, par exemple, la matrice*

$$P_k = \begin{pmatrix} 1 & 1 \\ 0 & 1/k \end{pmatrix}.$$

La suite $(P_k)_{k\in\mathbb{N}}$ converge mais sa limite n'est pas inversible...
Une autre matrice de passage possible est

$$Q_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

et cette fois, la suite $(Q_k)_{k\in\mathbb{N}}$ est divergente (car pas bornée).

[2.] On suppose ici que toutes les matrices de passage appartiennent au groupe orthogonal $O_n(\mathbb{R})$. Comme le groupe $O_n(\mathbb{R})$ est compact, il existe une suite extraite $(P_{\varphi(j)})_{j\in\mathbb{N}}$ qui converge vers une matrice $P\in O_n(\mathbb{R})$.

En tant que suites extraites de suites convergentes, les deux suites $(A_{\varphi(j)})_{j\in\mathbb{N}}$ et $(B_{\varphi(j)})_{j\in\mathbb{N}}$ convergent vers A et B respectivement.

Comme la multiplication matricielle est une opération continue et que

$$\forall\, j\in\mathbb{N},\quad A_{\varphi(j)}P_{\varphi(j)} = P_{\varphi(j)}B_{\varphi(j)},$$

on en déduit par passage à la limite que

$$AP = PB$$

et comme P est une matrice orthogonale, on en déduit que les limites A et B sont encore orthogonalement semblables.

rms130-544

Soit $(E, \|\cdot\|)$, un espace vectoriel normé réel. On considère un compact K (non vide) de E et une application $f : K \to K$ telle que

$$\forall\, x \neq y, \qquad \big\|f(x) - f(y)\big\| < \|x - y\|.$$

[1.] Démontrer que l'application f admet un unique point fixe, qu'on notera ω.
[2.] On considère une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de K telle que

$$\forall\, n \in \mathbb{N}, \qquad x_{n+1} = f(x_n).$$

Démontrer que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ω.

[1.]

✍ *Le vecteur $\omega \in K$ est un point fixe de f si, et seulement si, $f(\omega) = \omega$, c'est-à-dire $\|f(\omega) - \omega\| = 0$. Comme*

$$\forall\, x \in K, \qquad \big\|f(x) - x\big\| \geqslant 0,$$

on peut voir les points fixes de f comme les points de K où $\big\|f(x) - x\big\|$ atteint la plus petite valeur possible.
Vous avez dit minimum *?*

❧ Pour tout $x \in K$, on pose

$$g(x) = f(x) - x.$$

En tant que différence de deux fonctions lipschitziennes, la fonction g est lipschitzienne.
En tant que fonction continue sur le compact $K \neq \varnothing$, elle atteint un minimum : il existe $z_0 \in K$ tel que

$$\forall\, x \in K, \quad \big\|f(x) - x\big\| \geqslant \big\|f(z_0) - z_0\big\|.$$

❧ Comme $f : K \to K$, alors $z_1 = f(z_0) \in K$ et $\boxed{\text{si}}$ $z_1 \neq z_0$, alors

$$\big\|f(z_1) - f(z_0)\big\| < \|z_1 - z_0\|$$

c'est-à-dire

$$\big\|f(z_1) - z_1\big\| < \big\|f(z_0) - z_0\big\| = \min_{x\in K} \big\|f(x) - x\big\|.$$

Comme $z_1 \in K$, c'est une contradiction manifeste et par suite $z_1 = z_0$: le point z_0 est donc un point fixe.

❧ Si z_0 et y_0 étaient deux points fixes *distincts*, alors

$$\begin{aligned} \big\|z_0 - y_0\big\| &= \big\|f(z_0) - f(y_0)\big\| && \text{(points fixes)} \\ &< \big\|z_0 - y_0\big\| && \text{(car } z_0 \neq y_0\text{)} \end{aligned}$$

ce qui est à nouveau contradictoire.
L'application f admet donc un unique point fixe dans K.

✍ *L'hypothèse de compacité est nécessaire. La fonction* sin $:]0, 1[\to]0, 1[$ *vérifie les hypothèses de l'énoncé mais n'a pas de point fixe dans l'intervalle considéré.*

[2.] Comme $f : K \to K$, la suite $(x_n)_{n\in\mathbb{N}}$ est constituée de points de K.

✍ *Bien que K soit compact, une suite d'éléments de K n'a aucune raison de converger...*
En revanche, une suite qui prend ses valeurs dans un compact et n'a qu'une seule valeur d'adhérence est convergente.

❧ Si l'un des x_n est égal au point fixe ω, alors la suite est stationnaire : à partir d'un certain rang, *tous* les termes sont égaux à ω.

❧ On suppose donc que les x_n sont tous distincts de ω. Par conséquent, pour tout $n \in \mathbb{N}$,

$$\forall\, n \in \mathbb{N}, \quad 0 \leqslant \|x_{n+1} - \omega\| = \big\|f(x_n) - f(\omega)\big\| < \|x_n - \omega\|.$$

Une suite décroissante et positive converge vers une limite $\theta \geqslant 0$:

$$\lim_{n\to+\infty} \|x_n - \omega\| = \theta.$$

❧ Comme les x_n appartiennent tous au compact K, on peut en extraire une suite convergente $(x_{n_k})_{k\in\mathbb{N}}$ qui converge vers $\ell \in K$. On en déduit que

$$\|\ell - \omega\| = \lim_{k\to+\infty} \|x_{n_k} - \omega\| = 0$$

mais aussi que

$$\big\|f(\ell) - f(\omega)\big\| = \lim_{k\to+\infty} \big\|f(x_{n_k}) - f(\omega)\big\|$$

(par continuité de f) et donc que

$$\big\|f(\ell) - \omega\big\| = \lim_{n\to+\infty} \|x_{n_k+1} - \omega\| = 0.$$

Ainsi,

$$\|\ell - \omega\| = \big\|f(\ell) - f(\omega)\big\|$$

ce qui prouve que $\ell = \omega$ (puisque l'inégalité n'est pas stricte).

❧ On vient ainsi de démontrer que ω était la seule valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$. Or une suite à valeurs *dans une partie compacte* qui n'a qu'une seule valeur d'adhérence est convergente, donc la suite $(x_n)_{n\in\mathbb{N}}$ est convergente.

rms130-566

[1.] *Soit* $f : [0,1] \to \mathbb{R}$*, une application dérivable. On suppose que*

$$\forall\, x \in [0,1], \qquad \big(f(x), f'(x)\big) \neq (0,0).$$

Démontrer que l'ensemble des zéros de f *est fini.*

[2.] *Soit* $q : [0,1] \to \mathbb{R}$*, une fonction continue. On considère une fonction* $f \in \mathscr{C}^2([0,1])$*, non identiquement nulle, qui vérifie l'équation différentielle*

$$\forall\, x \in [0,1], \quad y''(x) + q(x)y(x) = 0.$$

Démontrer que f *n'a qu'un nombre fini de zéros.*

[1.] On raisonne par l'absurde, en supposant qu'il existe une infinité $(x_n)_{n\in\mathbb{N}}$ de zéros *distincts* dans $[0,1]$ pour la fonction f.

D'après le Théorème de Bolzano-Weierstrass, on peut extraire une suite $(x_{\varphi(k)})_{k\in\mathbb{N}}$ qui converge vers $\ell \in [0,1]$ et qui vérifie aussi

$$\forall\, k \in \mathbb{N}, \qquad f(x_{\varphi(k)}) = 0.$$

✍ *On peut aussi se permettre un peu de pédanterie et présenter* $[0,1]$ *comme une partie compacte de* $\mathbb{R}$*.*

❧ Comme les $x_{\varphi(k)}$ sont deux à deux distincts, seul l'un d'entre eux peut être égal à la limite ℓ et, à partir d'un certain rang, tous les $x_{\varphi(k)}$ sont distincts de ℓ.

❧ Par continuité de f,

$$f(\ell) = \lim_{k\to+\infty} f(x_{\varphi(k)}) = 0.$$

Par dérivabilité de f (nouvelle application du Théorème de composition des limites) :

$$f'(\ell) = \lim_{x\to\ell} \frac{f(x) - f(\ell)}{x - \ell} = \lim_{k\to+\infty} \frac{f(x_{\varphi(k)}) - f(\ell)}{x_{\varphi(k)} - \ell} = 0.$$

On a donc $f(\ell) = f'(\ell) = 0$. Mais par hypothèse, l'équation

$$\big(f(x), f'(x)\big) = (0,0)$$

n'a pas de solution. Donc la fonction f n'admet qu'un nombre FINI de zéros sur $[0,1]$.

[2.] La fonction $F = (f, f') : [0,1] \to \mathbb{R}^2$ est une solution de l'équation différentielle linéaire homogène canonique

$$\forall\, x \in [0,1], \quad Y'(x) = A(x)Y(x) \qquad \text{avec} \qquad A(x) = \begin{pmatrix} 0 & 1 \\ -q(x) & 0 \end{pmatrix}.$$

Comme $q : [0,1] \to \mathbb{R}$ est continue, alors $A : [0,1] \to \mathfrak{M}_2(\mathbb{R})$ est continue et le Théorème de Cauchy-Lipschitz s'applique : pour toute condition initiale $\big(x_0, (y_0, v_0)\big) \in [0,1] \times \mathbb{R}^2$, l'équation canonique admet une, et une seule solution Y telle que $Y(x_0) = (y_0, v_0)$.

En particulier, s'il existe $x_0 \in [0,1]$ tel que $Y(x_0) = (0,0)$, alors Y est l'unique solution associée à la condition initiale $\big(x_0, (0,0)\big)$. Comme la fonction identiquement nulle est une solution évidente associée à cette condition initiale, on en déduit que $Y(x) = (0,0)$ pour tout $x \in [0,1]$.

Comme f n'est pas identiquement nulle, la fonction F n'est pas identiquement nulle et, par contraposée,

$$\forall\, x \in [0,1], \qquad F(x) = \big(f(x), f'(x)\big) \neq (0,0).$$

D'après la question précédente, la fonction f n'a qu'un nombre fini de zéros sur le segment $[0,1]$.

rms130-574

Soient a et b, deux réels strictement positifs. Existence et calcul de l'intégrale

$$\int_0^{+\infty} \frac{e^{-at} - e^{-bt}}{t}\,dt.$$

Avec $0 < a < b$, la fonction f définie par

$$\forall\, t > 0, \quad f(t) = \frac{e^{-at} - e^{-bt}}{t}$$

est continue sur $I =]0, +\infty[$.

Elle tend vers $(b - a)$ au voisinage de 0, donc elle est intégrable au voisinage de 0.

Lorsque t tend vers $+\infty$,

$$f(t) \sim \frac{e^{-at}}{t} = o(e^{-at})$$

et comme $a > 0$, on en déduit que f est aussi intégrable au voisinage de $+\infty$.

La fonction f est donc intégrable sur $]0, +\infty[$.

✍ *Le piège arrive immédiatement : aucune des deux fonctions*

$$\left[t \mapsto \frac{e^{-at}}{t}\right] \quad \textit{et} \quad \left[t \mapsto \frac{e^{-bt}}{t}\right]$$

n'est intégrable au voisinage de 0 et il est donc impossible d'invoquer la linéarité de l'intégrale ! L'expression

$$\int_0^{+\infty} \frac{e^{-at}}{t}\,dt - \int_0^{+\infty} \frac{e^{-bt}}{t}\,dt$$

est une forme indéterminée en $\infty - \infty$...

❧ Considérons $\varepsilon > 0$. Les fonctions considérées ici sont intégrables sur le sous-intervalle $[\varepsilon, +\infty[$ donc, par linéarité de l'intégrale,

$$\int_\varepsilon^{+\infty} f(t)\,dt = \int_\varepsilon^{+\infty} \frac{e^{-at}}{t}\,dt - \int_\varepsilon^{+\infty} \frac{e^{-bt}}{t}\,dt.$$

On effectue alors les changements de variable affine $u = at$ (première intégrale) et $u = bt$ (deuxième intégrale)

$$\int_\varepsilon^{+\infty} f(t)\,dt = \int_{a\varepsilon}^{+\infty} \frac{e^{-u}}{u}\,du - \int_{b\varepsilon}^{+\infty} \frac{e^{-u}}{u}\,du$$

La relation de Chasles nous donne alors

$$\int_\varepsilon^{+\infty} f(t)\,dt = \int_{a\varepsilon}^{b\varepsilon} \frac{e^{-u}}{u}\,du.$$

Par convexité de la fonction exp, on sait que

$$\forall\, u > 0, \qquad \frac{1-u}{u} \leqslant \frac{e^{-u}}{u} \leqslant \frac{1}{u}.$$

En intégrant cet encadrement, on obtient que

$$\forall\, \varepsilon > 0, \quad \ln\frac{b}{a} + (b-a)\varepsilon \leqslant \int_\varepsilon^{+\infty} f(t)\,dt \leqslant \ln\frac{b}{a}$$

et par conséquent

$$\int_0^{+\infty} f(t)\,dt = \lim_{\varepsilon \to 0} \int_\varepsilon^{+\infty} f(t)\,dt = \ln\frac{b}{a}.$$

✍ *On pourrait être tenté de conclure en appliquant le théorème d'intégration des relations de comparaison mais je ne vois pas comment m'en sortir : pour résoudre les formes indéterminées qui apparaissent, il faudrait démontrer une variante du théorème, ce qui serait infiniment plus long que les calculs précédents.*

❧ **Variante.** Merci MG !

On peut considérer la fonction de deux variables définies par

$$\forall\,(a,b)\in U,\qquad F(a,b)=\int_0^{+\infty}\frac{e^{-at}-e^{-bt}}{t}\,dt$$

où l'ouvert U est $\mathbb{R}_+^*\times\mathbb{R}_+^*$.

On vérifie que le Théorème de dérivation sous $\int$ peut s'appliquer (la domination est très facile à vérifier) et donne

$$\forall\,(a,b)\in U,\qquad \frac{\partial F}{\partial a}(a,b)=\frac{-1}{a}\quad\text{et}\quad\frac{\partial F}{\partial b}(a,b)=\frac{1}{b}.$$

Ces deux dérivées partielles sont continues sur U (en tant que fonctions du couple (a,b)), donc F est de classe $\mathscr{C}^1$ sur U et il existe $C\in\mathbb{R}$ telle que

$$\forall\,(a,b)\in U,\qquad F(a,b)=\ln b-\ln a+C.$$

Il est évident que $F(a,a)=0$ pour tout $a>0$, donc la constante C est nulle.

Soit $u : \mathbb{R}_+ \to \mathbb{R}$, une fonction continue et intégrable sur $\mathbb{R}_+$. On considère une solution f de l'équation différentielle

$$\forall x \in \mathbb{R}_+, \quad y''(x) + \big(1 + u(x)\big)y(x) = 0$$

et on pose

$$\forall x \in \mathbb{R}_+, \quad g(x) = f(x) + \int_0^x \sin(x-t)u(t)f(t)\,dt.$$

[**1.**] *Former une équation différentielle linéaire vérifiée par g.*
[**2.**] *Démontrer qu'il existe $c > 0$ tel que*

$$\forall x \in \mathbb{R}_+, \quad \big|f(x)\big| \leqslant x + \int_0^x \big|u(t)f(t)\big|\,dt.$$

[**3.**] *Démontrer que la fonction f est bornée.*

[**1.**] On commence par un peu de trigonométrie pour y voir plus clair :

$$g(x) = f(x) + \sin x \int_0^x \cos t\,u(t)f(t)\,dt - \cos x \int_0^x \sin t\,u(t)f(t)\,dt.$$

Comme les fonctions

$$[t \mapsto \cos t\,u(t)f(t)] \quad \text{et} \quad [t \mapsto \sin t\,u(t)f(t)]$$

sont continues, on peut appliquer le Théorème fondamental et en déduire que g est de classe $\mathscr{C}^1$ avec

$$g'(x) = f'(x) + \cos x \int_0^x \cos t\,u(t)f(t)\,dt + \sin x \int_0^x \sin t\,u(t)f(t)\,dt$$

(en dérivant les deux produits, il apparaît deux termes qui se compensent exactement). Le Théorème fondamental, toujours lui, nous dit alors que g est en fait de classe $\mathscr{C}^2$ avec

$$\begin{aligned} g''(x) = f''(x) &- \sin x \int_0^x \cos t\,u(t)f(t)\,dt + \cos x \int_0^x \sin t\,u(t)f(t)\,dt \\ &+ (\cos^2 x + \sin^2 x)u(x)f(x). \end{aligned}$$

Or, par hypothèse,

$$\forall x \geqslant 0, \quad f''(x) = -f(x) - u(x)f(x)$$

donc

$$\forall x \geqslant 0, \quad g''(x) = -g(x).$$

[**2.**] Comme g est une solution de l'équation différentielle $y'' + y = 0$ (équation du pendule harmonique), on en déduit que g est bornée : il existe $c \in \mathbb{R}$ tel que

$$\forall x \geqslant 0, \quad \big|g(x)\big| \leqslant c.$$

Par définition de g et par inégalité triangulaire,

$$\begin{aligned} \forall x \geqslant 0, \quad \big|f(x)\big| &= \left|g(x) - \int_0^x \sin(x-t)u(t)f(t)\,dt\right| \\ &\leqslant \big|g(x)\big| + \int_0^x \big|\sin(x-t)u(t)f(t)\big|\,dt \\ &\leqslant c + \int_0^x \big|u(t)f(t)\big|\,dt. \end{aligned}$$

[**3.**] Nous allons maintenant démontrer l'**inégalité de GRÖNWALL**.
❧ Pour tout $x \geqslant 0$, on pose

$$\Phi(x) = \left[c + \int_0^x |u(t)|\,|f(t)|\,dt\right] \cdot \exp\left(-\int_0^x |u(t)|\,dt\right).$$

D'après le Théorème fondamental, la fonction Φ est de classe $\mathscr{C}^1$ et

$$\Phi'(x) = |u(x)| \cdot \left[|f(x)| - c - \int_0^x |u(t)|\,|f(t)|\,dt\right] \cdot \exp\left(-\int_0^x |u(t)|\,dt\right).$$

Le premier et le dernier facteur sont évidemment positifs ; le second facteur est négatif (d'après ce qui précède) ; par conséquent, la fonction Φ est décroissante sur $\mathbb{R}_+$ et en particulier

$$\forall\, x \geqslant 0, \quad \Phi(x) \leqslant \Phi(0) = c.$$

On déduit alors de la question précédente que

$$\forall\, x \geqslant 0, \quad |f(x)| \leqslant c \cdot \exp\left(\int_0^x |u(t)|\,dt\right).$$

Or la fonction u est supposée intégrable sur $\mathbb{R}_+$, donc la fonction croissante

$$\left[x \mapsto \int_0^x |u(t)|\,dt\right]$$

est bornée et comme exp est continue sur $\mathbb{R}$, on en déduit que la fonction f est bien bornée.

✍ *Cette démonstration est très astucieuse : comment peut-on penser à étudier cette fonction auxiliaire Φ si on ne connaît pas la démonstration ?*

❧ Reprenons l'encadrement établi à la question précédente et multiplions par $|u(x)|$ pour obtenir une inéquation différentielle :

$$\forall\, x \geqslant 0, \quad |u(x)||f(x)| \leqslant c|u(x)| + |u(x)| \int_0^x |u(t)||f(t)|\,dt.$$

Cette relation nous suggère d'étudier la fonction

$$\Psi = \left[x \mapsto \int_0^x |u(t)||f(t)|\,dt\right],$$

qui est de classe $\mathscr{C}^1$ avec

$$\forall\, x \geqslant 0, \quad \Psi'(x) = |u(x)||f(x)|.$$

Par conséquent,

$$\forall\, x \geqslant 0, \quad \Psi'(x) - |u(x)| \cdot \Psi(x) = m(x)$$

où m est une fonction telle que

$$\forall\, x \geqslant 0, \quad m(x) \leqslant c \cdot |u(x)|.$$

✍ *C'est assez astucieux, j'en conviens ! Une inéquation différentielle est en fait une équation différentielle pour laquelle le second membre n'est pas vraiment connu : on connaît seulement un majorant du second membre...*

On sait résoudre les équations différentielles linéaires du premier ordre : il existe une fonction dérivable K telle que

$$\forall\, x \geqslant 0, \quad \Psi(x) = K(x) \cdot \exp V(x)$$

où V est une primitive de $|u|$, par exemple :

$$\forall\, x \geqslant 0, \quad V(x) = \int_0^x |u(t)|\,dt.$$

Comme $\Psi(0) = 0$, on a $K(0) = 0$ et la méthode de variation de la constante nous dit que

$$\forall\, x \geqslant 0, \quad K'(x) = m(x) \cdot \exp\bigl[-V(x)\bigr].$$

On en déduit que

$$\begin{aligned}\forall\, x \geqslant 0, \quad \Psi(x) &= \exp V(x) \cdot \int_0^x \exp\bigl[-V(t)\bigr] \cdot m(t)\,dt \\ &\leqslant \exp V(x) \cdot \int_0^x \exp\bigl[-V(t)\bigr] \cdot c \cdot |u(t)|\,dt.\end{aligned}$$

Comme la fonction u est intégrable sur $\mathbb{R}_+$, la primitive V est bornée sur $\mathbb{R}_+$ et le produit $\exp[-V(t)] \cdot |u(t)|$ est intégrable sur $\mathbb{R}_+$ (produit d'une fonction bornée par une fonction intégrable).

On en déduit que Ψ est majorée et d'après la deuxième question,

$$\forall\, x \geqslant 0, \quad |f(x)| \leqslant c + \Psi(x)$$

donc la fonction f est bien bornée sur $\mathbb{R}_+$.

rms130-601

Soit $q : \mathbb{R}_+ \to \mathbb{R}$, une fonction continue et intégrable sur $\mathbb{R}_+$. On considère l'équation différentielle

$$\forall x \in \mathbb{R}_+, \quad y''(x) + q(x)y(x) = 0. \tag{E}$$

[**1.**] *On suppose que f est une solution bornée de (E). Démontrer que sa dérivée f' tend vers 0 au voisinage de $+\infty$.*
[**2.**] *Démontrer que (E) admet des solutions non bornées.*

[**1.**] Si f est une solution bornée de (E), alors $f'' = -qf$ est intégrable sur $\mathbb{R}_+$ (produit d'une fonction intégrable q par une fonction continue et bornée f). Or, comme f' est de classe $\mathscr{C}^1$,

$$\forall x \in \mathbb{R}_+, \qquad f'(x) = f'(0) + \int_0^x f''(t)\,dt$$

et par conséquent, la dérivée f' tend vers une limite finie au voisinage de $+\infty$.

Si f' tend vers une limite $\ell \neq 0$, alors on déduit du Théorème fondamental

$$\forall x \in \mathbb{R}_+, \qquad f(x) = f(0) + \int_0^x f'(t)\,dt$$

que $f(x)$ tend vers $+\infty$ (si $\ell > 0$) ou vers $-\infty$ (si $\ell < 0$) lorsque x tend vers $+\infty$.

✍ *En revanche, si f' tend vers 0 au voisinage de $+\infty$, l'intégrale est une forme indéterminée lorsque x tend vers $+\infty$ (cf les fonctions $[t \mapsto t^\alpha]$).*

Comme f est supposée bornée, on en déduit que la dérivée f' tend vers 0 au voisinage de $+\infty$.

[**2.**] Une fonction $y \in \mathscr{C}^2(\mathbb{R}_+, \mathbb{R})$ est solution de (E) si, et seulement si, la fonction $Y = (y, y') \in \mathscr{C}^1(\mathbb{R}_+, \mathbb{R}^2)$ est solution de l'équation réduite

$$\forall x \in \mathbb{R}_+, \quad Y'(x) = A(x)Y(x) \quad \text{avec} \quad A(x) = \begin{pmatrix} 0 & 1 \\ -q(x) & 0 \end{pmatrix}. \tag{R}$$

Comme $q : \mathbb{R}_+ \to \mathbb{R}$ est continue, alors $A : \mathbb{R}_+ \to \mathfrak{M}_2(\mathbb{R})$ est continue, donc l'équation réduite est une équation différentielle linéaire homogène du premier ordre à coefficients continus et on peut appliquer le Théorème de Cauchy-Lipschitz.

❧ L'ensemble des solutions de (R) est donc un plan vectoriel et on peut donc considérer une base (F, G) de ce plan. D'après la réduction de (E), il existe deux solutions f et g de (E) telles que $F = (f, f')$ et $G = (g, g')$.

Si toutes les solutions de (E) étaient bornées, alors f et g seraient bornées et, d'après la première question, leurs dérivées f' et g' tendraient vers 0 au voisinage de $+\infty$. Par conséquent, le wronskien

$$W(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix} = f(x)g'(x) - f'(x)g(x)$$

tendrait vers 0 au voisinage de $+\infty$ alors que

$$\forall x \in \mathbb{R}_+, \qquad W(x) \neq 0$$

puisque (F, G) est une base de l'espace des solutions de (R).

D'après le cours, tous les wronskiens de (R) vérifient l'équation différentielle

$$\forall x \in \mathbb{R}_+, \quad y'(x) = \big[\operatorname{tr} A(x)\big]y(x)$$

et comme, ici, $\operatorname{tr} A(x) = 0$ pour tout $x \in \mathbb{R}_+$, on en déduit que les wronskiens sont constants.

Une fonction constante non nulle ne peut tendre vers 0 au voisinage de $+\infty$, donc il est impossible que toutes les solutions de (E) soient bornées.

Il existe donc au moins une solution de (E) qui n'est pas bornée au voisinage de $+\infty$.

✍ *Il est intéressant de comparer ce résultat à l'équation du pendule harmonique : si la fonction q est strictement positive et constante, alors toutes les solutions de (E) sont bornées mais la fonction q n'est pas intégrable au voisinage de $+\infty$.*

rms130-624

Soient X et Y, deux variables aléatoires réelles indépendantes. On suppose qu'il existe une fonction $f : \mathbb{R} \to \mathbb{R}$ telle que $Y = f(X)$. Que dire de Y ?

Soit $E \subset \mathbb{R}$, le **support** de la loi de X, c'est-à-dire l'ensemble (fini ou dénombrable, puisque E est une variable aléatoire discrète) des réels x pour lesquels

$$\mathbf{P}(X = x) > 0.$$

✍ *Le support de la loi de X est l'ensemble des* valeurs qui sont vraiment prises *par la variable X. On peut considérer une variable de loi géométrique comme une variable aléatoire à valeurs dans $\mathbb{N}$ mais son support est $\mathbb{N}^*$. De même, on peut considérer une variable aléatoire de loi binomiale $\mathscr{B}(n, p)$ comme une variable aléatoire à valeurs dans $\mathbb{Z}$ mais son support est $[\![0, n]\!]$.*

Pour tout $\omega \in \Omega$, on a

$$Y(\omega) = f\big(X(\omega)\big)$$

et par conséquent

$$\forall\, x \in E, \quad X(\omega) = x \implies Y(\omega) = f(x).$$

Autrement dit,

$$\forall\, x \in E, \quad [X = x] \subset [Y = f(x)]$$

et donc

$$\forall\, x \in E, \quad [X = x] \cap [Y = f(x)] = [X = x].$$

❧ On en déduit que

$$\forall\, x \in E, \quad \mathbf{P}\big(X = x, Y = f(x)\big) = \mathbf{P}(X = x)$$

et par hypothèse d'indépendance

$$\forall\, x \in E, \quad \mathbf{P}(X = x)\,\mathbf{P}\big(Y = f(x)\big) = \mathbf{P}(X = x).$$

Or $\mathbf{P}(X = x) > 0$ pour tout $x \in E$ (par définition même du support), donc

$$\forall\, x \in E, \quad \mathbf{P}\big(Y = f(x)\big) = 1.$$

✍ *Si x n'est pas dans le support de la loi de X, alors on ne peut pas simplifier par $\mathbf{P}(X = x)$ (division par zéro).*

❧ Cette propriété signifie deux choses :
- la variable aléatoire Y n'est ni variable, ni aléatoire : elle est presque sûrement constante (loi de Dirac) ;
- la fonction f prend la même valeur (= la valeur de Y) en chaque point du support de la loi de X.

✍ *Cette conclusion n'est pas extraordinaire car l'hypothèse de départ était paradoxale : il est tout de même étonnant de supposer que X et Y soient* indépendantes *alors que $Y = f(X)$ est* déterminée *par X.*

rms130-630

Soit $(X_n)_{n\geqslant 1}$, une suite de variables aléatoires définies sur un même espace $(\Omega, \mathcal{A}, \mathbf{P})$, à valeurs dans $\mathbb{N}$, qui suivent toutes la même loi. Pour tout entier $n \geqslant 1$, on note $R_n(\omega)$, le cardinal de l'ensemble aléatoire

$$\{X_k(\omega),\ 1 \leqslant k \leqslant n\}.$$

[**1.**] *Démontrer que*

$$\forall\, a \in \mathbb{N}, \quad \mathbf{E}(R_n) \leqslant a + n\,\mathbf{P}(X_1 \geqslant a).$$

[**2.**] *En déduire que $\mathbf{E}(R_n) = o(n)$ lorsque n tend vers $+\infty$.*

[**3.**] *On suppose que les X_n sont des variables d'espérance finie. Démontrer que $\mathbf{E}(R_n) = o(\sqrt{n})$ lorsque n tend vers $+\infty$.*

✍ *On cherche ici à estimer le nombre de valeurs distinctes d'un échantillon. La démarche qui suit est "classique" en calcul des probabilités, mais elle me semble assez ardue à justifier dans le cadre du programme.*

On notera qu'on ne fait aucune hypothèse sur la loi de l'échantillon — il n'est par exemple pas utile de supposer que les X_k sont indépendantes.

[**1.**] Il est clair que les valeurs de R_n sont comprises entre 1 (cas où toutes les variables aléatoires X_k prennent la même valeur) et n (cas où toutes les variables aléatoires X_k prennent des valeurs distinctes).

L'ensemble $\mathbb{N}^n$ est dénombrable (produit cartésien d'un nombre fini d'ensembles dénombrables) et

$$\mathbb{N}^n = \bigsqcup_{1\leqslant k\leqslant n} E_k$$

où E_k est l'ensemble des n-uplets formés au moyen de k entiers naturels distincts.

Les ensembles E_k sont donc des ensembles dénombrables (ils sont contenus dans un ensemble dénombrable et manifestement ils sont infinis) et, pour tout entier $1 \leqslant k \leqslant n$,

$$R_n(\omega) = k \iff \exists\,(i_1,\dots,i_n) \in E_k, \quad \big(X_1(\omega),\dots,X_n(\omega)\big) = (i_1,\dots,i_n).$$

Autrement dit,

$$[R_n = k] = \bigsqcup_{(i_1,\dots,i_n)\in E_k} [X_1 = i_1] \cap \cdots \cap [X_n = i_n]$$

ce qui prouve que

$$\forall\, 1 \leqslant k \leqslant n, \qquad [R_n = k] \in \mathcal{A}$$

et donc que R_n est bien une variable aléatoire discrète sur $(\Omega, \mathcal{A})$.

✍ *Je viens de répondre à une question qui n'était pas posée. Néanmoins, il est important de se la poser et de savoir y répondre.*

- Comme R_n est une variable aléatoire bornée, elle est d'espérance finie.
- L'entier R_n est le cardinal d'un ensemble aléatoire : posons

$$\forall\, \omega \in \Omega, \qquad U(\omega) = \{X_k(\omega),\ 1 \leqslant k \leqslant n\}.$$

Pour $a > 0$, on peut définir une partition de $U(\omega)$:

$$U(\omega) = \big[U(\omega) \cap [\![0, a[\![\big] \sqcup \big[U(\omega) \cap [\![a, +\infty[\![\big],$$

si bien que

$$R_n(\omega) = \#\big[U(\omega) \cap [\![0, a[\![\big] + \#\big[U(\omega) \cap [\![a, +\infty[\![\big].$$

Toute partie de l'ensemble fini $[\![0, a[\![$ est un ensemble fini et son cardinal est inférieur à a. D'autre part,

$$\begin{aligned}\#\big[\{X_1(\omega),\dots,X_n(\omega)\} \cap [\![a, +\infty[\![\big] &= \#\left(\bigcup_{k=1}^{n} \{X_k(\omega)\} \cap [\![a, +\infty[\![\right)\\ &\leqslant \sum_{k=1}^{n} \#\big(\{X_k(\omega)\} \cap [\![a, +\infty[\![\big) = \sum_{k=1}^{n} \mathbb{1}_{[X_k\geqslant a]}(\omega).\end{aligned}$$

On a ainsi démontré que

$$\forall\, \omega \in \Omega, \qquad R_n(\omega) \leqslant a + \sum_{k=1}^{n} \mathbb{1}_{[X_k \geqslant a]}(\omega).$$

L'espérance conserve les inégalités presque sûres et est linéaire, donc :

$$\mathbf{E}(R_n) \leqslant a + \sum_{k=1}^{n} \mathbf{P}(X_k \geqslant a)$$

et comme les variables aléatoires X_k suivent toutes la même loi :

$$\forall\, n \geqslant 1, \qquad \mathbf{E}(R_n) \leqslant a + n\,\mathbf{P}(X_1 \geqslant a).$$

[**2.**] Soit $\varepsilon > 0$.
Comme

$$\lim_{a \to +\infty} \mathbf{P}(X_1 \geqslant a) = 0$$

par continuité décroissante, il existe $a_0 > 0$ tel que

$$0 \leqslant \mathbf{P}(X_1 \geqslant a) \leqslant \frac{\varepsilon}{2}.$$

Comme a_0/n tend vers 0, il existe aussi un rang n_0 tel que

$$\forall\, n \geqslant n_0, \quad 0 \leqslant \frac{a_0}{n} \leqslant \frac{\varepsilon}{2}.$$

D'après la première question,

$$\forall\, n \geqslant n_0, \quad 0 \leqslant \frac{\mathbf{E}(R_n)}{n} \leqslant \varepsilon$$

ce qui prouve que $\mathbf{E}(R_n) = o(n)$ lorsque n tend vers $+\infty$.

✍ *Cette démonstration est dans le cours : c'est ainsi qu'on démontre que le théorème de sommation des relations de comparaison avec o.*

[**3.**] Si X est une variable aléatoire d'espérance finie à valeurs dans $\mathbb{N}$, alors la série $\sum n\,\mathbf{P}(X = n)$ est convergente. En particulier, pour tout $N \geqslant 1$,

$$\sum_{n=N}^{+\infty} \mathbf{P}(X = n) \geqslant N \sum_{n=N}^{+\infty} \mathbf{P}(X = n) = N\,\mathbf{P}(X \geqslant N) \geqslant 0$$

et comme le reste d'une série convergente tend vers 0, on en déduit que

$$\lim_{N \to +\infty} \mathbf{P}(X \geqslant N) = o\Big(\frac{1}{N}\Big).$$

❧ Soit $\varepsilon > 0$. D'après ce qui précède, il existe un rang n_0 tel que

$$\forall\, n \geqslant n_0,\ \forall\, a > 0, \quad 0 \leqslant \mathbf{E}(R_n) \leqslant a + \frac{n\varepsilon}{a}.$$

En étudiant les variations de

$$\Big[a \mapsto a + \frac{n\varepsilon}{a}\Big]$$

sur $]0, +\infty[$, on constate que le minimum est atteint pour $a = \sqrt{n\varepsilon}$ et que ce minimum est égal à $2\sqrt{n\varepsilon}$. On en déduit donc que

$$\forall\, n \geqslant n_0, \quad 0 \leqslant \mathbf{E}(R_n) \leqslant 2\sqrt{n\varepsilon}$$

et donc que $E(R_n) = o(\sqrt{n})$ lorsque n tend vers $+\infty$.

✍ *Cette démonstration est classique : elle permet de déduire l'inégalité de Chernoff de l'inégalité de Markov.*

rms135-375

Soient $v_1, \dots, v_n$, des vecteurs unitaires d'un espace euclidien E. Démontrer qu'il existe une famille $(\varepsilon_k)_{1\leqslant k\leqslant n} \in \{-1, +1\}^n$ telle que

$$\left\| \sum_{k=1}^{n} \varepsilon_k v_k \right\| \leqslant \sqrt{n}.$$

Version géométrique.

On procède par récurrence sur le nombre n de vecteurs : le résultat est évident pour $n = 1$.

Supposons le résultat établi pour n vecteurs : on dispose donc d'un vecteur

$$S_n = \sum_{k=1}^{n} \varepsilon_k v_k \in E$$

tel que $\|S_n\| \geqslant \sqrt{n}$ et d'un vecteur unitaire v_{n+1}.

Avec $\varepsilon_{n+1} = \pm 1$, on a

$$\|S_n + \varepsilon_{n+1} v_{n+1}\|^2 = \|S_n\|^2 + \varepsilon_{n+1}^2 \|v_{n+1}\|^2 + 2\varepsilon_{n+1} \langle S_n \,|\, v_{n+1} \rangle \geqslant n + 1 + 2\varepsilon_{n+1} \langle S_n \,|\, v_{n+1} \rangle .$$

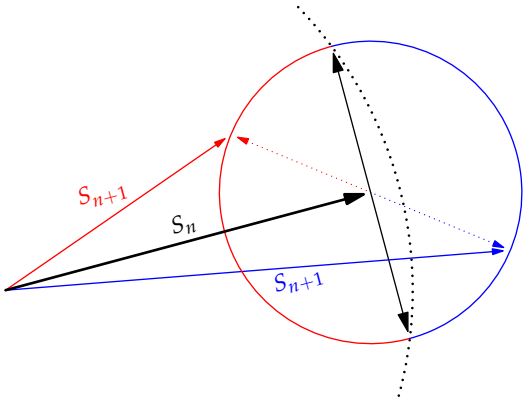
On choisit alors $\varepsilon_{n+1} = \pm 1$ de telle sorte que

$$2\varepsilon_{n+1} \langle S_n \,|\, v_{n+1} \rangle \geqslant 0,$$

ce qui nous donne un vecteur $S_{n+1} = S_n + \varepsilon_{n+1} v_{n+1}$ tel que $\|S_{n+1}\| \geqslant \sqrt{n+1}$.

✍ *En procédant par récurrence, on se ramène à un problème de géométrie plane : tout se déroule dans le sous-espace* $\mathrm{Vect}(S_n, v_{n+1})$.

Si S_n et v_{n+1} sont orthogonaux, on peut choisir indifféremment $\varepsilon_{n+1} = \pm 1$. Sinon, un seul choix est possible et la figure ci-dessous montre qu'il s'agit de choisir le bon angle, c'est-à-dire de choisir le bon signe pour le produit scalaire.



Le cercle en pointillés est le cercle de centre 0 et de rayon $r_n = \sqrt{\|S_n\|^2 + 1}$. Le cercle bicolore est le cercle de centre S_n et de rayon 1 : le demi-cercle bleu est constitué des vecteurs dont la norme est supérieure à r_n ; le demi-cercle rouge, des vecteurs dont la norme est inférieure à r_n.

Quel que soit le vecteur unitaire v_{n+1}, on peut choisir ε_{n+1} de telle sorte que le vecteur $S_n + \varepsilon_{n+1} v_{n+1}$ soit "assez grand".

Version Probabiliste.

Considérons une famille $(X_k)_{1\leqslant k\leqslant n}$ de variables aléatoires indépendantes qui suivent toutes la loi de Rademacher :

$$\forall\, 1 \leqslant k \leqslant n, \qquad \mathbf{P}(X_k = 1) = \mathbf{P}(X_k = -1) = \frac{1}{2}$$

et la variable aléatoire

$$S_n = \sum_{k=1}^{n} X_k \cdot v_k.$$

D'après les propriétés du produit scalaire,

$$\|S_n\|^2 = \sum_{k=1}^{n} X_k^2 \|v_k\|^2 + 2 \sum_{1\leqslant i<j\leqslant n} X_i X_j \langle v_i \,|\, v_j \rangle = n + 2 \sum_{1\leqslant i<j\leqslant n} X_i X_j \langle v_i \,|\, v_j \rangle$$

puisque les vecteurs v_k sont unitaires et que $\mathbf{P}(X_k^2 = 1) = 1$.

Par linéarité de l'espérance,

$$\mathbf{E}\big(\|S_n\|^2\big) = n + 2 \sum_{1\leqslant i<j\leqslant n} \mathbf{E}(X_i X_j) \langle v_i \,|\, v_j \rangle = n.$$

Or les variables aléatoires X_i sont centrées et indépendantes, donc

$$\forall\, i \neq j, \qquad \mathbf{E}(X_i X_j) = \mathbf{Cov}(X_i, X_j) = 0$$

et finalement

$$\mathbf{E}\big(\|S_n\|^2\big) = n.$$

Si $\|S_n\| < \sqrt{n}$ pour tout $(\varepsilon_k)_{1\leqslant k\leqslant n} \in \{-1,1\}^n$, alors

$$\mathbf{P}\big(\|S_n\|^2 < n\big) = 1$$

et on aurait alors

$$\mathbf{E}\big(\|S_n\|^2\big) < n.$$

On en déduit que

$$\mathbf{P}\big(\|S_n\|^2 \geqslant n\big) > 0$$

et donc qu'il existe au moins une famille $(\varepsilon_k)_{1\leqslant k\leqslant n} \in \{-1,1\}^n$ telle que $\|S_n\| \geqslant \sqrt{n}$.

✍ *La démonstration géométrique nous donne en particulier*

$$\mathbf{P}(\|S_n\| \geqslant \sqrt{n}) \geqslant \frac{1}{2^n} \qquad \textit{et} \qquad \mathbf{P}(\|S_n\| \leqslant \sqrt{n}) \geqslant \frac{1}{2^n}.$$

Si n n'est pas trop grand, on peut donc compter sur la "méthode de Monte-Carlo" pour exhiber une famille $(\varepsilon_k)_{1\leqslant k\leqslant n}$ convenable.