COLLÈGE SCIENCES ET TECHNOLOGIES

Éléments d'algèbre linéaire et hermitienne

Notes de cours de l'UE 4TPY303U

Elise Goujard

Année 2020-2021

Partie 0 : Rappels, Notions ensemblistes

1 Quelques notions mathématiques utiles

1.1 Quantificateurs

$\forall a$	Pour tout a / Soit a / Etant donné a / Quelque soit a
$\exists b$	Il existe b
$p \Rightarrow q$	Si p alors q / p implique q / p est une condition suffisante pour q
$p \Leftarrow q$	Si q alors p / q seulement si p / p est une condition nécessaire pour q
$p \Leftrightarrow q$	p si et seulement si q / p est équivalent à q

Exemple. "$\forall x \in \mathbb{R}, \exists y \in \mathbb{R}, y = x+1$" signifie "pour tout x dans $\mathbb{R}$, il existe y dans $\mathbb{R}$ tel que $y = x+1$". Cette phrase mathématique est donc vraie (le y qui convient est $x+1$).

Par contre "$\exists y \in \mathbb{R}, \forall x \in \mathbb{R}, y = x+1$" qui signifie "il existe y réel tel que pour tout x réel on a $y = x+1$" est bien sûr fausse ! L'ordre dans lequel on écrit les quantificateurs est donc primordial.

"$\exists!$" est parfois utilisé pour signifier "il existe un unique" (attention ce n'est pas une notation standard).

1.2 Ensembles

Ensemble : collection d'éléments.

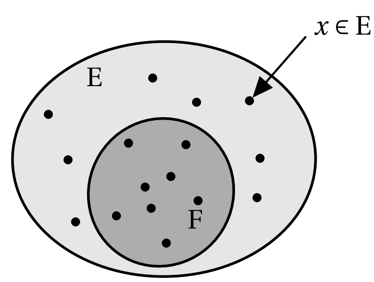
"$x \in \mathrm{E}$" signifie que l'élément x appartient à l'ensemble E.

F est une *partie* de E (ou un *sous-ensemble*) si tous ses éléments sont également des éléments de E.

C'est-à-dire $\forall x \in \mathrm{F}, x \in \mathrm{E}$.

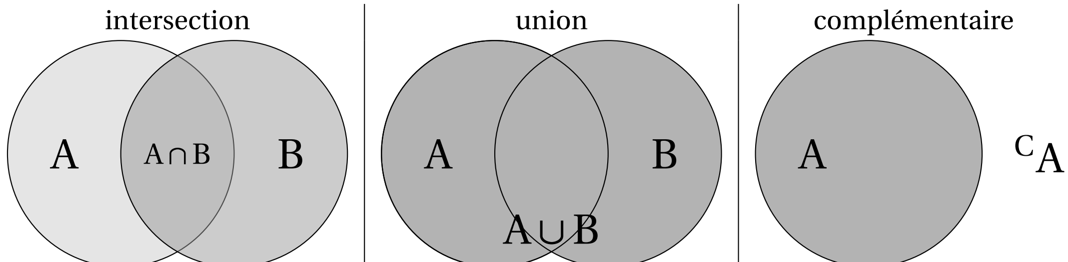
Notation $\mathrm{F} \subset \mathrm{E}$.

$\emptyset$ est l'ensemble vide, il ne contient aucun élément.



Opérations sur les ensembles :

intersection | union | complémentaire



$x \in \mathrm{A} \cup \mathrm{B} \iff x \in \mathrm{A}$ et $x \in \mathrm{B}$ | $x \in \mathrm{A} \cap \mathrm{B} \iff x \in \mathrm{A}$ ou $x \in \mathrm{B}$ | $x \in {}^{\mathrm{C}}\mathrm{A} \iff x \notin \mathrm{A}$

Le produit $\mathrm{E} \times \mathrm{F}$ de deux ensembles E et F est l'ensemble des couples d'éléments (x, y) avec $x \in \mathrm{E}$ et $y \in \mathrm{F}$.

Autrement dit $(x, y) \in \mathrm{E} \times \mathrm{F} \iff x \in \mathrm{E}, y \in \mathrm{F}$.

On note aussi $\mathrm{E}^2 = \mathrm{E} \times \mathrm{E}$ ou plus généralement $\mathrm{E}^n = \underbrace{\mathrm{E} \times \mathrm{E} \times \cdots \times \mathrm{E}}_{n \text{ fois}}$.

Récapitulatif :

$\in$	appartient à / est dans
$\subset$	est inclus dans / est une partie de
$\cap$	intersection (intersecté avec)
$\cup$	union
$\times$	produit (“croix")
$\emptyset$	ensemble vide

Ecriture d'un ensemble à l'aide de {} : Pour donner la liste des éléments d'un ensemble on utilise :

$$\{\text{élément}_1, \dots, \text{élément}_n\}.$$

Pour les ensembles décrits par des conditions on utilise la syntaxe suivante (l'ordre est important) :

$$\{\text{éléments}\textbf{,}\ \text{conditions}\}\ \text{ou}\ \{\text{éléments}\ |\ \text{conditions}\}\ \text{ou}\ \{\text{éléments}\ \textbf{;}\ \text{conditions}\}$$

Exemples.
- 1 est un élément de l'ensemble $\mathbb{R}$. $\{1\}$ est un ensemble, en particulier un sous-ensemble de $\mathbb{R}$.
- L'ensemble $C = \{(x, y) \in \mathbb{R}^2,\ x^2 + y^2 = 1\}$ est l'ensemble des couples (x, y) de réels satisfaisant la condition $x^2 + y^2 = 1$ (un cercle donc!). C'est un sous-ensemble de $\mathbb{R}^2$. L'ensemble $\{x^2 + y^2,\ (x, y) \in \mathbb{R}^2\}$ est l'ensemble des éléments s'écrivant $x^2 + y^2$ avec x et y réels. Il s'agit donc de $\mathbb{R}_+ = [0, +\infty[$, c'est un sous-ensemble de $\mathbb{R}$! Et pour faire encore plus tordu, l'ensemble $\{x^2 + y^2 = 1,\ (x, y) \in \mathbb{R}^2\}$ est juste composé de l'élément 1. On écrit $\{x^2 + y^2 = 1,\ (x, y) \in \mathbb{R}^2\} = \{1\}$.

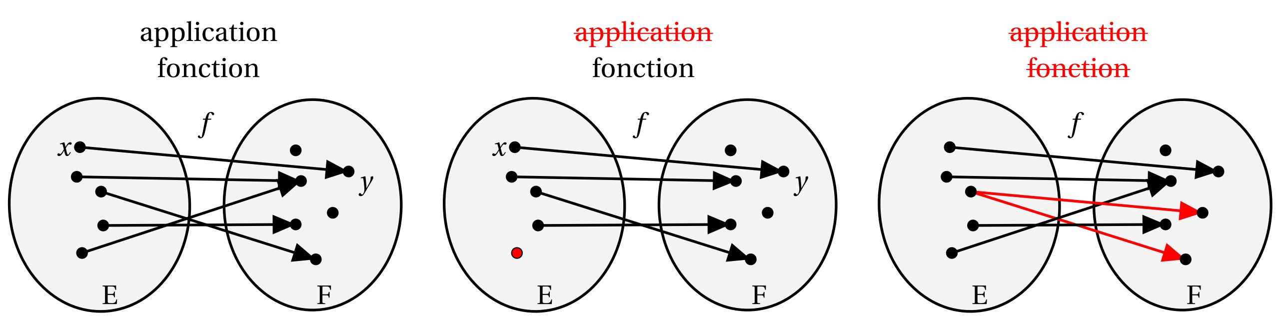
1.3 Applications

Définition 1. Soient E et F des ensembles. Une *application* f de E dans F est une correspondance entre les éléments de E et de F telle qu'à chaque élément x de E il correspond **exactement un** élément y de F (on note alors $f(x) = y$ et on appelle x l'antécédent de y et y l'image de x).

Notation :

$$\begin{array}{cccc} f & : & E & \rightarrow & F \\ & & x & \mapsto & f(x) \end{array}$$

Remarque. La définition d'une *fonction* est identique en remplaçant “**exactement un**" par “**au plus un**".



Définition 2. Image directe et image réciproque par une application :

Image *directe* d'un ensemble $A \subset E$

$f(A) := \{f(a), a \in A\}$

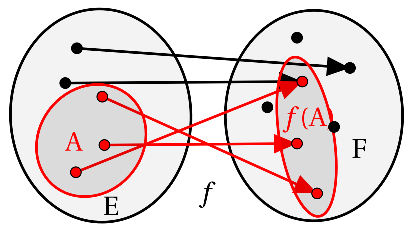
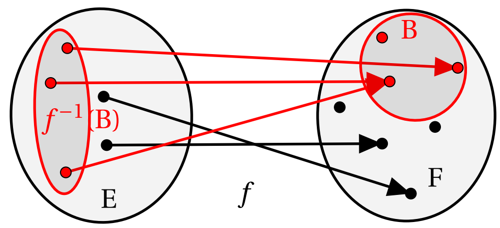


Image *réciproque* d'un ensemble $B \subset F$

$f^{-1}(B) := \{x \in E, f(x) \in B\}$



Définition 3. Soient $f : E \to F$ et $g : F \to G$ deux applications. On appelle *composée* de g par f et on note $g \circ f$ l'application

$$\begin{array}{rcccl} g \circ f & : & E & \to & G \\ & & x & \mapsto & (g \circ f)(x) := g(f(x)) \end{array}$$

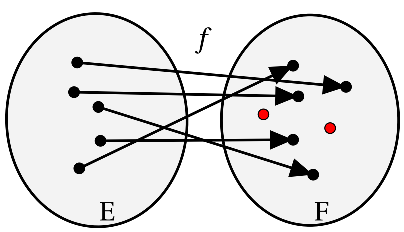
Remarque. L'élément $g(f(x))$ est obtenu en application d'abord f à x puis g au résultat :

$$x \overset{f}{\longmapsto} f(x) \overset{g}{\longmapsto} g(f(x))$$

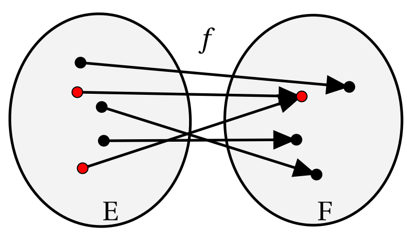
Définition 4. Une application $f : E \to F$ est dite *injective* si tout élément y de F a **au plus** un antécédent par f.

Une application $f : E \to F$ est dite *surjective* si tout élément y de F a **au moins** un antécédent par f.

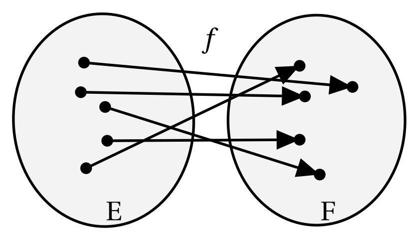
injective
~~surjective~~



~~injective~~
surjective



injective
et surjective
= **bijective**



Exemples.

- L'application $\begin{array}{rcccl} f & : & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & & (x,y) & \mapsto & (x,0) \end{array}$ n'est pas injective, car $(1,0)$ par exemple a au moins deux antécédents : $(1,0)$ et $(1,1)$ (il en a beaucoup plus que ça!). Elle n'est pas surjective car par exemple $(0,1)$ n'a pas d'antécédent par f. Son ensemble image est la droite horizontale d'équation $y = 0$.
- L'application $\begin{array}{rcccl} f & : & \mathbb{R} & \to & \mathbb{R} \\ & & x & \mapsto & x^2 \end{array}$ est-elle surjective? injective? Changer les espaces de départ et d'arrivée pour obtenir une fonction bijective.

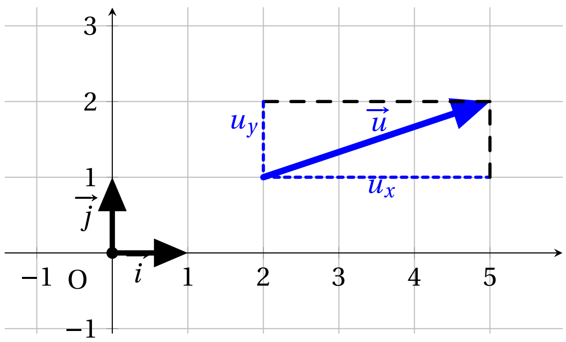
Partie I : Algèbre linéaire

2 Espaces vectoriels

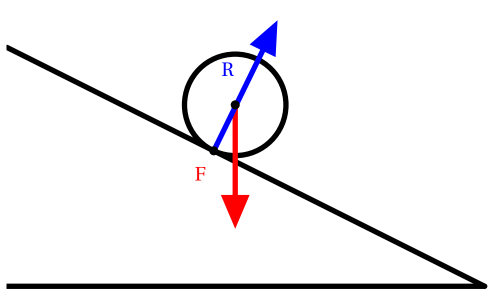
2.1 Introduction : vecteurs du plan

Rappel : un vecteur du plan (ou de l'espace) est caractérisé par une direction, un sens, et une longueur.
Dans un repère du plan $(O, \overrightarrow{i}, \overrightarrow{j})$ (on reviendra sur cette notion) un vecteur $\overrightarrow{u}$ est donné par deux coordonnées u_x et u_y. On note

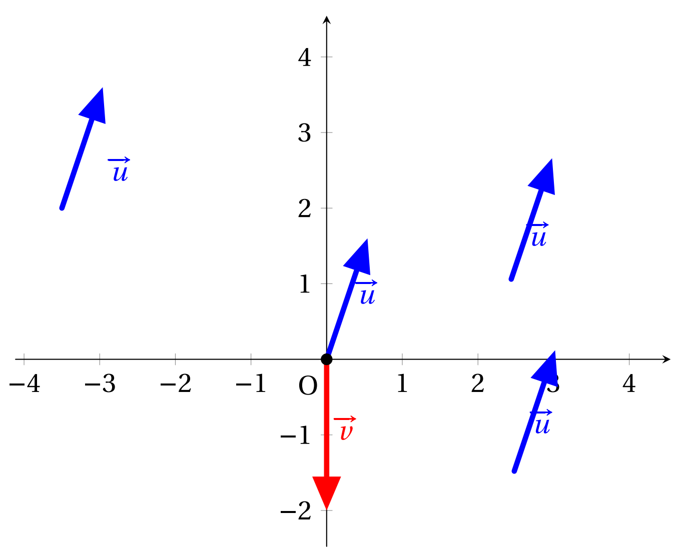
$$\overrightarrow{u} = u_x \overrightarrow{i} + u_y \overrightarrow{j} \text{ ou } \overrightarrow{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix} \text{ ou } \overrightarrow{u} = (u_x, u_y)$$



Remarque. Attention un vecteur n'est jamais "attaché" à un point d'origine (ne pas confondre avec bipoint). Cependant, dans les problèmes de physique, les vecteurs étant des grandeurs physiques (par exemple des vecteurs vitesse, des forces, des champs électriques ou magnétiques en un point, etc), ils sont souvent représentés en leur point d'application naturel. Dans un problème purement mathématique, on ne se préoccupe pas de la signification physique des vecteurs (ce sont des objets abstraits), ils peuvent être représentés n'importe où dans le plan, et souvent ils seront représentés à l'origine du repère.



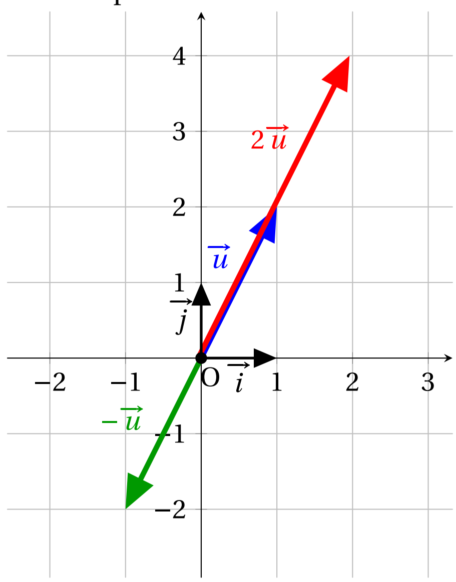
Des vecteurs en physique



Des vecteurs en maths

Vous savez déjà

multiplier un vecteur
par un nombre réel

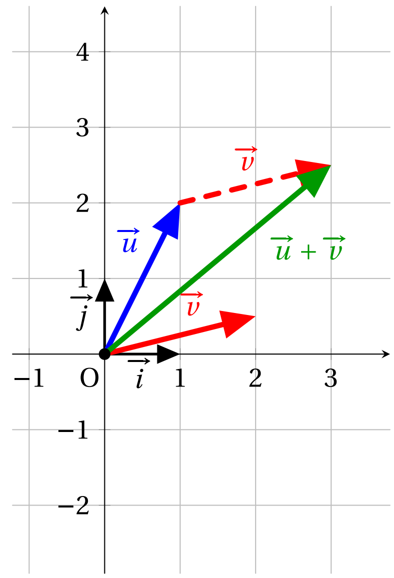


$\overrightarrow{u} = \begin{pmatrix}1\\2\end{pmatrix}, 2\overrightarrow{u} = \begin{pmatrix}2\\4\end{pmatrix}, -\overrightarrow{u} = \begin{pmatrix}-1\\-2\end{pmatrix}$

$\overrightarrow{u} = \overrightarrow{i} + 2\overrightarrow{j}, \quad 2\overrightarrow{u} = 2\overrightarrow{i} + 4\overrightarrow{j}$

$-\overrightarrow{u} = -\overrightarrow{i} - 2\overrightarrow{j}$

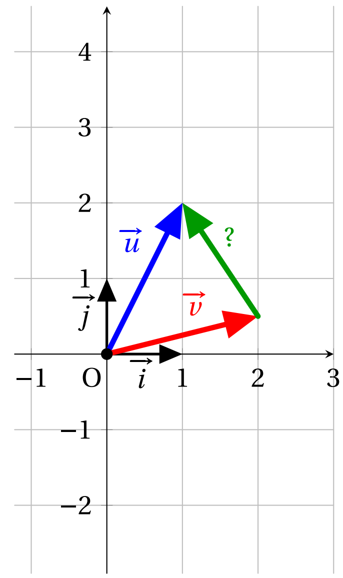
sommer deux vecteurs



$\overrightarrow{v} = \begin{pmatrix}2\\0.5\end{pmatrix}, \overrightarrow{u} + \overrightarrow{v} = \begin{pmatrix}3\\2.5\end{pmatrix},$

$\overrightarrow{u} = \overrightarrow{i} + 2\overrightarrow{j}, \quad \overrightarrow{v} = 2\overrightarrow{i} + 0.5\overrightarrow{j}$

$\overrightarrow{u} + \overrightarrow{v} = 3\overrightarrow{i} + 2.5\overrightarrow{j}$



Quel est le vecteur vert

par rapport à $\overrightarrow{u}$ et $\overrightarrow{v}$?

et vous connaissez également les règles de calcul pour combiner plusieurs de ces opérations (vérifiez en calculant par exemple $2 \times \left(\left(-3 \times \frac{1}{2}\right)\overrightarrow{v} + \overrightarrow{u}\right) + (2-1)(\overrightarrow{u} - (\overrightarrow{v} - \overrightarrow{u}))$)

Ces deux opérations ainsi que les règles de calcul associées définissent une structure d'*espace vectoriel* sur l'ensemble des vecteurs du plan.

2.2 Structure d'espace vectoriel sur un corps $\mathbb{K}$

Remarque. Attention nous n'allons pas définir la notion de *corps* dans ce cours. Vous pouvez penser à $\mathbb{R}$ ou à $\mathbb{C}$. Informellement, il s'agit d'un ensemble muni de deux opérations $+$ et $\times$, qui ont des inverses ($-$ et $\div$), et des éléments “neutres" (0 pour $+$ et 1 pour $\times$), sur lequel on peut travailler avec les règles de calcul habituelles (distributivité, etc). A titre d'exemple l'ensemble des entiers relatifs $\mathbb{Z}$ n'est pas un corps car si on divise 1 par 2 on obtient $1/2$ qui n'est pas dans $\mathbb{Z}$. Mais l'ensemble des rationnels $\mathbb{Q}$ est un corps, ainsi que $\mathbb{R}$ et $\mathbb{C}$. Il en existe beaucoup d'autres mais pour les applications à des problèmes physiques il suffit souvent de considérer $\mathbb{R}$ ou $\mathbb{C}$.

Définition 5. Un espace vectoriel $(\mathrm{E}, +, \cdot)$ sur un corps $\mathbb{K}$ (en pratique $\mathbb{R}$ ou $\mathbb{C}$) est un ensemble non vide E muni de deux opérations :

- une loi de composition interne notée + (addition des vecteurs) :

$$\begin{array}{ccc} \mathrm{E} \times \mathrm{E} & \to & \mathrm{E} \\ (u, v) & \mapsto & u + v \end{array}$$

- une loi de composition externe notée $\cdot$ (multiplication par un scalaire) :

$$\begin{array}{rcl} \mathbb{K} \times \mathrm{E} & \to & \mathrm{E} \\ (\lambda, u) & \mapsto & \lambda \cdot u \end{array}$$

satisfaisant les propriétés suivantes :

- Propriétés de l'addition :
  — commutativité : $\forall (u, v) \in \mathrm{E}^2, u + v = v + u$
  — associativité : $\forall (u, v, w) \in \mathrm{E}^3, (u + v) + w = u + (v + w)$
  — existence d'un zéro (vecteur nul) : $\exists 0 \in \mathrm{E}, \forall u \in \mathrm{E}, u + 0 = 0 + u = u$
  — existence de l'opposé : $\forall u \in \mathrm{E}, \exists u' \in \mathrm{E}, u + u' = u' + u = 0$ (on notera $u' = -u$)
- Propriétés de la multiplication par un scalaire :
  — neutre : $\forall u \in \mathrm{E}, 1 \cdot u = u$
  — associativité : $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall u \in \mathrm{E}, (\lambda \times \mu) \cdot u = \lambda \cdot (\mu \cdot u)$
- Interaction entre les deux opérations (distributivité) :
  — $\forall \lambda \in \mathbb{K}, \forall (u, v) \in \mathrm{E}^2, \lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$
  — $\forall (\lambda, \mu) \in \mathbb{K}^2, \forall u \in \mathrm{E}, (\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$

Les éléments d'un espace vectoriel sont appelés *vecteurs* (cela généralise la notion de vecteurs du plan ou de l'espace), les éléments du corps de base $\mathbb{K}$ (en général $\mathbb{R}$ ou $\mathbb{C}$) sont appelés *scalaires*. Un espace vectoriel sur un corps $\mathbb{K}$ est aussi appelé $\mathbb{K}$-espace vectoriel. Le symbole $\cdot$ peut-être oublié quand le contexte est clair. On notera aussi l'espace vectoriel E à la place de $(\mathrm{E}, +, \cdot)$ quand les lois sont connues.

Exemples.
- $(\mathbb{R}^n, +, \cdot)$, l'ensemble des n-uplets de réels (ou des vecteurs dans l'espace réel de dimension n) est un $\mathbb{R}$-espace vectoriel.
- L'ensemble $\mathscr{F}(\mathbb{R})$ des applications de $\mathbb{R}$ dans $\mathbb{R}$ est naturellement muni d'une loi $+$ (étant donné f et g dans $\mathscr{F}(\mathbb{R})$, on peut définir la fonction $f + g$ ainsi : $\forall x \in \mathbb{R}, (f + g)(x) = f(x) + g(x)$, c'est bien un élément de $\mathscr{F}(\mathbb{R})$), et de la multiplication par un scalaire ($\forall \lambda \in \mathbb{R}, \forall f \in \mathbb{R}$, la fonction $\lambda \cdot f$ est définie par $\forall x \in \mathbb{R}$, $(\lambda \cdot f)(x) = \lambda \cdot f(x)$. Vérifier que c'est bien un $\mathbb{R}$-espace vectoriel. *[il s'agit de vérifier chacun des axiomes de la définition en les traduisant en terme de fonctions]*.
- $(\mathbb{C}[\mathrm{X}], +, \cdot)$ l'ensemble des polynômes à coefficients complexes peut être vu comme un espace vectoriel sur $\mathbb{R}$ ou sur $\mathbb{C}$.

Il découle des axiomes de la définition d'un espace vectoriel les propriétés suivantes :

Proposition 1. Soit E un $\mathbb{K}$-espace vectoriel. Alors :
- Soient u, v, et w dans E, si $u + v = u + w$ alors $v = w$
- Soient $\lambda \in \mathbb{K}$ et $u \in \mathrm{E}$. On a $\lambda \cdot u = 0$ si et seulement si $\lambda = 0$ ou $u = 0$.
- Soient $\lambda \in \mathbb{K}$ et $u \in \mathrm{E}$, alors $(-\lambda) \cdot u = -(\lambda \cdot u) = \lambda \cdot (-u)$.

2.3 Sous-espaces vectoriels

Définition 6. Un sous-espace vectoriel F d'un espace vectoriel E est un sous-ensemble de E qui est un espace vectoriel pour les lois de E (sur le même corps donc).

La plupart des axiomes d'espace vectoriel vérifiés par E sont automatiquement vérifiés pour un sous-ensemble. On a donc la caractérisation suivante :

Proposition 2. Soit $(\mathrm{E},+,\cdot)$ un $\mathbb{K}$-espace vectoriel et F une partie de E.

$$\mathrm{F} \text{ est un sous-espace vectoriel de } \mathrm{E} \Leftrightarrow \begin{cases} \mathrm{F} \neq \emptyset \\ \forall (u,v) \in \mathrm{F}^2, \forall (\lambda,\mu) \in \mathbb{K}^2, \lambda \cdot u + \mu \cdot v \in \mathrm{F} \end{cases}$$

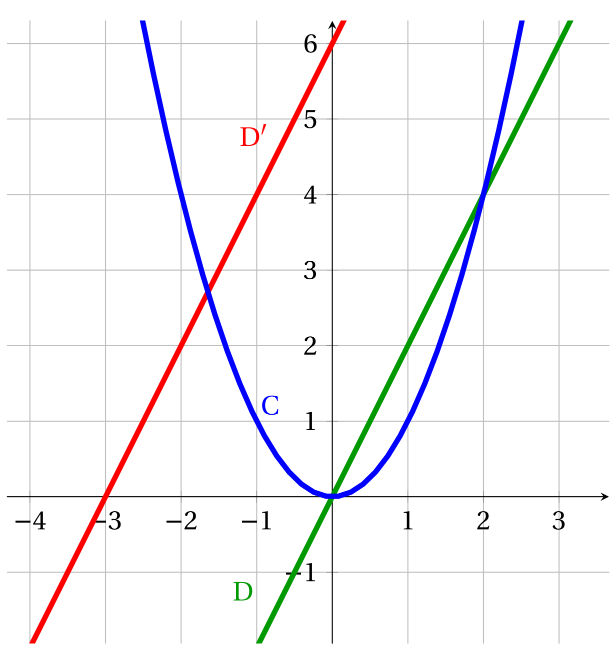
Remarque.

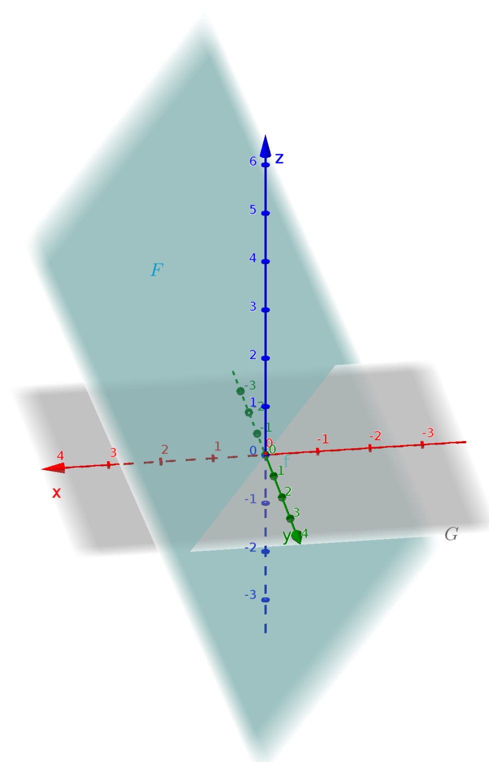
- L'axiome "$\mathrm{F} \neq \emptyset$" peut être remplacé par "$0 \in \mathrm{F}$" (pourquoi ?)
- Le deuxième axiome peut être décomposé en deux axiomes :
  - $\forall (u,v) \in \mathrm{F}^2$, $u+v \in \mathrm{F}$ (stabilité par la loi +)
  - $\forall u \in \mathrm{F}, \forall \lambda \in \mathbb{K}$, $\lambda u \in \mathrm{F}$ (stabilité par la loi externe $\cdot$)

Exemples.

- La droite $\mathrm{D} = \{(x,y) \in \mathbb{R}^2, y = 2x\}$ est un sous-espace vectoriel de l'espace vectoriel $\mathbb{R}^2$ (sur $\mathbb{R}$). En effet
  - $(0,0) \in \mathrm{D}$
  - Soient $(x_1,y_1) \in \mathrm{D}$ et $(x_2,y_2) \in \mathrm{D}$, alors $y_1 = 2x_1$, $y_2 = 2x_2$. Étant donné λ et μ dans $\mathbb{R}$, $\lambda(x_1,y_1)+\mu(x_2,y_2) = (\underbrace{\lambda x_1+\mu x_2}_{}, \underbrace{\lambda y_1+\mu y_2}_{})$ vérifie $\underbrace{(\lambda y_1+\mu y_2)}_{} = 2\underbrace{(\lambda x_1+\mu x_2)}_{}$ donc c'est bien un élément de D.

  La droite $\mathrm{D}' = \{(x,y) \in \mathbb{R}^2, y = 2x+6\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. (pourquoi ?)
  L'ensemble $\mathrm{C} = \{(x,y) \in \mathbb{R}^2, y = x^2\}$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$.
  Par exemple $(-1,1) \in \mathrm{C}$ et $(1,1) \in \mathrm{C}$ mais $(-1,1)+(1,1) = (0,1) \notin \mathrm{C}$.
  Les plans $\mathrm{F} = \{(x,y,z) \in \mathbb{R}^3, z = 2x-y\}$ et $\mathrm{G} = \{(x,y,z) \in \mathbb{R}^3, z = 0\}$ sont des sous-espaces vectoriels de $\mathbb{R}^3$.





- L'ensemble $F_0 = \{f \in \mathscr{F}(\mathbb{R}), f(0) = 0\}$ est un sous-espace vectoriel de $\mathscr{F}(\mathbb{R})$. (exercice)
- L'ensemble des solutions d'une equation différentielle linéaire homogène

$$a_n \frac{\mathrm{d}^n f}{\mathrm{d}t^n}(t) + a_{n-1}\frac{\mathrm{d}^{n-1} f}{\mathrm{d}t^{n-1}}(t) + \cdots + a_0 f(t) = 0,$$

  où $a_0, \ldots, a_n$ sont des constantes réelles, est un sous-espace vectoriel de $\mathscr{F}(\mathbb{R})$. (exercice)

2.4 Combinaisons linéaires et sous-espaces engendrés par une partie

Définition 7. Soit $(E, +, \cdot)$ un espace vectoriel sur un corps $\mathbb{K}$. Soient $u_1, \ldots, u_n$ des éléments de E, et $a_1, \ldots, a_n$ des éléments de $\mathbb{K}$, alors

$$u = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n = \sum_{i=1}^{n} a_i u_i$$

est un élément de E, appelé *combinaison linéaire* de vecteurs $u_1, \ldots u_n$. Les scalaires a_i sont appelés *coefficients* de la combinaison linéaire.

Définition 8. Soit A une partie non vide d'un espace vectoriel E. On appelle *sous-espace vectoriel engendré par* A l'ensemble des combinaisons linéaires d'éléments de A. Il se note Vect(A). Si $A = \{u_1, \ldots, u_n\}$ on a :

$$\mathrm{Vect}(A) = \left\{ \sum_{i=1}^{n} a_i u_i,\ a_i \in \mathbb{K}, u_i \in A \right\}.$$

On dit que les vecteurs $u_1, \ldots, u_n$ *engendrent* Vect(A) ou encore que la famille $(u_1, \ldots, u_n)$ est *génératrice* du sous-espace Vect(A).

Remarque. Il existe d'autres façon de définir ce sous-espace vectoriel : il s'agit également du plus petit sous-espace vectoriel contenant A (plus petit au sens de l'inclusion : tous les sous-espaces vectoriels contenant A le contiennent). C'est également l'intersection de tous les espaces vectoriels qui contiennent A.

Exemples.
- Dans les exemples précédent $D = \mathrm{V}ect(\overrightarrow{u})$ où $\overrightarrow{u} = \overrightarrow{i} + 2\overrightarrow{j}$, $F = \mathrm{Vect}\big((0,1,-1),(1,0,2)\big)$ et $G = \mathrm{Vect}\big((1,0,0),(0,1,0)\big)$.
- Tous les vecteurs du plan sont des combinaisons linéaires de $\overrightarrow{i}$ et $\overrightarrow{j}$. On peut donc écrire $\mathbb{R}^2 = \mathrm{Vect}(\overrightarrow{i}, \overrightarrow{j})$.
- Dans $\mathbb{C}[X]$, $\mathrm{Vect}(1, X, X^2) = \{a + bX + cX^2, (a,b,c) \in \mathbb{C}^2\}$, sera noté $\mathbb{C}_2[X]$.

2.5 Opérations sur les sous-espaces

Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E. Comme ce sont des sous-ensembles de E, on peut parler de leur intersection, leur union, leur complémentaire. MAIS :

- $F \cup G$ n'est en général pas un sous-espace vectoriel de E. (Ce n'est vrai que si $F \subset G$ ou $G \subset F$)
- ${}^{C}F$ n'est jamais un sous-espace vectoriel de E. (Il ne contient pas 0 !)

Comme contre-exemple pour la première affirmation, on peut prendre $F = \{(x, y, z) \in \mathbb{R}^3, z = 2x - y\}$ et $G = \{(x, y, z) \in \mathbb{R}^3, z = 0\}$ comme dans l'exemple précédent. Le vecteur $(0, 1, -1)$ appartient à F donc à $F \cup G$, le vecteur $(1, 0, 0)$ appartient à G donc à $F \cup G$, mais le vecteur $(0, 1, -1) + (1, 0, 0) = (1, 1, -1)$ n'appartient ni à F ni à G, donc n'appartient pas à $F \cup G$, ce n'est donc pas un sous-espace vectoriel de $\mathbb{R}^3$.

En revanche :

Proposition 3. Si F et G sont deux sous-espaces vectoriels d'un espace vectoriel E, alors $F \cap G$ est un sous-espace vectoriel de E.

Preuve : On suppose que F et G sont des sous-espaces vectoriels de E. On va utiliser la Propriété 2 pour montrer que $F \cap G$ est un sous-espace vectoriel de E.

- $0 \in F$ et $0 \in G$ donc $0 \in F \cap G$ et donc $F \cap G \neq \emptyset$
- Soient u, v dans $F \cap G$ et λ, μ dans $\mathbb{K}$. Comme u et v appartiennent à F, $\lambda u + \mu v \in F$. De même comme u et v appartiennent à G, $\lambda u + \mu v \in G$. Donc $\lambda u + \mu v \in F \cap G$ et le deuxième axiome est vérifié.

□

On peut également utiliser les opérations d'espace vectoriel pour définir de nouveaux sous-espaces vectoriels.

Proposition 4. Soient F et G deux sous-espaces vectoriels d'un $\mathbb{K}$-espace vectoriel E. On définit l'ensemble $F + G = \{u + v, u \in F, v \in G\}$ Alors

- C'est un sous-espace vectoriel de E.
- On a $F + G = \mathrm{Vect}(F \cup G)$.

La somme est appelée *directe* et notée $F \oplus G$ quand $F \cap G = \{0\}$.

Preuve : La preuve du premier point est laissée en exercice. Pour le deuxième point, on raisonne par double inclusion. Montrons d'abord que $F + G \subset \mathrm{Vect}(F \cup G)$. Prenons un élément de $F + G$, il s'écrit $u + v$ avec $u \in F$ et $v \in G$. Il s'agit bien d'une combinaison linéaire (avec coefficients 1) des vecteurs u et v qui appartiennent à $F \cup G$. L'inclusion est montrée.

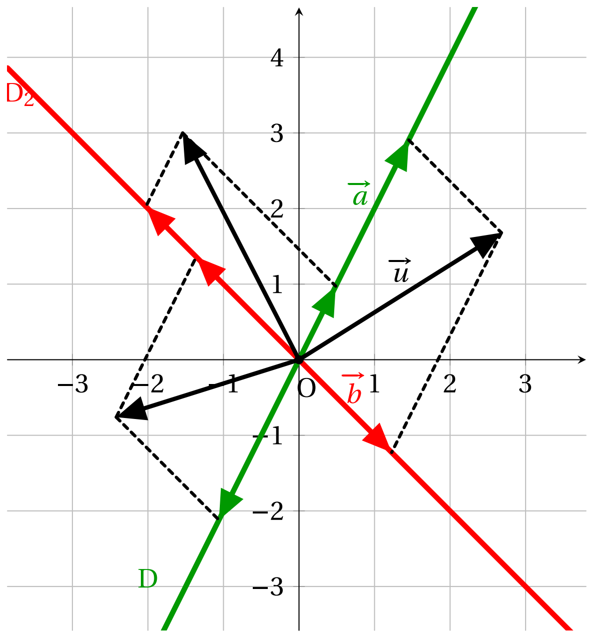
Montrons maintenant que $\mathrm{Vect}(F \cup G) \subset F + G$. Prenons un vecteur dans $\mathrm{Vect}(F \cup G)$. Il s'écrit $\sum_{i=1}^n a_i u_i$ avec $a_i \in \mathbb{K}$ et $u_i \in F \cup G$. Chaque u_i appartient à $F \cup G$, donc appartient à F ou à G. On peut alors décomposer la somme en deux : les vecteurs qui appartiennent à F et ceux qui appartiennent à G. On a $\sum_{i=1}^n a_i u_i = \underbrace{\sum_{i \text{ t.q. } u_i \in F} a_i u_i}_{\in F} + \underbrace{\sum_{i \text{ t.q. } u_i \in G} a_i u_i}_{\in G}$ donc le vecteur appartient bien à $F + G$ et le résultat est prouvé.

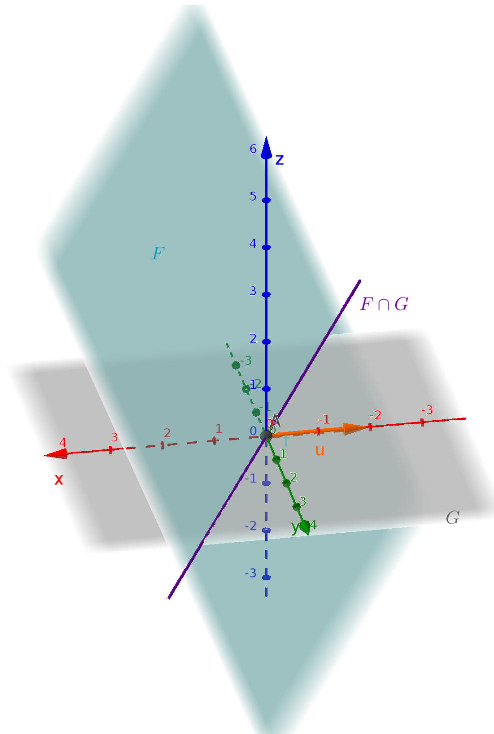
□

Remarque. Et pour la loi externe ? Pour $\lambda \in \mathbb{K}$ on peut définir $\lambda F = \{\lambda u, u \in F\}$. Quel est cet ensemble ?

Exemples.
- Dans $\mathbb{R}^2$, on considère les sous-espaces vectoriels $D = \{(x, y) \in \mathbb{R}^2, y = 2x\}$ et $D_2 = \{(x, y) \in \mathbb{R}^2, y = -x\}$. On a $D + D_2 = \mathbb{R}^2$. Cela signifie que tout vecteur du plan peut s'écrire comme une somme de vecteurs de D et de D_2. Par exemple le vecteur $\overrightarrow{u} = \overrightarrow{a} + \overrightarrow{b}$. Comme $D \cap D_2 = \emptyset$, D et D_2 sont en somme directe et on peut écrire $D \oplus D_2 = \mathbb{R}^2$.

- Dans $\mathbb{R}^3$, $F+G=\mathbb{R}^3$, mais F et G ne sont pas en somme directe : $F\cap G$ est une droite (en violet sur la figure). Par contre, la droite $\mathrm{Vect}(\overrightarrow{u})$ où $\overrightarrow{u}=(-2,0,0)$ n'intersecte F qu'en 0 . On a $F\oplus \mathrm{Vect}(\overrightarrow{u})=\mathbb{R}^3$. Quel est l'espace $(F\cap G)+\mathrm{Vect}(\overrightarrow{u})$?





Ces définitions se généralisent à plus de deux sous-espaces :

Proposition 5. Soient $F_1,\dots,F_k$ des sous-espaces vectoriels d'un espace vectoriel E. Alors

- $F_1+\dots+F_k := \{u_1+\dots,+u_k,\, u_i\in F_i\} = \mathrm{Vect}(\bigcup_{i=1}^k F_i)$.
- La somme est dit directe et notée $F_1\oplus\dots\oplus F_k=\bigoplus_{i=1}^k$ si pour tout i, $F_i\cap \mathrm{Vect}(\bigcup_{j\neq i} F_j)=\{0\}$.
- On a l'équivalence

$$F_1\oplus\dots\oplus F_k = E \Leftrightarrow \text{tout } u \text{ de E se décompose de manière UNIQUE}$$
$$\text{en } u=u_1+\dots,u_k,\ \text{avec } u_i\in F_i$$

Exemples.

- Dans l'exemple précédent D et D_2 sont en somme directe, mais pas D, D_2 et $D_3=\{[x,y)\in\mathbb{R}^2,\ y=0\}$. En effet $D_3\cap(\mathrm{Vect}(D\cup D_2))=D_3\cap\mathbb{R}^2=D_3\neq\emptyset$.
- Dans $\mathbb{R}^3$, les droites $\mathrm{Vect}((1,1,0)$, $\mathrm{Vect}((1,0,1)$, $\mathrm{Vect}((0,1,1)$ sont en somme directe.
- Dons $\mathbb{C}[X]$, en notant $F=\{X^3P(X),\ P\in\mathbb{C}[X]\}$ (c'est bien un sous-espace vectoriel), on a $\mathbb{C}_2[X]\oplus F=\mathbb{C}[X]$.

2.6 Familles génératrices, familles libres, bases

Remarque. La notion de *famille génératrice* vue à la Définition 8 se généralise à des familles infinies. On dit que la famille $(e_i)_{i\in I}$ est génératrice de l'espace vectoriel E si $\mathrm{Vect}(\{e_i,\, i\in I\})=E$. Attention chacun des éléments de $\mathrm{Vect}(\{e_i, i \in I\})$ est une combinaison linéaire faisant intervenir un nombre **fini** de vecteurs de la famille $(e_i)_{i\in I}$.

Exemples.
- Dans $\mathbb{R}^2$, la famille (e_1, e_2, e_3) où $e_1 = (1,0)$, $e_2 = (0,1)$ et $e_3 = (-1,1)$ est génératrice, mais également la famille (e_1, e_2) ou (e_1, e_3), ou (e_2, e_3).
- La famille de polynômes $(X^n)_{n\in\mathbb{N}}$ est une famille génératrice de $\mathbb{R}[X]$.

Définition 9. Une famille $(e_i)_{i\in I}$ est dite *libre* dans un espace vectoriel E si pour toute famille de scalaires $(\lambda_i)_{i\in I}$ nuls sauf un nombre fini d'entre eux, on a

$$\sum_{i\in I} \lambda_i e_i = 0 \Rightarrow \forall i \in I, \lambda_i = 0.$$

Les vecteurs e_i sont alors dits *indépendants*. Une famille non libre est appelée *liée*.

Remarque. Autrement dit une famille est libre si la seule combinaison linéaire des éléments de la famille qui donne 0 est celle avec tous les coefficients nuls.

Dans le cas d'un famille finie $(e_i)_{1\leq i\leq n}$, la condition devient :

$$\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \big(\lambda_1 e_1 + \cdots + \lambda_n e_n = 0 \Rightarrow \lambda_1 = \cdots = \lambda_n = 0\big).$$

Proposition 6. Une famille est liée si et seulement si il existe un vecteur s'exprimant comme combinaison linéaire des autres vecteurs de la famille.

Preuve : La preuve sera rédigée ici pour une famille finie. Supposons que la famille $(e_1, \dots, e_n)$ est liée. Cela signifie qu'il existe $\lambda_1, \dots, \lambda_n$ des scalaires non tous nuls tels que $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$. On peut supposer quitte à échanger les rôles que λ_1 n'est pas nul. On peut alors écrire $e_1 = -\frac{\lambda_2}{\lambda_1} e_2 - \cdots - \frac{\lambda_n}{\lambda_1} e_n$ et e_1 est une combinaison linéaire des autres vecteurs de la famille. Inversement, si par exemple $e_i = \sum_{j\neq i} \mu_j e_j$, alors la famille est liée car $e_i - \sum_{j\neq i} \mu_j e_j = 0$ et les coefficients de cette combinaison linéaires ne sont pas tous nuls : il y a le coefficient 1. $\square$

Exemples.
- Dans $\mathbb{R}^2$, la famille (e_1, e_2, e_3) considérée dans l'exemple précédent est liée. En effet, on voit que $e_3 = -e_1 + e_2$, autrement dit $e_1 - e_2 + e_3 = 0$, et les coefficients de cette combinaison linéaire (égaux à 1, -1, 1) ne sont pas nuls. Ainsi la famille n'est pas libre. Par contre la famille (e_1, e_2) est libre, tout comme la famille (e_2, e_3) ou la famille (e_1, e_3).
- Dans $\mathbb{R}[X]$, la famille $(X^n)_{n\in\mathbb{N}}$ est libre. (exercice).
- Dans $\mathbb{R}_2[X]$, les familles $(1, X)$, $(1, X^2)$, $(1, X, X^2)$ sont libres, la famille $(1, X, X^2, 1+X)$ est liée.

Définition 10. Une *base* d'un espace vectoriel E est une famille libre et génératrice.

Dans une base $(e_i)_{i\in I}$, tout vecteur $u \in E$ se décompose de façon **unique** en combinaison linéaire des vecteurs de base $u = \sum_{i\in I} \lambda_i e_i$. Les λ_i sont appelés *coordonnées* du vecteur u dans la base $(e_i)_{i\in I}$.

Exemples.
- On considère à nouveau les vecteurs $e_1 = (1,0)$, $e_2 = (0,1)$ et $e_3 = (-1,1)$ de $\mathbb{R}^2$. Les coordonnées du vecteur $u = 2e_1 - e_2$ sont $(2,-1)$ dans la base (e_1, e_2) (appelée base canonique), elles sont égales à $(1,-1)$ dans la base (e_1, e_3) et égales à $(1,-2)$ dans la base (e_2, e_3).

- $(X^n)_{n\in\mathbb{N}}$ est une base de $\mathbb{R}[X]$ et $P(X) = X^2 - 1$ a pour coordonnées $(-1, 0, 1, 0, 0, \dots)$ dans cette base.

Il n'est pas toujours évident de construire des bases d'espaces vectoriels (comment feriez-vous pour l'espace vectoriel $\mathscr{F}(\mathbb{R})$?). Les techniques mathématiques habituelles consistent à partir de familles génératrices et à enlever des vecteurs jusqu'à obtenir une base, ou à partir de familles libres et à ajouter des vecteurs jusqu'à obtenir une base. Ce genre de techniques permet de montrer le résultat suivant (admis dans ce cours) :

Théorème 1. Soit E un espace vectoriel. Alors E admet une base, et toutes ses bases sont de même cardinal (ont le même nombre d'éléments). Ce nombre (possiblement infini) est appelé *dimension* de l'espace vectoriel E et noté dim(E).

Exemples.
- $\mathbb{R}^2$ est de dimension 2, $\mathbb{R}^3$ de dimension 3, et plus généralement $\mathbb{R}^n$ est de dimension n.
- $\mathbb{C}[X]$ est de dimension infinie. Quelle est la dimension de $\mathbb{C}_2[X]$?
- $\mathscr{F}(\mathbb{R})$ est de dimension infinie (encore plus grande que la dimension de $\mathbb{C}[X]$! La preuve du théorème précédent dans ce cas fait appel à l'axiome du choix qui n'est pas utilisé par tou·te·s les mathématicien·ne·s).

Remarque. Attention la dimension d'un espace vectoriel dépend du corps utilisé. Par exemple $\mathbb{C}$ peut être vu comme un espace vectoriel sur $\mathbb{C}$ de dimension 1 car (1) est une base, ou un espace vectoriel sur $\mathbb{R}$ de dimension 2 car alors $(1, i)$ est une base.

Proposition 7. Soit E un espace vectoriel de dimension n. Alors $(e_1, \dots, e_n)$ est une base de E si et seulement si $(e_1, \dots, e_n)$ est libre si et seulement si $(e_1, \dots e_n)$ est génératrice.

Comme les sous-espaces vectoriels sont eux-mêmes des espaces vectoriels, on peut également parler de leur dimension. En voici quelques propriétés utiles.

Proposition 8. Soit E un espace vectoriel de dimension finie.
- Si F est un sous-espace vectoriel de E, alors F est de dimension finie et $\dim(F) \leq \dim(E)$.
- Si F et G sont deux sous-espaces vectoriels de E, alors

$$F \subset G \Rightarrow \dim(F) \leq \dim(G),$$

et

$$(F \subset G \text{ et } \dim(F) = \dim(G)) \Rightarrow F = G.$$

Remarque.
- Un sous-espace vectoriel de dimension 1 est appelé *droite vectorielle*. On peut l'écrire $\mathrm{Vect}(u)$ où u est un vecteur non nul de E. Cette notion généralise la notion de droite du plan ou de l'espace. Par exemple, $\mathbb{R}_0[X]$, l'ensemble des polynômes constants de $\mathbb{R}[X]$, forme une droite vectorielle dans l'espace $\mathbb{R}[X]$.

- Un sous-espace vectoriel de dimension 2 est appelé *plan vectoriel*. On peut l'écrire $\mathrm{Vect}(u, v)$ où u et v sont deux vecteurs non nuls et non colinéaires de E (ils forment une famille libre). Cette notion généralise la notion de plan dans l'espace. Par exemple, $\mathbb{R}_1[\mathrm{X}]$, l'ensemble des polynômes de degré au plus 1 de $\mathbb{R}[\mathrm{X}]$, forme un plan vectoriel dans l'espace $\mathbb{R}[\mathrm{X}]$.

Pour finir, quand un espace vectoriel est une somme directe de sous-espace vectoriels, on peut en construire une base par “recollement" de bases des sous-espaces vectoriels. Cette base est adaptée à la décomposition de l'espace selon ces sous-espaces, cette propriété sera très utile pour la diagonalisation.

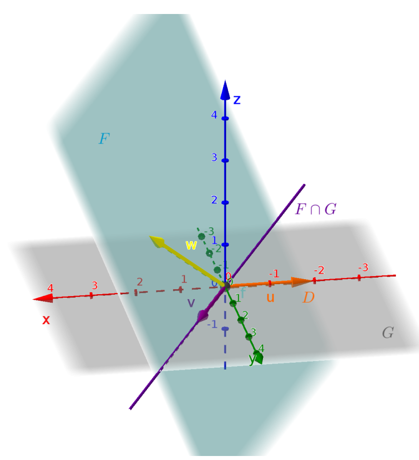
Proposition 9. Soient $\mathrm{F}_1, \ldots, \mathrm{F}_p$ des sous-espaces vectoriels de E, en somme directe, et soit $\mathrm{F} = \mathrm{F}_1 \oplus \cdots \oplus \mathrm{F}_p$. Si $(e_i)_{i\in \mathrm{I}_1}, \ldots, (e_i)_{i\in \mathrm{I}_p}$ sont des bases de $\mathrm{F}_1, \ldots, \mathrm{F}_p$, alors $(e_i)_{i\in \mathrm{I}_1\cup\cdots\cup \mathrm{I}_p}$ est une base de F.

On a alors $\dim(\mathrm{F}) = \dim(\mathrm{F}_1 \oplus \cdots \oplus \mathrm{F}_p) = \sum_{i=1}^{p} \dim(\mathrm{F}_i)$.

Preuve : Pour donner une idée de la preuve, on va traiter le cas de deux sous-espaces. On considère F_1 et F_2 en somme directe et $(e_i)_{i\in \mathrm{I}_1}$ et $(e_i)_{i\in \mathrm{I}_2}$ des bases de F_1 et F_2. Pour montrer que $(e_i)_{i\in \mathrm{I}_1\cup \mathrm{I}_2}$ est une base de $\mathrm{F} = \mathrm{F}_1 \oplus \mathrm{F}_2$, on montre que cette famille est libre et génératrice. La famille est clairement génératrice par définition de la somme de deux sous-espaces vectoriels. Prouvons maintenant qu'elle est libre. Soient λ_i pour $i \in \mathrm{I}_1 \cup \mathrm{I}_2$ des scalaires tels que $\sum_{i\in \mathrm{I}_1\cup \mathrm{I}_2} \lambda_i e_i = 0$. En séparant la somme en deux selon I_1 et I_2, et en passant des termes de l'autre côté, on obtient $\sum_{i\in \mathrm{I}_1} \lambda_i e_i = -\sum_{i\in \mathrm{I}_2} \lambda_i e_i$. Le membre de gauche est un vecteur de F_1 en tant que combinaison linéaire des vecteurs de base de F_1. De même le membre de droite est un vecteur de F_2. Comme ils sont égaux, il s'agit d'un élément de $\mathrm{F}_1 \cap \mathrm{F}_2$, c'est-à-dire 0 d'après la condition de somme directe. On a donc $\sum_{i\in \mathrm{I}_1} \lambda_i e_i = 0$ et comme $(e_i)_{i\in \mathrm{I}_1}$ est libre dans F_1, on conclut que tous les λ_i sont nuls pour $i \in \mathrm{I}_1$. De même pour F_2. Ainsi la famille totale $(e_i)_{i\in \mathrm{I}_1\cup \mathrm{I}_2}$ est libre dans F.

□

Exemples. • Dans $\mathbb{R}^3$, on considère à nouveau $\mathrm{F} = \{(x, y, z) \in \mathbb{R}^3,\ z = 2x - y\}$. On pose $u = (-2, 0, 0)$, $v = (1, 2, 0)$ et $w = (2, 2, 2)$, et $\mathrm{D} = \mathrm{Vect}(u)$.



Clairement (u) est une base de D. On peut montrer que (v, w) est une base de F (exercice). On a vu que $\mathrm{F} \oplus \mathrm{D} = \mathbb{R}^3$. Ainsi (u, v, w) est une base de $\mathbb{R}^3$, adaptée à la décomposition $\mathbb{R}^3 = \mathrm{D} \oplus \mathrm{F}$.

On note que (v) est aussi une base de la droite $\mathrm{F} \cap \mathrm{G}$.

• On a vu précédemment que l'ensemble des solutions d'une équation différentielle linéaire homogène d'ordre n :

$$a_n \frac{\mathrm{d}^n f}{\mathrm{d}t^n}(t) + a_{n-1}\frac{\mathrm{d}^{n-1} f}{\mathrm{d}t^{n-1}}(t) + \cdots + a_0 f(t) = 0,$$

où $a_0, \ldots, a_n$ sont des constantes réelles, est un sous-espace vectoriel de $\mathscr{F}(\mathbb{R})$. Un des buts de ce cours est de montrer que ce sous-espace est de dimension n. En reformulant les résultats que vous connaissez déjà en terme d'algèbre linéaire, on a :

- Pour $n = 1$ l'ensemble des solutions est la droite vectorielle $\mathrm{Vect}(t \to e^{-\frac{a_0}{a_1}t})$.
- Pour $n = 2$ l'ensemble des solutions est le plan vectoriel $\mathrm{Vect}(t \to e^{\lambda_1 t}, t \to e^{\lambda_2 t})$ dans le cas où l'équation caractéristique a deux solutions réelles distinctes λ_1 et λ_2, etc.

3 Applications linéaires

Une application linéaire entre deux espaces vectoriels est une application qui “préserve" la structure d'espace vectoriel, c'est-à-dire qui préserve les deux opérations :

Définition 11. Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $\varphi : \mathrm{E} \to \mathrm{F}$ est dite *linéaire* si

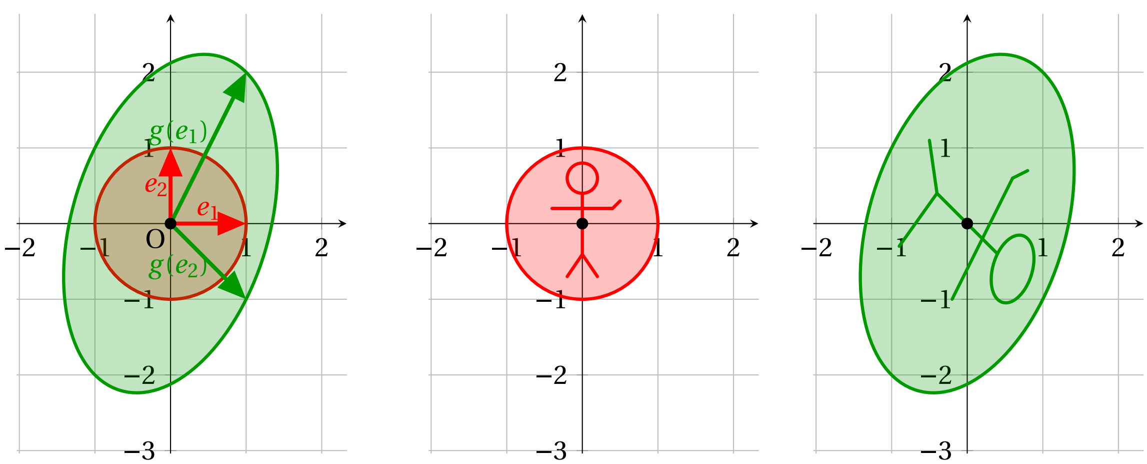
$$\forall (u,v) \in \mathrm{E}^2, \forall (\lambda,\mu) \in \mathbb{K}^2,\ \varphi(\lambda u + \mu v) = \lambda\varphi(u) + \mu\varphi(v).$$

On note $\mathscr{L}(\mathrm{E},\mathrm{F})$ l'ensemble des applications linéaires de E dans F, et $\mathscr{L}(\mathrm{E}) = \mathscr{L}(\mathrm{E},\mathrm{E})$.

Exemples.

- Les applications $f : \mathbb{R}^2 \to \mathbb{R}, (x,y) \mapsto x$ et $g : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \to (x+y, 2x-y)$ sont linéaires (exercice).
- L'application $\phi : \mathscr{F}_d(\mathbb{R}) \to \mathscr{F}(\mathbb{R}), f \mapsto f'$ est linéaire ($\mathscr{F}_d(\mathbb{R})$ désigne ici l'ensemble des fonctions dérivables). En effet, soient f et g dans $\mathscr{F}(\mathbb{R})$ et λ, μ dans $\mathbb{R}$. Alors pour tout $x \in \mathbb{R}$, on a $(\lambda f + \mu g)'(x) = \lambda f'(x) + \mu g'(x)$, c'est-à-dire $(\lambda f + \mu g)' = \lambda f' + \mu g'$.

Remarque. Si l'espace vectoriel E est de dimension finie, une application linéaire est entièrement déterminée par son action sur une base de E. Par exemple l'application $g : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \to (x+y, 2x-y)$ est entièrement déterminée par $g(e_1) = e_1 + 2e_2$ et $g(e_2) = e_1 - e_2$, où (e_1, e_2) est la base canonique de $\mathbb{R}^2$. En effet on obtient l'image de tout élément $xe_1 + ye_2$ par linéarité : $g(xe_1 + ye_2) = x(e_1 + 2e_2) + y(e_1 - e_2) = (x+y)e_1 + (2x-y)e_2$.



Action de g sur la base canonique   Action globale de g :   départ (rouge), arrivée (vert)

Proposition 10. Soit $\varphi : \mathrm{E} \to \mathrm{F}$ une application linéaire entre espaces vectoriels. Alors

- $\varphi(0) = 0$
- $\forall u \in \mathrm{E}, \varphi(-u) = -\varphi(u)$
- $\forall (u_1, \dots, u_n) \in \mathrm{E}^n, \forall (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \varphi\left(\sum_{i=1}^n \lambda_i u_i\right) = \sum_{i=1}^n \lambda_i \varphi(u_i)$.

3.1 Noyau et image d'une application linéaire

Proposition 11. Soient E et F deux $\mathbb{K}$-espaces vectoriels et φ une application linéaire de E dans F. Le *noyau* de φ est l'ensemble défini par

$$\mathrm{Ker}(\varphi) = \{u \in \mathrm{E},\ \varphi(u) = 0\} = \varphi^{-1}(\{0\}).$$

C'est un sous-espace vectoriel de E.

Preuve : Montrons que $\mathrm{Ker}(\varphi)$ est un sous-espace vectoriel de E. On va utiliser la Proposition 2.

- $0 \in \mathrm{Ker}(\varphi)$ car $\varphi(0) = 0$, donc $\mathrm{Ker}(\varphi) \neq \emptyset$.
- Soient u et v des éléments de $\mathrm{Ker}(\varphi)$, alors pour tous λ et μ dans $\mathbb{K}$, $\varphi(\lambda u + \mu v) = \lambda\varphi(u) + \mu\varphi(v) = 0$, donc $\lambda u + \mu v \in \mathrm{Ker}(\varphi)$.

Ainsi $\mathrm{Ker}(\varphi)$ est bien un sous-espace vectoriel de E. □

Proposition 12. Soit $\varphi : \mathrm{E} \to \mathrm{F}$ une application linéaire entre espaces vectoriels. Alors :

$$\big(\varphi \text{ est injective}\big) \Leftrightarrow \big(\mathrm{Ker}(\varphi) = \{0\}\big).$$

Preuve : Supposons φ injective. Alors tout élément de F a au plus un antécédent par φ. Or l'élément 0 de E est un antécédent de l'élément 0 de F. Cela doit donc être le seul et $\mathrm{Ker}(\varphi) = \{0\}$.

Réciproquement, supposons $\mathrm{Ker}(\varphi) = \{0\}$. Soit v un élément de F et u_1 et u_2 deux antécédents possibles de v par φ. Alors $u_1 - u_2 \in \mathrm{Ker}(\varphi)$, car $\varphi(u_1 - u_2) = \varphi(u_1) - \varphi(u_2)$ par linéarité. Par hypothèse on a donc $u_1 - u_2 = 0$ soit $u_1 = u_2$. L'élément v ne peut donc admettre qu'au plus un antécédent par φ, et φ est injective. □

Proposition 13. Soient E et F deux $\mathbb{K}$-espaces vectoriels et φ une application linéaire de E dans F. L'*image* de φ est l'ensemble défini par

$$\mathrm{Im}(\varphi) = \{\varphi(u),\ u \in \mathrm{E}\} = \varphi(\mathrm{E}).$$

C'est un sous-espace vectoriel de F.

Si $(e_i)_{i\in \mathrm{I}}$ est une base de E alors $\mathrm{Im}(\varphi) = \mathrm{Vect}(\{\varphi(e_i), i \in \mathrm{I}\})$.

Remarque. Par définition, une application linéaire φ est surjective si et seulement si $\mathrm{Im}(\varphi) = \mathrm{F}$.

Exemples.

- L'application $f : \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x$ a pour noyau

$$\mathrm{Ker}(f) = \{(x, y) \in \mathbb{R}^2, x = 0\} = \{(0, y), y \in \mathbb{R}\} = \{0\} \times \mathbb{R},$$

  et pour image $\mathbb{R}$. On note que $\mathrm{Ker}(f)$ est de dimension 1 et $\mathrm{Im}(f)$ est de dimension 1.
- Déterminons l'image et le noyau de l'application $h : \mathbb{R}^3 \to \mathbb{R}^2, (x, y, z) \mapsto (x - y, y - z)$. On a

$$\begin{aligned}(x, y, z) \in \mathrm{Ker}(h) \quad &\Longleftrightarrow \quad x - y = 0 \text{ et } y - z = 0\\ &\Longleftrightarrow \quad x = y = z\end{aligned}$$

D'où $\mathrm{Ker}(h) = \{(x,x,x), x \in \mathbb{R}\} = \mathrm{Vect}\big((1,1,1)\big)$. Pour calculer $\mathrm{Im}(h)$ on calcule les images des vecteurs de base. Si (e_1, e_2, e_3) est la base canonique de $\mathbb{R}^3$, alors $h(e_1) = (1,0)$ (car en coordonnées $e_1 = (1,0,0)$), $h(e_2) = (-1,1)$, $h(e_3) = (0,-1)$. On remarque que la famille $\big((0,1),(-1,1),(0,-1)\big)$ est génératrice de $\mathbb{R}^2$, donc $\mathrm{Im}(h) = \mathbb{R}^2$. On note que $\mathrm{Ker}(h)$ est de dimension 1 et $\mathrm{Im}(h)$ est de dimension 2.

- Quel est le noyau et l'image de $g : \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \to (x+y, 2x-y)$? Quelles sont leurs dimensions ?

Théorème 2 (Théorème du rang)**.** Soit $f : \mathrm{E} \to \mathrm{F}$ une application linéaire entre espaces vectoriels. Si E est de dimension finie, alors $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont de dimension finie et on a

$$\dim(\mathrm{E}) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)).$$

L'entier $\dim(\mathrm{Im}(f))$ aussi noté $\mathrm{rg}(f)$ est appelé *rang* de f.

Remarque. Par convention $\dim(\{0\}) = 0$.

Preuve : Il existe plusieurs façons de prouver ce résultat, en voici une qui repose sur la construction de base. On pose $\dim(\mathrm{Ker}(f)) = k$ et $\dim(\mathrm{Im}(f)) = l$. Soit $(e_1, \dots, e_k)$ une base de $\mathrm{Ker}(f)$ et $(u_1, \dots, u_l)$ une base de $\mathrm{Im}(f)$ (elles existent grâce au théorème 1). Comme chaque u_i est dans l'image de f on peut trouver $e'_i \in \mathrm{E}$ tel que $u_i = f(e'_i)$. Montrons que $(e_1, \dots, e_k, e'_1, \dots, e'_l)$ est une base de E.

- Cette famille est génératrice : Soit $u \in \mathrm{E}$. On a $f(u) = \lambda_1 u_1 + \cdots + \lambda_l u_l$ en décomposant $f(u)$ dans la base de $\mathrm{Im}(f)$. On pose $v = u - \lambda_1 e'_1 - \cdots - \lambda_k e'_k$. Montrons que $v \in \mathrm{Ker}(f)$. Il suffit de calculer $f(v) = f(u - \lambda_1 e'_1 - \cdots - \lambda_k e'_k) = f(u) - \lambda_1 f(e'_1) - \cdots - \lambda_k f(e'_k) = f(u) - \lambda_1 u_1 - \cdots - \lambda_k u_k = 0$. On peut donc décomposer v dans la base de $\mathrm{Ker}(f)$: $v = \mu_1 e_1 + \cdots + \mu_k e_k$. Ainsi on a écrit u comme combinaison linéaire des vecteurs de $(e_1, \dots, e_k, e'_1, \dots, e'_l)$:

$$u = \mu_1 e_1 + \cdots + \mu_k e_k + \lambda_1 e'_1 + \cdots + \lambda_l e'_l.$$

- Cette famille est libre. Prenons $a_1, \dots a_k, b_1, \dots, b_l$ tels que $a_1 e_1 + \cdots + a_k e_a + b_1 e'_1 + \dots b_l e'_l = 0$. Alors

$$\begin{aligned} f(a_1 e_1 + \cdots + a_k e_a + b_1 e'_1 + \dots b_l e'_l) &= f(0) \\ a_1 f(e_1) + \dots a_k f(e_k) + b_1 f(e'_1) + \cdots + b_l f(e'_l) &= 0 \\ b_1 u_1 + \cdots + b_l u_l &= 0 \end{aligned}$$

  Comme $(u_1, \dots u_l)$ est libre cela implique que tous les b_i sont nuls. On a alors $a_1 e_1 + \cdots + a_k e_k = 0$ et comme $(e_1, \dots e_k)$ est libre, tous les a_i sont nuls, ce qu'il fallait montrer.

□

Une application linéaire est parfois appelée *morphisme* ou *homomorphisme* d'espaces vectoriels (le terme morphisme signifie qu'on préserve la structure, ici la structure d'espace vectoriel, c'est-à-dire les deux lois + et ·).

Définition 12. Soit f une application linéaire de E dans F.

- f est un *isomorphisme* si f est bijective ("iso"= égal)
- f est un *endomorphisme* si $\mathrm{F} = \mathrm{E}$ ("endo"= dans)
- f est un *automorphisme* si f est bijective et $\mathrm{F} = \mathrm{E}$ ("auto"= être soi-même).

Exemples.

- L'application $f : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x, 0)$ est un endomorphisme de $\mathbb{R}^2$. Elle n'est pas bijective (pourquoi ?) donc ce n'est pas un automorphisme de $\mathbb{R}^2$.
- L'application $g : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + y, 2x - y)$ est un automorphisme de $\mathbb{R}^2$.

En combinant les propositions 12 et 13 avec les propriétés sur les dimensions (proposition 8), on peut montrer les caractérisations suivantes :

Proposition 14. Soit $\varphi \in \mathscr{L}(\mathrm{E}, \mathrm{F})$ une application linéaire entre espaces vectoriels. Les propriétés suivantes sont équivalentes :

- φ est un isomorphisme
- tout élément v de F admet un unique antécédent u dans E par φ (i.e. $\varphi(u) = v$)
- $\mathrm{Ker}(\varphi) = \{0\}$ et $\mathrm{Im}(\varphi) = \mathrm{F}$
- $\mathrm{Ker}(\varphi) = \{0\}$ et $\dim(\mathrm{E}) = \dim(\mathrm{F})$
- $\mathrm{Im}(\varphi) = \mathrm{F}$ et $\dim(\mathrm{E}) = \dim(\mathrm{F})$
- Il existe une base $(e_i)_{i\in\mathrm{I}}$ de E telle que $(\varphi(e_i))_{i\in\mathrm{I}}$ soit une base de F
- Pour toute base $(e_i)_{i\in\mathrm{I}}$ de E, $(\varphi(e_i))_{i\in\mathrm{I}}$ est une base de F.

Quand il existe un isomorphisme entre deux espaces vectoriels, on dit qu'ils sont *isomorphes*. La propriété précédente implique qu'ils sont nécessairement de même dimension.

Exemples.

- $\mathbb{R}^n$ et $\mathbb{R}^m$ ne sont isomorphes que si $n = m$
- Montrer que $\mathbb{R}^2$ et $\mathbb{R}_1[\mathrm{X}]$ sont isomorphes. (on pourra chercher une application qui envoie une base de $\mathbb{R}^2$ sur une base de $\mathbb{R}_1[\mathrm{X}]$).

L'exercice précédent se généralise à tous les espaces vectoriels de même dimension :

Théorème 3. Tout $\mathbb{K}$-espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

4 Matrices

Définition 13. Une *matrice* à n lignes et p colonnes sur un corps $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$) est un tableau à n lignes et p colonnes d'éléments de $\mathbb{K}$ (tableau de nombres). Elle s'écrit sous la forme

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} \quad \text{ou} \quad (a_{i,j})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$$

Le scalaire $a_{i,j}$ est appelé *coefficient* de i-ième ligne et j-ième colonne.

L'ensemble des matrices à n lignes et p colonnes (ou encore de taille $n \times p$) sur $\mathbb{K}$ se note $\mathrm{M}_{n,p}(\mathbb{K})$ (matrices *rectangulaires*). Si $n = p$ on note $\mathrm{M}_n(\mathbb{K}) = \mathrm{M}_{n,n}(\mathbb{K})$ (matrices *carrées*).

Exemples. $\begin{pmatrix} 2 & -1 \\ 3.2 & \pi \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$, $\begin{pmatrix} -i & 3+i & 1 \\ -4 & 5.6-\sqrt{2}i & 0 \end{pmatrix} \in \mathrm{M}_{2,3}(\mathbb{C})$, $\begin{pmatrix} -1 \\ 0.5 \\ 3 \end{pmatrix} \in \mathrm{M}_{3,1}(\mathbb{R})$.

De façon générale un vecteur de taille n à coefficients dans $\mathbb{K}$ peut-être considéré comme un élément de $\mathrm{M}_{n,1}(\mathbb{K})$.

Deux matrices sont égales si et seulement si tous leurs coefficients $a_{i,j}$ sont égaux (cela implique en particulier qu'elles ont la même taille). La matrice nulle est celle dont tous les coefficients sont nuls.

Définition 14. Une matrice *diagonale* est une matrice carrée dont tous les coefficients en dehors de la diagonale (descendante) sont nuls, c'est-à-dire de la forme $\begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}$ (notée aussi $\operatorname{diag}(a_{1,1}, \dots, a_{n,n})$).

Une matrice *triangulaire supérieure* (resp. *inférieure*) est une matrice carrée dont tous les coefficients en-dessous (resp. au dessus) de la diagonale sont nuls, c'est-à-dire de la forme $\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}$.

La matrice *identité* de taille n notée I_n est la matrice diagonale de $\mathrm{M}_n(\mathbb{K})$ dont tous les coefficients diagonaux sont égaux à 1, c'est-à-dire $\mathrm{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$.

4.1 Opérations sur les matrices

On peut additionner et multiplier des matrices par des scalaires, tout comme on le fait avec les vecteurs, c'est-à-dire coefficient par coefficient.

Définition 15. Soient $\mathrm{A} = (a_{i,j})_{\substack{1\le i\le n \\ 1\le j\le p}}$ et $\mathrm{B} = (b_{i,j})_{\substack{1\le i\le n \\ 1\le j\le p}}$ deux éléments de $\mathrm{M}_{n,p}(\mathbb{K})$, et $\lambda \in \mathbb{K}$. On définit

- $\mathrm{A} + \mathrm{B} = (a_{i,j} + b_{i,j})_{\substack{1\le i\le n \\ 1\le j\le p}}$
- $\lambda \cdot \mathrm{A} = (\lambda a_{i,j})_{\substack{1\le i\le n \\ 1\le j\le p}}$ (souvent noté $\lambda\mathrm{A}$)

Exemples.

- $\begin{pmatrix} 0 & -1.2 & 3 \\ 1 & 2 & -10 \end{pmatrix} + \begin{pmatrix} -2.5 & 0 & 1 \\ 0 & 3 & 5 \end{pmatrix} = \begin{pmatrix} -2.5 & -1.2 & 4 \\ 1 & 5 & -5 \end{pmatrix}$.

- $3\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ -6 & 3 \end{pmatrix}$, $\frac{1}{2}\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ -1 & 1/2 \end{pmatrix}$, $-\begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 2 & -1 \end{pmatrix}$.

Comme pour les vecteurs, ces deux opérations munissent l'ensemble des matrices de taille $n \times p$ d'une structure d'espace vectoriel.

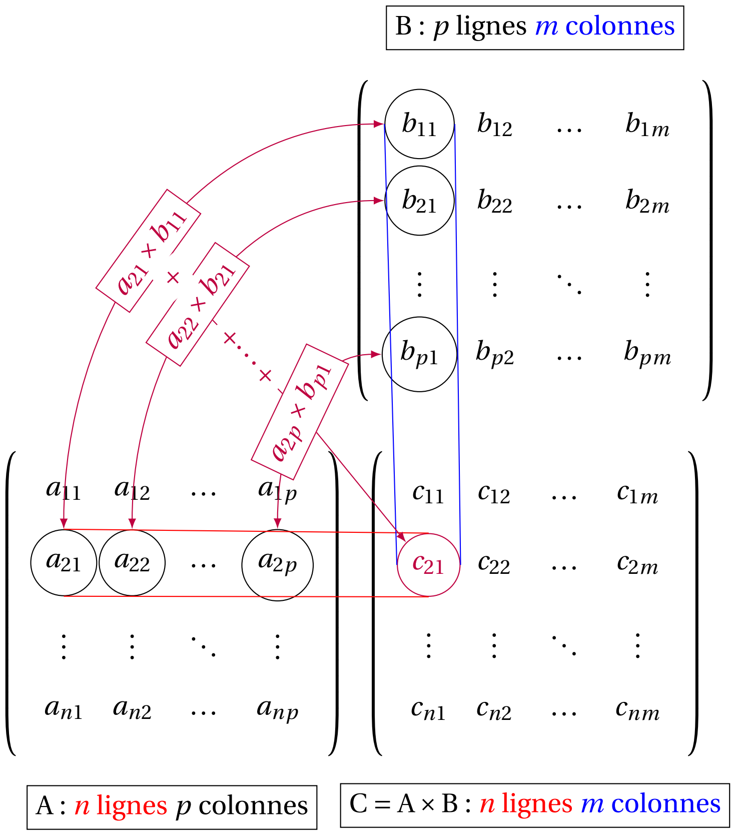
Proposition 15. $\mathrm{M}_{n,p}(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel de dimension $n \times p$.

Idée de preuve : On montre comme pour les vecteurs que les axiomes d'espace vectoriel sont vérifiés, car ils sont vérifiés pour chaque coefficient. On montre ensuite que la famille $(\mathrm{E}_{i,j})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ forme une base de $\mathrm{M}_{n,p}(\mathbb{K})$, où $\mathrm{E}_{i,j}$ est la matrice dont tous les coefficients sont nuls sauf celui à la place (i,j) qui vaut 1. Il y a exactement $n \times p$ éléments dans cette famille. □

Définition 16. Soit $\mathrm{A} = (a_{i,j})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}} \in \mathrm{M}_{n,p}(\mathbb{K})$ et $\mathrm{B} = (b_{i,j})_{\substack{1\leq i\leq p\\ 1\leq j\leq m}} \in \mathrm{M}_{p,m}(\mathbb{K})$. On définit

$$\mathrm{A} \times \mathrm{B} = \left(\sum_{k=1}^{p} a_{i,k} b_{k,j}\right)_{\substack{1\leq i\leq n\\ 1\leq j\leq m}} \in \mathrm{M}_{n,m}(\mathbb{K}).$$

Cette opération peut être visualisée en "faisant tomber les colonnes de B sur les lignes de A" comme sur la figure ci-dessous.



Exemple. $\begin{pmatrix} 0 & -1.2 & 3 \\ 1 & 2 & -10 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 11 \end{pmatrix}$.

Remarque. Attention, même pour des matrices carrées, **on n'a pas en général** $A \times B = B \times A$.

Par exemple $\begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} -3 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 3 & 1 \end{pmatrix} \neq \begin{pmatrix} -1 & -1 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} -3 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$.

L'opération de multiplication de matrices ainsi définie est compatible avec les opérations + et $\cdot$ dans le sens suivant :

Proposition 16. Soient A, B, C des matrices de $M_n(\mathbb{K})$ et $\lambda \in K$. Alors

- $(A \times B) \times C = A \times (B \times C)$,
- $A \times (B + C) = A \times B + A \times C, \quad (A + B) \times C = A \times C + B \times C$,
- $\lambda \cdot (A \times B) = (\lambda \cdot A) \times B = A \times (\lambda \cdot B)$
- $A \times I_n = I_n \times A = A$ où I_n est la matrice identité de taille n.

Dans la suite du chapitre nous allons surtout nous intéresser aux matrices carrées.

Définition 17. Soit $A \in M_n(\mathbb{K})$. On appelle *inverse* de A et on note A^{-1} l'unique matrice B, **si elle existe**, telle que $A \times B = B \times A = I_n$.

Si une telle matrice existe A est dite *inversible*.

Exemple.

- La matrice $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ n'est pas inversible car pour tout $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, on a
$$A \times B = \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

- La matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ est inversible d'inverse $A^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$. En effet on a bien $A \times A^{-1} = I_3 = A^{-1} \times A$.

Remarque. On notera souvent AB au lieu de $A \times B$ dans la suite (comme on le fait avec la multiplication de scalaires).

4.2 Matrices et systèmes linéaires

Tout système linéaire de n équations à n inconnues

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \cdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n \end{cases}$$

peut s'écrire sous forme matricielle :
$$AX = B$$

où $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix} \in M_n(\mathbb{K})$, $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(\mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in M_{n,1}(\mathbb{K})$.

Ainsi si A est inversible, **l'unique solution** du système linéaire est le vecteur
$$X = A^{-1}B$$

(obtenue en multipliant l'équation matricielle précédente par A^{-1} **à gauche**).

Une des méthodes pour inverser une matrice A est obtenue en analogie avec la méthode de résolution des systèmes linéaires par réduction à un système échelonné. Il s'agit de l'algorithme du *pivot de Gauss*.

Algorithme du pivot de Gauss pour déterminer si une matrice est inversible et le cas échéant calculer son inverse :

1. Écrire la matrice A ainsi que la matrice identité I_n de même taille dans deux colonnes.
2. Dans la colonne de gauche, échelonner la matrice A à l'aide des opérations suivantes :

- permutations de lignes $L_i \leftarrow L_j$
- ajout à une ligne d'une combinaison linéaire des autres $L_i \leftarrow L_i + \sum_{j\neq i} \lambda_j L_j$

A chaque étape, effectuer les mêmes opérations sur la matrice de la colonne de droite.

3. Si la matrice échelonnée à gauche a tous ses coefficients diagonaux non nuls, alors A est inversible et on continue l'algorithme. Sinon A n'est pas inversible et on arrête l'algorithme.
4. Dans la colonne de gauche, utiliser les opérations élémentaires $L_i \leftarrow L_i + \sum_{j\neq i} \lambda_j L_j$ jusqu'à se ramener à une matrice diagonale. Dans la colonne de droite appliquer les mêmes opérations élémentaires sur la matrice.
5. A l'aide d'opérations de type $L_i \leftarrow \lambda L_i$, se ramener à gauche à la matrice identité, et appliquer les mêmes opérations à la matrice dans la colonne de droite.

A la fin de l'algorithme, la matrice obtenue en bas de la colonne de droite est la matrice A^{-1}.

Exemple. Algorithme du pivot de Gauss appliqué à la matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

	$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$L_2 \leftarrow L_2 - L_1$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$L_3 \leftarrow L_3 + L_2$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$
$L_2 \leftarrow L_2 - \frac{1}{2}L_3$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & 1 & 1 \end{pmatrix}$
$L_1 \leftarrow L_1 + L_2$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & 1 & 1 \end{pmatrix}$
$L_2 \leftarrow -L_2$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -1 & 1 & 1 \end{pmatrix}$
$L_3 \leftarrow \frac{1}{2}L_3$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

Application : le système linéaire

$$\begin{cases} x+y & = 2 \\ x+z & = -1 \\ y+z & = -2 \end{cases}$$

admet une solution unique donnée par $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 1/2 \\ 5/2 \end{pmatrix}$.

4.3 Polynômes de matrices

Définition 18. Si $A \in M_n(\mathbb{K})$, on note $A^k = \underbrace{A \times A \times \cdots \times A}_{k \text{ fois}}$, et par convention $A^0 = I_n$.
Si $P = a_p X^p + a_{p-1} X^{p-1} + \cdots + a_1 X + a_0 \in \mathbb{K}[X]$ est un polynôme à coefficients dans $\mathbb{K}$, on définit

$$P(A) = a_p A^p + a_{p-1} A^{p-1} + \cdots + a_1 A + a_0 I_n.$$

Il s'agit d'un élément de $M_n(\mathbb{K})$.
On dit que le polynôme P est *annulateur* de A quand $P(A) = 0$ (matrice nulle).

L'existence de polynômes annulateurs peut permettre de tester l'inversibilité de la matrice A dans certains cas, comme nous allons le voir sur les exemples suivants.

- La matrice $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ admet le polynôme annulateur suivant $P(X) = X(X-1)$. En effet, calculons $A(A - I_2) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. En posant $B = A - I_2$, cela signifie qu'on a trouvé une matrice B non nulle telle $AB = 0$, donc A ne peut pas être inversible, sinon en multipliant l'égalité précédente par A^{-1} à gauche on obtiendrait $B = A^{-1}0 = 0$, ce qui est faux.
- La matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ admet le polynôme $P(X) = X^3 - 2X^2 - X + 2$ comme polynôme annulateur. En effet on calcule

$$A^3 - 2A^2 - A + 2I_3 = \begin{pmatrix} 3 & 3 & 2 \\ 3 & 2 & 3 \\ 2 & 3 & 3 \end{pmatrix} - 2 \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

  En réécrivant cette égalité sous la forme $A\left(-\frac{1}{2}A^2 + A + \frac{1}{2}I_3\right) = I_3 = \left(-\frac{1}{2}A^2 + A + \frac{1}{2}I_3\right)A$, on en déduit que A est inversible d'inverse $A^{-1} = -\frac{1}{2}A^2 + A + \frac{1}{2}I_3$.

4.4 Matrices d'applications linéaires

Définition 19. Soit $\varphi \in \mathscr{L}(E, F)$ une application linéaire entre deux $\mathbb{K}$-espaces vectoriels E et F de dimensions finies p et n respectivement. Soit $\mathscr{B} = (e_j)_{1 \le j \le p}$ une base de E et $\mathscr{B}' = (\varepsilon_i)_{1 \le i \le n}$. Alors la matrice de φ relativement aux bases $\mathscr{B}$ et $\mathscr{B}'$ est la matrice $M_{\mathscr{B},\mathscr{B}'}(\varphi) := \left(a_{i,j}\right)_{\substack{1 \le i \le n \\ 1 \le j \le p}}$ où $a_{i,j}$ est le coefficient de $\varphi(e_j)$ selon ε_i.
Autrement dit la j-ième colonne de $M_{\mathscr{B},\mathscr{B}'}(\varphi)$ donne les coordonnées de $\varphi(e_j)$ dans la base $\mathscr{B}'$.

$$\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi) = \begin{array}{c c} \begin{array}{cccc} \varphi(e_1) & \varphi(e_2) & & \varphi(e_p) \end{array} & \\ \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix} & \begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \\ \varepsilon_n \end{array} \end{array}$$

Exemple.

- La matrice de l'application $g : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x+y, 2x-y)$ dans la base canonique $\mathscr{B} = (e_1, e_2)$ est $\mathrm{M}_{\mathscr{B},\mathscr{B}}(g) = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$.
- La matrice de l'application $\phi : \mathbb{R}_2[\mathrm{X}] \to \mathbb{R}_1[\mathrm{X}], \mathrm{P} \mapsto \mathrm{P}'$ relative aux bases $\mathscr{B} = (1, \mathrm{X}, \mathrm{X}^2)$ de $\mathbb{R}_2[\mathrm{X}]$ et $\mathscr{B}' = (1, \mathrm{X})$ de $\mathbb{R}_1[\mathrm{X}]$ est

$$\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\phi) = \begin{array}{c c} \begin{array}{ccc} \phi(1) & \phi(\mathrm{X}) & \phi(\mathrm{X}^2) \end{array} & \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} & \begin{array}{c} 1 \\ \mathrm{X} \end{array} \end{array}.$$

- L'application "identité" $\varphi : \mathbb{R}^2 \to \mathbb{R}^2, u \mapsto u$ a pour matrice $\mathrm{M}_{\mathscr{B},\mathscr{B}}(\varphi) = \mathrm{I}_2$ dans la base canonique $\mathscr{B}$, mais en posant $\mathscr{B}' = (e_1 + e_2, e_1 - e_2)$, on a $\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$, $\mathrm{M}_{\mathscr{B}',\mathscr{B}}(\varphi) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ et $\mathrm{M}_{\mathscr{B}',\mathscr{B}'}(\varphi) = \mathrm{I}_2$.

Toutes les opérations usuelles sur les applications linéaires se traduisent en terme de matrices une fois qu'on a fixé une base de chaque espace vectoriel.

Proposition 17. Soient E,F G des $\mathbb{K}$-espaces vectoriels de dimension finie, de bases respectives $\mathscr{B}, \mathscr{B}'$ et $\mathscr{B}''$.

- $\forall (\lambda, \mu) \in \mathbb{K}^2$, $\forall (\varphi, \psi) \in \mathscr{L}(\mathrm{E}, \mathrm{F})$, $\lambda\varphi + \mu\psi \in \mathscr{L}(\mathrm{E}, \mathrm{F})$ et

$$\mathrm{M}_{\mathscr{B},\mathscr{B}'}\big(\lambda\varphi + \mu\psi\big) = \lambda \mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi) + \mu \mathrm{M}_{\mathscr{B},\mathscr{B}'}(\psi).$$

- $\forall \varphi \in \mathscr{L}(\mathrm{E}, \mathrm{F}), \forall \Phi \in \mathscr{L}(\mathrm{F}, \mathrm{G}), \Phi \circ \varphi \in \mathscr{L}(\mathrm{E}, \mathrm{G})$ et

$$\mathrm{M}_{\mathscr{B},\mathscr{B}''}\big(\Phi \circ \varphi\big) = \mathrm{M}_{\mathscr{B}',\mathscr{B}''}(\Phi) \times \mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi).$$

- Si $\dim(\mathrm{E}) = \dim(\mathrm{F})$ et $\varphi \in \mathscr{L}(\mathrm{E}, \mathrm{F})$ alors

$$\big(\varphi \text{ est bijective}\big) \iff \big(\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi) \text{ est inversible}\big).$$

Et dans ce cas $\mathrm{M}_{\mathscr{B}',\mathscr{B}}(\varphi^{-1}) = \big(\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi)\big)^{-1}$.

- Soient $u \in \mathrm{E}$ de coordonnées $\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ et $\varphi \in \mathscr{L}(\mathrm{E},\mathrm{F})$. Les coordonnées de $\varphi(u)$ dans la base $\mathscr{B}'$ sont $\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi) \times \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$.

Exemple. En reprenant les exemples précédents (applications $g : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + y, 2x - y)$ et $\varphi : \mathbb{R}^2 \to \mathbb{R}^2, u \mapsto u)$, on a

- $\mathrm{M}_{\mathscr{B},\mathscr{B}}(2g + \varphi) = \begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix}$
- $\mathrm{M}_{\mathscr{B}',\mathscr{B}}(g \circ \varphi) = \mathrm{M}_{\mathscr{B},\mathscr{B}}(g)\mathrm{M}_{\mathscr{B}',\mathscr{B}}(\varphi) = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$.
- $\mathrm{M}_{\mathscr{B},\mathscr{B}}(g)$ est inversible d'inverse $\left(\mathrm{M}_{\mathscr{B},\mathscr{B}}(g)\right)^{-1} = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix}$ donc g est bijective et la matrice de g^{-1} est $\mathrm{M}_{\mathscr{B},\mathscr{B}}(g^{-1}) = \left(\mathrm{M}_{\mathscr{B},\mathscr{B}}(g)\right)^{-1}$.
- Soit $u = e_1 + 2e_2$. Les coordonnées de $\varphi(u)$ dans la base $\mathscr{B}' = (e_1 + e_2, e_1 - e_2)$ sont $\mathrm{M}_{\mathscr{B},\mathscr{B}'}(\varphi) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ -1/2 \end{pmatrix}$. On peut vérifier qu'en effet $u = e_1 + 2e_2 = \frac{3}{2}(e_1 + e_2) - \frac{1}{2}(e_1 - e_2)$.

Proposition 18. Soient $\mathscr{B}_\mathrm{E}$ et $\mathscr{B}'_\mathrm{E}$ des bases de E, et $\mathscr{B}_\mathrm{F}$ et $\mathscr{B}'_\mathrm{F}$ des bases de F. Soit $\varphi \in \mathscr{L}(\mathrm{E},\mathrm{F})$. Alors

$$\mathrm{M}_{\mathscr{B}'_\mathrm{E},\mathscr{B}'_\mathrm{F}}(\varphi) = \mathrm{M}_{\mathscr{B}_\mathrm{F},\mathscr{B}'_\mathrm{F}}(\mathrm{Id}_\mathrm{F}) \times \mathrm{M}_{\mathscr{B}_\mathrm{E},\mathscr{B}_\mathrm{F}}(\varphi) \times \mathrm{M}_{\mathscr{B}'_\mathrm{E},\mathscr{B}_\mathrm{E}}(\mathrm{Id}_\mathrm{E}).$$

Idée de preuve : On considère le diagramme suivant

$$\begin{array}{ccc} \mathrm{E}_{\mathscr{B}_\mathrm{E}} & \xrightarrow{\varphi} & \mathrm{F}_{\mathscr{B}_\mathrm{F}} \\ \mathrm{Id}_\mathrm{E} \uparrow & & \downarrow \mathrm{Id}_\mathrm{F} \\ \mathrm{E}_{\mathscr{B}'_\mathrm{E}} & \xrightarrow{\varphi} & \mathrm{F}_{\mathscr{B}'_\mathrm{F}} \end{array}$$

et on applique la proposition précédente pour la composée $\varphi = \mathrm{Id}_\mathrm{F} \circ \varphi \circ \mathrm{Id}_\mathrm{E}$. □

Exemple. La matrice de $g : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + y, 2x - y)$ dans la base $\mathscr{B}' = (e_1 + e_2, e_1 - e_2)$ (au départ et à l'arrivée), peut être calculée comme

$$\mathrm{M}_{\mathscr{B}',\mathscr{B}'}(\varphi) = \mathrm{M}_{\mathscr{B},\mathscr{B}'}(\mathrm{Id}) \times \mathrm{M}_{\mathscr{B},\mathscr{B}}(g) \times \mathrm{M}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3/2 & -3/2 \\ 1/2 & -3/2 \end{pmatrix},$$

ce qui peut être vérifié également directement (exercice).

5 Déterminants

5.1 Déterminant d'une famille de vecteurs

Proposition 19. Soit E un $\mathbb{K}$-espace vectoriel de dimension n, et $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E. On appelle application *déterminant* (relativement à la base $\mathscr{B}$) l'application

$$\begin{array}{cccc} \det_{\mathscr{B}} & : & \mathrm{E}^n & \to & \mathbb{K} \\ & & (u_1, \ldots, u_n) & \mapsto & \det_{\mathscr{B}}(u_1, \ldots, u_n) \end{array}$$

telle que

- $\det_{\mathscr{B}}$ est linéaire par rapport à chacune de ses n variables
- $\det_{\mathscr{B}}(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_n) = -\det_{\mathscr{B}}(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_n)$
- $\det_{\mathscr{B}}(e_1, \ldots, e_n) = 1$.

Idée de preuve : On montre que n'importe quelle application f satisfaisant ces 3 propriétés est entièrement déterminée. Pour exprimer $f(u_1, \ldots, u_n)$, par linéarité on se ramène à des combinaisons de $f(e_i, \ldots, e_j)$ (vecteurs de base pas forcément dans l'ordre), et en échangeant les vecteurs deux par deux on peut se ramener à un multiple de $f(e_1, \ldots e_n)$. Si cette valeur est fixée à 1, alors $f(u_1, \ldots u_n)$ est entièrement déterminé. $\square$

Quand il n'y a pas d'ambiguïté possible sur le choix de la base $\mathscr{B}$ (base canonique par exemple), on note simplement $\det(u_1, \ldots, u_n)$.

Si $u_i = \begin{pmatrix} u_{1,i} \\ u_{2,i} \\ \vdots \\ u_{n,i} \end{pmatrix}$ sont les coordonnées de u_i dans la base $\mathscr{B}$, on note souvent

$$\det(u_1, \ldots, u_n) = \begin{vmatrix} u_{1,1} & u_{1,2} & \ldots & u_{1,n} \\ u_{2,1} & u_{2,2} & \ldots & u_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n,1} & u_{n,2} & \ldots & u_{n,n} \end{vmatrix}.$$

Exemples.

$n = 1 : \det(u) = u$

$n = 2 : \det(u, v) = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} = u_1 v_2 - u_2 v_1$. (règle de Sarrus)

$$n = 3 : \det(u, v, w) = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = u_1 v_2 w_3 + u_2 v_3 w_1 + u_3 v_1 w_2 - u_3 v_2 w_1 - u_1 v_2 w_2 - u_2 v_3 w_1.$$
$$\begin{array}{ccc} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_3 \end{array}$$

(règle de Sarrus)

$n \geq 4$: Attention, la règle de Sarrus ne s'applique plus ! Nous allons voir d'autres techniques pour calculer ces déterminants.

Proposition 20.

- $\det(u_1, \ldots, u, \ldots, u, \ldots, u_n) = 0$
- $\det(u_1, \ldots, 0, \ldots, u_n) = 0$
- $\forall \lambda_1, \ldots, \lambda_n \in \mathbb{K}, \det(u_1, \ldots, u_i + \sum_{j \neq i} \lambda_j u_j, \ldots, u_n) = \det(u_1, \ldots, u_i, \ldots, u_n)$
- $\forall \lambda \in \mathbb{K}, \det(u_1, \ldots, \lambda u, \ldots, u_n) = \lambda \det(u_1, \ldots, u, \ldots, u_n)$
- $\forall \lambda \in \mathbb{K}, \det(\lambda u_1, \ldots, \lambda u_n) = \lambda^n \det(u_1, \ldots, u_n)$.

Idée de preuve : Ces propriétés découlent directement de la définition. Par exemple, pour la première, si deux vecteurs colonnes sont égaux (u et u), en les échangeant, on doit avoir $\det(u_1, \ldots, u, \ldots, u, \ldots, u_n) = -\det(u_1, \ldots, u, \ldots, u, \ldots, u_n)$. Ainsi la seule possibilité est que $\det(u_1, \ldots, u, \ldots, u, \ldots, u_n) = 0$. □

Proposition 21. Le développement par rapport à la j-ième colonne est

$$\det(u_1, \ldots, u_n) = \sum_{i=1}^{n} (-1)^{i+j} u_{i,j} \mathrm{D}_{i,j},$$

où $\mathrm{D}_{i,j}$ est le déterminant obtenu en effaçant la i-ème ligne et la j-ième colonne.

Autrement dit on a :

$$\begin{vmatrix} u_{1,1} & \ldots & u_{1,j} & \ldots & u_{1,n} \\ u_{2,1} & \ldots & u_{2,j} & \ldots & u_{2,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{n,1} & \ldots & u_{n,j} & \ldots & u_{n,n} \end{vmatrix} = (-1)^{1+j} u_{1,j} \begin{vmatrix} u_{1,1} & \ldots & u_{1,j} & \ldots & u_{1,n} \\ u_{2,1} & \ldots & u_{2,j} & \ldots & u_{2,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{n,1} & \ldots & u_{n,j} & \ldots & u_{n,n} \end{vmatrix} + (-1)^{2+j} u_{2,j} \begin{vmatrix} u_{1,1} & \ldots & u_{1,j} & \ldots & u_{1,n} \\ u_{2,1} & \ldots & u_{2,j} & \ldots & u_{2,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{n,1} & \ldots & u_{n,j} & \ldots & u_{n,n} \end{vmatrix}$$

$$+ \cdots + (-1)^{n+j} u_{n,j} \begin{vmatrix} u_{1,1} & \ldots & u_{1,j} & \ldots & u_{1,n} \\ u_{2,1} & \ldots & u_{2,j} & \ldots & u_{2,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ u_{n,1} & \ldots & u_{n,j} & \ldots & u_{n,n} \end{vmatrix}$$

$$= (-1)^{1+j} u_{1,j} \begin{vmatrix} u_{2,1} & \ldots & u_{2,j-1} & u_{2,j+1} & \ldots & u_{2,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ u_{n,1} & \ldots & u_{n,j-1} & u_{n,j+1} & \ldots & u_{n,n} \end{vmatrix}$$

$$+ (-1)^{2+j} u_{2,j} \begin{vmatrix} u_{1,1} & \ldots & u_{1,j-1} & u_{1,j+1} & \ldots & u_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ u_{n,1} & \ldots & u_{n,j-1} & u_{n,j+1} & \ldots & u_{n,n} \end{vmatrix} + \ldots$$

$$+ (-1)^{n+j} u_{n,j} \begin{vmatrix} u_{1,1} & \ldots & u_{1,j-1} & u_{n,j+1} & \ldots & u_{1,n} \\ \vdots & \ddots & \ldots & \vdots & \ddots & \vdots \\ u_{n-1,1} & \ldots & u_{n-1,j-1} & u_{n-1,j+1} & \ldots & u_{n-1,n} \end{vmatrix}.$$

Exemple. Calcul du déterminant $\begin{vmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{vmatrix}$ à l'aide des différentes méthodes et propriétés :

- Méthode de Sarrus :

$$\begin{vmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{vmatrix} = 1\cdot 0\cdot 0 + 2\cdot 3\cdot 1 + (-1)\cdot(-1)\cdot 1 - (-1)\cdot 0\cdot 1 - 2\cdot(-1)\cdot 0 - 1\cdot 3\cdot 1 = 6+1-3 = 4.$$

- Développement par rapport à la deuxième colonne (on choisit en général une colonne contenant le maximum de zéros) :

$$\begin{vmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{vmatrix} = (-1)^{1+2}\cdot 2\cdot \begin{vmatrix} -1 & 3 \\ 1 & 0 \end{vmatrix} + (-1)^{2+2}\cdot 0\cdot \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} + (-1)^{3+2}\cdot 1\cdot \begin{vmatrix} 1 & -1 \\ -1 & 3 \end{vmatrix} = -2\cdot(-3) - (3-1) = 4.$$

- Opérations sur les colonnes (et développement par rapport à une colonne) :

$$\begin{aligned} \begin{vmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{vmatrix} &= \begin{vmatrix} 0 & 2 & -1 \\ -4 & 0 & 3 \\ 0 & 1 & 0 \end{vmatrix} \qquad (C_1 \leftarrow C_1 - C_2 - C_3) \\ &= (-1)^{2+1}(-4)\begin{vmatrix} 2 & -1 \\ 1 & 0 \end{vmatrix} = 4 \quad \text{(développement par rapport à } C_1\text{).} \end{aligned}$$

Proposition 22. Soit E un $\mathbb{K}$-espace vectoriel de dimension n, $(u_i)_{1\le i\le n}$ une famille de n vecteurs de E et $\mathscr{B}$ une base de E.

- $\big(\det_{\mathscr{B}}(u_1,\dots,u_n) = 0\big) \Rightarrow \big(\forall \mathscr{B}' \text{ base de E}, \det_{\mathscr{B}'}(u_1,\dots,u_n) = 0\big)$.
- $\big(\det_{\mathscr{B}}(u_1,\dots,u_n) = 0\big) \iff \big((u_i)_{1\le i\le n} \text{ est liée.}\big)$
- $\big(\det_{\mathscr{B}}(u_1,\dots,u_n) \neq 0\big) \iff \big((u_i)_{1\le i\le n} \text{ est libre. (donc est une base de E).}\big)$

Idée de preuve : Ces propriétés résultent du fait que $\det_{cB'}$ et $\det\mathscr{B}$ sont deux applications satisfaisant les deux premiers axiomes de la proposition 19. Elles sont donc proportionnelles : il existe $\lambda \in \mathbb{K}$ tel que pour toute famille (u_i), $\det_{\mathscr{B}'}(u_1,\dots,u_n) = \lambda \det_{\mathscr{B}}(u_1,\dots,u_n)$. La valeur de λ est alors déterminée par la normalisation choisie (troisième axiome). Autrement dit pour $(u_1,\dots u_n) = (e_1,\dots,e_n)$ on doit avoir $\det_{\mathscr{B}'}(e_1,\dots,e_n) = \lambda \det_{\mathscr{B}}(e_1,\dots,e_n) = \lambda$ d'où

$$\forall (u_i)_{1\le i\le n}, \quad \det_{\mathscr{B}'}(u_1,\dots,u_n) = \det_{\mathscr{B}'}(e_1,\dots,e_n)\det_{\mathscr{B}}(u_1,\dots,u_n).$$

Si la famille (u_i) est liée, en utilisant la Proposition 20, on montre que le déterminant est nul.

Si la famille $(u_i)_{1\le i\le n}$ est libre, c'est une base (car elle a le bon nombre de vecteurs, cf Prop.7), en notant $\mathscr{B}'$ cette base, on a, en utilisant l'éqalité précédente,

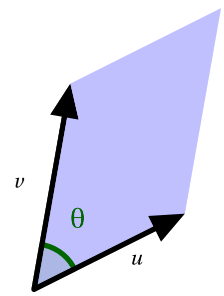
$$\det_{\mathscr{B}'}(u_1,\dots,u_n) = 1 = \det_{\mathscr{B}'}(e_1,\dots,e_n)\det_{\mathscr{B}}(u_1,\dots,u_n),$$

d'où $\det_{\mathscr{B}}(u_1,\dots,u_n) \neq 0$. □

5.2 Aire et volume

Proposition 23. Soient u et v deux vecteurs de $\mathbb{R}^2$, alors

$$|\det(u,v)| = \|u\|\cdot\|v\|\cdot|\sin(\theta)| = \text{ Aire du parallélogramme formé par } u \text{ et } v.$$



Proposition 24. Soient u, v et w trois vecteurs de $\mathbb{R}^3$, alors

$$|\det(u,v,w)| = \text{ Volume du parallélépipède formé par } u, v \text{ et } v.$$

Remarque. Le signe du déterminant peut être vu comme une orientation de l'aire ou du volume considéré. Le déterminant de n vecteurs peut être vu comme une généralisation du volume (orienté) à n dimensions.

5.3 Déterminant et changement de variables

Théorème 4 (Changement de variables pour l'intégration)**.** Soit U un domaine de $\mathbb{R}^n$ et soit F une application de classe C^1 de U dans $\mathbb{R}^n$, injective. Si $g : U \to \mathbb{R}_+$ est intégrable, alors

$$\int_{F(U)} g(y_1,\ldots,y_n)dy_1\ldots dy_n = \int_U g(F(x_1,\ldots,x_n))\left|\det\left(\frac{\partial F}{\partial x_1},\ldots,\frac{\partial F}{\partial x_n}\right)\right| dx_1\ldots dx_n.$$

Exemple (Passage aux coordonnées polaires)**.** . On considère

$$\begin{array}{cccc} F & : & \mathbb{R}_+\times[0,2\pi[& \to & \mathbb{R}^2 \\ & & (r,\theta) & \mapsto & (r\cos(\theta), r\sin(\theta)) \end{array}$$

On a $\det\left(\frac{\partial F}{\partial r}(r,\theta), \frac{\partial F}{\partial \theta}(r,\theta)\right) = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r(\cos^2(\theta)+\sin^2(\theta)) = r.$

Ainsi l'élément d'aire élémentaire en coordonnées euclidiennes $dxdy$ s'écrit $rdrd\theta$ en cordonnées polaires.

Exercice : dans $\mathbb{R}^3$ calculer les éléments de volume en coordonnées cylindriques et sphériques.

5.4 Déterminant d'une matrice

Définition 20. Le déterminant d'une matrice est le déterminant de ces vecteurs colonnes.

Proposition 25. Soit $A \in M_n(\mathbb{K})$, $A = (a_{i,j})_{1\le i,j\le n}$. La *transposée* de A, notée ${}^tA$ est la matrice $(a_{j,i})_{1\le i,j\le n}$ (symétrisée par rapport à la diagonale).

On a $\det(A) = \det({}^tA)$.

En conséquence, toutes les propriétés du déterminant relatives aux colonnes s'appliquent également aux lignes (ex : opérations élémentaires sur les lignes, développement par rapport à une ligne).

Exemple. $\begin{vmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ -1 & 3 & 0 \end{vmatrix}$.

Proposition 26. Le déterminant d'une matrice triangulaire supérieure est le produit des termes diagonaux :

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{vmatrix} = a_{1,1} \cdot a_{2,2} \cdot \dots \cdot a_{n,n}.$$

Cela est vrai en particulier pour les matrices diagonales !

Idée de preuve : Ce résultat se montre par récurrence sur n la taille de la matrice, en considérant le développement par rapport à la première colonne (ou la dernière ligne). □

Proposition 27.
- $\forall A, B \in M_n(\mathbb{K}), \det(A \times B) = \det(A)\det(B)$.
- $\forall A \in M_n(\mathbb{K})$, $\big(\text{A est inversible}\big) \iff \big(\det(A) \neq 0\big)$.
  Dans ce cas $\det(A^{-1}) = \frac{1}{\det(A)}$, et $A^{-1} = \frac{1}{\det(A)}\,{}^t\mathrm{C}o(A)$ où $\mathrm{C}o(A) = \big((-1)^{i+j} D_{i,j}\big)_{1\le i,j\le n}$ est la *comatrice* de A, et $D_{i,j}$ comme dans la Prop. 21.
- Des matrices A et B de $M_n(\mathbb{K})$ sont dites *semblables* si il existe une matrice inversible $P \in M_n(\mathbb{K})$ telle que $A = PBP^{-1}$.
  Dans ce cas $\det(A) = \det(B)$.

Remarque. La dernière propriété résulte des deux premières. Elle implique en particulier par la Proposition 18 que si $\varphi \in \mathcal{L}(E)$, et $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, $\det(M_{\mathcal{B},\mathcal{B}}(\varphi)) = \det(M_{\mathcal{B}',\mathcal{B}'}(\varphi))$. Dans la suite on notera $M_{\mathcal{B}}(\varphi)$ pour $M_{\mathcal{B},\mathcal{B}}(\varphi)$ (il est usuel de choisir la même base au départ et à l'arrivée pour des endomorphismes, on dira que φ a pour matrice $M_{\mathcal{B}}(\varphi)$ dans la base $\mathcal{B}$).

Exemple. Soit $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{pmatrix}$. On a vu que $\det(A) = 4 \neq 0$. Cela signifie que A est inversible.

Calculons A^{-1} à l'aide de la formule $\frac{1}{\det(A)}{}^t\mathrm{C}o(A)$. On a

$$\mathrm{C}o(A) = \begin{pmatrix} +\begin{vmatrix} 0 & 3 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} -1 & 3 \\ 1 & 0 \end{vmatrix} & +\begin{vmatrix} -1 & 0 \\ 1 & 1 \end{vmatrix} \\ -\begin{vmatrix} 2 & -1 \\ 1 & 0 \end{vmatrix} & +\begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} \\ +\begin{vmatrix} 2 & -1 \\ 0 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & -1 \\ -1 & 3 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -3 & 3 & -1 \\ -1 & 1 & 1 \\ 6 & -2 & 2 \end{pmatrix}.$$

Ainsi on obtient l'inverse en passant à la transposée et en divisant par le déterminant :

$$A^{-1} = \frac{1}{4}\begin{pmatrix} -3 & -1 & 6 \\ 3 & 1 & -2 \\ -1 & 1 & 2 \end{pmatrix}.$$

Remarque. La formule de l'inverse pour $n = 2$ est tout simplement

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

On l'utilise très souvent !

Attention par contre, la formule de l'inverse devient très pénible à calculer dès que $n \geq 3$. La méthode par pivot de Gauss est à privilégier dans ce cas (c'est une des méthodes implantées dans python pour calculer l'inverse de façon efficace).

Définition 21. Soit $\varphi \in \mathscr{L}(E)$ et $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E. On définit le déterminant de φ par $\det(\varphi) := \det(M_{\mathscr{B}}(\varphi)) = \det(\varphi(e_1), \ldots, \varphi(e_n))$. Il ne dépend pas de la base choisie $\mathscr{B}$.

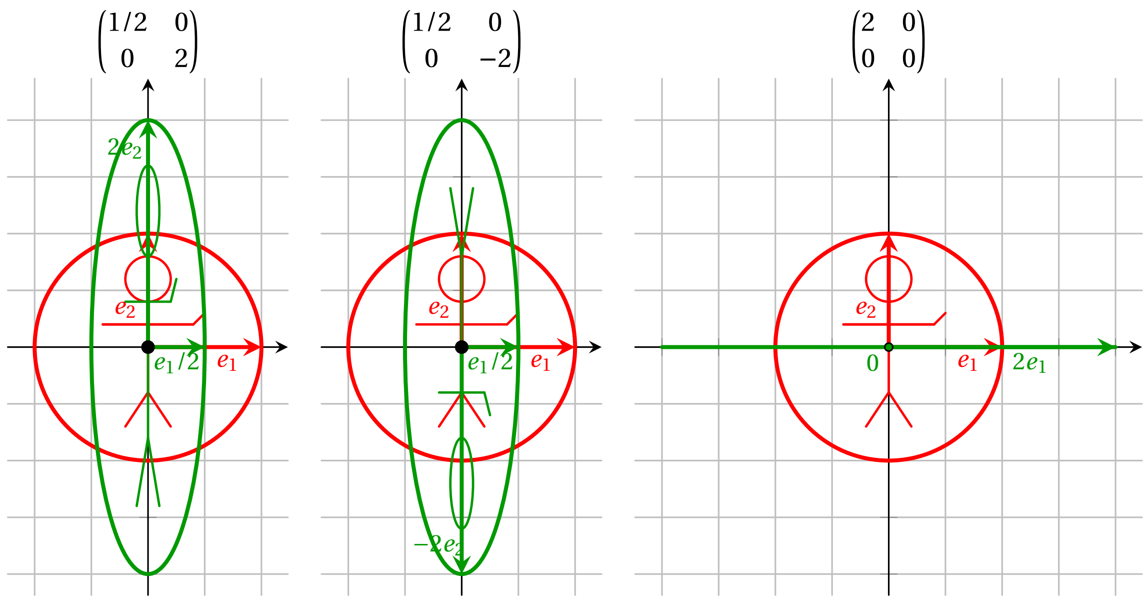
6 Diagonalisation

Dans ce chapitre, E désigne un espace vectoriel de dimension finie sur un corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

6.1 Principe et intérêt de la diagonalisation

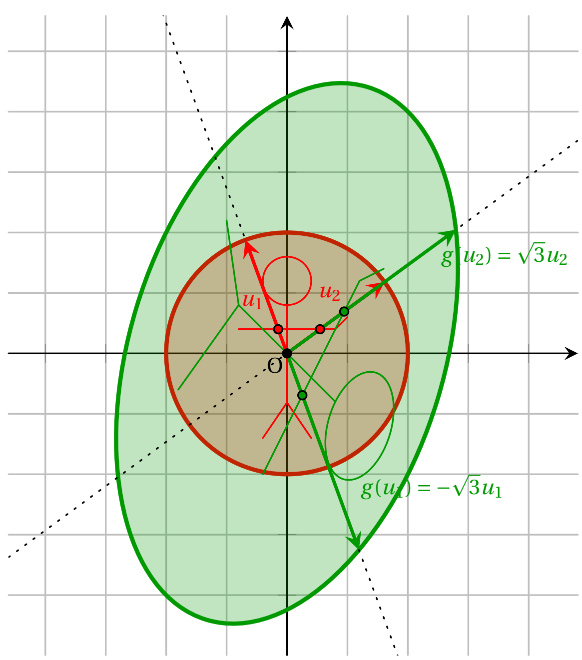
On a vu précédemment que tout endomorphisme pouvait être encodé par une matrice, une fois fixé un choix de base de l'espace vectoriel. Nous allons nous intéresser aux applications linéaires dites "diagonalisables", c'est-à-dire celles pour lesquelles il existe un choix de base tel que la matrice représentative de l'application est diagonale.

Un endomorphisme représenté par une matrice diagonale est particulièrement facile à comprendre. Voici par exemple l'action d'endomorphismes de $\mathbb{R}^2$ dont les matrices dans la base canonique sont diagonales.



En particulier pour de tels endomorphismes φ, la puissance n-ième de l'endomorphisme (pour la composition) est très facile à calculer, puisque la puissance n-ième d'une matrice diagonale se calcule très facilement.

A titre d'exemple, l'endomorphisme g de la section 3 est diagonalisable. En effet, dans la base de vecteurs $u_1 = (1-\sqrt{3}, 2)$, $u_2 = (1+\sqrt{3}, 2)$, la matrice de g prend la forme $\begin{pmatrix} -\sqrt{3} & 0 \\ 0 & \sqrt{3} \end{pmatrix}$.

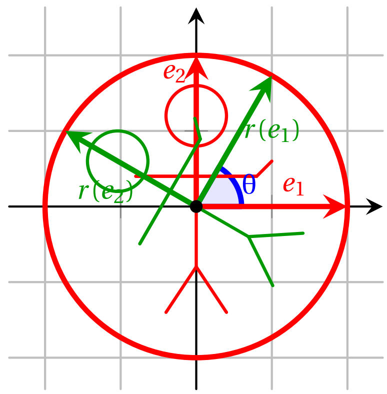


Les vecteurs u_1 et u_2 jouent donc un rôle particulier vis à vis de l'application g, on parlera de "vecteurs propres" : les droites engendrées par ces vecteurs sont préservées par g ("espaces propres").

Attention, il existe des applications qui ne peuvent être représentées par des matrices diagonales dans aucune base, on verra pourquoi. En voici quelques exemples dans $\mathbb{R}^2$.

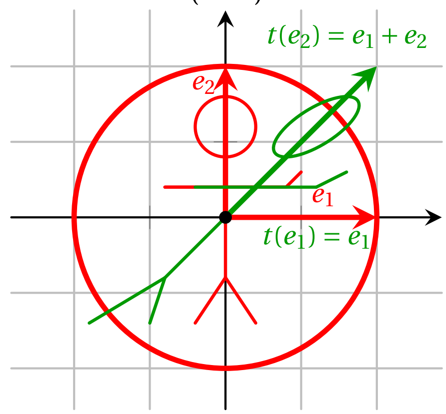
rotation r

$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$



transvection t

$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$



6.2 Valeurs propres et vecteurs propres

Définition 22. Soit E un $\mathbb{K}$-espace vectoriel et $\varphi \in \mathscr{L}(\mathrm{E})$. On dit que $\lambda \in \mathbb{K}$ est *valeur propre* de φ s'il existe un vecteur u non nul tel que $\varphi(u) = \lambda u$.

Un tel vecteur $u \neq 0$ est alors appelé *vecteur propre* de φ associé à la valeur propre λ.

Le sous-espace vectoriel $\mathrm{E}_\lambda(u) := \mathrm{Ker}(\varphi - \lambda \mathrm{I}d) = \{u \in \mathrm{E}, \varphi(u) = \lambda u\}$ est le *sous-espace propre* associé à λ.

L'ensemble des valeurs propres de φ est appelé *spectre* de φ et noté $\mathrm{Sp}(\varphi)$.

Remarque. L'égalité $\varphi(u) = \lambda u$ se traduit en termes matriciels (après avoir choisi une base $\mathscr{B}$) en $\mathrm{MU} = \lambda \mathrm{U}$ où $\mathrm{M} = \mathrm{M}_{\mathscr{B}}(\varphi)$ est la matrice de φ dans la base $\mathscr{B}$, et U est le vecteur colonne des coordonnées de u dans la base $\mathscr{B}$. On parle également de vecteur propre et de valeur propre pour la matrice M. L'espace propre en terme matriciels est donc $\mathrm{Ker}(\mathrm{M} - \lambda \mathrm{I}_n)$.

Exemples.

- L'endomorphisme φ canoniquement associé à la matrice $\begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}$ (c'est-à-dire celui dont la matrice dans la base canonique est M), a pour valeurs propres $1/2$ et 2. Le vecteur e_1 est un vecteur propre associé à la valeur propre $1/2$, il en est de même pour tous ses multiples. Le vecteur e_2 est un vecteur propre associé à la valeur propre 2. On a $\mathrm{E}_{1/2}(\varphi) = \mathrm{Vect}(e_1)$ et $\mathrm{E}_2(\varphi) = \mathrm{Vect}(e_2)$.
- L'application $g : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + y, 2x - y)$ a pour valeurs propres $\sqrt{3}$ (vecteur propre associé $(1+\sqrt{3}, 2)$)et $-\sqrt{3}$, (vecteur propre associé $(1-\sqrt{3}, 2)$).
- La rotation r d'angle 60° ou $\pi/3$ (de matrice $\begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$ dans la base canonique) n'a aucune valeur propre réelle. (tous les vecteurs sont tournés par r !)
- La matrice $\mathrm{B} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}$ a pour valeur propre 2. En effet Le vecteur $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ est propre (cela peut se voir car par exemple la somme des coefficients sur chaque ligne est constante égale à 2). Cherchons

$E_2(B) = \mathrm{Ker}(B - 2I_3)$. Pour cela on écrit la matrice $B - 2I_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$ et on cherche son noyau, c'est-à-dire les vecteurs $U = \begin{pmatrix} x \\ y \\ y \end{pmatrix}$ tels que $(B - 2I_3)U = 0$. On obtient l'égalité

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ x-y \\ x-y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

On a donc

$$E_2(B) = \mathrm{Ker}(B - 2I_3) = \left\{ \begin{pmatrix} a \\ a \\ b \end{pmatrix}, a, b \in \mathbb{R} \right\} = \mathrm{Vect}\left(\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right).$$

et l'espace propre est de dimension 2.

- De même on peut montrer que la matrice $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{pmatrix}$ admet la valeur propre 2 et que l'espace propre associé $E_2(A) = \mathrm{Vect}(\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\})$ est de dimension 1.

6.3 Polynôme caractéristique

Avant de définir le polynôme caractéristique d'un endomorphisme ou d'une matrice, on va énoncer quelques propriétés à connaître sur les polynômes à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$.

Propriétés des polynômes sur $\mathbb{R}$ et $\mathbb{C}$. Soit $P(X) = a_0 + a_1 X + \cdots + a_n X^n \in \mathbb{K}[X]$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Si $a_n \neq 0$ on dit que P est de *degré n*. Un $\lambda \in \mathbb{K}$ tel que $P(\lambda) = 0$ est appelé *racine* de P dans $\mathbb{K}$. On a l'équivalence :

$$\big(\lambda \text{ racine de } P\big) \iff \big(\exists Q \in \mathbb{K}[X],\ P(X) = (X - \lambda)Q(X)\big).$$

Si le facteur $(X - \lambda)$ se factorise plusieurs fois dans P on parle de racine avec multiplicité (en opposition à racine *simple* quand il n'y a pas de multiplicité) :

$$\big(\lambda \text{ racine de } P \text{ de multiplicité } k\big) \iff \big(\exists Q \in \mathbb{K}[X],\ P(X) = (X - \lambda)^k Q(X) \text{ et } Q(\lambda) \neq 0\big).$$

- Un polynôme de $\mathbb{C}[X]$ se factorise toujours en produit de polynômes de degré 1. S'il est de degré n il a donc exactement n racines (comptées avec multiplicités).
- Un polynôme de $\mathbb{R}[X]$ se factorise toujours en produit de polynômes de degré 1 et de polynômes de degré 2 $aX^2 + bX + c$ sans racines réelles (de discriminant $b^2 - 4ac$ négatif, les racines correspondantes sont complexes conjuguées). S'il est de degré n il a donc au plus n racines (comptées avec multiplicités).

On dit que P est *scindé* sur $\mathbb{K}$ s'il se décompose en polynômes de degré 1 dans $\mathbb{K}[X]$. Un polynôme à coefficients complexes est toujours scindé, ce qui n'est pas forcément le cas pour un polynôme à coefficients réels.

En pratique pour factoriser un polynôme de degré 3 on cherche s'il a une racine évidente λ pour factoriser le terme en $X-\lambda$ et on étudie le facteur restant de degré 2, on considérant éventuellement le discriminant.

Exemples.
- Le polynôme $P(X) = X^3 - 5X^2 + 8X - 4 = (X-1)(X-2)^2$ est scindé sur $\mathbb{R}$ et a pour racines dans $\mathbb{R}$ (donc également dans $\mathbb{C}$) 2 de multiplicité 2 et 1 (racine simple)
- Le polynôme $Q(X) = X^3 - X^2 + X - 1 = (X-1)(X^2+1)$ n'est pas scindé sur $\mathbb{R}$, il a 1 comme racine réelle. Sur $\mathbb{C}$ il est scindé, et s'écrit $(X-1)(X-i)(X+i)$, il n'a donc que des racines simples qui sont 1, i et $-i$.

Définition 23. Soit $\varphi \in \mathscr{L}(E)$. Le *polynôme caractéristique* de φ est

$$\chi_\varphi(X) = \det(\varphi - X\,\mathrm{Id}).$$

Si M est la matrice de φ dans une base quelconque $\mathscr{B}$, alors

$$\chi_\varphi(X) = \chi_M(X) := \det(M - XI_n).$$

Remarque. Le polynôme caractéristique ne dépend pas de la base choisie $\mathscr{B}$. En effet, si M' est la matrice de φ dans une base $\mathscr{B}'$, alors M et M' sont semblables : il existe P matrice inversible telle que $M' = PMP^{-1}$ (où $P = M_{\mathscr{B}',\mathscr{B}}(\mathrm{Id})$ est la matrice de passage de $\mathscr{B}'$ à $\mathscr{B}$). Mais alors $M' - XI_n = PMP^{-1} - XI_n = PMP^{-1} - XPP^{-1} = P(M - XI_n)P^{-1}$ donc les matrices $M' - XI_n$ et $M - XI_n$ sont aussi semblables, elles ont donc le même déterminant, c'est-à-dire que M et M' ont le même polynôme caractéristique. On a montré au passage le résultat suivant : si deux matrices sont semblables elles ont le même polynôme caractéristique.

Exemples.
- $\chi_B(X) = \begin{vmatrix} 2-X & 0 & 0 \\ 1 & 1-X & 0 \\ 1 & -1 & 2-X \end{vmatrix} = (2-X)(1-X)(2-X)$ (simple à calculer car la matrice est triangulaire inférieure)
- $\chi_A(X) = \begin{vmatrix} 1-X & 2 & -1 \\ -1 & -X & 3 \\ 1 & 1 & -X \end{vmatrix} = \begin{vmatrix} 1-X & 2 & -1 \\ 0 & -X+1 & 3-X \\ 1 & 1 & -X \end{vmatrix}$ $(L2 \leftarrow L2 + L3)$
  $= (1-X)\begin{vmatrix} 1-X & 3-X \\ 1 & -X \end{vmatrix} + \begin{vmatrix} 2 & -1 \\ 1-X & 3-X \end{vmatrix} = (1-X)(-X+X^2-3+X) + 2(3-X) + (1-X)$
  $= -X^3 + X^2 + 4 = -(X-2)(X^2+X+2)$.

Théorème 5. Les valeurs propres d'un endomorphisme φ d'un $\mathbb{K}$-espace vectoriel E sont exactement les racines du polynôme caractéristique $\chi_\varphi(X)$ dans $\mathbb{K}$.

Remarques.
- Si un matrice est triangulaire supérieure (ou inférieure), il découle que ses valeurs propres sont les coefficients sur la diagonale. Par exemple pour la matrice $B = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}$ on a

calculé $\chi_B(X) = (2-X)^2(1-X)$ de racine 2 (de multiplicité 2) et 1 : ce sont exactement les coefficients diagonaux de B.

- Le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ sur lequel on travaille est important. Par exemple $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{pmatrix}$ est vue a priori comme une matrice de $M_3(\mathbb{R})$ relative à un endomorphisme du $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$. On cherche donc ses valeurs propres réelles, donc les racines réelles du polynôme caractéristique $\chi_A(X) = -(X-2)(X^2+X+2)$ c'est-à-dire 2.

  Il est possible d'interpréter A comme une matrice de $M_3(\mathbb{C})$, on la voit alors comme la matrice d'un endomorphisme du $\mathbb{C}$-espace vectoriel $\mathbb{C}^3$. Dans ce cas, ses valeurs propres sont les racines complexes du polynôme caractéristique, c'est-à-dire 2, $\frac{-1+i\sqrt{7}}{2}$, et $\frac{-1-i\sqrt{7}}{2}$. En pratique on précisera donc toujours si on cherche les valeurs propres réelles ou complexes.
- Un endomorphisme sur un espace vectoriel de dimension n ou une matrice carrée de taille n a un polynôme caractéristique de degré n, il a donc au plus n racines dans $\mathbb{K}$, et l'endomorphisme a au plus n valeurs propres.

Le polynôme caractéristique d'un endomorphisme ou d'une matrice est annulateur de cet endomorphisme ou de cette matrice :

Théorème 6 (Théorème de Cayley-Hamilton)**.** Soit $\varphi \in \mathscr{L}(E)$. Alors $\chi_\phi(\phi) = 0$.

Si $M \in M_n(\mathbb{K})$, alors $\chi_M(M) = 0$.

Remarque. Ce théorème peut par exemple être utilisé pour calculer l'inverse d'une matrice, comme cela a été vu à la section 4.3.

6.4 Endomorphismes et matrices diagonalisables

Définition 24. Soit $\varphi \in \mathscr{L}(E)$. φ est dit *diagonalisable* s'il existe une base de E formée de vecteurs propres. Autrement dit, la matrice représentative de φ dans cette base est diagonale, et les coefficients diagonaux sont les valeurs propres de φ.

Soit $M \in M_n(\mathbb{K})$. M est dite *diagonalisable* si elle représente dans la base canonique un endomorphisme φ diagonalisable. Autrement dit il existe une matrice inversible P (c'est la matrice de passage $M_{\mathscr{B}',\mathscr{B}}(\mathrm{Id})$ de la base de vecteurs propres $\mathscr{B}'$ dans la base canonique $\mathscr{B}$), et une matrice diagonale D (c'est la matrice $M_{\mathscr{B}'}(\varphi)$), telle que
$$M = PDP^{-1}.$$

Exemples.

- L'endomorphisme $g : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \to (x+y, 2x-y)$ de la section 6.1 est diagonalisable.

  Il est représenté par la matrice $M = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ dans la base canonique $\mathscr{B} = (e_1, e_2)$, et par la matrice $D = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & -\sqrt{3} \end{pmatrix}$ dans la base $\mathscr{B}' = (u_1, u_2)$ où $u_1 = (1+\sqrt{3}, 2)$ et $u_2 = (1-\sqrt{3}, 2)$, où $\sqrt{3}$ et $-\sqrt{3}$ sont les valeurs propres de g. La matrice de passage $\mathrm{P} = \mathrm{M}_{\mathscr{B}',\mathscr{B}}(\mathrm{Id})$ est donc $\mathrm{P} = \begin{pmatrix} 1+\sqrt{3} & 1-\sqrt{3} \\ 2 & 2 \end{pmatrix}$, et on a $\mathrm{M} = \mathrm{PDP}^{-1}$. On peut dire également que M est diagonalisable.

- La rotation r d'angle $\pi/3$ sur $\mathbb{R}^2$ n'est pas diagonalisable (on travaille sur $\mathbb{R}$ ici) : elle n'a pas de valeur propre réelle donc a fortiori pas de vecteurs propres donc il ne peut pas exister une base formée de vecteurs propres.

Remarque. On peut montrer que des sous-espaces propres associés à des valeurs propres distinctes sont toujours en somme directe. En effet soient λ et $\mu \in \mathbb{K}$ tels que $\lambda \neq \mu$; considérons $\mathrm{E}_\lambda = \mathrm{Ker}(\varphi - \lambda\,\mathrm{Id})$ et $\mathrm{E}_\mu = \mathrm{Ker}(\varphi - \mu\,\mathrm{Id})$. Si $u \in \mathrm{E}_\lambda \cap \mathrm{E}_\mu$ alors $\lambda u = \varphi(u) = \mu u$ donc $u = 0$ nécessairement. Ainsi $\mathrm{E}_\lambda \cap \mathrm{E}_\mu = \{0\}$.

La définition de diagonalisabilité peut alors se traduire par le fait que $\mathrm{E} = \mathrm{E}_{\lambda_1} \oplus \cdots \oplus \mathrm{E}_{\lambda_r}$, où $\lambda_1, \dots \lambda_r$ sont les valeurs propres distinctes de φ. En effet, en recollant les bases de chaque sous-espace propre (au sens de la proposition 9), on obtient une base de E (formée de vecteurs propres par construction).

Voici quelques caractérisations utiles de la diagonalisabilité.

Proposition 28. Soit $\varphi \in \mathscr{L}(\mathrm{E})$. Alors on a les équivalences suivantes (conditions nécessaires et suffisantes) :

$$\begin{aligned}\big(\varphi \text{ est diagonalisable}\big) \iff &\big(\chi_\varphi(\mathrm{X}) \text{ est scindé et pour toute valeur propre } \lambda,\\ &\dim \mathrm{E}_\lambda = n_\lambda \text{ multipicité de la racine } \lambda \text{ de } \chi_\varphi(\mathrm{X})\big)\\ \iff &\big(\text{il existe un polynôme annulateur scindé à racines simples}\big)\end{aligned}$$

On a également la condition suffisante suivante :

$$\big(\chi_\varphi(\mathrm{X}) \text{ est scindé à racines simples}\big) \Rightarrow \big(\varphi \text{ est diagonalisable}\big).$$

Idée de preuve : Le premier critère découle de la caractérisation $\mathrm{E} = \mathrm{E}_{\lambda_1} \oplus \cdots \oplus \mathrm{E}_{\lambda_r}$ de la diagonalisabilité. En effet, si les λ_i sont les racines du polynômes caractéristiques dans $\mathbb{K}$, donc les valeurs propres de φ, alors les sous-espaces propres E_{λ_i} sont en somme directe. Pour que leur somme soit égale à E il suffit de vérifier qu'elle a la bonne dimension. On a $\dim\big(\mathrm{E}_{\lambda_1} \oplus \cdots \oplus \mathrm{E}_{\lambda_r}\big) = \sum_{i=1}^r \dim(\mathrm{E}_{\lambda_i})$ et d'autre part comme $\chi_\varphi(\mathrm{X})$ a pour degré $n = \dim(\mathrm{E})$, et que chaque racine λ_i a multiplicité n_{λ_i} dans χ_φ, on a $n = \sum_i n_{\lambda_i}$. A vous de finir la preuve...

Le deuxième critère nécessite de faire un peu plus d'algèbre !

La dernière condition suffisante découle du premier critère. □

Exemples.

- On reprend la matrice $\mathrm{B} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}$. On a déjà vu dans la section 6.2 que 2 était valeur propre et on a calculé $\mathrm{E}_2(\mathrm{B})$ qui est de dimension 2. On a déjà calculé $\chi_\mathrm{B}(\mathrm{X}) = (2-\mathrm{X})^2(1-\mathrm{X})$. Ainsi 1 est l'autre valeur propre. Comme il existe au moins un vecteur propre (donc non nul) associé, on sait que $\dim \mathrm{E}_1(\mathrm{B}) \geq 1$, et comme $\mathrm{E}_1(\mathrm{B}) \oplus \mathrm{E}_2(\mathrm{B}) \subset \mathrm{E}$, on a $\dim(\mathrm{E}_1(\mathrm{B}) \oplus \mathrm{E}_2(\mathrm{B})) = \dim \mathrm{E}_1(\mathrm{B}) + \dim \mathrm{E}_2(\mathrm{B}) \leq \dim \mathrm{E} = 3$. Ainsi nécessairement $\dim \mathrm{E}_1(\mathrm{B}) = 1$, et B est diagonalisable.

On peut calculer $E_1(B)$ explicitement :

$$U = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_1(B) \iff BU = U \iff \begin{cases} 2x & = x \\ x + y & = y \\ x - y + 2z & = z \end{cases} \iff U = \begin{pmatrix} 0 \\ y \\ y \end{pmatrix}.$$

On a donc $E_1(B) = \mathrm{Vect}(\left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\})$.

En posant $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ (matrice de passage : matrice des vecteurs propres en colonnes), on a $B = PDP^{-1}$.

- On reprend la matrice $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 0 & 3 \\ 1 & 1 & 0 \end{pmatrix}$. On a vu que $\chi_A(X) = (X-2)(X^2+X+2)$ n'est pas scindé sur $\mathbb{R}$, donc la matrice A n'est pas diagonalisable (sur $\mathbb{R}$, c'est-à-dire en tant que matrice de $M_3(\mathbb{R})$).
- On pose $C = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. On peut calculer $\chi_C(X) = (2-X)^2(1-X)$ (simple à calculer car la matrice est triangulaire. On peut noter que les valeurs propres des matrices triangulaires se lisent toujours sur la diagonale). Pour l'instant on ne peut pas conclure que C est diagonalisable ou non. Il faut calculer $E_2(C)$:

$$U = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E_2(C) \iff CU = 2U \iff \begin{cases} 2x + y & = 2x \\ 2y & = 2y \\ z & = 2z \end{cases} \iff U = \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}.$$

Donc $E_2(C) = \mathrm{Vect}(\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\})$ est de dimension 1 donc strictement inférieure à la multiplicité (égale à 2) de la valeur propre 2. C n'est donc pas diagonalisable.

- Un projecteur est un endomorphisme f de E vérifiant $f^2 = f$. Il est toujours diagonalisable puisque que le polynôme $P(X) = X^2 - X = X(X-1)$ est annulateur de f, et il est scindé à racines simples. On peut montrer que f a pour valeur propre 0 d'espace propre associé $E_0(f) = \mathrm{Ker}(f)$ et 1 d'espace propre associé $E_1(f) = \mathrm{Im}(f)$. (On a vu en TD que $\mathrm{Ker}(f) \oplus \mathrm{Im}(f) = E$.)

Remarque. En pratique les méthodes de diagonalisation reposent sur les deux critères énoncés ci-dessus : on applique l'un ou l'autre selon les cas :

- Si l'endomorphisme est représenté par une matrice explicite, ou qu'on part directement d'une matrice, alors on calcule le polynôme caractéristique, qu'on factorise pour obtenir les valeurs propres (les racines).
  - → S'il n'est pas scindé, on conclut que la matrice (ou l'endomorphisme) n'est pas diagonalisable.
  - → S'il est scindé,
    - → Si ces racines sont simples, la matrice est diagonalisable.

→ Sinon, on continue l'étude en calculant les espaces propres associés aux valeurs propres multiples. Cela revient à résoudre un système linéaire.

→ Si au moins l'un des espaces propres n'a pas la bonne dimension (égale à la multiplicité de la valeur propre), l'endomorphisme n'est pas diagonalisable

→ Sinon, il l'est!

On a alors déterminé si l'endomorphisme était diagonalisable ou non. On peut finir par déterminer tous les espaces propres pour calculer la matrice de passage P.

- Si l'endomorphisme n'est pas donné explicitement ou si on sait qu'il vérifie une relation de la forme P(M) = 0 où P est un polynôme, on peut considérer le deuxième critère. Si le polynôme P est scindé à racines simples, alors la matrice (ou l'endomorphisme associé) est diagonalisable.

7 Applications de la diagonalisation

Plusieurs problèmes linéaires peuvent être résolus efficacement par la mise sous forme matricielle et la diagonalisation de la matrice du système, comme nous allons le voir sur quelques exemples.

7.1 Suites récurrentes linéaires

• Un *système de suites récurrentes linéaires* (d'ordre 1 à coefficients constants) est un système de relations de récurrence de la forme

$$\begin{cases} u_{n+1} &= a_{1,1}u_n + a_{1,2}v_n + \cdots + a_{1,k}z_n \\ v_{n+1} &= a_{2,1}u_n + a_{2,2}v_n + \cdots + a_{2,k}z_n \\ \vdots & \vdots \\ z_{n+1} &= a_{k,1}u_n + a_{k,2}v_n + \cdots + a_{k,k}z_n, \end{cases}$$

pour tout $n \in \mathbb{N}$.

Il se met sous forme matricielle de la façon suivante : $\mathrm{U}_{n+1} = \mathrm{AU}_n$ avec $\mathrm{U}_n = \begin{pmatrix} u_n \\ v_n \\ \vdots \\ z_n \end{pmatrix}$ et $\mathrm{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,k} \\ a_{2,1} & a_{2,2} & \dots & a_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \dots & a_{k,k} \end{pmatrix}$,

où les coefficients de A ne dépendent pas de n.

Il se résout donc en $\mathrm{U}_n = \mathrm{A}^n\mathrm{U}_0$ (par récurrence). Si A est diagonalisable, on sait alors calculer facilement A^n (voir exercices en TD).

• Une *suite récurrente linéaire d'ordre p* (à coefficients constants) est donnée par une relation de récurrence de la forme

$$u_{n+p} = a_0u_n + a_1u_{n+1} + \cdots + a_{p-1}u_{n+p-1}.$$

Cette relation peut également se traduire matriciellement sous la forme suivante : $\mathrm{U}_{n+1} = \mathrm{AU}_n$ où $\mathrm{U}_n = \begin{pmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{n+p-1} \end{pmatrix}$ et $\mathrm{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 \\ a_0 & a_1 & a_2 & \dots & a_{p-1} \end{pmatrix}$.

On procède de même pour la résolution : on a $\mathrm{U}_n = \mathrm{A}^n\mathrm{U}_0$ et on calcule A^n facilement si A est diagonalisable.

7.2 Systèmes d'équations différentielles linéaires d'ordre 1

Un système différentiel linéaire (d'ordre 1 à coefficients constants) est un système de la forme :

$$\begin{cases} x_1' &= a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p + g_1 \\ x_2' &= a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p + g_2 \\ \vdots & \vdots \\ x_p' &= a_{p,1}x_1 + a_{p,2}x_2 + \cdots + a_{p,p}x_p + g_p, \end{cases}$$

où les x_i et la g_i sont des fonctions de classe C^1 d'un intervalle $I \subset \mathbb{R}$ à valeur dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, les $a_{i,j}$ sont des constantes de $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Dans ce système on suppose les fonctions g_i connues, ainsi que les constantes $a_{i,j}$, et les fonctions x_i sont les fonctions à déterminer.

Ce système est équivalent à l'équation matricielle $X' = AX+B$ où $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$, $B = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_p \end{pmatrix}$, et $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p,1} & a_{p,2} & \dots & a_{p,p} \end{pmatrix}$.

Attention ici X est une fonction vectorielle de I dans $\mathbb{K}^p$: c'est la fonction $t \mapsto \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_p(t) \end{pmatrix}$. De même B est une fonction vectorielle de I dans $\mathbb{K}^p$: c'est la fonction $t \mapsto \begin{pmatrix} g_1(t) \\ g_2(t) \\ \vdots \\ g_p(t) \end{pmatrix}$. Par contre A est une matrice de $M_p(\mathbb{K})$, ses coefficients ne dépendent pas de t.

Si la matrice A est diagonalisable, alors ce système différentiel linéaire admet des solutions, qu'on peut construire grâce aux méthodes suivantes.

Méthode 1 : se ramener à des équations différentielles indépendantes

On suppose donc A diagonalisable et on note D la matrice diagonale des valeurs propres $\lambda_1, \dots \lambda_p$, P la matrice de passage inversible telle que $A = PDP^{-1}$. On a

$$X' = AX + B \iff X' = PDP^{-1}X + B \iff P^{-1}X' = DP^{-1}X + P^{-1}B.$$

Ainsi, en posant les nouvelles fonctions vectorielles $Y = P^{-1}X = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ et $H = P^{-1}B = \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}$ on a obtenu l'équation matricielle $Y' = DY + H$ qui se traduit par le système suivant

$$\begin{cases} y_1' &= \lambda_1 y_1 + h_1 \\ y_2' &= \lambda_2 y_2 + h_2 \\ \vdots & \vdots \\ y_p' &= \lambda_p y_p + h_p, \end{cases}$$

Chaque équation est bien indépendante des autres, et se résout donc séparément. On détermine ainsi Y puis $X = PY$.

Exemple. On considère le système différentiel suivant :

$$\begin{cases} x_1' = 2x_1 + x_2 - 2x_3 + e^t \\ x_2' = x_1 + e^t \\ x_3' = x_2 + e^t \end{cases}.$$

Il se met sous la forme $X' = AX + B$ avec $A = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $B : t \mapsto \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}$. En appliquant les méthodes d'étude de la section précédente, on montre que A est diagonalisable de valeurs propres $-1, 1$ et 2. En posant $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $P = \begin{pmatrix} 1 & 1 & 4 \\ -1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$ on a $A = PDP^{-1}$, où $P^{-1} = \begin{pmatrix} 1/6 & -1/2 & 1/3 \\ -1/2 & 1/2 & 1 \\ 1/3 & 0 & -1/3 \end{pmatrix}$. En posant $Y = P^{-1}X$ et $H = P^{-1}B$ le système est donc équivalent à l'équation $Y' = DY + H$ c'est-à-dire

$$\begin{cases} y_1' = y_1 \\ y_2' = y_2 + e^t \\ y_3' = 2y_3. \end{cases}$$

On résout alors les trois équations indépendamment.

- $y_1'(t) = y_1(t)$ est homogène et se résout en $y_1(t) = c_1 e^{-t}$.
- De même $y_3'(t) = 2y_3(t)$ est homogène et se résout en $y_3(t) = c_3 e^{2t}$.
- Les solutions de $y_2'(t) = y_2(t) + e^t$ s'écrivent sous la forme $y_{2,hom} + y_{2,part}$ avec $y_{2,hom} = c_2 e^t$ et $y_{2,part} = te^t$ (trouvée par variation de la constante par exemple).

Après application de la matrice P, on trouve donc les solutions suivantes

$$\begin{cases} x_1(t) = c_1 e^{-t} + c_2 e^t + 4c_3 e^{2t} + te^t \\ x_2(t) = -c_1 e^{-t} + c_2 e^t + 2c_3 e^{2t} + te^t \\ x_3(t) = c_1 e^{-t} + c_2 e^t + c_3 e^{2t} + te^t, \end{cases}$$

avec $c_1, c_2, c_3 \in \mathbb{R}$. On note que les constantes c_1, c_2, c_3 peuvent être déterminées grâce à des conditions initiales, la solution du système est alors unique.

La deuxième méthode est une reformulation vectorielle du résultat obtenu par la première méthode. On fait toutes les variations des constantes ensemble sous forme vectorielle.

Méthode 2 : variation de la constante, version vectorielle

La solution de $X' = AX + B$ est de la forme $X = X_{hom} + X_{part}$ où

- $X_{hom} = \sum_i c_i e^{\lambda_i t} V_i$ est la solution de l'équation homogène $X' = AX$, où les λ_i sont les valeurs propres de A (répétées avec multiplicité), les V_i sont les vecteurs propres correspondants (tels que $(V_i)_i$ forme une base de $\mathbb{K}^p$), et les c_i sont des constantes de $\mathbb{K}$.
- X_{part} est une solution particulière de $X' = AX + B$. Elle s'obtient par variation des constantes, c'est-à-dire en la cherchant sous la forme $X_{part} = \sum_i c_i(t) e^{\lambda_i t} V_i$ et en déterminant les fonctions $c_i(t)$.

Exemple. On reprend le même exemple que précédemment avec la seconde méthode. On a déjà vu que les valeurs propres de A sont $\lambda_1 = -1, \lambda_2 = 1$ et $\lambda_3 = 2$ de vecteurs propres correspondants $V_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $V_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, et $V_3 = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$.

Ainsi la solution homogène est $X_{hom} = c_1 e^{-t} V_1 + c_2 e^t V_2 + c_3 e^{2t} V_3$.

On cherche une solution particulière sous la forme $X_{part}(t) = c_1(t) e^{-t} V_1 + c_2(t) e^t V_2 + c_3(t) e^{2t} V_3$. En dérivant et en injectant dans l'équation $X' = AX + B$ on obtient

$$\big(c_1'(t) - c_1(t)\big) e^{-t} V_1 + \big(c_2' + c_2(t)\big)(t) e^t V_2 + \big(c_3'(t) + 2c_3(t)\big) e^{2t} V_3 = A\big(c_1(t) e^{-t} V_1 + c_2(t) e^t V_2 + c_3(t) e^{2t} V_3\big) + B.$$

En utilisant que $AV_i = \lambda_i V_i$ car les V_i sont des vecteurs propres, l'équation se simplifie en

$$c_1'(t) e^{-t} V_1 + c_2'(t) e^t V_2 + c_3'(t) e^{2t} V_3 = B.$$

se qui se réécrit (comme les V_i sont les colonnes de P) comme

$$P \begin{pmatrix} c_1'(t) e^{-t} \\ c_2'(t) e^t \\ c_3'(t) e^{2t} \end{pmatrix} = B$$

c'est-à-dire

$$\begin{pmatrix} c_1'(t) e^{-t} \\ c_2'(t) e^t \\ c_3'(t) e^{2t} \end{pmatrix} = P^{-1} B = \begin{pmatrix} 0 \\ e^t \\ 0 \end{pmatrix},$$

donc on peut choisir comme solution particulière celle où $c_1(t) = 0, c_2(t) = t, c_3(t) = 0$ ce qui donne $X_{part} = te^t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, et les solutions de l'équation générale sont de la forme

$$X = c_1 e^{-t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + c_2 e^t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_3 e^{2t} \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} + te^t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

avec $c_1, c_2, c_3 \in \mathbb{R}$.

7.3 Equations différentielles linéaires d'ordre supérieur

De même que pour les suites récurrentes linéaires, une équation linéaire d'ordre p peut se mettre sous forme matricielle et se résoudre par des méthodes similaires à celles vues précédemment.

Une équation différentielle linéaire d'ordre p est une équation de la forme

$$x^{(p)}(t) + a_{p-1} x^{(p-1)}(t) + \cdots + a_0 x(t) = b(t).$$

Elle se met sous la forme matricielle suivante :

$$X'(t) = AX(t) + B(t)$$

où $\mathrm{X}(t) = \begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(p-1)}(t) \end{pmatrix}$, $\mathrm{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{p-1} \end{pmatrix}$ et $\mathrm{B}(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}$.

Si A est diagonalisable, cette équation se résout avec les méthodes vues précédemment.

Exercice 1. Redémontrer les résultats que vous connaissez pour les équations différentielles linéaires d'ordre 2 grâce à l'écriture matricielle. Que représente l'équation caractéristique dans ce formalisme ?

Partie II : Algèbre hermitienne

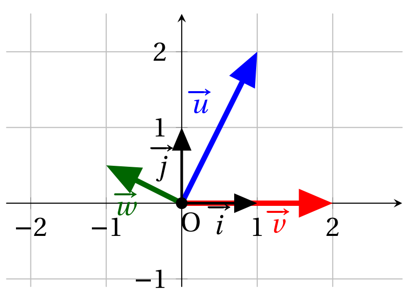
8 Espaces vectoriels munis de produits scalaire

8.1 Rappel : produit scalaire de vecteurs du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$

- Dans un repère orthonormé $(O, \vec{i}, \vec{j})$ du plan, le produit scalaire de deux vecteurs $\vec{u}$ et $\vec{v}$ de coordonnées $\begin{pmatrix} x \\ y \end{pmatrix}$ et $\begin{pmatrix} x' \\ y' \end{pmatrix}$ est $\vec{u} \cdot \vec{v} = xx' + yy'$.
- La norme de $\vec{u}$ est $\|\vec{u}\| = \sqrt{x^2 + y^2}$, on a $\|\vec{u}\|^2 = \vec{u} \cdot \vec{u}$.
- $\vec{u}$ et $\vec{v}$ sont orthogonaux si et seulement si $\vec{u} \cdot \vec{v} = 0$.

(Définitions similaires dans l'espace $\mathbb{R}^3$)

Exemple :



- $\vec{u} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} -1 \\ 1/2 \end{pmatrix}$
- $\vec{u} \cdot \vec{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 1 \times 2 + 2 \times 0 = 2$.

$\vec{u} \cdot \vec{w} = 0$ ($\vec{u}$ et $\vec{w}$ sont orthogonaux)

$\vec{v} \cdot \vec{w} = -2$

- $\|\vec{u}\| = \sqrt{1^2 + 2^2} = \sqrt{5}$,

$\|\vec{v}\| = 2$,

$\|\vec{w}\| = \frac{\sqrt{5}}{2}$.

On remarque que le produit scalaire de vecteurs du plan satisfait quelques propriétés :

- $\forall \vec{u}, \vec{v} \in \mathbb{R}^2, \quad \vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- $\forall \vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^2, \forall \lambda, \mu \in \mathbb{R}, \quad \vec{u} (\cdot \lambda \vec{v} + \mu \vec{w}) = \lambda \vec{u} \cdot \vec{v} + \mu \vec{u} \cdot \vec{w}$
- $\forall \vec{u} \in \mathbb{R}^2, \quad \vec{u} \cdot \vec{u} = \|\vec{u}\|^2 \geq 0$
- $\forall \vec{u} \in \mathbb{R}^2, \quad \left(\|\vec{u}\| = 0 \Rightarrow \vec{u} = \vec{0}\right)$.

Dans ce chapitre, nous allons voir comment généraliser la notion de produit vectoriel (et de norme, et d'orthogonalité) à $\mathbb{R}^n$ et plus généralement à tous les $\mathbb{R}$-espaces vectoriels, et voir également quelle est la bonne notion pour définir un produit scalaire sur un $\mathbb{C}$-espace vectoriel.

8.2 Produit scalaire sur un $\mathbb{K}$-espace vectoriel

Définition 25. Soit E un $\mathbb{K}$-espace vectoriel. Un *produit scalaire* sur E est une application

$$\begin{array}{ccc} E \times E & \to & \mathbb{K} \\ (u, v) & \mapsto & (u|v) \end{array}$$

vérifiant les propriétés suivantes :

- Symétrie hermitienne : $\forall (u,v) \in E^2, (u|v) = \overline{(v|u)}$
- Linéarité par rapport à la deuxième variable : $\forall (u, v_1, v_2) \in E^3, \forall (\lambda, \mu) \in \mathbb{K}^2, \quad (u|\lambda v_1 + \mu v_2) = \lambda(u|v_1) + \mu(u|v_2)$
- Positivité : $\forall u \in E, (u|u) \geq 0$
- Définie positivité : $\forall u \in E, \big((u|u) = 0 \Rightarrow u = 0\big)$

Un espace vectoriel muni d'un produit scalaire est appelé espace *préhilbertien*. S'il est de plus de dimension finie, et que

- $\mathbb{K} = \mathbb{C}$, on l'appelle espace *hermitien*
- $\mathbb{K} = \mathbb{R}$, on l'appelle espace *euclidien*

Remarques.

- Si $\mathbb{K} = \mathbb{R}$, la première propriété devient une symétrie simple : $(u|v) = (v|u)$
- Les deux premières propriétés impliquent que le produit scalaire est “semi-linéaire" par rapport à la première variable : $\forall (u_1, u_2, v) \in E^2, \forall (\lambda, \mu) \in \mathbb{K}^2, \quad (\lambda u_1 + \mu u_2|v) = \overline{\lambda}(u_1|v) + \overline{\mu}(u_2|v)$. Dans le cas où $\mathbb{K} = \mathbb{R}$ on a bien la linéarité simple par rapport à la première variable.
- La symétrie hermitienne implique que $(u|u) = \overline{(u|u)}$ donc c'est toujours un réel. On peut donc le comparer à 0 et demander à ce qu'il soit positif (troisième propriété)
- Le produit scalaire peut se noter de plusieurs façons, comme $\langle u, v\rangle$, $\langle u|v\rangle$, (u, v), etc.

Exemples.

- Le produit scalaire des vecteurs du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$ est bien un produit scalaire au sens de la définition 25. $\mathbb{R}^2$ et $\mathbb{R}^3$ sont des espaces euclidiens.
- Sur $\mathbb{C}^3$ on définit $(u|v) = {}^t\overline{U}V = \overline{x}x' + \overline{y}y' + \overline{z}z'$ pour $U = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ et $V = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$ les coordonnées de u et v dans la base canonique. C'est bien un produit scalaire sur le $\mathbb{C}$-espace vectoriel $\mathbb{C}^3$, qui est un espace hermitien.
- Soit $C^0_T(\mathbb{R}, \mathbb{C})$ l'espace des fonctions continues de $\mathbb{R}$ dans $\mathbb{C}$ périodiques de période T (espace important pour l'analyse de Fourier des signaux par exemple). On peut vérifier qu'il s'agit d'un $\mathbb{C}$-espace vectoriel. Pour f et g dans $C^0_T(\mathbb{R}, \mathbb{C})$, on pose
$$(f|g) = \frac{1}{T}\int_{-T/2}^{T/2} \overline{f(t)}g(t)dt,$$
c'est bien un produit scalaire. Ainsi $C^0_T(\mathbb{R}, \mathbb{C})$ est un espace préhilbertien (il n'existe de pas de base finie). A titre d'exemple nous allons calculer $(f|g)$ pour $f(t) = e^{in\omega t}$ et $g(t) = e^{im\omega t}$, avec $\omega = 2\pi/T$.

Compléments d'analyse.

Pour définir l'intégrale une fonction à valeurs complexes, on intègre séparément la partie réelle et la partie imaginaire. Si $f : \mathbb{R} \to \mathbb{C}$ se décompose en $f(t) = g(t) + ih(t)$ avec g et h fonctions continues d'un intervalle $[a, b]$ dans $\mathbb{R}$, alors on définit

$$\int_a^b f(t)dt = \int_a^b g(t)dt + i\int_a^b h(t)dt.$$

Exemple important : l'exponentielle complexe. On pose $\omega \neq 0$.

$$\begin{aligned}\int_a^b e^{i\omega t}dt &= \int_a^b \cos(\omega t)dt + i\int_a^b \sin(\omega t)dt \\ &= \left[\frac{1}{\omega}\sin(\omega t)\right]_a^b + i\left[\frac{1}{\omega}-\cos(\omega t)\right]_a^b = \left[\frac{1}{i\omega}\big(\cos(\omega t)+i\sin(\omega t)\big)\right]_a^b = \left[\frac{1}{i\omega}e^{i\omega t}\right]_a^b\end{aligned}$$

Ainsi, calculons $(f|g)$ pour $f(t)=e^{in\omega t}$ et $g(t)=e^{im\omega t}$, avec $\omega = 2\pi/\mathrm{T}$ et $n\neq m$.

$$\begin{aligned}(f|g) &= \frac{1}{\mathrm{T}}\int_{-\mathrm{T}/2}^{\mathrm{T}/2} e^{-in\omega t}e^{im\omega t}dt = \frac{1}{\mathrm{T}}\int_{-\mathrm{T}/2}^{\mathrm{T}/2} e^{i(m-n)\omega t}dt \\ &= \frac{1}{\mathrm{T}}\left[\frac{1}{i(n-m)\omega}e^{i(n-m)\omega t}\right]_{-\mathrm{T}/2}^{\mathrm{T}/2} = \frac{1}{i(m-n)}\left(e^{i(m-n)\pi}-e^{-i(m-n)\pi}\right)=0.\end{aligned}$$

8.3 Représentation matricielle

Dans le cas des espaces de dimension finie, on va pouvoir encoder le produit scalaire à l'aide d'une matrice.

Proposition 29. Soit E un $\mathbb{K}$-espace vectoriel de dimension finie , $(\cdot|\cdot)$ un produit scalaire sur E, et $\mathscr{B} = (e_i)_i$ une base de E. Alors pour tout u et v de E, on a

$$(u|v) = {}^t\overline{\mathrm{U}}\mathrm{M}\mathrm{V}$$

où U et V sont les vecteurs colonnes des coordonnées de u et v dans la base $\mathscr{B}$, M est la matrice de coefficient $m_{i,j} = (e_i|e_j)$: c'est la matrice du produit scalaire dans la base $\mathscr{B}$ et elle vérifie ${}^t\overline{\mathrm{M}} = \mathrm{M}$.

Idée de preuve : Le résultat est obtenu en décomposant u et v sur la base $\mathscr{B}$ et en utilisant les propriétés du produit scalaire. □

Exemples.

- La matrice du produit scalaire sur $\mathbb{C}^3$ considéré précédemment est la matrice identité I_3.
- On considère $\mathrm{F} = \mathrm{Vect}(u,v)$ le sous-espace vectoriel de $\mathrm{C}^0_{2\pi}(\mathbb{R},\mathbb{C})$, muni du produit scalaire $(f|g) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{f(t)}g(t)dt$, engendré par les fonctions $e_1(t)=\cos(t)$ et $e_2(t)=\sin(t)$ (qui sont bien indépendantes). Calculons la matrice $(\cdot|\cdot)$ sur F. Il faut calculer ses coefficients $(e_1|e_1)$, $(e_2|e_2)$ et $(e_1|e_2)$ (le dernier coefficient étant obtenu comme $(e_2|e_1) = \overline{(e_1|e_2)}$). On a

$$\begin{aligned}(e_1|e_1) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos^2(t)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1+\cos(2t)}{2}dt = \frac{1}{2},\\ (e_2|e_2) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sin^2(t)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}(1-\cos^2(t))dt = \frac{1}{2},\\ (e_1|e_2) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(t)\sin(t)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\sin(2t)}{2} = 0\end{aligned}$$

La matrice M du produit scalaire associée à la base $\mathscr{B} = (e_1, e_2)$ de F est donc $M = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$. Grâce à cette matrice on peut donc calculer le produit scalaire de n'importe quelles combinaisons de cos et sin, par exemple $(e^{it}|e^{-it}) = \overline{\begin{pmatrix} 1 & i \end{pmatrix}} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} 1 & -i \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = 0$ ou $(i\cos - \sin|2\cos) = \overline{\begin{pmatrix} i & -1 \end{pmatrix}} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = -i$.

Notation “Bra" et “Ket" de Dirac :

C'est une notation utilisée en physique pour le produit scalaire dans le cas d'un espace de dimension finie. En reprenant les notations de la Proposition 29, on note (notation “Bra-Ket" en rouge) :

$$(u|v) = {}^t\overline{U}MV = \langle u|M|v\rangle,$$

où

- $|v\rangle = V$ (vecteur colonne des coordonnées de v dans $\mathscr{B}$), est appelé “Ket" de v
- $\langle u| = {}^t\overline{U}$ (vecteur ligne des conjugués des coordonnées de u dans $\mathscr{B}$), est appelé “Bra" de u

La terminologie vient du fait que pour la matrice identité on retrouve la notation usuelle $\langle u|I_n|v\rangle = \langle u|v\rangle$ avec des crochets $\langle\rangle$ (“brackets" en anglais).

8.4 Norme associée à un produit scalaire

Définition 26. Soit E un $\mathbb{K}$-espace vectoriel muni d'un produit scalaire $(\cdot|\cdot)$. La norme associée au produit scalaire $(\cdot|\cdot)$ est l'application

$$\begin{array}{rccl} \|\cdot\| : & E & \longrightarrow & \mathbb{R}^+ \\ & u & \longmapsto & \|u\| = \sqrt{(u|u)}. \end{array}$$

Exemples.

- La norme du vecteur $u = (1, i, 1+2i)$ dans $\mathbb{C}^3$ associée au produit scalaire défini précédemment est $\|u\| = \sqrt{1^2 + (-i)i + (1-2i)(1+2i)} = \sqrt{1+1+5} = \sqrt{7}$.
- Dans $C^0_T(\mathbb{R}, \mathbb{C})$, la norme de f où $f(t) = e^{in\omega t}$ se calcule comme

$$\|f\|^2 = (f|f) = \frac{1}{T}\int_{-T/2}^{T/2} \overline{f(t)} f(t)\,dt = \frac{1}{T}\int_{-T/2}^{T/2} |f(t)|^2\,dt = \frac{1}{T}\int_{-T/2}^{T/2} 1\,dt = 1,$$

d'où $\|f\| = \sqrt{1} = 1$.

Proposition 30 (Inégalité de Cauchy-Schwartz)**.** Soit E un espace préhilbertien de produit scalaire $(\cdot|\cdot)$ et de norme associée $\|\cdot\|$. Alors

$$\forall u, v \in E, \quad |(u|v)| \leq \|u\| \cdot \|v\|,$$

de plus, on sait caractériser le cas d'égalité :

$$\big(|(u|v)| = \|u\| \cdot \|v\|\big) \iff \big(u \text{ et } v \text{ sont liés}\big).$$

Remarque. Pour les vecteurs du plan (et de l'espace), on a l'égalité $(u|v) = \|u\| \cdot \|v\| \cos(\theta)$, où θ est l'angle entre les deux vecteurs. Cela implique directement le résultat précédent. En ce sens le produit scalaire permet de généraliser la notion d'angle entre des éléments d'espaces vectoriels plus généraux.

Preuve :

- On suppose $u \neq 0$ (sinon l'inégalité est trivialement vraie). Alors $\|u\| \neq 0$ et on peut considérer le scalaire $\frac{(u|v)}{\|u\|^2}$. On a

$$\begin{aligned}0 \leq \left\| v - \frac{(u|v)}{\|u\|^2} u \right\|^2 &= \left(v - \frac{(u|v)}{\|u\|^2} u \,\middle|\, v - \frac{(u|v)}{\|u\|^2} u \right) \\ &= (v|v) - \frac{(u|v)}{\|u\|^2} \cdot (v|u) - \frac{\overline{(u|v)}}{\|u\|^2} \cdot (u|v) + \frac{(u|v)}{\|u\|^2} \cdot \frac{\overline{(u|v)}}{\|u\|^2} \cdot (u|u) \quad \text{(par lin. et sym. herm.)} \\ &= \|v\|^2 - \frac{2(u|v)\overline{(u|v)}}{\|u\|^2} + \frac{(u|v)\overline{(u|v)}}{\|u\|^4} \cdot \|u\|^2 \quad \text{(après regroupement des termes)} \\ &= \|v\|^2 - \frac{|(u|v)|^2}{\|u\|^2} \quad \text{(par définition du module).}\end{aligned}$$

  D'où l'inégalité principale, obtenue après multiplication par $\|u\|^2 > 0$.

- On considère maintenant le cas d'égalité. Le sens $\Leftarrow$ est facile car si u et v sont liés alors $u = \lambda v$ avec $\lambda \in \mathbb{K}$ ou alors $v = \lambda u$ avec $\lambda \in \mathbb{K}$. Dans tous les cas, l'égalité se déduit par linéarité et symétrie hermitienne.
  Pour le sens $\Rightarrow$, supposons $u \neq 0$ (sinon le résultat est trivialement vrai : 0 est lié avec n'importe quel vecteur). D'après le premier point on a l'égalité si et seulement si $\left\| v - \frac{(u|v)}{\|u\|^2} u \right\|^2 = 0$, c'est-à-dire $v - \frac{(u|v)}{\|u\|^2} u = 0$. C'est exactement une relation de liaison entre u et v.

$\square$

Proposition 31. La norme $\|\cdot\|$ associé à un produit scalaire sur un espace vectoriel E satisfait les propriétés suivantes :

- $\forall u \in E, \big(\|u\| = 0\big) \iff \big(u = 0\big)$
- $\forall u \in E, \forall \lambda \in \mathbb{K}, \|\lambda u\| = |\lambda| \cdot \|u\|$
- $\forall (u, v) \in E^2, \|u + v\| \leq \|u\| + \|v\|$.

Preuve : Les deux premiers points découlent facilement des propriétés du produit scalaire.

Pour le troisième point, on développe :

$$\begin{aligned}\|u + v\|^2 &= (u + v|u + v) = (u|u) + (u|v) + (v|u) + (v|v) \\ &= \|u\|^2 + 2\Re((u|v)) + \|v\|^2 \leq \|u\|^2 + 2|\Re((u|v))| + \|v\|^2 \\ &\leq \|u\|^2 + 2|(u|v)| + \|v\|^2 \leq \|u\|^2 + 2\|u\| \cdot \|v\| + \|v\|^2 \quad \text{(par Cauchy-Schwartz)} \\ &= \big(\|u\| + \|v\|\big)^2\end{aligned}$$

d'où le résultat. $\square$

Remarque. Un vecteur de norme 1 est appelé vecteur *normé* ou *unitaire*. Un vecteur non nul u peut toujours être "renormalisé" : le vecteur $\frac{u}{\|u\|}$ colinéaire à u est de norme 1.

Par exemple, $2\cos$ est un vecteur normé dans $C^0_{2\pi}(\mathbb{R}, \mathbb{C})$.

8.5 Orthogonalité dans un espace préhilbertien

Définition 27. Deux vecteurs u et v d'un espace préhilbertien E sont dit *orthogonaux* si $(u|v) = 0$. On note alors $u \perp v$.

C'est la généralisation directe de l'orhtogonalité de vecteurs dans le plan ou l'espace.

Proposition 32. Toute famille $(u_i)_i$ de vecteurs non nuls et orthogonaus deux à deux est libre.

Une base de E (ou famille génératrice) formée de vecteurs orthogonaux deux à deux est appelée *base orthogonale*.

Une base orthogonale où les vecteurs sont normés est appelée *base orthonormée*. Si l'espace est de dimension finie, la matrice représentative du produit scalaire dans une base orthonormée est la matrice identité.

Preuve : Soient des scalaires λ_i tels que $\sum_{i=1}^{N} \lambda_i u_i = 0$. Pour chaque k on calcule

$$0 = \left(u_k \middle| \sum_{i=1}^{N} \lambda_i u_i \right) = \sum_{i=1}^{N} \lambda_i (u_k|u_i) = \lambda_k (u_k|u_k) = \lambda_k \|u_k\|^2$$

(par linéarité du produit scalaire à droite, puis orthogonalité de la famille). Comme $u_k \neq 0$ alors $\|u_k\|^2 \neq 0$ donc $\lambda_k = 0$. En répétant l'opération pour chaque k on obtient donc que tous les coefficients sont nuls. $\square$

Exemple. La base canonique (e_1, e_2, e_3) de $\mathbb{C}^3$ est une base orthonormale pour le produit scalaire canonique $(u|v) = {}^t\overline{u}v$. En effet $\|e_1\|^2 = (e_1|e_1) = 1$, de même pour e_2 et e_3. D'autre part $(e_1|e_2) = \overline{\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0$, de même pour $(e_1|e_3)$ et $(e_2|e_3)$.

Remarque. On s'intéressera particulièrement aux bases orthonormées dans la suite. Elles sont particulièrement pratiques pour faire des calculs de produit scalaire, comme c'est le cas dans le plan. Un des but du chapitre est de voir à quelles conditions un endormophisme se diagonalise en base orthonormée.

Proposition 33 (Généralisation du théorème de Pythagore)**.** Soient u et v deux vecteurs d'un espace préhilbertien E. Alors

$$\big(u \perp v\big) \Rightarrow \big(\|u+v\|^2 = \|u\|^2 + \|v\|^2\big).$$

La réciproque est fausse en général.

Preuve : On procède comme dans la preuve de la proposition 31 : on a

$$\|u+v\|^2 = \|u\|^2 + 2\Re(u|v) + \|v\|^2 = \|u\|^2 + \|v\|^2.$$

Voici un contre-exemple pour la réciproque : dans $\mathbb{C}^2$ muni du produit scalaire canonique, on considère $u = (i, 2i)$ et $v = (1, 0)$. Alors $(u|v) = -i \neq 0$ donc u et v ne sont pas orthogonaux, mais $\|u+v\|^2 = 6 = 5+1 = \|u\|^2 + \|v\|^2$. $\square$

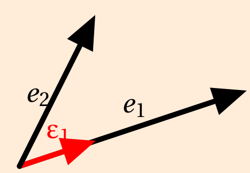
Définition 28. Deux sous-espaces vectoriels F et G d'un espace préhilbertien E sont dits *orthogonaux* si $\forall u \in \mathrm{F}, \forall v \in \mathrm{G}, u \perp v$.

p sous-espaces vectoriels $\mathrm{F}_1, \dots \mathrm{F}_p$ sont dits orthogonaux s'ils sont orthogonaux deux à deux. Dans ce cas leur somme est directe.

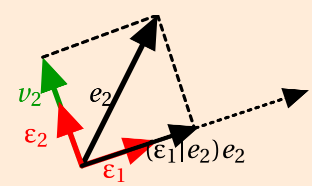
Enfin nous allons voir que tout espace préhilbertien de dimension finie admet une base orthonormée, qu'on peut construire explicitement à partir d'une base quelconque (procédé d'orthonormalisation de Gram-Schmidt).

Procédé d'orthonormalisation de Gram-Schmidt. Soit E un espace préhilbertien de dimension finie, et $\mathscr{B} = (e_i)_{i=1..n}$ une base de E. Le procédé de Gram-Schmidt permet de construire une base orthonormée $\mathscr{B}' = (\varepsilon_i)_{i=1..n}$ à partir des vecteurs de $\mathscr{B}$.

Etape 1 : on normalise e_1 : $\varepsilon_1 = \frac{e_1}{\|e_1\|}$



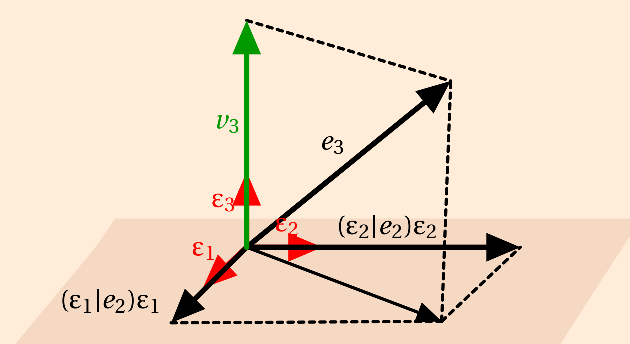
Etape 2 : on pose $v_2 = e_2 - (\varepsilon_1|e_2)\varepsilon_1$, alors $v_2 \perp \varepsilon_1$
on normalise : $\varepsilon_2 = \frac{v_2}{\|v_2\|}$



Etape 3 : on pose $v_3 = e_3 - (\varepsilon_1|e_3)\varepsilon_1 - (\varepsilon_2|e_3)\varepsilon_2$, alors $v_3 \perp \varepsilon_1$ et $v_3 \perp \varepsilon_2$.
on normalise : $\varepsilon_2 = \frac{v_2}{\|v_2\|}$

On itère jusqu'au n-ième vecteur :
$v_n = e_n - \sum_{i=1}^{n-1} (\varepsilon_i|e_n)\varepsilon_i$, $\varepsilon_n = \frac{v_n}{\|v_n\|}$

En conclusion (grâce au Théorème 1) :



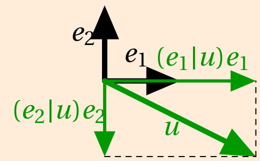
Proposition 34. Tout espace préhilbertien de dimension finie admet au moins une base orthonormée.

Exemple. Dans $\mathbb{C}^2$ on considère la base $e_1 = (i, 1), e_2 = (0, i)$ (du $\mathbb{C}$-espace vectoriel). Le procédé de Gram-Schmidt donne $\varepsilon_1 = \frac{e_1}{\|e_1\|} = \frac{e_1}{\sqrt{2}} = \left(\frac{i}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. Puis $v_2 = e_2 - (\varepsilon_1|e_2)\varepsilon_1 = e_2 - \frac{i}{\sqrt{2}}\varepsilon_1 = \left(\frac{1}{2}, \frac{i}{2}\right)$. Enfin $\varepsilon_2 = \frac{v_2}{\|v_2\|} = \sqrt{2}v_2 = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}\right)$.

Dans une base orthonormée, il est simple d'exprimer les coordonnées de n'importe quel vecteur, ou de calculer le produit scalaire de deux vecteurs (on retrouve d'ailleurs les formules connues pour les vecteurs de $\mathbb{R}^2$ ou $\mathbb{R}^3$).

Proposition 35. Soit E un espace préhilbertien de dimension finie muni d'une base orthonormée $\mathscr{B} = (e_i)_{i=1..n}$. Alors

$$\forall u \in \mathrm{E}, \quad u = \sum_{i=1}^{n} (e_i|u)e_i$$



$$\forall (u,v) \in \mathrm{E}^2, \quad (u|v) = \sum_{i=1}^{n} \overline{x_i} y_i = {}^t\overline{\mathrm{U}}\mathrm{V} \text{ et } \|u\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \text{ où } u = \sum_{i=1}^{n} x_i e_i, v = \sum_{i=1}^{n} y_i e_i, \mathrm{U} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ et } \mathrm{V} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Autrement dit la matrice du produit scalaire dans la base $\mathscr{B}$ est la matrice identité.

Preuve : Pour la première égalité on pose x_i les coordonnées de u dans la base $\mathscr{B}$: $u = \sum_{i=1}^{n} x_i e_i$. En calculant $(e_j|u) = \left(e_j | \sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i (e_j|e_i) = x_j$ on obtient le résultat.
Pour la deuxième égalité il suffit de remarquer que la matrice M du produit scalaire $(\cdot|\cdot)$ dans la B.O.N $\mathscr{B}$ est la matrice identité (car ses coefficients $m_{i,j}$ sont les $(e_i|e_j)$ qui valent 1 si $i = j$ et 0 sinon). □

Exemple. Dans la BON $\mathscr{B} = (\varepsilon_1, \varepsilon_2)$ de $\mathbb{C}^2$ de l'exemple précédent, où $\varepsilon_1 = \left(\frac{i}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ et $\varepsilon_2 = \left(\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}\right)$, le vecteur $u = (1,0)$ se décompose en $u = (\varepsilon_1|u)\varepsilon_1 + (\varepsilon_2|u)\varepsilon_2 = -\frac{i}{\sqrt{2}}\varepsilon_1 + \frac{1}{\sqrt{2}}\varepsilon_2$.

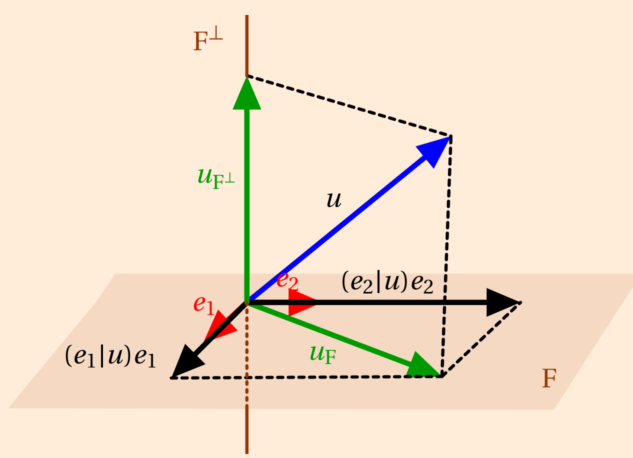
Projection orthogonale sur un sous-espace vectoriel en dimension finie

Soit E un espace préhilbertien de dimension finie et F un sous-espace de E. Soit $\mathscr{B}_\mathrm{F} = (e_i)_{i=1..p}$ une base orthonormée de F. On peut la compléter par $n-p$ vecteurs pour en faire une B.O.N. de E, notée $\mathscr{B} = (e_i)_{i=1..n}$ (grâce à la Proposition 34). On note $\mathrm{F}^\perp = \mathrm{Vect}\big((e_i)_{i=p+1..n}\big)$. $\mathrm{F}^\perp$ est un sous-espace vectoriel tel que $\mathrm{E} = \mathrm{F} \oplus \mathrm{F}^\perp$ et $\mathrm{F} \perp \mathrm{F}^\perp$. Il est appelé *orthogonal* de F.

Pour tout $u \in \mathrm{E}$, la *projection orthogonale* de u sur F est le vecteur $u_\mathrm{F} = \sum_{i=1}^{p} (e_i|u)e_i \in \mathrm{F}$.

u se décompose en $u = u_\mathrm{F} + u_{\mathrm{F}^\perp}$ avec $u_{\mathrm{F}^\perp} = u - u_\mathrm{F} = \sum_{i=p+1}^{n} (e_i|u)e_i \in \mathrm{F}^\perp$.

L'application linéaire $p_\mathrm{F} : \mathrm{E} \to \mathrm{E}, u \mapsto u_\mathrm{F}$ est appelée *projection orthogonale* sur F.



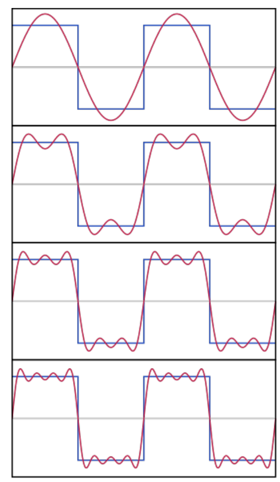
Le procédé de Gram-Schmidt se réinterprète alors de la façon suivante : à l'étape k, on enlève du vecteur e_k sa projection orthogonale sur le sous-espace engendré par les $(e_1,\dots,e_{k-1})$, puis on normalise le vecteur obtenu.

Exemple. Dans $C^0_{2\pi}(\mathbb{R},\mathbb{C})$, on considère les fonctions $e_k(t) = e^{ikt}$ pour $k \in Z$ et $F_n = \mathrm{Vect}(e_{-n},\dots,e_n)$ pour $n \in \mathbb{N}$. Chaque famille $(e_{-n},\dots e_n)$ forme bien une base orthonormée de F_n (voir calculs dans les exemples de 8.2 et 8.4). La projection orthogonale de $u \in F_2$ avec $u(t) = \cos(2t) - \sin(t)$ sur F_1 est donc $p_{F_1}(u) = (e_{-1}|u)e_{-1} + (e_0|u)e_0 + (e_1|u)e_1$. On calcule séparément

$$\begin{aligned}
(e_{-1}|u) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{it}(\cos(2t)-\sin(t))dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{it}\left(\frac{e^{2it}+e^{-2it}}{2} - \frac{e^{it}-e^{-it}}{2i}\right)dt = \frac{1}{2i}\\
(e_0|u) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(\cos(2t)-\sin(t))dt = 0\\
(e_1|u) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-it}(\cos(2t)-\sin(t))dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-it}\left(\frac{e^{2it}+e^{-2it}}{2} - \frac{e^{it}-e^{-it}}{2i}\right)dt = -\frac{1}{2i}
\end{aligned}$$

On a donc $p_F(u) = \frac{1}{2i}e_{-1} - \frac{1}{2i}e_1 = \sin$.

Remarque : La notion de projection orthogonale se généralise à des espaces de dimension infinie, comme $C^0_{2\pi}(\mathbb{R},\mathbb{C})$ ou même l'espace des fonctions continues par morceaux 2π-périodiques par exemple. Un des buts de l'analyse de Fourier est de comprendre à quelle condition un signal (par exemple signal en créneau) est bien approché par ses projections orthogonales sur les sous-espaces F_n quand n tend vers l'infini.
Sur la figure à droite est représenté un signal en créneau, (fonction u) en bleu, et ses projections orthogonales sur F_1, F_2, F_3 et F_4 successivement (c'est à dire les fonctions $p_{F_1}(u)$, etc) en rouge.



9 Opérateurs linéaires sur des espaces préhilbertiens

Les opérateurs linéaires généralisent la notion d'application linéaire.

Définition 29. Soient E et F des $\mathbb{K}$-espaces vectoriels. Un *opérateur linéaire* φ est une fonction φ dont l'ensemble de définition D_φ est un sous-espace vectoriel de F et telle que la restriction à D_φ de φ soit une application linéaire.

Exemple. Pour $E = \mathscr{F}(\mathbb{R})$, la fonction $D : f \mapsto f'$ est définie sur le sous-espace vectoriel des fonctions dérivables, c'est un opérateur linéaire (appelé opérateur de dérivation).

Pour $E = \mathscr{F}(\mathbb{R}^3,\mathbb{R})$ on a l'opérateur gradient : $\nabla : f \mapsto \begin{pmatrix} \frac{\partial f}{\partial x}\\ \frac{\partial f}{\partial y}\\ \frac{\partial f}{\partial z}\end{pmatrix}$, défini sur le sous-espace des fonctions admettant des dérivées partielles par rapport à chaque variable.

Définition 30. Soit E un espace préhilbertien, et φ un opérateur linéaire de E. Un *opérateur adjoint* de φ est un opérateur linéaire de E noté φ^* tel que

$$\forall u \in \mathrm{D}_\varphi, \forall v \in \mathrm{D}_{\varphi^*}, \big(\varphi(u)|v\big) = \big(u|\varphi^*(v)\big).$$

Attention a priori rien ne garantit l'existence et l'unicité d'un tel opérateur.

Exemple. On considère $\mathrm{E} = \mathrm{C}^1_\mathrm{T}(\mathbb{R},\mathbb{C})$ l'espace des fonctions de classe C^1 de $\mathbb{R}$ dans $\mathbb{C}$ et T-périodiques, muni du produit scalaire $(f|g) = \frac{1}{\mathrm{T}}\int_{-\mathrm{T}/2}^{\mathrm{T}/2} \overline{f(t)}g(t)dt$, et l'opérateur D de dérivation. Soient f et g dans E, alors par intégration par partie

$$\begin{aligned}(\mathrm{D}f|g) &= \frac{1}{\mathrm{T}}\int_{-\mathrm{T}/2}^{\mathrm{T}/2}\overline{f'(t)}g(t)dt = \frac{1}{\mathrm{T}}\left(\left[\overline{f(t)}g(t)\right]_{-\mathrm{T}/2}^{\mathrm{T}/2} - \int_{-\mathrm{T}/2}^{\mathrm{T}/2}\overline{f(t)}g'(t)\right)\\ &= -\frac{1}{\mathrm{T}}\int_{-\mathrm{T}/2}^{\mathrm{T}/2}\overline{f(t)}g'(t) = (f|-\mathrm{D}g),\end{aligned}$$

d'où $\mathrm{D}^* = -\mathrm{D}$.

Proposition 36. Soit E un espace préhilbertien de dimension finie. Alors

- Si $\varphi \in \mathscr{L}(\mathrm{E})$, φ^* existe et est unique
- Si $\mathscr{B}$ est un base orthonormée de E, et $\mathrm{M} = \mathrm{M}_\mathscr{B}(\varphi)$, alors la matrice $\mathrm{M}^* = \mathrm{M}_\mathscr{B}(\varphi^*)$ de φ^* est $\mathrm{M}^* = {}^t\overline{\mathrm{M}}$.

Preuve : Le premier point est admis. On note $\mu_{i,j}$ les coefficients de M^* et $m_{i,j}$ les coefficients de M. On a $\mu_{i,j} = (e_i|\varphi^*(e_j))$ car $\mu_{i,j}$ est la i-ème coordonnée de $\varphi^*(e_j)$ et $\mathscr{B} = (e_i)_i$ est orthonormée (Proposition 35). Donc $\mu_{i,j} = (e_i|\varphi^*(e_j)) = (\varphi(e_i)|e_j) = \overline{(e_j|\varphi(e_i))} = \overline{m_{j,i}}$. On obtient donc $\mathrm{M}^* = {}^t\overline{\mathrm{M}}$. □

Exemple. Dans $\mathbb{C}^2$ muni du produit scalaire canonique (celui pour laquelle la base canonique est orthonormée), on considère l'application linéaire $\varphi : (x, y) \mapsto (x + iy, 2ix - y)$. Comme E est de dimension finie φ^* existe et est unique, déterminons-le. Pour cela on va passer par les représentations matricielles. On construit donc la matrice M de φ dans la base $\mathscr{B}$ (ne pas confondre avec la matrice du produit scalaire qui est ici I_2 comme $\mathscr{B}$ est orthonormée pour ce produit scalaire). On a $\varphi(e_1) = \varphi((1,0)) = (1,2i) = e_1 + 2ie_2$ et de même $\varphi(e_2) = ie_1 - e_2$. Ainsi $\mathrm{M} = \begin{pmatrix}1 & i\\ 2i & -1\end{pmatrix}$. D'après le résultat précédent, la matrice de φ^* est donc $\mathrm{M}^* = {}^t\overline{\mathrm{M}} = \begin{pmatrix}1 & -2i\\ -i & -1\end{pmatrix}$, et les coordonnées de $\varphi^*((x,y))$ se calculent par $\mathrm{M}^*\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}x-2iy\\ -ix-y\end{pmatrix}$, donc $\varphi^* : (x, y) \mapsto (x - 2iy, -ix - y)$.

Définition 31. Un opérateur linéaire est dit *autoadjoint* ou *hermitien* si il est égal à son adjoint.

Exemple. • L'opérateur $\mathrm{D}^2 = \mathrm{D}\circ\mathrm{D}$ sur $\mathrm{C}^2_{\mathrm{T}}(\mathbb{R},\mathbb{R})$ par exemple est autoadjoint, car $\mathrm{D}^* = -\mathrm{D}$ d'où $(\mathrm{D}^2)^* = \mathrm{D}^2$:

$$\forall (f,g)\in \mathrm{C}^2_{\mathrm{T}}(\mathbb{R},\mathbb{C})^2,\quad \left(\mathrm{D}^2 f\middle| g\right) = \left(\mathrm{D}f\middle| -\mathrm{D}g\right) = \left(f\middle| -\mathrm{D}(-\mathrm{D}g)\right) = \left(f\middle|\mathrm{D}^2 g\right).$$

- L'opérateur linéaire *Hamiltonien* $\hat{\mathrm{H}}$ sur l'espace des fonctions de $\mathbb{R}^3$ dans $\mathbb{C}$ est défini comme $\hat{\mathrm{H}} = \frac{-\hbar}{2m}\Delta + \hat{\mathrm{V}}$ où $\hat{\mathrm{V}}$ est l'opérateur de multiplication par le réel $\mathrm{V}(x,y,z)$ (représentant un potentiel). On peut vérifier que $\hat{\mathrm{H}}$ est autoadjoint (exercice). Pour résoudre l'équation de Schrödinger on s'intéresse aux vecteurs propres (états propres) pour cet opérateur (en notation "Ket", les $|\varphi\rangle$ tels que $\hat{\mathrm{H}}|\varphi\rangle = \mathrm{E}|\varphi\rangle$), les valeurs propres E étant alors les niveaux d'énergie des états propres correspondants.

10 Matrices particulières

Dans cette section on va passer en revue les propriétés de quelques matrices aux symétries particulières.

Définition 32. Un matrice M de $\mathrm{M}_n(\mathbb{C})$ est dite :
hermitienne si ${}^t\overline{\mathrm{M}} = \mathrm{M}$
unitaire si ${}^t\overline{\mathrm{M}} = \mathrm{M}^{-1}$

Un matrice de $\mathrm{M}_n(\mathbb{R})$ est dite :
symétrique si ${}^t\mathrm{M} = \mathrm{M}$
orthogonale si ${}^t\mathrm{M} = \mathrm{M}^{-1}$

Rien n'interdit d'utiliser les termes "hermitienne " et "unitaire" pour une matrice à coefficients réels, C'est alors équivalent à "symétrique" et "orthogonale".

Proposition 37. Un matrice hermitienne a
- ses coefficients diagonaux réels
- son déterminant réel
- ses valeurs propres réelles
- ses espaces propres orthogonaux

La matrice d'un endomorphisme hermitien en base orthonormée est hermitienne, ainsi les propriétés précédentes (déterminant, valeurs propres, espaces propres) sont valables pour tout endomorphisme hermitien (ou autoadjoint).

Preuve : On a :
- ${}^t\overline{\mathrm{M}} = \mathrm{M}$ donc les termes diagonaux vérifient $\overline{m_{i,i}} = m_{i,i}$ donc sont réels ;
- $\det(\mathrm{M}) = \det({}^t\overline{\mathrm{M}}) = \det(\overline{\mathrm{M}}) = \overline{\det(\mathrm{M})}$ donc $\det(\mathrm{M})$ est réel ;
- si φ est l'endomorphisme canoniquement associé à M ($\varphi^* = \varphi$ car $\mathrm{M}^* = \mathrm{M}$), et u un vecteur propre associé à la valeur propre λ, alors

$$\overline{\lambda}\|u\|^2 = \overline{\lambda}(u|u) = (\lambda u|u) = (\varphi(u)|u) = (u|\varphi^*(u)) = (u|\varphi(u)) = (u|\lambda u) = \lambda(u|u) = \lambda\|u\|^2,$$

  et comme $u \neq 0$, on en déduit que $\lambda = \overline{\lambda}$ est réel ;

- si de plus v est un vecteur propre associé à une valeur propre $\mu \neq \lambda$, alors comme $\lambda \in \mathbb{R}$,

$$\lambda(u|v) = \overline{\lambda}(u|v) = (\lambda u|v) = (\varphi(u)|v) = (u|\varphi^*(v)) = (u|\varphi(v)) = (u|\mu v) = \mu(u|v),$$

d'où $(\lambda - \mu)(u|v) = 0$ et comme $\lambda \neq \mu$, on en déduit $u \perp v$.

Le reste découle de la proposition 36. □

Exemple. La matrice $M = \begin{pmatrix} 0 & 1+i \\ 1-i & -1 \end{pmatrix}$ est hermitienne, ses coefficients diagonaux sont réels, son déterminant vaut -2, ses valeurs propres sont 1 et -2 et les espaces propres correspondants sont $\mathrm{Vect}(\begin{pmatrix} 1+i \\ 1 \end{pmatrix})$ et $\mathrm{Vect}(\begin{pmatrix} 1+i \\ -2 \end{pmatrix})$. On peut vérifier que $\overline{\begin{pmatrix} 1+i & 1 \end{pmatrix}} \begin{pmatrix} 1+i \\ -2 \end{pmatrix} = 0$ donc les espaces propres sont orthogonaux.

Proposition 38. Un matrice M unitaire a :

- ses vecteurs colonnes qui forment une base orthonormée pour le produit scalaire canonique
- ses vecteurs lignes également
- son déterminant de module 1
- ses valeurs propres de module 1
- la propriété suivante : $\forall X \in \mathbb{C}^n$, $\|MX\| = \|X\|$.

Toute matrice unitaire est une matrice de passage d'une base orthonormée à une autre (pour le produit scalaire canonique), et réciproquement.

Preuve : Soit M unitaire.

- On note C_i la i-ème colonne de M. Alors l'égalité $M^{-1} = {}^t\overline{M}$ indique que la i-ème ligne de M^{-1} est exactement ${}^t\overline{C_i}$. Ainsi quand on effectue le calcul matriciel $M^{-1}M = I_n$, le coefficient (i, j) à droite (qui vaut 1 si $i = j$ et 0 sinon) est égal à ${}^t\overline{C_i}C_j$ à gauche, qui est exactement le produit scalaire des vecteurs colonnes i et j de M.
- Le même raisonnement appliqué aux lignes et au calcul $MM^{-1} = I_n$ donne le résultat pour les lignes.
- On a $|\det(M)|^2 = \overline{\det(M)}\det(M) = \det(\overline{M})\det(M) = \det({}^t\overline{M})\det(M) = \det(M^{-1})\det(M) = \det(M^{-1}M) = \det(I_n) = 1$.
- $\forall X \in \mathbb{C}^n$, $\|MX\|^2 = {}^t(\overline{MX})MX = {}^t\overline{X}{}^t\overline{M}MX = {}^t\overline{X}X = \|X\|^2$.
- Si X est un vecteur propre (colonne) associé à la valeur propre λ, alors $\|X\|^2 = \|MX\|^2 = {}^t(\overline{MX})MX = {}^t(\overline{\lambda X})\lambda X = \overline{\lambda}\lambda\|X\|^2$, d'où $|\lambda| = 1$.

Comme les vecteurs colonnes d'une matrice hermitienne forment une base orthonormée par rapport au produit scalaire canonique, il s'agit bien d'une matrice de passage en base orthonormée. Réciproquement si M est une matrice de passage de base orthonormée $\mathscr{B}$ à base orthonormée $\mathscr{B}'$, ses colonnes C_i forment les coordonnées des vecteurs de la base orthonormée $\mathscr{B}'$ exprimées dans la base $\mathscr{B}$, et les coefficients de ${}^t\overline{M}M$ sont exactement ${}^t\overline{C_i}C_j$ qui valent 1 si $i = j$ et 0 sinon (il s'agit bien du produit scalaire des vecteurs C_i et C_j calculé dans la base $\mathscr{B}$ orthonormée). On retrouve donc ${}^t\overline{M}M = I_n$. □.

Exemple. La matrice $P = \begin{pmatrix} \frac{1+i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$ issue de l'exemple précédent (matrice de passage à la base de vecteurs propres renormalisés) est unitaire : son inverse vaut $P^{-1} = -\frac{\sqrt{2}}{1+i} \begin{pmatrix} -\frac{2}{\sqrt{6}} & -\frac{1+i}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1+i}{\sqrt{3}} \end{pmatrix} = {}^t\overline{P} = \begin{pmatrix} \frac{1-i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1-i}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{pmatrix}$.

11 Le théorème spectral

On a vu précédennement que les matrices hermitiennes (ou les endomorphismes autoadjoints) avaient des espaces propres orthogonaux. Le résultat suivant indique qu'elles sont de plus diagonalisables :

Théorème 7. Tout endormophisme autoadjoint d'un espace vectoriel préhilbertien **de dimension finie** est diagonalisable en base orthonormée.

Matriciellement, cela signifie :

- Toute matrice hermitienne complexe se décompose sous la forme $M = PDP^{-1} = PD{}^t\overline{P}$ avec D diagonale et P unitaire.
- Toute matrice symétrique réelle se décompose sous la forme $M = PDP^{-1} = PD{}^tP$ avec D diagonale et P orthogonale.

Preuve : On montre le résultat dans le cas complexe pour simplifier. La preuve dans le cas réel peut être adaptée de cette démonstration.

On raisonne par récurrence sur la dimension n de l'espace E.

- Si $n = 1$, n'importe quel vecteur u non nul forme une base de E, et $e_1 = \frac{u}{\|u\|}$ est une base orthonormée de E qui diagonalise n'importe quel endomorphisme (une matrice de taille 1 est toujours diagonale !)
- Soit $n \geq 1$. Supposons le résultat vrai pour tous les espaces préhilbertiens de dimension n, et considérons E un espace préhilbertien de dimension $n+1$, et φ un endomorphisme autoadjoint de E. Comme on travaille sur $\mathbb{C}$ le polynôme caractéristique de φ a au moins une racine complexe λ. On considère un vecteur propre associé u et la droite engendré par u : $F = \mathbb{C}u = \mathrm{Vect}(u)$. Elle est stable par φ (cela signifie que pour tout $v \in F$, $\varphi(v) \in F$), car u est propre. Donc $F^\perp$ est également stable par φ. En effet si $v \in F^\perp$, alors pour tout $w = \alpha u \in F$, $(\varphi(v)|w) = (v|\varphi^*(w)) = (v|\varphi^*(\alpha u)) = (v|\alpha\varphi^*(u)) = \alpha(v|\varphi^*(u) = \alpha(v|\varphi(u)) = \alpha(v|\lambda u) = \alpha\lambda(v|u) = 0$.

  On considère alors le nouvel endomorphisme $\varphi_{F^\perp} : F^\perp \to F^\perp, v \mapsto \varphi(v)$, il est autoadjoint (car φ l'est), et défini sur un espace préhilbertien de dimension n. Par hypothèse de récurrence il existe donc une base orthonormée de vecteurs propres pour $\varphi_{F^\perp}$, pour le produit scalaire induit sur $F^\perp$. Cette famille de vecteurs forme une famille libre de vecteurs orthonormés pour le produit scalaire sur E, et en lui adjoignant le vecteur $\frac{u}{\|u\|}$, on obtient une base orthonormée de $E = F \oplus F^\perp$ formée de vecteurs propres pour φ.

□

Remarque. Les espaces propres d'un endomorphisme autoadjoint ou d'une matrice hermitienne étant toujours orthogonaux, pour trouver une base orthonormée de vecteurs propres en pratique il suffit d'orthonormaliser une base de chaque espace propre, grâce au procédé de Gram-Schmidt.

Théorème 8 (Théorème spectral pour les opérateurs)**.** Soit φ un opérateur hermitien sur un espace vectoriel de dimension finie n, de valeurs propres $\lambda_1, \ldots, \lambda_r$. Alors $\varphi = \sum_{i=1}^r \lambda_i p_i$ où p_i est la projection orthogonale sur le sous-espace-propre E_{λ_i}.
Ces projections vérifient $\sum_{i=1}^r p_i = Id_E$ et $\forall i \neq j, p_i \circ p_j = 0$.

Idée de preuve : Donnons l'idée générale de la preuve sur un exemple. Soit φ un endomorphisme hermitien de $\mathbb{C}^3$ de valeurs propres λ de multiplicité 2 et $\mu \neq \lambda$ de multiplicité 1. On se donne une base orthonormée de vecteurs propres $\mathscr{B}' = (e'_1, e'_2, e'_3)$ associés. Alors la matrice de φ dans la base $\mathscr{B}'$ est

$$M_{\mathscr{B}'}(\varphi) = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mu \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \lambda P_1 + \mu P_2.$$

On remarque que P_1 est la matrice de la projection orthogonale p_1 sur $E_\lambda = \mathrm{Vect}(e'_1, e'_2)$ dans la base $\mathscr{B}'$ et P_2 est la matrice de la projection orthogonale p_2 sur $E_\mu = \mathrm{Vect}(e'_3)$ dans la base $\mathscr{B}'$. En effet $\forall (x, y, z) \in \mathbb{C}^3$,
$p_1(xe'_1 + ye'_2 + ze'_3) = P_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = xe'_1 + ye'_2.$

Dans le cas d'un opérateur hermitien général on fait la même étude sur le sous-espace vectoriel de définition de φ. $\square$

A Code python pour les exemples du cours

Pour faire de l'algèbre linéaire en python, commencer par charger la bibliothèque `numpy` et `scipy.linalg` (doc en ligne `https://docs.scipy.org/doc/scipy/reference/tutorial/linalg.html`)

```
>>> import numpy as np
>>> from scipy import linalg
```

Exemples de la section 2 : Espaces vectoriels

```
>>> u=np.array([1,2])
>>> u
array([1, 2])
>>> print(u)
[1 2]
>>> print(2*u)
[2 4]
>>> print(-u)
[-1 -2]
>>> v=np.array([2,0.5])
>>> print(u+v)
[ 3.   2.5]
```

Exemples de la section 4 : Matrices

Opérations d'addition et de multiplication par un scalaire sur les matrices :

```
>>> A=np.array([[0,1.2,3],[1,2,-10]])
>>> print(A)
[[  0.    1.2   3. ]
 [  1.    2.  -10. ]]
>>> B=np.array([[-2.5, 0, 1] ,[0,3,5]])
>>> print(B)
[[-2.5  0.   1. ]
 [ 0.   3.   5. ]]
>>> A+B
array([[-2.5,  1.2,  4. ],
       [ 1. ,  5. , -5. ]])
>>> C=np.array([[1,0],[-2,1]])
>>> print(C)
[[ 1  0]
 [-2  1]]
>>> print(3*C)
[[ 3  0]
 [-6  3]]
>>> print(-C)
[[-1  0]
 [ 2 -1]]
>>> C*0.5
array([[ 0.5,  0. ],
       [-1. ,  0.5]])
```

Multiplication matricielle :

```
>>> A=np.array([[0,1.2,3],[1,2,-10]])
>>> A
array([[  0. ,   1.2,   3. ],
       [  1. ,   2. , -10. ]])
>>> U=np.array([[1],[0],[-1]])
>>> U
array([[ 1],
       [ 0],
       [-1]])
>>> A.dot(U)
array([[ -3.],
       [ 11.]])
```

Non commutativité de la multiplication :

```
>>> A=np.array([[0,1],[-1,2]])
>>> B=np.array([[-3,1],[0,-1]])
>>> A
array([[ 0,  1],
       [-1,  2]])
>>> B
array([[-3,  1],
       [ 0, -1]])
>>> A.dot(B)
array([[ 0, -1],
       [ 3, -3]])
>>> B.dot(A)
array([[-1, -1],
       [ 1, -2]])
```

Inversibilité :

```
>>> A=np.array([[0,0],[0,1]])
>>> linalg.inv(A)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  File "/usr/lib/python2.7/dist-packages/scipy/linalg/basic.py", line 687, in inv
    raise LinAlgError("singular matrix")
numpy.linalg.linalg.LinAlgError: singular matrix
>>> A=np.array([[1,1,0],[1,0,1],[0,1,1]])
>>> linalg.inv(A)
array([[ 0.5,  0.5, -0.5],
       [ 0.5, -0.5,  0.5],
       [-0.5,  0.5,  0.5]])
>>> A.dot(linalg.inv(A)) #verification
array([[ 1.,  0.,  0.],
       [ 0.,  1.,  0.],
       [ 0.,  0.,  1.]])
```

Résolution de systèmes linéaires :

```
>>> A=np.array([[1,1,0],[1,0,1],[0,1,1]])
>>> A
array([[1, 1, 0],
       [1, 0, 1],
       [0, 1, 1]])
>>> B=np.array([[2],[1],[-2]])
>>> linalg.inv(A).dot(B) #lent
array([[ 2.5],
       [-0.5],
       [-1.5]])
>>> X=np.linalg.solve(A,B) #plus rapide
>>> X
array([[ 2.5],
       [-0.5],
       [-1.5]])
>>> A.dot(X)-B #verification
array([[ 0.],
       [ 0.],
       [ 0.]])
```

Puissances et polynômes de matrices :

```
>>> A=np.array([[0,0],[0,1]])
>>> A
array([[0, 0],
       [0, 1]])
>>> I2=np.identity(2)
>>> A.dot(A-I2)
array([[ 0.,  0.],
       [ 0.,  0.]])
>>> A=np.array([[1,1,0],[1,0,1],[0,1,1]])
>>> I3=np.identity(3)
>>> np.linalg.matrix_power(A,3)-2*A.dot(A)-A+2*I3
array([[ 0.,  0.,  0.],
       [ 0.,  0.,  0.],
       [ 0.,  0.,  0.]])
```

Exemples de la section 5 : Déterminants

Calcul de déterminant :

```
>>> A=np.array([[1,2,-1],[-1,0,3],[1,1,0]])
>>> linalg.det(A)
4.0
>>> B=np.array([[0,2,-1],[-4,0,3],[0,1,0]]) #meme chose en remplacant C1 par C1-C2-C3
>>> linalg.det(B)-linalg.det(A)
0.0
```

Transposée :

```
>>> A=np.array([[1,2,-1],[-1,0,3],[1,1,0]])
>>> A.T #transposee
array([[ 1, -1,  1],
       [ 2,  0,  1],
       [-1,  3,  0]])
>>> linalg.det(A)-linalg.det(A.T)
0.0
```

Inverse : représentative

```
>>> A=np.array([[1,2,-1],[-1,0,3],[1,1,0]])
>>> linalg.inv(A)
array([[-0.75, -0.25,  1.5 ],
       [ 0.75,  0.25, -0.5 ],
       [-0.25,  0.25,  0.5 ]])
```

Exemples de la section 6 : Diagonalisation

Valeurs propres et vecteurs propres :

```
>>> G=np.array([[1,1],[2,-1]])
>>> G
array([[ 1,  1],
       [ 2, -1]])
>>> linalg.eig(G)
(array([ 1.73205081+0.j, -1.73205081+0.j]),
 array([[ 0.80689822, -0.34372377],
        [ 0.59069049,  0.9390708 ]]))
>>> np.sqrt(3)
 1.7320508075688772
>>> U=np.array([[0.80689822],[0.59069049]]) #verif vecteur propre
>>> G.dot(U)
array([[1.39758871],
       [1.02310595]])
>>> G.dot(U)-np.sqrt(3)*U
array([[-3.57688967e-09],
       [ 9.77224435e-09]])
>>> R=np.array([[1/2, -np.sqrt(3)/2],[np.sqrt(3)/2, 1/2]])
>>> linalg.eig(R) #les valeurs propres sont complexes
(array([0.5+0.8660254j, 0.5-0.8660254j]),
 array([[0.70710678+0.j        , 0.70710678-0.j        ],
        [0.        -0.70710678j, 0.        +0.70710678j]]))
>>> B=np.array([[2,0,0],[1,1,0],[1,-1,2]])
>>> linalg.eig(B)
(array([2.+0.j, 1.+0.j, 2.+0.j]),
 array([[0.        , 0.        , 0.70710678],
        [0.        , 0.70710678, 0.70710678],
        [1.        , 0.70710678, 0.        ]]))
```

Polynôme caractéristique :

```
>>> A=np.array([[1,2,-1],[-1,0,3],[1,1,0]])
```

```
>>> linalg.eig(A)[0]
array([-0.5+1.32287566j, -0.5-1.32287566j,  2. +0.j        ])
>>> coeffpoly=[1,1,2] #coeffs polynome X^2+X+2
>>> np.roots(coeffpoly) #verif racines
array([-0.5+1.32287566j, -0.5-1.32287566j])
>>> np.sqrt(7)/2
1.3228756555322954
```

Matrices diagonalisables :

```
>>> G=np.array([[1,1],[2,-1]])
>>> D=np.diag([np.sqrt(3), -np.sqrt(3)])
>>> P=np.array([[1+np.sqrt(3),1-np.sqrt(3)],[2,2]])
>>> G-P.dot(D.dot(linalg.inv(P)))
array([[  2.22044605e-16,   2.22044605e-16],
       [  2.22044605e-16,   0.00000000e+00]])
>>> A=np.array([[1,2,-1],[-1,0,3],[1,1,0]])
>>> linalg.eig(A)
(array([-0.5+1.32287566j, -0.5-1.32287566j,  2.0+0.j        ]),
 array([[-0.74161985+0.j        , -0.74161985-0.j        ,  0.57735027+0.j
],
        [ 0.47193990-0.35675303j,  0.47193990+0.35675303j,  0.57735027+0.j
],
        [-0.16854997+0.26756478j, -0.16854997-0.26756478j,  0.57735027+0.j
]]))
>>> C=np.array([[2,1,0],[0,2,0],[0,0,1]])
>>> linalg.eig(C)
(array([ 2.+0.j,  2.+0.j,  1.+0.j]),
 array([[  1.00000000e+00,  -1.00000000e+00,   0.00000000e+00],
        [  0.00000000e+00,   4.44089210e-16,   0.00000000e+00],
        [  0.00000000e+00,   0.00000000e+00,   1.00000000e+00]]))
```

(attention pour C on dirait que c'est diagonalisable, mais c'est dû à la précision machine (le deuxième vecteur propre est égal à moins le premier à une erreur de 10^{-16} près.)

Exemples de la partie II : Algèbre hermitienne

Produit scalaire du plan ou de l'espace :

```
>>> u=np.array([1,2])
>>> v=np.array([2,0])
>>> np.vdot(u,v)
2
```

Produit scalaire canonique dans $\mathbb{C}^n$:

```
>>> u=np.array([1,1.j,1+2.j])
>>> np.vdot(u,u)
(7+0j)
>>> u=np.array([1.j,2.j])
>>> v=np.array([1,0])
>>> np.vdot(u,v)
```

```
-1j
>>> np.vdot(u+v,u+v)
(6+0j)
>>> np.vdot(u,u)+np.vdot(v,v)
(6+0j)
```

Procédé de Gram-Schmidt : il est encodé sur numpy comme la décomposition QR (les colonnes de la matrice Q sont alors les vecteurs de la B.O.N., et il faut entrer les vecteurs à normaliser comme les colonnes d'une matrice)

```
>>> def gs(X):
                Q, R = np.linalg.qr(X)
                return Q
>>> A=np.array([[1.j,0],[1,1.j]])
>>> gs(A)
array([[  0.00000000e+00 -7.07106781e-01j,
          7.07106781e-01 -1.11022302e-16j],
       [ -7.07106781e-01 +0.00000000e+00j,
         -1.11022302e-16 +7.07106781e-01j]])
>>> import math
>>> 1/math.sqrt(2)
0.7071067811865475
```

Matrice adjointe :

```
>>> M=np.array([[1,1.j],[2.j , -1]])
>>> M.conj().T
array([[ 1.-0.j,  0.-2.j],
       [ 0.-1.j, -1.-0.j]])
```

Diagonalisation d'une matrice hermitienne :

```
>>> M=np.array([[0,1+1.j],[1-1.j, -1]])
>>> np.linalg.eigh(M) #retourne la table des valeurs propres et la BON de vecteurs propres
(array([-2.,  1.]), array([[ 0.57735027+0.j        , -0.81649658+0.j        ],
        [-0.57735027+0.57735027j, -0.40824829+0.40824829j]]))
>>> 1/math.sqrt(3)
0.5773502691896258
>>> 1/math.sqrt(6)
0.4082482904638631
>>> 2/math.sqrt(6)
0.8164965809277261
```