MATH 223 - Formes Quadratiques et Géométrie (UPMC 2005/06)

Jan Nekovář

0. Introduction

(0.0) Que fait-on dans ce cours ?

Dans le cours LM 125 (Espaces Vectoriels) on a considéré des objects linéaires (espaces vectoriels, applications linéaires, formes linéaires). Par exemple, $3x-4y$ est une **forme linéaire** sur un espace vectoriel de dimension deux muni de coordonnées x et y.

Dans le cours actuel on va étudier les **formes quadratiques** et des objets associés. Par exemple, $q = 2x^2 - 8xy + 9y^2$ est une forme quadratique. On va fréquemment aborder des questions de **diagonalisation** de formes quadratiques; voici une transformation qui diagonalise la forme quadratique q :

$$2x^2 - 8xy + 9y^2 = 2(x-2y)^2 + y^2 = 2x'^2 + y'^2, \qquad (x' = x - 2y,\ y' = y). \tag{0.0.0}$$

Attention : il y a au moins 3 notions distinctes de diagonalisation :

(1) Diagonalisation d'applications linéaires (voir §1.6 ci-dessous).
(2) Diagonalisation "facile" de formes quadratiques, dont (0.0.0) est un cas particulier (voir §3.3 ci-dessous).
(3) Diagonalisation "difficile" de formes quadratiques (voir §5.1 ci-dessous).

(0.1) Où peut-on rencontrer des formes quadratiques ?

(0.1.0) Géométrie euclidienne : la longeur euclidienne $\|x\|$ d'un vecteur $x \in \mathbf{R}^n$ (à coordonnées $x_1, \ldots, x_n$) est donnée par le théorème de Pythagore

$$\|x\|^2 = x_1^2 + \cdots + x_n^2 = (x \mid x)$$

(où l'on a noté $(x \mid y) = x_1y_1 + \cdots + x_ny_n$ le produit scalaire euclidien usuel).

L'équation $\|x\|^2 = 1$ définit la sphère unité en $\mathbf{R}^n$; par exemple, lorsque $n = 2$, on obtient le cercle unité $x_1^2 + x_2^2 = 1$ en $\mathbf{R}^2$.

On peut considérer des équations analogues définies par des formes quadratiques plus générales, telles que

$$2x^2 - 8xy + 9y^2 = 1 \quad \text{(ellipse)}, \qquad 2x^2 - 8xy - 9y^2 = 1 \quad \text{(hyperbole)}.$$

En $\mathbf{R}^3$, les équations

$$x_1^2 + x_2^2 - x_3^2 = 1, \qquad x_1^2 - x_2^2 - x_3^2 = 1$$

définissent un hyperboloïde à une nappe (resp., à deux nappes).

(0.1.1) Géométrie non-euclidienne (= pseudo-euclidienne) : dans la théorie de la relativité, on remplace l'espace usuel $\mathbf{R}^3$ par l'espace de Minkowski ("espace-temps") $\mathbf{R}^{1,3}$. Un élément $x \in \mathbf{R}^{1,3}$ a quatre coordonnées x_0, x_1, x_2, x_3; ici, x_1, x_2, x_3 sont les coordonnées usuelles dans $\mathbf{R}^3$ et la coordonnée x_0 est égale à $x_0 = ct$, où t est le temps et c est la vitesse de la lumière. La longeur pseudo-euclidienne est définie par la formule

$$\|x\|_{Min}^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2.$$

L'équation

$$0 = \|x\|_{Min}^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2$$

définit le **cône de lumière**, qui est balayé par les rayons de la lumière.

1. Algèbre Linéaire

Dans ce chapitre on rappele les résultats de l'algèbre linéaire qui seront utilisés dans la suite.

Il y a deux versions de l'algèbre linéaire :

Version abstraite, où l'on travaille avec les espace vectoriels, les vecteurs et les applications linéaires abstrait(e)s.

Version matricielle, qui est indispensable pour les calculs : dès qu'on choisit des bases (= des systèmes de coordonnées) de tous les espaces vectoriels en question (supposés d'être de dimension finie), on écrit tous les objects abstraits en termes matriciels.

Voici le début du dictionnaire entre les deux langages :

Objet abstrait	Objet matriciel
espace vectoriel E	K^n
vecteur $x \in E$	vecteur colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
application linéaire $\alpha : E \longrightarrow F$	matrice $A \in M_{p,n}(K)$

1.1 Les définitions de base

(1.1.1) Corps des scalaires. On travaille avec les espaces vectoriels sur un **corps des scalaires** K. En pratique, on utilise surtout $K = \mathbf{R}$ (le corps des **nombres réels**), $K = \mathbf{C}$ (le corps des **nombres complexes**) ou $K = \mathbf{F}_q$ (**le corps fini ayant** $q = p^r$ **éléments**, où p est un nombre premier; par exemple, les éléments de $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ sont les classes résiduelles (= classes des restes) modulo p.

(1.1.2) Espaces vectoriels. Un élément d'un K**-espace vectoriel** E (abréviation : K-ev) s'appelle un **vecteur**. Deux vecteurs admettent une somme; un vecteur peut être multiplié par un scalaire quelconque. Les opérations

$$x, y \in E \mapsto x + y \in E, \qquad x \in E,\ \lambda \in K \mapsto \lambda x \in E$$

satisfont aux propriétés **"usuelles"** (pour tous $x, y, z \in E$ et $\lambda, \mu \in K$) :

$$\begin{gathered} x + y = y + x, \qquad x + (y + z) = (x + y) + z, \\ \lambda(\mu x) = (\lambda\mu)x, \qquad \lambda(x + y) = \lambda x + \lambda y, \qquad (\lambda + \mu)x = \lambda x + \mu x, \\ 1x = x, \qquad 0x = \vec{0}, \qquad x + \vec{0} = x. \end{gathered}$$

Il est commode d'écrire $-x$ au lieu de $(-1)x$, et $x - y$ au lieu de $x + (-y)$.

Attention : il faut distinguer le **scalaire nul** $0 \in K$ et le **vecteur nul** $\vec{0} \in E$, même si l'on utilise souvent le même symbole.

(1.1.3) Exemples d'espaces vectoriels : (0) L'espace vectoriel nul $0 = \{\vec{0}\}$ qui ne contient que le vecteur nul.

(1) L'espace des vecteurs colonnes de dimension $m \geq 1$:

$$K^m = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \middle| a_i \in K \right\}$$

(2) L'espace des matrices de type $m \times n$ à coefficients dans K ($m =$ le nombre de lignes, $n =$ le nombre de colonnes) :

$$M_{m,n}(K) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \middle| a_{ij} \in K \right\}$$

L'espace des **matrices carrées** (lorsque $m = n$) est noté $M_n(K) = M_{n,n}(K)$.

(1.1.4) Sous-espaces vectoriels. Soit E un K-ev. Un (K-)**sous-espace vectoriel** (abréviation : sev) de E est un sous-ensemble non vide $F \subset E$ tel que

$$\begin{aligned} \forall x, y \in F \qquad & x + y \in F \\ \forall \lambda \in K \ \forall x \in F \qquad & \lambda x \in F \end{aligned}$$

(rappelons que le symbole "$\forall$" signifie **"pour tout"**). Ceci entraîne que F est un K-ev ayant le vecteur nul $\overrightarrow{0}$ en commun avec E.

(1.1.5) Sous-espace vectoriel engendré. Soit S un sous-ensemble d'un K-ev E. Il existe **le plus petit sous-espace vectoriel** de E contenant S; on l'appelle le **sous-espace (vectoriel) engendré par** S et l'on le note $\mathrm{vect}(S)$. Les éléments de $\mathrm{vect}(S)$ sont les **combinaisons linéaires** des éléments de S, à savoir les sommes **finies**

$$\sum_{s \in S} \lambda_s \, s \qquad\qquad (\lambda_s \in K).$$

En particulier, si l'ensemble $S = \{x_1, \ldots, x_n\}$ est fini, alors on a

$$\mathrm{vect}(x_1, \ldots, x_n) = \{\lambda_1 x_1 + \cdots + \lambda_n x_n \mid \lambda_i \in K\}.$$

(1.1.6) Espaces vectoriels de dimension finie. Un K-ev E est dit de **dimension finie** s'il existe un ensemble fini de vecteurs $S = \{x_1, \ldots, x_n\}$ qui engendre E ($\iff$ tel que $\mathrm{vect}(S) = E$). Si c'est le cas, la cardinalité (= le nombre d'éléments) minimum d'un tel ensemble S s'appelle **la dimension de** E; on la note $\dim(E)$.

(1.1.7) (In)dépendance linéaire. On dit qu'une famille (finie) de vecteurs $x_1, \ldots, x_n$ d'un K-ev E est **linéairement dépendante (ou liée)** s'il existe une relation linéaire

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = \overrightarrow{0} \qquad\qquad (\lambda_i \in K) \qquad\qquad (1.1.7.1)$$

non-triviale ($\iff$ les scalaires λ_i ne sont pas tous nuls).

Si, par contre, toute relation linéaire (1.1.7.1) entraîne $\lambda_1 = \cdots = \lambda_n = 0$, on dit que les vecteurs $x_1, \ldots, x_n$ sont **linéairement indépendants**.

(1.1.8) Bases (le cas de dimension finie). Un ensemble fini (ordonné) de vecteurs $\mathcal{B} = \{e_1, \ldots, e_n\}$ d'un K-ev E est une **base de** E si les vecteurs de $\mathcal{B}$ sont linéairement indépendants et si $\mathrm{vect}(\mathcal{B}) = E$. Dans ce cas, tout vecteur $x \in E$ s'écrit d'une manière **unique**

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n \qquad\qquad (\lambda_i \in K).$$

On appelle les scalaires $\lambda_1, \ldots, \lambda_n$ **les coordonnées de** x **dans la base** $\mathcal{B}$.

Selon un résultat fondamental, tout K-ev de dimension finie admet une base, et la cardinalité de chaque base est égale à $\dim(E)$.
Un **choix** de base $\mathcal{B}$ donne lieu à une **identification**

$$M_{\mathcal{B}} : E \xrightarrow{\sim} K^n \qquad\qquad (n = \dim(E)),$$

qui dépend de $\mathcal{B}$: on associe à chaque vecteur

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n \in E$$

le vecteur colonne

$$M_{\mathcal{B}}(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n$$

des coordonnées de x dans la base $\mathcal{B}$. On dit que le vecteur colonne $M_{\mathcal{B}}(x)$ **représent** x **dans la base** $\mathcal{B}$. Plus généralement, on associe à chaque famille finie de vecteurs $x_1, \dots, x_p \in E$ la matrice

$$M_{\mathcal{B}}(x_1, \dots, x_p) = (M_{\mathcal{B}}(x_1) \mid \cdots \mid M_{\mathcal{B}}(x_p)) \in M_{n,p}(K)$$

dont les colonnes représentent les vecteurs $x_1, \dots, x_p$ dans la base $\mathcal{B}$.
Exemple : La **base canonique** de K^n contient les vecteurs colonnes

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \qquad \cdots \qquad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Dans ce cas, on a

$$\lambda_1 e_1 + \cdots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

(1.1.9) Le cas de dimension infinie. On peut aussi définir la notion de base (algébrique) pour les K-ev de dimension infinie. Cependant, dans les espaces qu'on rencontre le plus fréquemment (tels que les espaces des fonctions continues ou intégrables $f : [a,b] \longrightarrow \mathbf{R}$), il est utile de considérer aussi des **combinaisons linéaires infinies** $\sum_{i=1}^{\infty} \lambda_i f_i$, à condition qu'elles existent. Ceci conduit à des notions diverses de bases analytiques, qui apparaissent, par exemple, en analyse de Fourier ou dans la théorie des ondelettes.

1.2 Application linéaires et leurs représentations matricielles

(1.2.1) Applications linéaires. Une application $f : E \longrightarrow E'$ entre deux K-ev est dite **linéaire** lorsqu'elle conserve les opérations :

$$\forall x, y \in E \qquad f(x+y) = f(x) + f(y)$$
$$\forall \lambda \in K \ \forall x \in E \qquad f(\lambda x) = \lambda\, f(x)$$

(ce qui entraîne $f(\vec{0}) = \vec{0}$). **Notation :** on note $\mathscr{L}(E, E')$ l'ensemble de toutes les applications linéaires de E vers E'. $\mathscr{L}(E, E')$ est un K-ev, dont les opérations sont données par les formules

$$(f+g)(x) = f(x) + g(x), \qquad (\lambda f)(x) = \lambda(f(x)) \qquad\qquad (f, g \in \mathscr{L}(E, E'),\ x \in E,\ \lambda \in K).$$

D'ici jusqu'à fin de §1.2 on ne considère que des K-ev de dimension finie.

(1.2.2) Exemples d'applications linéaires : (0) L'application nulle $f = 0$ ($\forall x \in E \quad f(x) = \overrightarrow{0}$).
(1) L'application identité $f = \mathrm{Id}_E$, lorsque $E = E'$ ($\forall x \in E \quad f(x) = x$).
(2) On suppose que $f : K^3 \longrightarrow K^2$ est une application linéaire telle que

$$f(e_1) = f(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad f(e_2) = f(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -1 \\ 4 \end{pmatrix}, \quad f(e_3) = f(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 3 \\ 5 \end{pmatrix}.$$

Par définition, on a

$$f(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = f(x\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}) = x\,f(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) + y\,f(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}) + z\,f(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}) =$$

$$= x\begin{pmatrix} 2 \\ 1 \end{pmatrix} + y\begin{pmatrix} -1 \\ 4 \end{pmatrix} + z\begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 2x - y + 3z \\ x + 4y + 5z \end{pmatrix} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 4 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Autrement dit, l'application f est donnée par produit matriciel (à gauche) par la matrice

$$\begin{pmatrix} 2 & -1 & 3 \\ 1 & 4 & 5 \end{pmatrix} = (f(e_1) \mid f(e_2) \mid f(e_3)) \in M_{2,3}(K),$$

dont les colonnes sont égales aux valeurs de f en les vecteurs e_1, e_2, e_3 de la base canonique de K^3.
(3) En général, une application linéaire $f : E \longrightarrow E'$ est déterminée par les valeurs $f(e_1), \ldots, f(e_n)$, où $\{e_1, \ldots, e_n\}$ est une base quelconque de E, puisqu'on a

$$f(\lambda_1 e_1 + \cdots + \lambda_n e_n) = \lambda_1\, f(e_1) + \cdots + \lambda_n\, f(e_n).$$

(1.2.3) Produit matriciel. Rappelons que le produit matriciel

$$\begin{array}{c} M_{m,n}(K) \times M_{n,p}(K) \longrightarrow M_{m,p}(K) \\ A = (a_{ij}),\ B = (b_{kl}) \ \mapsto\ AB = C = (c_{il}) \end{array}$$

est défini par la formule

$$c_{il} = \sum_{j=1}^{n} a_{ij}\, b_{jl} = \text{ le produit scalaire de la } i\text{−ième ligne de } A \text{ et la } l\text{−ième colonne de } B.$$

Attention : le produit AB n'est défini que si le nombre de colonnes de A et le nombre de lignes de B coïncident.

Un cas particulier : Lorsque $p = 1$, on a $M_{n,1} = K^n$ et $M_{m,1} = K^m$, d'où le produit

$$M_{m,n}(K) \times K^n \longrightarrow K^m, \qquad A, x \mapsto Ax. \tag{1.2.3.1}$$

On peut reconstituer chaque matrice $A \in M_{m,n}(K)$ à partir des produits Ae_i (où $e_1, \ldots, e_n$ est la base canonique de K^n (cf. 1.1.8)), car on a

$$Ae_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix} = \text{ the } i\text{−th column of } A. \tag{1.2.3.2}$$

Autrement dit, on peut écrire

$$A = (Ae_1 \mid \cdots \mid Ae_n). \tag{1.2.3.3}$$

(1.2.4) Matrices = applications linéaires. L'exemple 1.2.2(2) est un cas particulier de l'énoncé suivant :

$$\mathscr{L}(K^n, K^m) = M_{m,n}(K) \tag{1.2.4.1}$$

(**attention :** l'ordre de m et n est renversé).
En effet, une application linéaire quelconque $f : K^n \longrightarrow K^m$ vérifie

$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = f(x_1e_1 + \cdots + x_ne_n) = x_1\, f(e_1) + \cdots + x_n\, f(e_n) = (f(e_1) \mid \cdots \mid f(e_n)) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

où l'on a noté

$$(f(e_1) \mid \cdots \mid f(e_n)) \in M_{m,n}(K) \tag{1.2.4.2}$$

la matrice dont les colonnes sont égales à $f(e_i) \in K^m$.
Réciproquement, la multiplication à gauche par une matrice $A \in M_{m,n}(K)$ définit, d'après (1.2.3.1), une application linéaire

$$\begin{aligned} f : K^n &\longrightarrow K^m \\ X &\mapsto AX \end{aligned} \tag{1.2.4.3}$$

telle que l'on a

$$A = (Ae_1 \mid \cdots \mid Ae_n) = (f(e_1) \mid \cdots \mid f(e_n)) \tag{1.2.4.4}$$

(d'après (1.2.3.3)). On note cette application

$$K^n \xrightarrow{A} K^m. \tag{1.2.4.5}$$

(1.2.5) Représentation matricielle des applications linéaires. Soit $f : E \longrightarrow E'$ une application linéaire (entre deux K-ev de dimension finie). Étant donné des bases respectives $\mathcal{B} = \{e_1, \ldots, e_n\}$ et $\mathcal{B}' = \{e'_1, \ldots, e'_p\}$ de E et E' (où $n = \dim(E)$ et $p = \dim(E')$), les identifications

$$M_{\mathcal{B}} : E \xrightarrow{\sim} K^n, \qquad M_{\mathcal{B}'} : E' \xrightarrow{\sim} K^p$$

que l'on a définies en 1.1.8 nous permettent de transporter f à une application linéaire de K^n vers K^p, qui s'écrit forcément sous la forme (1.2.4.5), où $A \in M_{n,p}(K)$:

$$\begin{array}{rccc} M_{\mathcal{B}} : & E & \xrightarrow{\sim} & K^n \\ & \Big\downarrow f & & \Big\downarrow A \\ M_{\mathcal{B}'} : & E' & \xrightarrow{\sim} & K^p \end{array}$$

On dit que la matrice

$$M_{\mathcal{B},\mathcal{B}'}(f) = A \in \mathscr{L}(K^n, K^p) = M_{n,p}(K)$$

représent f dans les bases $\mathcal{B}$ et $\mathcal{B}'$.

Formulaire: (1) Lorsqu'on note $X = M_{\mathcal{B}}(x) \in K^n$ et $Y = M_{\mathcal{B}'}(y) \in K^p$ les vecteurs colonnes qui représentent des vecteurs $x \in E$ et $y \in E'$ dans les bases respectives $\mathcal{B}$ et $\mathcal{B}'$, on a

$$y = f(x) \iff Y = AX.$$

(2) En particulier, lorsque $x = e_i \in \mathcal{B}$ est un élément de la base $\mathcal{B}$ et $y = f(e_i)$, alors

$$X = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{(avec 1 en la } i\text{−ième ligne);}$$

c'est aussi l'élément e_i de la base canonique de K^n, donc on a

$$Y = M_{\mathcal{B}'}(y) = M_{\mathcal{B}'}(f(e_i)) = Ae_i = \text{ la } i\text{−ième colonne de } A,$$

d'après (1.2.3.2). En utilisant (1.2.3.3), on obtient la formule

$$A = (Ae_1 \mid \cdots \mid Ae_n) = (M_{\mathcal{B}'}(f(e_1)), \ldots, M_{\mathcal{B}'}(f(e_n))) = M_{\mathcal{B}'}(f(\mathcal{B})).$$

Autrement dit, les colonnes de A sont les coordonnées (dans la base $\mathcal{B}'$) des images des vecteurs de la base $\mathcal{B}$ par l'application f.

(3) Lorsque $E = E'$, $\mathcal{B} = \mathcal{B}'$ et $f = \mathrm{Id}_E$, alors la matrice

$$A = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in M_n(K)$$

est égale à la matrice identité.

(1.2.6) Produit matriciel = application composée. Soient E, E', E'' des K-ev de dimensions respectives n, p, q. Étant donné des applications linéaires $f : E \longrightarrow E'$ et $g : E' \longrightarrow E''$, on note $g \circ f : E \longrightarrow E''$ leur composée

$$g \circ f : E \xrightarrow{f} E' \xrightarrow{g} E'' \qquad\qquad (g \circ f(x) = g(f(x)))$$

(le symbole $g \circ f$ signifie **"g suit f"**).

Étant donné des bases respectives $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ de E, E', E'', on obtient les identifications

$$M_{\mathcal{B}} : E \xrightarrow{\sim} K^n, \qquad M_{\mathcal{B}'} : E' \xrightarrow{\sim} K^p, \qquad M_{\mathcal{B}''} : E'' \xrightarrow{\sim} K^q,$$

qui transforment f et g en des applications "multiplication à gauche par A et B", respectivement

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^p \\ X & \mapsto & AX, \end{array} \qquad \begin{array}{ccc} K^p & \xrightarrow{B} & K^q \\ Y & \mapsto & BY, \end{array}$$

dont la composée est donnée par la multiplication à gauche par la matrice BA :

$$\begin{array}{ccccc} K^n & \xrightarrow{A} & K^p & \xrightarrow{B} & K^q \\ X & \mapsto & AX & \mapsto & BAX \end{array}$$

Autrement dit, la composée $g \circ f$ des applications g et f correspond au produit matriciel

$$M_{q,p}(K) \times M_{p,n}(K) \longrightarrow M_{q,n}(K), \qquad B, A \mapsto BA.$$

En particulier,

$$f \text{ est inversible} \iff A \text{ est inversible;}$$

si c'est le cas, alors

$$g = f^{-1} \iff B = A^{-1}.$$

(1.2.7) Noyau, Image, Rang. Soit $f : E \longrightarrow E'$ une application linéaire. Le **noyau** de f est le sev de E

$$\mathrm{Ker}(f) = \{x \in E \mid f(x) = 0\} \subset E.$$

L'**image** de f est le sev de E'

$$\mathrm{Im}(f) = f(E) = \{f(x) \mid x \in E\} \subset E'.$$

Le **rang** de f est l'entier

$$\mathrm{rg}(f) = \dim(\mathrm{Im}(f)).$$

Exemple : si $f : K^2 \longrightarrow K^2$ est définie par la formule

$$f(\begin{pmatrix} x \\ y \end{pmatrix}) = f(xe_1 + ye_2) = \begin{pmatrix} x \\ 0 \end{pmatrix} = xe_1,$$

alors le noyau (resp., l'image) de f est égal(e) à la droite verticale (resp., horizontale) $\mathrm{Ker}(f) = \mathrm{vect}(e_2)$ (resp., $\mathrm{Im}(f) = \mathrm{vect}(e_1)$); d'où $\mathrm{rg}(f) = 1$.

La formule du rang : pour toute application linéaire $f : E \longrightarrow E'$, on a

$$\dim(E) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(\mathrm{Ker}(f)) + \mathrm{rg}(f). \tag{1.2.7.1}$$

Formulation matricielle : une application linéaire $f : K^n \longrightarrow K^p$ s'écrit $f(X) = AX$ ($A \in M_{p,n}(K)$), ce qui entraîne que le noyau

$$\mathrm{Ker}(f) = \{X \in K^n \mid AX = 0\} \subset K^n$$

est l'ensemble des solutions du système des équations linéaires $AX = 0$. D'autre part,

$$f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = f(x_1 e_1 + \cdots + x_n e_n) = x_1\, Ae_1 + \cdots + x_n\, Ae_n;$$

comme Ae_i coïncide avec la i-ième colonne de A, on a

$$\mathrm{Im}(f) = \{AX \mid X \in K^n\} = \mathrm{vect}(\text{les colonnes de } A) \subset K^p.$$

En particulier, $\mathrm{rg}(f)$ est égal au **rang de la matrice** A, qui est défini comme

$$\mathrm{rg}(A) = \dim(\text{l'espace vectoriel engendré par les colonnes de } A)$$

(voir 1.4 ci-dessous pour le calcul du rang de A). La formule (1.2.7.1) s'écrit
La formule du rang pour des matrices : Si $A \in M_{p,n}(K)$, alors on a

$$n = \dim(\{X \in K^n \mid AX = 0\}) + \mathrm{rg}(A). \tag{1.2.7.2}$$

Par exemple, si $A \in M_n(K)$ est une matrice carrée, alors on a :

$$\text{les colonnes de } A \text{ forment une base de } K^n \iff \mathrm{rg}(A) = n \iff \det(A) \neq 0, \tag{1.2.7.3}$$

ce qui nous permet de **vérifier si n vecteurs donnés de K^n forment une base.**

(1.2.8) Applications linéaires surjectives, injectives, bijectives. On dit qu'une application linéaire $f : E \longrightarrow E'$ est

(a) **surjective** lorsque $f(E) = E'$;
(b) **injective** lorsque $f(x) \neq f(y)$ si $x \neq y$ $(x, y \in E)$, ce qui équivaut (exercice!) à $\mathrm{Ker}(f) = \{\vec{0}\}$;
(c) **bijective** (ou un **isomorphisme**) lorsque f est à la fois surjective and injective, ce qui équivaut à l'existence de l'application inverse $g = f^{-1} : E' \longrightarrow E$ vérifiant $g(f(x)) = x$ et $f(g(y)) = y$, pour tous $x \in E$ et $y \in E'$ (l'application g est encore linéaire). Si c'est le cas, alors les espaces E et E' se comportent de la même façon; par exemple, f envoie un ensemble de vecteurs de E linéairement indépendants vers un ensemble de vecteurs de E' linéairement indépendants, d'où $\dim(F) = \dim f(F)$, pour tout sev $F \subset E$;
(d) **un endomorphisme de** E lorsque $E = E'$. Dans ce cas (plus généralement, lorsqu'on a $\dim(E) = \dim(E')$), la formule du rang (1.2.7.1) entraîne que les conditions (a) $\iff$ (b) $\iff$ (c) ci-dessus sont équivalentes.

Exemples : (1) L'application $M_{\mathcal{B}} : E \xrightarrow{\sim} K^n$ définie en 1.1.8 est un isomorphisme.
(2) L'application $M_{\mathcal{B},\mathcal{B}'} : \mathscr{L}(E, E') \longrightarrow \mathscr{L}(K^n, K^p) = M_{p,n}(K)$ définie en 1.2.5 est un isomorphisme, donc

$$\dim(\mathscr{L}(E, E')) = \dim(M_{p,n}(K)) = np = \dim(E)\dim(E').$$

Formulation matricielle : Une application linéaire $f : K^n \longrightarrow K^p$ s'écrit $f(X) = AX$ $(A \in M_{p,n}(K))$; dans ce cas on a

f est surjective $\iff$ $\mathrm{rg}(A) = p$;
f est injective $\iff$ $X = 0 \in K^n$ est l'unique solution du système linéaire $AX = 0$.
f is bijective $\iff$ $n = p = \mathrm{rg}(A)$ $\iff$ $n = p$ et $\det(A) \neq 0$ $\iff$ $n = p$ et A est inversible (voir (1.2.7.2) et la remarque (d) ci-dessus).

1.3 Changement de base

Dans le paragraphe §1.3 on ne considère que des K-ev de dimension finie.

(1.3.1) Changement de base : représentation des vecteurs. Soit E un K-ev de dimension $n = \dim(E)$ muni de deux bases $\mathcal{B}_1$ (= "ancienne base") et $\mathcal{B}_2$ (= "nouvelle base"). On appelle la matrice

$$P = M_{\mathcal{B}_1}(\mathcal{B}_2) \in M_n(K)$$

(dont les colonnes sont les coordonnées des vecteurs de la nouvelle base $\mathcal{B}_2$ dans l'ancienne base $\mathcal{B}_1$) **la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$.** Si l'on utilise la notation que l'on a introduite en 1.2.5, le diagramme

$$\begin{array}{cccc} M_{\mathcal{B}_2} : & E & \xrightarrow{\sim} & K^n \\ & \Big\downarrow{\scriptstyle \mathrm{Id}_E} & & \Big\downarrow{\scriptstyle P} \\ M_{\mathcal{B}_1} : & E & \xrightarrow{\sim} & K^n \end{array} \tag{1.3.1.1}$$

montre que l'on a

$$P = M_{\mathcal{B}_2,\mathcal{B}_1}(\mathrm{Id}_E).$$

Lorsqu'on échange les rôles de $\mathcal{B}_1$ et $\mathcal{B}_2$, on déduit de 1.2.6 que la matrice P est inversible et que son inverse est égal à $P^{-1} = M_{\mathcal{B}_1,\mathcal{B}_2}(\mathrm{Id}_E)$ (= la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_1$). Réciproquement, lorsqu'on fixe $\mathcal{B}_1$, chaque matrice inversible $P \in M_n(K)$ est obtenue comme la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$, pour une base unique $\mathcal{B}_2$.

Formulaire: (1) Si $x \in E$, on note $X_1 = M_{\mathcal{B}_1}(x) \in K^n$ et $X_2 = M_{\mathcal{B}_2}(x) \in K^n$ les vecteurs colonnes qui représentent x dans les bases respectives $\mathcal{B}_1$ et $\mathcal{B}_2$. Le diagramme (1.3.1.1) montre que l'on a

$$X_1 = PX_2, \qquad X_2 = P^{-1}X_1. \tag{1.3.1.2}$$

(1.3.2) Exemples : (1) Lorsque $E = K^n$ et $\mathcal{B}_1 = \{e_1, \dots, e_n\}$ est la base canonique de K^n, alors les colonnes de P et les éléments de $\mathcal{B}_2$ coïncident.
(2) Par exemple, si $E = \mathbf{R}^2$ et

$$\mathcal{B}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathcal{B}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \end{pmatrix},$$

alors on a

$$P = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}.$$

Pour vérifier que les vecteurs de $\mathcal{B}_2$ forment une base de $\mathbf{R}^2$, il suffit de noter que le déterminant

$$\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} \neq 0$$

ne s'annule pas (cf. (1.2.7.3)).
(3) Soient $E = \mathbf{R}^2$ et

$$\mathcal{B}_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathcal{B}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Comme

$$\begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} \neq 0 \neq \begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix},$$

on a bien deux bases de $\mathbf{R}^2$. Pour déterminer la matrice de passage $P = M_{\mathcal{B}_2,\mathcal{B}_1}(\mathrm{Id}_E)$ de $\mathcal{B}_1$ à $\mathcal{B}_2$, on considère aussi la base canonique $\mathcal{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. D'après (1), on sait que la matrice de passage P_1 (resp., P_2) de $\mathcal{B}$ à $\mathcal{B}_1$ (resp., à $\mathcal{B}_2$), est égale à

$$P_1 = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix}.$$

Le comparaison des deux diagrammes

$$\begin{array}{cccc}
M_{\mathcal{B}_2}: & E & \xrightarrow{\sim} & K^n \\
& \Big\downarrow \mathrm{Id}_E & & \Big\downarrow P \\
M_{\mathcal{B}_1}: & E & \xrightarrow{\sim} & K^n \\
& \Big\downarrow \mathrm{Id}_E & & \Big\downarrow P_1 \\
M_{\mathcal{B}}: & E & \xrightarrow{\sim} & K^n
\end{array}
\qquad\qquad
\begin{array}{cccc}
M_{\mathcal{B}_2}: & E & \xrightarrow{\sim} & K^n \\
& \Big\downarrow \mathrm{Id}_E & & \Big\downarrow P_2 \\
M_{\mathcal{B}}: & E & \xrightarrow{\sim} & K^n
\end{array}$$

montre (en utilisant 1.2.6) que l'on a

$$P_1 P = P_2,$$

d'où

$$P = P_1^{-1} P_2 = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ -8 & -5 \end{pmatrix}.$$

Pour vérifier le résultat, on calcule

$$5 \begin{pmatrix} 2 \\ 3 \end{pmatrix} - 8 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \qquad 4 \begin{pmatrix} 2 \\ 3 \end{pmatrix} - 5 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Voici une autre méthode qui nous permet de déduire la relation $P_1 P = P_2$: si l'on représent un vecteur $x \in \mathbf{R}^2$ dans les bases respectives $\mathcal{B}, \mathcal{B}_1, \mathcal{B}_2$, on obtient trois vecteurs colonnes $X, X_1, X_2 \in \mathbf{R}^2$ tels que

$$X = P_1 X_1, \qquad X = P_2 X_2, \qquad X_1 = P X_2,$$

d'après (1.3.1.2) (bien sûr, $X = x$, car $\mathcal{B}$ est la base canonique). On en déduit

$$P_1 P X_2 = P_1 X_1 = P_2 X_2 \implies P_1 P = P_2.$$

(1.3.3) Changement de base : représentation des applications linéaires. Soit $f : E \longrightarrow E'$ une application linéaire; on note $n = \dim(E)$ et $p = \dim(E')$. On suppose E (resp., E') muni de deux bases $\mathcal{B}_1, \mathcal{B}_2$ (resp., $\mathcal{B}'_1, \mathcal{B}'_2$). On considère les matrices suivantes :

La matrice de passage $P = M_{\mathcal{B}_1}(\mathcal{B}_2) \in M_n(K)$ de $\mathcal{B}_1$ à $\mathcal{B}_2$.
La matrice de passage $P' = M_{\mathcal{B}'_1}(\mathcal{B}'_2) \in M_p(K)$ de $\mathcal{B}'_1$ à $\mathcal{B}'_2$.
La matrice $A_1 = M_{\mathcal{B}'_1}(f(\mathcal{B}_1)) \in M_{p,n}(K)$ qui représent f dans les bases $\mathcal{B}_1$ et $\mathcal{B}'_1$.
La matrice $A_2 = M_{\mathcal{B}'_2}(f(\mathcal{B}_2)) \in M_{p,n}(K)$ qui représent f dans les bases $\mathcal{B}_2$ et $\mathcal{B}'_2$.

Le comparaison des deux diagrammes

$$\begin{array}{cccc}
M_{\mathcal{B}_2}: & E & \xrightarrow{\sim} & K^n \\
& \Big\downarrow \mathrm{Id}_E & & \Big\downarrow P \\
M_{\mathcal{B}_1}: & E & \xrightarrow{\sim} & K^n \\
& \Big\downarrow f & & \Big\downarrow A_1 \\
M_{\mathcal{B}'_1}: & E' & \xrightarrow{\sim} & K^p \\
& \Big\uparrow \mathrm{Id}_{E'} & & \Big\uparrow P' \\
M_{\mathcal{B}'_2}: & E' & \xrightarrow{\sim} & K^p
\end{array}
\qquad\qquad
\begin{array}{cccc}
M_{\mathcal{B}_2}: & E & \xrightarrow{\sim} & K^n \\
& \Big\downarrow f & & \Big\downarrow A_2 \\
M_{\mathcal{B}'_2}: & E' & \xrightarrow{\sim} & K^p
\end{array}$$

montre (en utilisant 1.2.6) que l'on a

$$A_2 = (P')^{-1} A_1 P. \tag{1.3.3.1}$$

Un cas particulier : lorsque $E = E'$, $\mathcal{B}_1 = \mathcal{B}'_1$ et $\mathcal{B}_2 = \mathcal{B}'_2$, alors on a $n = p$, $P' = P$ et $A_1, A_2 \in M_n(K)$, donc la formule (1.3.3.1) s'écrive

$$A_2 = P^{-1} A_1 P. \tag{1.3.3.2}$$

1.4 Détermination du rang : opérations élémentaires

Soit $A \in M_{m,n}$ une matrice de type $m \times n$; on note $L_1, \ldots, L_m$ (resp., $C_1, \ldots, C_n$) les lignes (resp., les colonnes) de A. On peut agir sur A par les opérations suivantes.

(1.4.1) Opérations élémentaires sur les lignes :

- Échanger deux lignes : $L_i \longleftrightarrow L_j$.
- Remplacer L_i par $L_i + \lambda L_j$ $(\lambda \in K)$.
- Multiplier L_i par un scalaire non nul $\lambda \in K$, $\lambda \neq 0$.

(1.4.2) Opérations élémentaires sur les colonnes :

- Échanger deux colonnes : $C_i \longleftrightarrow C_j$.
- Remplacer C_i par $C_i + \lambda C_j$ $(\lambda \in K)$.
- Multiplier C_i par un scalaire non nul $\lambda \in K$, $\lambda \neq 0$.

(1.4.3) Propriétés de base : (0) Les opérations élémentaires sont inversibles; l'inverse d'une opération élémentaire sur les lignes (resp., sur les colonnes) est une opération élémentaire sur les lignes (resp., sur les colonnes).

(1) Les opérations élémentaires conservent le rang (voir 1.4.6 ci-dessous).

(2) En utilisant une succession d'opérations élémentaires, on peut transformer une matrice quelconque $A \in M_{m,n}$ à une matrice A' de la forme

$$A' = \begin{pmatrix} d_1 & * & \cdots & * & * & \cdots & * \\ 0 & d_2 & \cdots & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & * & \cdots & * \\ 0 & 0 & \cdots & d_r & * & \cdots & * \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (r \geq 0;\ d_1, \ldots, d_r \neq 0)$$

(où tous les éléments en-dessous de la diagonale et en-dessous de la r-ième ligne sont nuls).

(3) Plus précisement, une autre suite d'opérations élémentaires transforme A' à la matrice suivante

$$A'' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

(4) Les propriétés (1)–(3) entraînent

$$\mathrm{rg}(A) = \mathrm{rg}(A') = \mathrm{rg}(A'') = r.$$

(1.4.4) Exemple ($K = \mathbf{R}$): le calcul

$$A = \begin{pmatrix} 3 & 2 & -1 & 2 \\ 1 & -2 & 5 & 6 \\ 5 & 1 & 3 & 8 \end{pmatrix} \xrightarrow{[L_1 \longleftrightarrow L_2]} \begin{pmatrix} 1 & -2 & 5 & 6 \\ 3 & 2 & -1 & 2 \\ 5 & 1 & 3 & 8 \end{pmatrix} \xrightarrow{[L_2 := L_2 - 3L_1]} \begin{pmatrix} 1 & -2 & 5 & 6 \\ 0 & 8 & -16 & -16 \\ 5 & 1 & 3 & 8 \end{pmatrix} \xrightarrow{[L_2 := L_2/8]}$$

$$\xrightarrow{[L_2 := L_2/8]} \begin{pmatrix} 1 & -2 & 5 & 6 \\ 0 & 1 & -2 & -2 \\ 5 & 1 & 3 & 8 \end{pmatrix} \xrightarrow{[L_3 := L_3 - 5L_1]} \begin{pmatrix} 1 & -2 & 5 & 6 \\ 0 & 1 & -2 & -2 \\ 0 & 11 & -22 & -22 \end{pmatrix} \xrightarrow{[L_3 := L_3 - 11L_2]}$$

$$\xrightarrow{[L_3 := L_3 - 11L_2]} \begin{pmatrix} 1 & -2 & 5 & 6 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = A'$$

montre que l'on a $\mathrm{rg}(A) = \mathrm{rg}(A') = 2$. On peut continuer, cette fois-ci en utilisant des opérations sur les colonnes :

$$\begin{pmatrix} 1 & -2 & 5 & 6 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{[C_2 := C_2 + 2C_1]} \begin{pmatrix} 1 & 0 & 5 & 6 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{[C_3 := C_3 - 5C_1]} \begin{pmatrix} 1 & 0 & 0 & 6 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{[C_4 := C_4 - 6C_1]}$$

$$\xrightarrow{[C_4 := C_4 - 6C_1]} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{[C_3 := C_3 + 2C_2]} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{[C_4 := C_4 + 2C_2]} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = A''$$

(1.4.5) Formulation matricielle . (1) Les opérations élémentaires sur les lignes (exemple : le cas d'une matrice à deux lignes) :

$$\begin{pmatrix} L_1 + 3L_2 \\ L_2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}, \qquad \begin{pmatrix} \lambda L_1 \\ L_2 \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}, \qquad \begin{pmatrix} L_2 \\ L_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}.$$

(2) Les opérations élémentaires sur les colones (exemple : le cas d'une matrice à trois colonnes) :

$$(C_1 \mid C_2 \mid C_3 + 2C_1) = (C_1 \mid C_2 \mid C_3) \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$(C_1 \mid \lambda C_2 \mid C_3) = (C_1 \mid C_2 \mid C_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (C_2 \mid C_1 \mid C_3 + 2C_1) = (C_1 \mid C_2 \mid C_3) \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(3) Les deux exemples précédents se généralisent facilement au cas général. On en déduit qu'une succession d'opérations élémentaires transforme une matrice $A \in M_{m,n}(K)$ à une matrice

$$gAh,$$

où $g \in M_m(K)$ (resp., $h \in M_n(K)$) est une matrice inversible qui représent une suite d'opérations élémentaires sur les lignes (resp., sur les colonnes).

(1.4.6) Proposition. *Les opérations élémentaires conservent le rang.*

Démonstration. En utilisant la notation de 1.4.5(3), il faut montrer que les deux rangs

$$\mathrm{rg}(A) = \dim(\mathrm{vect}(\text{les colonnes de } A))$$
$$\mathrm{rg}(gAh) = \dim(\mathrm{vect}(\text{les colonnes de } gAh))$$

coïncident. Les opérations élémentaires sur les colonnes ne changent pas l'espace (vect(les colonnes de A)), donc $\mathrm{rg}(gAh) = \mathrm{rg}(gA)$. D'autre part, la multiplication à gauche par g définit une application linéaire

$$g : \mathrm{vect}(\text{les colonnes de } A) \longrightarrow \mathrm{vect}(\text{les colonnes de } gA)$$

qui est bijective (son inverse est donné par la multiplication à gauche par g^{-1}); il en résulte que les deux espaces ont la même dimension, c'est-à-dire que $\mathrm{rg}(A) = \mathrm{rg}(gA)$.

(1.4.7) La matrice transposée. Rapellons que la **transposée** d'une matrice $A \in M_{m,n}(K)$ est la matrice ${}^tA \in M_{n,m}(K)$ de coefficients

$$({}^tA)_{ij} = A_{ji}.$$

Autrement dit, les lignes (resp., les colonnes) de A correspondent aux colonnes (resp., aux lignes) de ${}^tA$. Par exemple,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad {}^tA = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Attention : la transposée renverse l'ordre du produit :

$${}^t(AB) = {}^tB\,{}^tA.$$

(1.4.8) Proposition. *Pour toute matrice $A \in M_{m,n}(K)$, on a $\mathrm{rg}(A) = \mathrm{rg}({}^tA)$. Autrement dit, les dimensions des deux espaces suivants coïncident :*

$$\mathrm{vect}(\text{les colonnes de } A) \subset K^m$$
$$\mathrm{vect}(\text{les colonnes de } {}^tA) = {}^t(\mathrm{vect}(\text{les lignes de } A)) \subset K^n.$$

Démonstration. Il existe une suite d'opérations élémentaires qui transforme A à

$$B = gAh = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(K).$$

Lorsqu'on fait agir la transposée, on obtient une suite d'opérations élémentaires qui transforme ${}^tA$ à

$${}^tB = {}^th\,{}^tA\,{}^tg = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_{n,m}(K).$$

Il est clair que l'on a $\mathrm{rg}(B) = \mathrm{rg}({}^tB) = r$. D'autre part, 1.4.6 entraîne $\mathrm{rg}(A) = \mathrm{rg}(B)$ et $\mathrm{rg}({}^tA) = \mathrm{rg}({}^tB)$, d'où $\mathrm{rg}(A) = r = \mathrm{rg}({}^tA)$.

(1.4.9) Détermination de l'inverse. Si le rang d'une matrice carrée $A \in M_n(K)$ est égal à $\mathrm{rg}(A) = n$ (= le rang maximum), alors on peut transformer A à la matrice $A'' = I_n$ en n'utilisant qu'une suite d'opérations élémentaires **sur les lignes**. En exprimant cette suite d'opérations en termes de la multiplication à gauche par une matrice (inversible) $B \in M_n(K)$, l'équalité $BA = I_n$ entraîne $B = A^{-1}$. Il en résulte que la même suite d'opérations élémentaires transforme la matrice identité I_n à la matrice $BI_n = B = A^{-1}$. En pratique, on met les matrice A et I_n côté à côté et l'on fait agir les opérations à la fois sur les lignes des deux matrices.

Exemple : Déterminer l'inverse de la matrice

$$A = \begin{pmatrix} 4 & 2 & 3 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

$$\left(\begin{array}{ccc|ccc} 4 & 2 & 3 & 1 & 0 & 0 \\ 6 & 5 & 4 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{[L_1:=L_1-L_3]} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & -1 \\ 6 & 5 & 4 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{[L_2:=L_2-2L_3]} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{[L_3:=L_3-3L_1]} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 2 & -5 & -3 & 0 & 4 \end{array}\right) \xrightarrow{[L_3:=L_3-2L_2]} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 0 & -9 & -3 & -2 & 8 \end{array}\right) \xrightarrow{[L_3:=-L_3/9]}$$

$$\xrightarrow{[L_3:=-L_3/9]} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 0 & 1 & 1/3 & 2/9 & -8/9 \end{array}\right) \xrightarrow{[L_1:=L_1-2L_3]} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & -4/9 & 7/9 \\ 0 & 1 & 2 & 0 & 1 & -2 \\ 0 & 0 & 1 & 1/3 & 2/9 & -8/9 \end{array}\right) \xrightarrow{[L_2:=L_2-2L_3]}$$

$$\xrightarrow{[L_2:=L_2-2L_3]} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/3 & -4/9 & 7/9 \\ 0 & 1 & 0 & -2/3 & 5/9 & -2/9 \\ 0 & 0 & 1 & 1/3 & 2/9 & -8/9 \end{array}\right)$$

Autrement dit, on a

$$A^{-1} = \begin{pmatrix} 1/3 & -4/9 & 7/9 \\ -2/3 & 5/9 & -2/9 \\ 1/3 & 2/9 & -8/9 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 3 & -4 & 7 \\ -6 & 5 & -2 \\ 3 & 2 & -8 \end{pmatrix}.$$

Pour vérifier le résultat, on calcule

$$\begin{pmatrix} 4 & 2 & 3 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & -4 & 7 \\ -6 & 5 & -2 \\ 3 & 2 & -8 \end{pmatrix} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = 9I_3.$$

Attention : Si le calcul échoue, c'est-à-dire si on obtient une colonne nulle, on conclu que $\mathrm{rg}(A) < n$, donc la matrice A n'est pas iversible.

Petits dimensions : Lorsque $n = 2$ on peut écrire directement

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{1.4.9.1}$$

(à condition que $ad - bc \neq 0$, ce qui entraîne que la matrice est inversible).

1.5 Déterminants

(1.5.1) Rappelons que le **déterminant** d'une **matrice carrée** $A \in M_n(K)$ est le scalaire $\det(A) = |A| \in K$ défini par récurrence de la façon suivante :

(1) Lorsque $n = 1$, on a $A = (a_{11})$ et $\det(A) = a_{11}$.
(2) Lorsque $n = 2$, on a

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

(3) Plus généralement, pour tout $n \geq 2$ on peut développer le déterminant $\det(A)$ par rapport à chaque ligne (resp., chaque colonne). Par exemple, en développant le déterminant suivant par rapport à la seconde ligne, on obtient

$$\begin{vmatrix} 1 & 4 & -2 \\ 7 & 3 & 5 \\ 1 & -3 & 2 \end{vmatrix} = -7 \begin{vmatrix} 4 & -2 \\ -3 & 2 \end{vmatrix} + 3 \begin{vmatrix} 1 & -2 \\ 1 & 2 \end{vmatrix} - 5 \begin{vmatrix} 1 & 4 \\ 1 & -3 \end{vmatrix} = -7 \cdot 2 + 3 \cdot 4 - 5(-7) = 33.$$

Voici la règle générale : si l'on pose (pour tous $1 \leq i, j \leq n$)

$$M_{ij} = (-1)^{i+j} \det(\text{la matrice obtenue de } A \text{ en enlèvant la } i\text{−ième ligne et la } j\text{−ième colonne}),$$

alors on a, pour tout $k = 1, \ldots, n$,

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} a_{kj}\, M_{kj} \qquad (\text{développement par rapport à la } k\text{−ième ligne}) \\ &= \sum_{i=1}^{n} a_{ik}\, M_{ik} \qquad (\text{développement par rapport à la } k\text{−ième colonne}) \end{aligned}$$

(4) En particulier, si

$$A = \begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

est une matrice triangulaire supérieure ($\Longleftrightarrow$ tous les éléments en-dessous de la diagonale sont nuls), alors le développement par rapport à la première colonne montre, par récurrence, que le déterminant de A

$$\begin{vmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{vmatrix} = \cdots = a_{11}a_{22}\cdots a_{nn}$$

est égal au produit des éléments diagonaux de A; de même pour les matrices triangulaires inférieures (dont tous les éléments en-dessus de la diagonale sont nuls).

(1.5.2) Propriétés des déterminants

(1) $\det({}^tA) = \det(A)$.
(2) $\det(AB) = \det(A)\det(B)$.
(3) On peut reformuler les règles de développement (voir 1.5.1(3)) en termes de la matrice $M = (M_{ij}) \in M_n(K)$ de la façon suivante :

$${}^tM\, A = \det(A) I_n.$$

En particulier, si $\det(A) \neq 0$, alors la matrice A est inversible et son inverse est égal à la matrice

$$A^{-1} = \frac{1}{\det(A)}\, {}^tM$$

(ce qui généralise la formule (1.4.9.1)). Cependant, cette méthode n'est pas utile pour la détermination de l'inverse en pratique.

(4) A est inversible $\iff$ $\det(A) \neq 0$ (on déduit les implications respectives "$\Longrightarrow$" et "$\Longleftarrow$" de (2) et (3)).

(1.5.3) Calcul du déterminant en utilisant des opérations élémentaires

Les opérations élémentaires transforment le déterminant selon les règles suivantes :

(1) En remplacant L_i (resp., C_i) par $L_i + \lambda L_j$ (resp., by $C_i + \lambda C_j$) on **ne change pas le déterminant.**
(2) En multipliant L_i (resp., C_i) par $\lambda \in K$, on **multiplie le déterminant par λ.**
(3) En échangeant deux lignes (resp., deux colonnes), on **multiplie le déterminant par -1.**

En utilisant les règles (1)–(3), il suffit qu'on transforme A à une matrice triangulaire, dont le déterminant est donné par 1.5.1(4). Par exemple, on calcule

$$\begin{vmatrix} 1 & 4 & -2 \\ 7 & 3 & 5 \\ 1 & -3 & 2 \end{vmatrix} \overset{[L_1:=L_1+L_3]}{=\!=\!=\!=} \begin{vmatrix} 2 & 1 & 0 \\ 7 & 3 & 5 \\ 1 & -3 & 2 \end{vmatrix} \overset{[C_1:=C_1-2C_2]}{=\!=\!=\!=} \begin{vmatrix} 0 & 1 & 0 \\ 1 & 3 & 5 \\ 7 & -3 & 2 \end{vmatrix}.$$

À ce point, on peut soit développer le déterminant par rapport à la première ligne :

$$\begin{vmatrix} 0 & 1 & 0 \\ 1 & 3 & 5 \\ 7 & -3 & 2 \end{vmatrix} = -1 \begin{vmatrix} 1 & 5 \\ 7 & 2 \end{vmatrix} = -(1 \cdot 2 - 5 \cdot 7) = 33,$$

soit effectuer encore des opérations élémentaires :

$$\begin{vmatrix} 0 & 1 & 0 \\ 1 & 3 & 5 \\ 7 & -3 & 2 \end{vmatrix} \overset{[C_1 \longleftrightarrow C_2]}{=\!=\!=\!=} - \begin{vmatrix} 1 & 0 & 0 \\ 3 & 1 & 5 \\ -3 & 7 & 2 \end{vmatrix} \overset{[C_3:=C_3-5C_2]}{=\!=\!=\!=} \begin{vmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -3 & 7 & -33 \end{vmatrix} = 33.$$

1.6 Vecteurs propres, valeurs propres; diagonalisation d'un endomorphisme

(1.6.1) Définition. *Soit $f : E \longrightarrow E$ un endomorphisme d'un K-ev E. On dit qu'un scalaire $\lambda \in K$ est une* **valeur propre de** f *s'il existe un vecteur non nul $x \in E$ tel que $f(x) = \lambda x$ (on dit que x est un* **vecteur propre de** f **de valeur propre** λ*). Si $E = K^n$ et $f(x) = Ax$ (où $A \in M_n(K)$), on parle de valeurs propres et de vecteurs propres de la matrice A.*

(1.6.2) Exemples : (1) Si $f : K^2 \longrightarrow K^2$ est défini par la formule $f(X) = AX$, où

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix},$$

alors on a

$$Ae_1 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2e_1, \qquad Ae_2 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 3e_2,$$

ce qui signifie que les vecteurs e_1 et e_2 sont des vecteurs propres de f, de valeurs propres respectives 2 et 3.
(2) Plus généralement, si $E = K^n$ et $f(X) = AX$, où la matrice A est diagonale

$$A = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

alors tous les éléments $e_1, \ldots, e_n$ de la base canonique de K^n sont des vecteurs propres de f, car $f(e_i) = Ae_i = d_i e_i$. En particulier, si $d_1 = \cdots = d_n = d$ ($\iff A = dI_n$ est un multiple de la matrice identité), alors on a $f(x) = dx$ pour tout $x \in K^n$, donc tout vecteur non nul est un vecteur propre de f, de valeur propre d.

(1.6.3) La condition $f(x) = \lambda x$ équivaut à

$$0 = f(x) - \lambda x = (f - \lambda \operatorname{Id}_E)(x),$$

donc

$$\lambda \text{ est une valeur propre de } f \iff \operatorname{Ker}(f - \lambda \operatorname{Id}_E) \neq 0.$$

Si c'est le cas, alors les vecteurs propres de f de valeur propre λ (+ le vecteur nul) forment le sous-espace

$$\operatorname{Ker}(f - \lambda \operatorname{Id}_E) \subset E,$$

que l'on appelle **l'espace propre de f de valeur propre λ.**

(1.6.4) Formulation matricielle. Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ ($n = \dim(E)$) une base de E; on note $A = M_{\mathcal{B},\mathcal{B}}(f) \in M_n(K)$ la matrice de f dans la base $\mathcal{B}$. Si l'on identifie

$$\begin{aligned} M_{\mathcal{B}} : E &\xrightarrow{\sim} K^n \\ x &\mapsto X, \end{aligned}$$

l'application f s'écrit $X \mapsto AX$, donc

$$\lambda \text{ est une valeur propre de } f \iff \operatorname{Ker}(A - \lambda I_n) \neq 0 \iff P_A(\lambda) = 0,$$

où

$$P_A(t) = \det(t I_n - A)$$

est le **polynôme caractéristique de** A. Autrement dit, les valeurs propres de A (= de f) et les racines de P_A coïncident. Si l'on remplace $\mathcal{B}$ par une autre base, la matrice A sera remplacée par $A' = P^{-1}AP$, où $P \in M_n(K)$ est une matrice inversible (voir (1.3.3.2)). La multiplicativité du déterminant entraîne

$$P_{A'}(t) = \det(t I_n - A') = \det(P^{-1}(t I_n - A)P) = \det(P)^{-1} \det(t I_n - A) \det(P) = \det(t I_n - A) = P_A(t),$$

ce qui signifie que le polynôme caractéristique est indépendant du choix de base; on peut écrire $P_f(t)$ au lieu de $P_A(t)$.

(1.6.5) Exemples : (1) $n = 2$: Le polynôme caractéristique de la matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

est égal à

$$P_A(t) = \begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix} = (t-a)(t-d) - (-b)(-c) = t^2 - (a+d)t + (ad-bc) = t^2 - \operatorname{Tr}(A)t + \det(A).$$

(2) En général, on a

$$P_A(t) = t^n - \mathrm{Tr}(A)t^{n-1} + \cdots + (-1)^n \det(A);$$

rappelons que la **trace** d'une matrix carrée $A \in M_n(K)$ est la somme de ses éléments diagonaux : $\mathrm{Tr}(A) = A_{11} + A_{22} + \cdots + A_{nn}$.
(3) Si A est une matrice triangulaire

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{pmatrix},$$

alors on a

$$P_A(t) = (t - A_{11}) \cdots (t - A_{nn}),$$

donc les valeurs propres de A sont précisement les éléments diagonaux $A_{11}, \ldots, A_{nn}$ de A.
(1.6.6) Endomorphismes diagonalisables. Soit $f : E \longrightarrow E$ un endomorphisme d'un K-ev E de dimension $\dim(E) = n$. On choisit une base $\mathcal{B}$ de E; soit $A \in M_n(K)$ la matrice (carrée) qui représent f dans la base $\mathcal{B}$.

Definition. *On dit que l'endomorphisme f (ou la matrice A) est* **diagonalisable (sur** K**)** *s'il existe une base $\mathcal{B}'$ de E dans laquelle la matrice A' de f ait diagonale :*

$$A' = D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n. \end{pmatrix}$$

Comme $A' = P^{-1}AP$, où $P \in M_n(K)$ est la matrice de passage (inversible) de $\mathcal{B}$ à $\mathcal{B}'$, on a

$$A \text{ est diagonalisable sur } K \iff \exists P \in M_n(K) \text{ inversible telle que } P^{-1}AP = D \text{ soit diagonale}$$

(le symbole "$\exists$" signifie **"il existe"**). On va reformuler cette condition en termes des colonnes $v_1, \ldots, v_n$ de la matrice $P = (v_1 \mid \cdots \mid v_n)$. D'abord, on a

$$P \text{ est inversible} \iff \text{les vecteurs } v_1, \ldots, v_n \text{ sont linéairement indépendants.}$$

Ensuite, la relation $P^{-1}AP = D$ équivaut $AP = PD$, puisque la matrice P est inversible. Les formules

$$AP = (Av_1 \mid \cdots \mid Av_n), \qquad PD = (d_1v_1 \mid \cdots \mid d_nv_n),$$

entraînent

$$AP = PD \iff Av_1 = d_1v_1, \ldots, Av_n = d_nv_n.$$

En résumé,

$$A \text{ est diagonalisable sur } K \iff \exists n \text{ vecteurs propres linéairement indépendants } v_1, \ldots, v_n \in K^n \text{ de } A. \tag{1.6.6.1}$$

Si c'est le cas, la matrice $P = (v_1 \mid \cdots \mid v_n) \in M_n(K)$ dont les colonnes sont égales aux vecteurs propres $v_1, \dots, v_n$ est inversible et la matrice $P^{-1}AP$ est diagonale (ses éléments diagonaux sont égaux aux valeurs propres respectives des vecteurs propres $v_1, \dots, v_n$).

Il peut arriver que la matrice A ne soit pas diagonalisable sur K, mais devienne diagonalisable sur un corps K' plus grand.

(1.6.7) Exemples : (1) On considère la matrice

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbf{R}).$$

Le polynôme caractéristique $P_A(t) = t^2 - \mathrm{Tr}(A)t + \det(A) = t^2 + 1 = (t-i)(t+i)$ n'a aucune racine réelle, donc A n'a aucun vecteur propre $x \in \mathbf{R}^2$. En particulier, A **n'est pas diagonalisable sur R**. D'autre part, les vecteurs propres complexes de A

$$v_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix} \in \mathrm{Ker}(A - iI), \qquad v_2 = \begin{pmatrix} 1 \\ i \end{pmatrix} \in \mathrm{Ker}(A + iI)$$

sont linéairement indépendants (en tant que des éléments du **C**-espace vectoriel $\mathbf{C}^2$), donc A **est diagonalisable sur C :**

$$\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}^{-1} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

(2) Le polynôme caractéristique de la matrice

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(K)$$

est égal à $P_B(t) = t^2$, donc il n'y a qu'une valeur propre $\lambda = 0$ de B. Par définition, le sous-espace propre associé $\mathrm{Ker}(B - \lambda I) = \mathrm{Ker}(B)$ est l'espace des solutions du système linéaire homogène

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

Cet espace est engendré par le vecteur $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, ce qui entraîne qu'il n'y a pas de deux vecteurs propres linéairement indépendants, donc la matrice B **n'est pas diagonalisable sur** K (le corps K étant quelconque).

(3) Plus généralement, si le polynôme caractéristique d'une matrice $A \in M_n(K)$ n'a qu'une racine à multiplicité n, c'est-à-dire si $P_A(t) = (t - \lambda)^n$ $(\lambda \in K)$, alors on a :

$$A \text{ est diagonalisable sur } K \iff A = \lambda\, I_n.$$

En effet, si A est diagonalisable (sur K), alors tous les éléments diagonaux de $P^{-1}AP$ sont des valeurs propres de A, donc égaux à λ : $P^{-1}AP = \lambda\, I_n$, d'où $A = P\lambda\, I_n P^{-1} = \lambda\, I_n$.

(1.6.8) Critères généraux de diagonalisabilité. Rappelons deux résultats fondamentaux ($A \in M_n(K)$):

(a) Tout ensemble de vecteurs propres de A dont les valeurs propres sont distinctes est linéairement indépendant.

(b) Pour toute valeur propre $\lambda \in K$ de A, la dimension de l'espace propre $\mathrm{Ker}(A - \lambda\, I_n)$ (la **multiplicité géométrique de** λ) est au plus égale à la multiplicité de λ en tant qu'une racine de $P_A(t)$ (la **multiplicité algébrique de** λ).

En particulier, on déduit de (a) le critère suivant :

Critère 1. *Si la matrice $A \in M_n(K)$ a n valeurs propres distinctes, toutes contenues dans K, alors A est diagonalisable sur K.*

Si le critère précédent ne s'applique pas, à savoir si le polynôme caractéristique $P_A(t)$ admet des racines multiples, il faut utiliser le critère suivant, qui résulte de (b) :

Criterion 2. *Une matrice $A \in M_n(K)$ est diagonalisable sur $K \iff$ toutes les racines de P_A sont contenues dans K et, pour toute racine λ de P_A, la multiplicité algébrique et géométrique de λ coïncident.*

(1.6.9) Exercice. (1) *Lesquelles, parmi les matrices suivantes, sont diagonalisables sur* **R** *(resp., sur* **C***) :*

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & -2 \\ 2 & -2 \end{pmatrix}, \quad \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

(2) *Déterminer toutes les valeurs $t \in \mathbf{R}$ pour lesquelles la matrice $\begin{pmatrix} -1 & t \\ 1 & 3 \end{pmatrix}$ soit diagonalisable sur* **R** *(resp., sur* **C***).*
(3) *Lesquelles, parmi les matrices suivantes, sont diagonalisables sur* **R** *(resp., sur* **C***) :*

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

(4) *Construire une matrice $A \in M_2(\mathbf{C})$ symétrique ($A = {}^tA$) qui ne soit pas diagonalisable sur* **C**.

1.7 L'espace dual *

(1.7.1) Qu'est-ce un vecteur ligne ? Si l'on choisit une base d'un K-ev E de dimension finie, alors on identifie les éléments de E aux **vecteurs colonnes**. À quoi correspondent les **vecteurs lignes** ?
Un vecteur ligne est un élément de l'espace

$$M_{1,n}(K) = \mathscr{L}(K^n, K) = \{\text{linear maps } K^n \longrightarrow K\},$$

ce qui nous conduit à la définition suivante.

(1.7.2) Définition. *Le* **dual** *d'un K-ev E est le K-ev*

$$E^* = \mathscr{L}(E, K) = \{\text{applications linéaires } E \longrightarrow K\};$$

on appelle un élément de E^* une **forme linéaire** *sur E.*

(1.7.3) Exemples : (1) Si $E = K^n = M_{n,1}(K)$, alors les éléments de

$$E^* = (K^n)^* = \mathscr{L}(K^n, K) = M_{1,n}(K) = \{(a_1, \ldots, a_n) \mid a_i \in K\}$$

sont les vecteurs lignes. La forme linéaire $f_a : K^n \longrightarrow K$ qui correspond à un vecteur ligne $a = (a_1, \ldots, a_n)$ s'écrit

$$f_a : x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto (a_1, \ldots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1 + \cdots + a_nx_n.$$

* **Ce paragraph est facultatif.**

(2) Par exemple, $(3,-2) \in (\mathbf{R}^2)^*$ est la forme linéaire

$$f(x) = f(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = (3,-2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 - 2x_2.$$

(3) Lorsque $\dim(E) < \infty$, la formule 1.2.8(2) entraîne $\dim(E^*) = \dim(E)\dim(K) = \dim(E)$.

(1.7.4) Base duale. On suppose que $n = \dim(E) < \infty$. Étant donné une base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de E, tout vecteur $x \in E$ s'écrit d'une manière unique

$$x = x_1 v_1 + \cdots + x_n v_n \qquad\qquad (x_i \in K).$$

Pour tout indice $i = 1, \ldots, n$, la fonction qui associe à x sa i-ième coordonnée dans la base $\mathcal{B}$ définit une forme linéaire $E \longrightarrow K$; on la note

$$\begin{aligned} v_i^* : E &\longrightarrow K \\ x &\mapsto x_i. \end{aligned}$$

Les **"formes coordonnées"** $v_1^*, \ldots, v_n^* \in E^*$ sont des éléments de l'espace dual; elles sont caractérisées par la propriété suivante :

$$v_i^*(v_j) = v_i(0 \cdot v_1 + \cdots + 1 \cdot v_j + \cdots + 0 \cdot v_n) = \delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j, \end{cases} \qquad (1.7.4.1)$$

ce qui entraîne que les éléments $v_1^*, \ldots, v_n^*$ sont linéairement indépendants; comme $\dim(E^*) = n$, ils forment une base $\mathcal{B}^*$ de E^*. On dit que $\mathcal{B}^*$ est **la base duale** de $\mathcal{B}$.

Exemple : Si $E = K^n$ et $\mathcal{B} = \{e_1, \ldots, e_n\}$ est la base canonique de K^n, alors on a

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \cdots + x_n e_n, \qquad x_i = (0, \ldots, 1, \ldots, 0)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = e_i^*(x),$$

donc la base duale $\mathcal{B}^*$ de $(K^n)^*$ contient les vecteurs lignes

$$e_1^* = (1, 0, \ldots, 0), \qquad e_2^* = (0, 1, 0, \ldots, 0), \qquad \cdots \qquad e_n^* = (0, \ldots, 0, 1).$$

(1.7.5) Détermination de la base duale. Étant donné une base $v_1, \ldots, v_n$ de K^n, on cherche à trouver la base duale $v_1^*, \ldots, v_n^*$ of $(K^n)^*$.

Si l'on condidère les matrices

$$A = (v_1 \mid \cdots \mid v_n) \in M_n(K), \qquad B = \begin{pmatrix} v_1^* \\ \vdots \\ v_n^* \end{pmatrix} \in M_n(K),$$

alors la formule (1.7.4.1) entraîne

$$BA = \begin{pmatrix} v_1^*(v_1) & \cdots & v_1^*(v_n) \\ \vdots & \ddots & \vdots \\ v_n^*(v_1) & \cdots & v_n^*(v_n) \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} = I_n,$$

donc les éléments $v_1^*, \ldots, v_n^*$ sont les lignes de la matrice $B = A^{-1}$.

Exemple : Trouver la base duale $v_1^*, v_2^* \in (\mathbf{R}^2)^*$ de la base

$$v_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 8 \\ 5 \end{pmatrix}.$$

Le calcul

$$A = \begin{pmatrix} 3 & 8 \\ 2 & 5 \end{pmatrix}, \qquad A^{-1} = \frac{1}{3 \cdot 5 - 8 \cdot 2} \begin{pmatrix} 5 & -8 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} -5 & 8 \\ 2 & -3 \end{pmatrix},$$

montre que l'on a $v_1^* = (-5, 8)$ and $v_2^* = (2, -3)$.

(1.7.6) À quoi sert la base duale ? La connaissance de la base duale nous permet de calculer les coordonnées d'un vecteur quelconque $x \in E$ dans la base $\mathcal{B} = \{v_1, \dots, v_n\}$:

$$x = v_1^*(x)\, v_1 + \cdots + v_n^*(x)\, v_n.$$

L'exemple précédent montre que l'on a

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (-5x_1 + 8x_2) \begin{pmatrix} 3 \\ 2 \end{pmatrix} + (2x_1 - 3x_2) \begin{pmatrix} 8 \\ 5 \end{pmatrix}.$$

(1.7.7) L'application duale. Soit $\alpha : E \longrightarrow F$ une application linéaire entre deux K-ev. **L'application duale** $\alpha^* : F^* \longrightarrow E^*$ est définie par la formule

$$\alpha^*(\lambda) = \lambda \circ \alpha : \, E \xrightarrow{\alpha} F \xrightarrow{\lambda} K.$$

Formulation matricielle : on suppose que $n = \dim(E)$, $p = \dim(F)$. Étant donné une base $\mathcal{B} = \{e_1, \dots, e_n\}$ (resp., $\mathcal{C} = \{f_1, \dots, f_p\}$) de E (resp., de F), on note $\mathcal{B}^* = \{e_1^*, \dots, e_n^*\}$ (resp., $\mathcal{C}^* = \{f_1^*, \dots, f_p^*\}$) la base duale de E^* (resp., de F^*). Soient

$$\begin{aligned} A &= M_{\mathcal{B},\mathcal{C}}(\alpha) = M_{\mathcal{C}}(\alpha(\mathcal{B})) \in M_{p,n}(K) \\ M &= M_{\mathcal{C}^*,\mathcal{B}^*}(\alpha^*) = M_{\mathcal{B}^*}(\alpha^*(\mathcal{C}^*)) \in M_{n,p}(K) \end{aligned}$$

les matrices représentatives de α et α^*, respectivement. Par définition, on a

$$\begin{aligned} A_{ij} &= (\text{la } i\text{−ième coordonée de } \alpha(e_j) \text{ dans la base } \mathcal{C}) = f_i^*(\alpha(e_j)) = (f_i^* \circ \alpha)(e_j) \qquad (1.7.7.1) \\ M_{ji} &= (\text{la } j\text{−ième coordonée de } \alpha^*(f_i^*) \text{ dans la base } \mathcal{B}^*) = (\alpha^*(f_i^*))(e_j). \end{aligned}$$

Comme

$$\alpha^*(f_i^*) = f_i^* \circ \alpha,$$

on en déduit que $A_{ij} = M_{ji}$ (pour tous i, j), donc **l'application duale est représentée (dans les bases duales) par la matrice transposée :**

$$M = {}^t A.$$

(1.7.8) Résumé : On peut ajouter quelques lignes au dictionnaire que l'on a considéré en début de ce chapitre :

Objet abstrait	Objet matriciel
espace vectoriel E	K^n
vecteur $x \in E$	vecteur colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
application linéaire $\alpha : E \longrightarrow F$	matrice $A \in M_{p,n}(K)$
$\mathscr{L}(E, F)$	$M_{p,n}(K)$
forme linéaire $f \in E^*$	vecteur ligne $(a_1, \dots, a_n)$
application duale $\alpha^* : F^* \longrightarrow E^*$	matrice transposée ${}^tA \in M_{n,p}(K)$

(1.7.9) Le bidual: l'évaluation en un vecteur $x \in E$ définit une forme linéaire sur E^*, donc un élément ev_x de E^{**}:

$$\begin{aligned} \mathrm{ev}_x : E^* &\longrightarrow K \\ \lambda &\mapsto \lambda(x). \end{aligned}$$

De plus, l'application

$$\begin{aligned} \mathrm{ev} : E &\longrightarrow E^{**} \\ x &\mapsto \mathrm{ev}_x \end{aligned}$$

est linéaire et injective. Si $n = \dim(E) < \infty$, alors ev est un isomorphisme, puisque $\dim(E^{**}) = \dim(E^*) = \dim(E)$. Si $\mathcal{B} = \{e_1, \dots, e_n\}$ est une base de E et $\mathcal{B}^* = \{e_1^*, \dots, e_n^*\}$ est la base duale de E^*, alors la base duale $\mathcal{B}^{**}$ de $\mathcal{B}^*$ est égale à $\mathrm{ev}(\mathcal{B}) = \{\mathrm{ev}(e_1), \dots, \mathrm{ev}(e_n)\}$. Dans ce cas il est commode d'identifier E à son bidual E^{**} (et la base $\mathcal{B}$ à $\mathcal{B}^{**}$) en utilisant l'isomorphisme ev.

2. Applications multilinéaires

2.1 Exemples

(2.1.1) Produit scalaire euclidien. Le produit scalaire euclidien usuel

$$(x \mid y) = x_1 y_1 + \cdots + x_n y_n \qquad (x, y \in \mathbf{R}^n)$$

est une fonction **bilinéaire** en les variables x et y, ce qui signifie qu'il est

- **linéaire en y, si l'on fixe x :**

$$\begin{aligned} (x \mid y + y') &= (x \mid y) + (x \mid y') \\ (x \mid \lambda y) &= \lambda\,(x \mid y) \end{aligned} \qquad (x, y, y' \in \mathbf{R}^n,\ \lambda \in \mathbf{R})$$

- **linéaire en x, si l'on fixe y :**

$$\begin{aligned} (x + x' \mid y) &= (x \mid y) + (x' \mid y) \\ (\lambda x \mid y) &= \lambda\,(x \mid y) \end{aligned} \qquad (x, x', y \in \mathbf{R}^n,\ \lambda \in \mathbf{R}).$$

(2.1.2) Définition. *Soient E, F des K-ev. Une application $f : \underbrace{E \times \cdots \times E}_{p-\text{fois}} \longrightarrow F$ (autrement dit, une fonction $f(x_1, \ldots, x_p)$ en p variables $x_i \in E$ à valeurs dans F) est dite* p**-linéaire** *(bilinéaire lorsque $p = 2$) si, chaque fois que l'on fixe toutes les variables sauf une, f est une application linéaire $E \longrightarrow F$ en la variable restée libre. Si $F = K$, on dit que f est une* **forme** p**-linéaire** *(une* **forme bilinéaire** *si $p = 2$).*

(2.1.3) Exemples : (1) Le **produit scalaire euclidien usuel** est une forme bilinéaire

$$f : \mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R}, \qquad f(x, y) = (x \mid y) = x_1 y_1 + \cdots + x_n y_n.$$

(2) Le **produit vectoriel** est une application bilinéaire

$$f : \mathbf{R}^3 \times \mathbf{R}^3 \longrightarrow \mathbf{R}^3, \quad f(x, y) = x \times y = z, \quad z_1 = \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix}, \quad z_2 = -\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix}, \quad z_3 = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

(3) Le **déterminant** d'une matrice de type $n \times n$, en tant qu'une fonction des colonnes de la matrice, est une forme n-linéaire :

$$f : \underbrace{K^n \times \cdots \times K^n}_{n-\text{fois}} \longrightarrow K, \qquad f(v_1, \ldots, v_n) = \det(v_1 \mid \cdots \mid v_n).$$

Si $K = \mathbf{R}$, alors $f(v_1, \ldots, v_n)$ est égal au volume algébrique du parallelipipède engendré par les vecteurs $v_1, \ldots, v_n$. Par exemple, lorsque $n = 2$, le déterminant

$$f(x, y) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

est égal, à un signe près, à l'aire du parallelogramme engendré par $x, y \in \mathbf{R}^2$.

2.2 Formes bilinéaires

(2.2.1) En particulier, une **forme bilinéaire sur un K-ev** E est une application $f : E \times E \longrightarrow K$ vérifiant

$$f(x+x',y) = f(x,y)+f(x',y) \qquad f(x,y+y') = f(x,y)+f(x,y') \qquad (x,x',y,y' \in E,\ \lambda \in K).$$
$$f(\lambda x,y) = \lambda\, f(x,y) \qquad f(x,\lambda y) = \lambda\, f(x,y)$$

(2.2.2) Définition. *Une forme bilinéaire $f : E \times E \longrightarrow K$ est dite*

$$\begin{array}{rcll} \textbf{symétrique} & \Longleftrightarrow & \forall x,y \in E & f(y,x) = f(x,y) \\ \textbf{anti}-\textbf{symétrique} & \Longleftrightarrow & \forall x,y \in E & f(y,x) = -f(x,y). \end{array}$$

(2.2.3) Exemples : (1) Le produit scalaire euclidien usuel est symétrique, puisqu'on a $(y \mid x) = (x \mid y)$ pour tous $x, y \in \mathbf{R}^n$.

(2) L'aire algébrique $f : \mathbf{R}^2 \times \mathbf{R}^2 \longrightarrow \mathbf{R}$ (voir 2.1.3(3) pour $n = 2$) est anti-symétrique :

$$f(y,x) = \begin{vmatrix} y_1 & x_1 \\ y_2 & x_2 \end{vmatrix} = -\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = -f(x,y).$$

2.3 Représentation matricielle des formes bilinéaires

(2.3.1) On suppose que $n = \dim(E) < \infty$. Un choix d'une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E nous permet d'identifier E à K^n (voir 1.1.8 ci-dessus). En particulier, soient $x = x_1e_1+\cdots+x_ne_n$ et $y = y_1e_1+\cdots+y_ne_n \in E$ des vecteurs de E; on note

$$X = M_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = M_{\mathcal{B}}(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in K^n$$

les vecteurs colonnes qui représentent x et y dans la base $\mathcal{B}$.

Si $f : E \times E \longrightarrow K$ est une forme bilinéaire, alors on a

$$f(x,y) = f(\sum_{i=1}^n x_ie_i, \sum_{j=1}^n y_je_j) = \sum_{i=1}^n x_i\, f(e_i, \sum_{j=1}^n y_je_j) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j\, f(e_i,e_j) = \sum_{i,j=1}^n f(e_i,e_j)\, x_iy_j. \quad (2.3.1.1)$$

(2.3.2) Définition. *Sous les hypothèses de 2.3.1, on appelle la matrice carrée $A = (A_{ij}) \in M_n(K)$, où $A_{ij} = f(e_i,e_j)$,* ***la matrice de f dans la base $\mathcal{B}$.***

(2.3.3) Écriture matricielle. En utilisant la matrice A, on peut exprimer (2.3.1.1) comme le produit matriciel

$$f(x,y) = \sum_{i,j=1}^n A_{ij}\, x_iy_j = (x_1 \cdots x_n) \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = {}^tXAY. \quad (2.3.3.1)$$

Réciproquement, pour toute matrice $A \in M_n(K)$, la formule (2.3.3.1) (où $X = M_{\mathcal{B}}(x)$ et $Y = M_{\mathcal{B}}(y)$) définit une forme bilinéaire $f : E \times E \longrightarrow K$ vérifiant $f(e_i,e_j) = A_{ij}$.

Résumé : La matrice $A = (f(e_i,e_j))$ est déterminée par la forme bilinéaire f; réciproquement, f s'écrit en termes de A (voir (2.3.3.1)). Voici une formulation plus scientifique: l'application qui associe à f la matrice $A = (f(e_i,e_j))$ définit un isomorphisme d'espaces vectoriels

$$\{\text{formes bilinéaires } f : E \times E \longrightarrow K\} \xrightarrow{\sim} M_n(K).$$

Cette correspondance dépend du choix de base.

(2.3.4) Exemples : (1) Si $n = 2$, $E = \mathbf{R}^2$ et $\mathcal{B} = \{e_1, e_2\}$ (la base canonique), la matrice $\begin{pmatrix} 3 & 2 \\ 4 & 5 \end{pmatrix}$ correspond à la forme bilinéaire

$$f(x,y) = f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1\ x_2)\begin{pmatrix} 3 & 2 \\ 4 & 5 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (x_1\ x_2)\begin{pmatrix} 3y_1 + 2y_2 \\ 4y_1 + 5y_2 \end{pmatrix} =$$
$$= 3x_1y_1 + 2x_1y_2 + 4x_2y_1 + 5x_2y_2.$$

(2) Dans la base canonique de $\mathbf{R}^n$, le produit scalaire euclidien usuel correspond à la matrice $A = ((e_i \mid e_j)) = I_n$, ce qui équvaut à la formule

$$(x \mid y) = x_1y_1 + \cdots + x_ny_n = (x_1 \cdots x_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (x_1 \cdots x_n)\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

(3) L'aire algébrique $f : \mathbf{R}^2 \times \mathbf{R}^2 \longrightarrow \mathbf{R}$ (voir 2.2.3(2)) s'écrit sous la forme matricielle

$$f(x,y) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1 = (x_1\ x_2)\begin{pmatrix} y_2 \\ -y_1 \end{pmatrix} = (x_1\ x_2)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

donc la matrice représentative de f dans la base canonique de $\mathbf{R}^2$ est égale à $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(2.3.5) Formes bilinéaires (anti)-symétriques et matrices (anti)-symétriques.
Soit

$$f(x,y) = {}^tXAY \qquad (x, y \in E) \tag{2.3.5.1}$$

une forme bilinéaire, écrite sous la forme matricielle dans une base $\mathcal{B}$ de E. Un scalaire $\lambda \in K$ peut être considéré comme une matrice de type 1×1, $\lambda \in M_1(K)$; dans ce cas il est égal à sa transposée $\lambda = {}^t\lambda$. Si l'on prend $\lambda = f(y,x)$, on obtient

$$f(y,x) = {}^tf(y,x) = {}^t({}^tYAX) = {}^tX{}^tAY \tag{2.3.5.2}$$

Comme la forme f est déterminée par la matrice A (et la réciproque est aussi vraie), le comparaison de (2.3.5.1) et (2.3.5.2) montre les équivalences

$$f \text{ est symétrique} \iff {}^tA = A$$
$$f \text{ est anti}-\text{symétrique} \iff {}^tA = -A.$$

La base $\mathcal{B}$ était quelconque; comme la propriété de (anti)-symétrie de f ne dépend pas du choix de base, on en déduit que l'on a

$$\begin{aligned} f \text{ est symétrique} &\iff {}^tA = A \text{ pour un choix de base} \iff {}^tA = A \text{ pour tout choix de base} \\ f \text{ est anti}-\text{symétrique} &\iff {}^tA = -A \text{ pour un choix de base} \iff {}^tA = -A \text{ pour tout choix de base.} \end{aligned} \tag{2.3.5.3}$$

(2.3.6) Changement de base. Soit $f : E \times E \longrightarrow K$ une forme bilinéaire; soient $\mathcal{B} = \{e_1, \ldots, e_n\}$ et $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ deux bases de E. On note

$$A = (A_{ij}) = (f(e_i, e_j)) = (\text{la matrice de } f \text{ dans la base } \mathcal{B}) \in M_n(K)$$
$$A' = (A'_{ij}) = (f(e'_i, e'_j)) = (\text{la matrice de } f \text{ dans la base } \mathcal{B}') \in M_n(K).$$

Soit $P \in M_n(K)$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Soient $x, y \in E$ deux vecteurs de E; on note $X = M_{\mathcal{B}}(x), Y = M_{\mathcal{B}}(y) \in K^n$ (resp., $X' = M_{\mathcal{B}'}(x), Y' = M_{\mathcal{B}'}(y) \in K^n$) les vecteurs colonnes qui représentent x et y dans la base $\mathcal{B}$ (resp., dans la base $\mathcal{B}'$). D'après (1.3.1.2), on a

$$X = PX', \qquad Y = PY'.$$

Le comparaison des formules matricielles

$$f(x, y) = {}^tX'A'Y'$$
$$f(x, y) = {}^tXAY = {}^t(PX')A(PY') = {}^tX'\,{}^tPAP\,Y',$$

montre que l'on a

$$A' = {}^tPAP. \tag{2.3.6.1}$$

Il résulte de cette formule que

$${}^tA' = {}^t({}^tPAP) = {}^tP\,{}^tAP,$$

ce qui signifie que

$$A' \text{ est (anti}-\text{)symétrique} \iff A \text{ est (anti}-\text{)symétrique},$$

en accord avec (2.3.5.3).

Exemple : Si $E = \mathbf{R}^2$, $\mathcal{B} =$ la base canonique, $\mathcal{B}' = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $A = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$, alors on a

$$x = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = X = x'_1 \begin{pmatrix} 3 \\ -2 \end{pmatrix} + x'_2 \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = PX',$$

d'où $x_1 = 3x'_1 + 2x'_2$, $x_2 = -2x'_1 - x'_2$ (et de même pour y). Il en résulte que l'on a

$$A' = {}^tPAP = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix},$$

donc

$$f(x, y) = -x'_1y'_2 + 2x'_2y'_1 + x'_2y'_2.$$

Pour vérifier le calcul précédent, on peut aussi calculer directement

$$f(\begin{pmatrix} 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \end{pmatrix}) = 0, \quad f(\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}) = -f(\begin{pmatrix} 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix}) = 1, \quad f(\begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \end{pmatrix}) = 2.$$

(2.3.7) Rang. On appelle le **rang** d'une forme bilinéaire $f : E \times E \longrightarrow K$ le rang de la matrice représentative $A = (f(e_i, e_j)) \in M_n(K)$ de f dans une base $\mathcal{B} = \{e_i\}$ de E :

$$\mathrm{rg}(f) = \mathrm{rg}(A).$$

Le rang est **bien-défini :** si l'on remplace $\mathcal{B}$ par une autre base $\mathcal{B}'$, alors la matrice A sera remplacée par $A' = {}^tPAP$, où $P \in M_n(K)$ est une matrice **inversible**. On sait (voir la preuve de 1.4.6) que la multiplication à gauche (resp., à droite) par une matrice inversible conserve le rang, donc $\mathrm{rg}(A) = \mathrm{rg}(A')$.

(2.3.8) Formes bilinéaires et l'espace dual. * Soit $f : E\times E \longrightarrow K$ une forme bilinéaire. Par définition, si l'on fixe $x \in E$, la formule

$$\begin{aligned} f(x,\cdot) : E &\longrightarrow K \\ y &\mapsto f(x,y) \end{aligned}$$

définit une forme linéaire sur E, donc un élément $f(x,\cdot) \in E^*$ de l'espace dual de E. De plus, on a

$$f(x+x',\cdot) = f(x,\cdot) + f(x',\cdot), \qquad f(\lambda x,\cdot) = \lambda f(x,\cdot),$$

ce qui signifie que l'application

$$\begin{aligned} f_1 : E &\longrightarrow E^* \\ x &\mapsto f(x,\cdot) \end{aligned}$$

est linéaire. Soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E; on note $\mathcal{B}^* = \{e_1^*, \ldots, e_n^*\}$ la base duale de E^*. On va exprimer la matrice $A_1 \in M_n(K)$ de f_1 dans les bases $\mathcal{B}, \mathcal{B}^*$ en termes de la matrice $A = (f(e_i, e_j)) \in M_n(K)$. En utilisant la formule (1.7.7.1) pour l'application f_1 (et l'identification de E^{**} à E; voir 1.7.9), on obtient

$$(A_1)_{ij} = e_i^{**}(f_1(e_j)) = (f_1(e_j))(e_i) = f(e_j, e_i) = A_{ji}, \qquad A_1 = {}^tA.$$

De même, si l'on fixe $y \in E$, alors la formule

$$\begin{aligned} f(\cdot,y) : E &\longrightarrow K \\ x &\mapsto f(x,y) \end{aligned}$$

définit une forme linéaire $f(\cdot,y) \in E^*$ et lapplication

$$\begin{aligned} f_2 : E &\longrightarrow E^* \\ y &\mapsto f(\cdot,y) \end{aligned}$$

est ausi linéaire. Si l'on note $A_2 \in M_n(K)$ la matrice de f_2 dans les bases $\mathcal{B}, \mathcal{B}^*$, on a

$$(A_2)_{ij} = e_i^{**}(f_2(e_j)) = (f_2(e_j))(e_i) = f(e_i, e_j) = A_{ij}, \qquad A_2 = A.$$

En particulier, on peut définir le rang de f de la façon abstraite suivante :

$$\mathrm{rg}(f) := \mathrm{rg}(f_1) \ (= \mathrm{rg}(f_2)).$$

*** Ce paragraphe est facultatif.**

3. Formes quadratiques

Hypothèse fondamentale : **Désormais**, on suppose que la caractéristique du corps K n'est pas égale à 2 ($\iff$ $2 \neq 0$ dans K $\iff$ 2 est inversible dans K).

Cette hypothèse **est** satisfaite si $K = \mathbf{R}, \mathbf{C}$ ou $\mathbf{F}_{p^n}$ pour $p \neq 2$; elle **n'est pas** satisfaite si $K = \mathbf{F}_{2^n}$.

3.1 Notions de base

(3.1.1) Exemple : Le produit scalaire euclidien usuel

$$(x \mid y) = x_1 y_1 + \cdots + x_n y_n$$

est une forme bilinéaire symétrique $\mathbf{R}^n \times \mathbf{R}^n \longrightarrow \mathbf{R}$. Le carré de la longeur

$$\|x\|^2 = (x \mid x) = x_1^2 + \cdots + x_n^2$$

est une **forme quadratique** sur $\mathbf{R}^n$. Cette exemple nous conduit à la définition générale suivante.

(3.1.2) Définition. *Soit $f : E \times E \longrightarrow K$ une forme bilinéaire* **symétrique** *sur un K-ev E. On appelle la fonction $q : E \longrightarrow K$ définie par la formule $q(x) = f(x,x)$* **la forme quadratique associée à** *f. [On a $\forall \lambda \in K\ \forall x \in E \quad q(\lambda x) = \lambda^2 q(x)$.]*

(3.1.3) Formulation matricielle. Sous les hypothèses de 3.1.2, soit $\mathcal{B} = \{e_1, \ldots, e_n\}$ une base de E. La matrice $A = (A_{ij}) \in M_n(K)$ de f dans la base $\mathcal{B}$ est symétrique, car $A_{ji} = f(e_j, e_i) = f(e_i, e_j) = A_{ij}$. Si l'on associe à un vecteur $x = \sum_{i=1}^n x_i e_i \in E$ le vecteur colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ qui représent x dans la base $\mathcal{B}$, on a

$$q(x) = f(x,x) = {}^t X A X = \sum_{i,j=1}^{n} A_{ij}\, x_i x_j = \sum_{i=1}^{n} A_{ii}\, x_i^2 + 2 \sum_{i<j} A_{ij}\, x_i x_j. \tag{3.1.3.1}$$

D'après 2.3.6, un changement de base donné par la formule $X = PX'$ (où P est la matrice de passage de $\mathcal{B}$ à une autre base $\mathcal{B}'$) remplace A par la matrice $A' = {}^t PAP$ (cette matrice est encore symétrique).

Réciproquement, pour toute matrice symétrique $A = {}^t A \in M_n(K)$, la formule (3.1.3.1) définit une forme quadratique sur E.

Exemple : Si $K = \mathbf{R}$, $E = \mathbf{R}^2$, $\mathcal{B} =$ la base canonique et $A = \begin{pmatrix} 3 & -1 \\ -1 & 5 \end{pmatrix}$, alors on a

$$q(x) = q\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = (x_1\ x_2) \begin{pmatrix} 3 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1\ x_2) \begin{pmatrix} 3x_1 - x_2 \\ -x_1 + 5x_2 \end{pmatrix} = 3x_1^2 - 2x_1 x_2 + 5x_2^2.$$

(3.1.4) Proposition ("f est déterminée par q"). *Toute forme quadratique $q : E \longrightarrow K$ est associée à une* **unique** *forme bilinéaire symétrique $f : E \times E \longrightarrow K$ (on appelle f* **la forme polaire de** *q). [On utilise l'hypothèse $2 \neq 0$.]*

Preuve. On utilise la formule bien connue $(a+b)^2 - a^2 - b^2 = 2ab$. Plus précisement, la propriété de symétrie de la forme f entraîne

$$q(x+y) = f(x+y, x+y) = f(x,x) + f(x,y) + f(y,x) + f(y,y) = q(x) + 2f(x,y) + q(y)$$

(pour tous $x, y \in E$), d'où

$$f(x,y) = \frac{1}{2}\left(q(x+y) - q(x) - q(y)\right).$$

(3.1.5) Détermination de la forme polaire (exemple) : Soit $q : \mathbf{R}^2 \longrightarrow \mathbf{R}$ la forme quadratique $q(x) = x_1^2 - 4x_1x_2 + 7x_2^2$. L'écriture matricielle de q

$$q(x) = x_1^2 - 4x_1x_2 + 7x_2^2 = x_1^2 - 2x_1x_2 - 2x_2x_1 + 7x_2^2 = (x_1\ x_2)\begin{pmatrix} 1 & -2 \\ -2 & 7 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

entraîne que l'on a

$$f(x,y) = (x_1\ x_2)\begin{pmatrix} 1 & -2 \\ -2 & 7 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 7x_2y_2.$$

3.2 Rang, noyau (= radical) d'une forme quadratique

(3.2.1) Définition. *Soit $q : E \longrightarrow K$ une forme quadratique; on note $f : E \times E \longrightarrow K$ la forme polaire. Soient $x, y \in E$ des vecteurs. On dit que x* **est orthogonal à** *y* **(par rapport à** *f***)** *si $f(x,y) = 0$* **(notation :** *$x \perp y$; parfois, on écrit $x \perp_f y$ ou $x \perp_q y$). [La forme f étant symétrique, on a $x \perp y \iff y \perp x$.]*
Le **rang de** *q est l'entier* $\mathrm{rg}(q) = \mathrm{rg}(f)$*; le* **noyau** *(= le* **radical***) de q est le sous-espace de E*

$$N(q) = \mathrm{rad}(q) = \{x \in E \mid \forall y \in E \quad x \perp y\} = \{x \in E \mid \forall y \in E \quad y \perp x\}.$$

(3.2.2) Formules matricielles
(3.2.2.1) Soit $\mathcal{B}$ une base de E (on suppose que $n = \dim(E) < \infty$). On écrit, en utilisant 2.3.1 et 3.1.3, $f(x,y) = {}^tXAY$ et $q(x) = {}^tXAX$ (où $A = {}^tA \in M_n(K)$ est une matrice symétrique). Par définition, on a

$$\mathrm{rg}(q) = \mathrm{rg}(f) = \mathrm{rg}(A).$$

(3.2.2.2) Proposition. *Lorsqu'on identifie $M_{\mathcal{B}} : E \xrightarrow{\sim} K^n$ en utilisant la base $\mathcal{B}$, le sous-espace $N(q) \subset E$ correspond au sous-espace $\{X \in K^n \mid AX = 0\} \subset K^n$. En particulier, on a*

$$\dim N(q) = \dim\{X \in K^n \mid AX = 0\} = n - \mathrm{rg}(A) = n - \mathrm{rg}(q).$$

Preuve. Par définition, $N(q)$ correspond au sous-espace suivant de K^n:

$$F = \{X \in K^n \mid \forall Y \in K^n \quad {}^tYAX = 0\} = \{X \in K^n \mid \forall i = 1, \dots, n \quad {}^te_iAX = 0\},$$

où $e_1, \dots, e_n$ est la base canonique de K^n. On note b le vecteur colonne $AX \in K^n$; comme ${}^te_ib$ est égale à la i-ième coordonnée de b, on en déduit que les éléments de F sont les vecteurs X vérifiant $b(= AX) = 0$. La dimension de cet espace est déterminée par la formule du rang (1.2.7.2).

(3.2.2.3) Définition. *La forme quadratique q est dite* **non-dégénérée** *si $N(q) = 0$.*

(3.2.2.4) Il résulte de 3.2.2.2 que les conditions suivantes sont équivalentes :

$$q \text{ est non} - \text{dégénérée} \iff N(q) = 0 \iff \mathrm{rg}(q) = \dim(E) \iff A \text{ is inversible} \iff \det(A) \neq 0.$$

(3.2.3) Exemples : (0) $q = 0$: dans ce cas on a $f = 0$, $A = 0$, $N(q) = E$, $\mathrm{rg}(q) = 0$.
(1) $E = K^n$, $q(x) = x_1^2 + \cdots + x_n^2$: dans ce cas on a $f(x,y) = x_1y_1 + \cdots + x_ny_n$, $A = I_n$, $N(q) = 0$, $\mathrm{rg}(q) = n$.

(2) $E = K^n$, $q(x) = x_1^2 + \cdots + x_r^2$ $(0 \leq r \leq n)$: dans ce cas on a $f(x,y) = x_1y_1 + \cdots + x_ry_r$,

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in M_n(K), \qquad N(q) = \{x \in K^n \mid Ax = 0\} = \{{}^t(0, \ldots, 0, x_{r+1}, \ldots, x_n) \mid x_i \in K\},$$

$\mathrm{rg}(q) = \mathrm{rg}(A) = r$, $\dim(N(q)) = n - r$.
(3) $E = \mathbf{R}^2$, $q(x) = x_1^2 - 4x_1x_2 + 7x_2^2$: dans ce cas on a

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 7 \end{pmatrix}.$$

Comme $\det(A) = 3 \neq 0$, on a $\mathrm{rg}(q) = \mathrm{rg}(A) = 2$ and $N(q) = 0$.
(4) $E = \mathbf{R}^2$, $q(x) = x_1^2 - 4x_1x_2 + 4x_2^2$: dans ce cas on a

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}.$$

Comme $\det(A) = 0$, on a $\mathrm{rg}(q) = \mathrm{rg}(A) < 2$. D'autre part, $A \neq 0$, donc $\mathrm{rg}(A) > 0$, ce qui entraîne $\mathrm{rg}(A) = 1$ et $\dim(N(q)) = 2 - 1 = 1$. Plus précisement, on a

$$N(q) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;\middle|\; \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \right\} = \mathrm{vect}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix} \right).$$

Pour déterminer $N(q)$, on peut aussi utiliser la formule $q(x) = (x_1 - 2x_2)^2$, qui entraîne $f(x,y) = (x_1 - 2x_2)(y_1 - 2y_2)$.

3.3 Diagonalisation "facile" d'une forme quadratique ("décomposition en carrés")

Le résultat de base de la théorie des formes quadratiques est le théorème suivant (qui utilise l'hypothèse $2 \neq 0 \in K$).

(3.3.1) Théorème. *Pour toute forme quadratique $q : E \longrightarrow K$ (où $n = \dim(E) < \infty$) il existe une base $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ de E dans laquelle la matrice de q est diagonale :*

$$q(x) = d_1 x_1'^2 + \cdots + d_n x_n'^2 \qquad \qquad (x = \sum_{i=1}^n x'_i e'_i)$$

(on dit que $\mathcal{B}'$ est une **base orthogonale** *par rapport à q).*
Formulation matricielle : *Pour toute matrice* **symétrique** *$A = {}^tA \in M_n(K)$ il existe une matrice inversible $P \in M_n(K)$ telle que*

$${}^tPAP = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}.$$

(3.3.2) Exemples : Dans les exemples suivants, $K = \mathbf{R}$, $E = \mathbf{R}^n$ et $x_1, \ldots, x_n$ sont les coordonnées de $\mathbf{R}^n$ dans la base canonique.
(1) $E = \mathbf{R}^2$, $q(x) = x_1^2 - 4x_1x_2 + 7x_2^2$: on a

$$q(x) = x_1^2 - 4x_1x_2 + 7x_2^2 = (x_1 - 2x_2)^2 + 3x_2^2 = x_1'^2 + 3x_2'^2,$$

où

$$X' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} x_1 - 2x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = P'X.$$

Ce changement de variable s'écrit, dans sens inverse, de la façon suivante :

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = PX',$$

ce qui entraîne (d'après (1.3.1.2)) que x_1', x_2' sont les coordonnées de x dans la base $\mathcal{B}'$ donnée par les colonnes de la matrice P :

$$\mathcal{B}' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1' \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2' \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Les matrices A et A' de q dans les bases respectives $\mathcal{B}$ et $\mathcal{B}'$ sont égales à

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 7 \end{pmatrix}, \qquad A' = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix},$$

d'où $\mathrm{rg}(q) = \mathrm{rg}(A') = 2$.

(2) $E = \mathbf{R}^2$, $q(x) = x_1^2 - 6x_1x_2 + 9x_2^2$: on a

$$q(x) = x_1^2 - 6x_1x_2 + 9x_2^2 = (x_1 - 3x_2)^2 = x_1'^2 = x_1'^2 + 0 \cdot x_2'^2,$$

où $x_1' = x_1 - 3x_2$. Il faut encore trouver x_2' (une forme linéaire en x_1 et x_2) pour que x_1' et x_2' soient les coordonnées dans $\mathbf{R}^2$ par rapport à une base convenable $\mathcal{B}'$. Autrement dit,

$$x_2' = cx_1 + dx_2, \qquad X' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = P'X$$

(pour un choix de $c, d \in \mathbf{R}$); il faut s'assurer que la matrice P' soit **inversible** (elle sera égale à la matrice de passage de $\mathcal{B}'$ à la base canonique $\mathcal{B}$. Dès que l'on choisit P', on calcule $P = (P')^{-1}$ (= la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$); les éléments de $\mathcal{B}'$ se trouvent dans les colonnes de la matrice P. Il vaut mieux choisir la matrice P' aussi simple que possible; par exemple, on peut prendre

$$x_2' = x_2, \qquad \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad P' = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}.$$

Dans ce cas on voit directement que $\det(P') = 1 \cdot 1 = 1 \neq 0$, donc la matrice P' est inversible. Le calcul analogue à celui déjà effectué en (1) montre que

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = PX',$$

donc $q(x) = x_1'^2$ dans la base

$$\mathcal{B}' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1' \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2' \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Les matrices A et A' de q dans les bases respectives $\mathcal{B}$ et $\mathcal{B}'$ sont égales à

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 9 \end{pmatrix}, \qquad A' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

d'où $\mathrm{rg}(q) = \mathrm{rg}(A') = 1$.

(3) $E = \mathbf{R}^3$, $q(x) = x_1^2 + 4x_1x_2 - 6x_1x_3 + 3x_2^2 + 2x_2x_3 - 40x_3^2$:

$$\begin{aligned} q(x) = & \underbrace{x_1^2 + 4x_1x_2 - 6x_1x_3}_{\text{une partie de } (x_1+2x_2-3x_3)^2 = x_1^2+4x_1x_2-6x_1x_3+4x_2^2-12x_2x_3+9x_3^2} +3x_2^2 + 2x_2x_3 - 40x_3^2 = \\ & = (x_1 + 2x_2 - 3x_3)^2 - x_2^2 + 14x_2x_3 - 49x_3^2 = (x_1 + 2x_2 - 3x_3)^2 - (x_2 - 7x_3)^2 = \\ & = x_1'^2 - x_2'^2 + 0 \cdot x_3'^2, \end{aligned}$$

où

$$X' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 - 3x_3 \\ x_2 - 7x_3 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & -7 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P'X.$$

Comme $\det(P') = 1 \cdot 1 \cdot 1 = 1 \neq 1$, la matrice P' est inversible. On peut calculer son inverse $P = (P')^{-1}$ soit en utilisant la méthode expliquée dans 1.4.9, soit par un calcul direct :

$$x_3 = x_3', \qquad x_2 = x_2' + 7x_3', \qquad x_1 = x_1' - 2(x_2' + 7x_3') + 3x_3' = x_1' - 2x_2' - 11x_3'$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & -11 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = x_1' \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2' \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + x_3' \begin{pmatrix} -11 \\ 7 \\ 1 \end{pmatrix},$$

donc $q(x) = x_1'^2 - x_2'^2 + 0 \cdot x_3'^2$ dans les coordonnées x_1', x_2', x_3' par rapport à la base

$$\mathcal{B}' = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -11 \\ 7 \\ 1 \end{pmatrix}.$$

Les matrices A et A' de q dans les bases respectives $\mathcal{B}$ et $\mathcal{B}'$ sont égales à

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 3 & 1 \\ -3 & 1 & -40 \end{pmatrix}, \qquad A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

d'où $\mathrm{rg}(q) = \mathrm{rg}(A') = 2$. En pratique, on devrait utiliser une méthode indépendante (par exemple, celle du §1.4) pour déterminer $\mathrm{rg}(A)$, afin de **vérifier le résultat**).

(4) $E = \mathbf{R}^3$, $q(x) = x_1x_2 + x_1x_3 + x_2x_3$. Dans ce cas il n'y a aucun élément diagonal x_j^2. On utilise l'identité suivante :

$$x_1x_2 = \frac{(x_1 + x_2)^2 - (x_1 - x_2)^2}{4} = \left(\frac{x_1 + x_2}{2}\right)^2 - \left(\frac{x_1 - x_2}{2}\right)^2.$$

On fait le changement de coordonnées

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} (x_1 + x_2)/2 \\ (x_1 - x_2)/2 \end{pmatrix}, \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1' + x_2' \\ x_1' - x_2' \end{pmatrix}$$

(en gardant x_3) et ensuite on utilise la méthode précédente; on obtient

$$
\begin{aligned}
q(x) &= (x_1' + x_2')(x_1' - x_2') + (x_1' + x_2')x_3 + (x_1' - x_2')x_3 = \\
&= x_1'^2 + 2x_1'x_3 - x_2'^2 = (x_1' + x_3)^2 - x_2'^2 - x_3^2 = \\
&= y_1^2 - y_2^2 - y_3^2,
\end{aligned}
$$

où

$$
Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1' + x_3 \\ x_2' \\ x_3 \end{pmatrix} = \begin{pmatrix} (x_1 + x_2)/2 + x_3 \\ (x_1 - x_2)/2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P'X.
$$

On a

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 + y_2 - y_3 \\ y_1 - y_2 - y_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = y_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + y_3 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}
$$

(on calcule l'inverse $(P')^{-1}$, ou on fait un calcul direct), ce qui entraîne que l'on a $q(x) = y_1^2 - y_2^2 - y_3^2$ dans les coordonnées y_1, y_2, y_3 par rapport à la base

$$
\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.
$$

(3.3.3) Preuve du Théorème 3.3.1 : Les méthodes utilisées dans les exemples précédents marchent en général. On fixe une base $\mathcal{B} = \{e_1, \dots, e_n\}$ de E; soit $A = {}^tA \in M_n(K)$ la matrice de q dans la base $\mathcal{B}$. Si $A = 0$, alors on prend $\mathcal{B}' = \mathcal{B}$. On suppose que $A \neq 0$. D'après (3.1.3.1), on a

$$
q(x) = \sum_{i,j=1}^{n} A_{ij}\, x_i x_j = \sum_{i=1}^{n} A_{ii}\, x_i^2 + 2 \sum_{i<j} A_{ij}\, x_i x_j \qquad\qquad (x = \sum_{i=1}^{n} x_i e_i).
$$

(a) Il existe un élément diagonal non nul $A_{ii} \neq 0$: Après un éxchange des coordonnées x_1 et x_i on peut supposer que $A_{11} \neq 0$. Comme

$$
q(x) = A_{11} x_1^2 + 2A_{12}\, x_1 x_2 + \cdots + 2A_{1n}\, x_1 x_n + q_1(x)
$$

(où la variable x_1 n'intervient pas dans $q_1(x)$), on écrit

$$
q(x) = A_{11} \left(x_1 + \frac{A_{12}}{A_{11}} x_2 + \cdots + \frac{A_{1n}}{A_{11}} x_n \right)^2 + q_2(x) = A_{11}\, x_1'^2 + q_2(x),
$$

où

$$
x_1' = x_1 + \frac{A_{12}}{A_{11}} x_2 + \cdots + \frac{A_{1n}}{A_{11}} x_n
$$

est la variable x_1 n'intervient pas dans q_2, c'est-à-dire que q_2 est une forme quadratic sur le sous-espace $E' = \mathrm{vect}(e_2, \dots, e_n) \subset E$ de dimension $n - 1$.

(b) Tous les termes diagonaux sont nuls $A_{11} = \cdots = A_{nn} = 0$: Il existe un élément non nul $A_{ij} \neq 0$, $i < j$. Quitte à échanger les coordonnées $x_1 \longleftrightarrow x_i$ et $x_2 \longleftrightarrow x_j$, on peut supposer que $A_{12} \neq 0$ (et que les termes diagonaux sont nuls), ce qui entraîne que la forme $q(x)$ s'écrit

$$q(x) = 0 \cdot x_1^2 + 2A_{12}\, x_1 x_2 + 0 \cdot x_2^2 + \cdots .$$

Après le changement de coordonnées introduit dans 3.3.2(4)

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} (x_1+x_2)/2 \\ (x_1-x_2)/2 \end{pmatrix}, \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1'+x_2' \\ x_1'-x_2' \end{pmatrix}$$

(on garde les coordonnées $x_3, \ldots, x_n$), on obtient

$$q(x) = 2A_{12}\, x_1'^2 + \cdots ,$$

donc le cas (a) s'applique à la forme q exprimée dans les coordonnées $x_1', x_2', x_3, \ldots, x_n$. En appliquant (a), on écrit

$$q(x) = 2A_{12}\, (x_1'')^2 + q_2(x),$$

où q_2 ne contient que les variables $x_2', x_3, \ldots, x_n$.
Ensuite on applique le même argument à la forme q_2, etc.

(3.3.4) Rang, noyau et diagonalisation : Sous l'hypothèses du Théorème 3.3.1, on peut choisir l'ordre des éléments de la base $\mathcal{B}'$ de telle façon que l'on ait

$$d_1, \ldots, d_r \neq 0, \qquad d_{r+1} = \cdots = d_n = 0 \qquad\qquad (0 \leq r \leq n),$$

ce qui signifie que la matrice A' de q dans la base $\mathcal{B}'$ est égale à

$$A' = \begin{pmatrix} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (d_1, \ldots, d_r \neq 0),$$

d'où

$$\mathrm{rg}(q) = r. \tag{3.3.4.1}$$

Le système des équations linéaires

$$A' \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = 0$$

équivaut à

$$x_1' = \cdots = x_r' = 0, \tag{3.3.4.2}$$

ce qui entraîne (d'après 3.2.2.2) que le noyau de q est défini, **dans les coordonnées par rapport à la base** $\mathcal{B}'$, par les équations (3.3.4.2). Autrement dit, on a

$$N(q) = \{\sum_{i=1}^{n} x_i' e_i' \mid x_1' = \cdots = x_r' = 0\} = \mathrm{vect}(e_{r+1}', \ldots, e_n'\}. \tag{3.3.4.3}$$

Exemples : (1) Dans l'exemple 3.3.2(1), on a $r = 2 = \dim(E)$, donc $N(q) = 0$.
(2) Dans l'exemple 3.3.2(2), on a $r = 1$ et

$$N(q) = \Big\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \,\Big|\, x_1 - 3x_2 = 0\Big\} = \mathrm{vect}\Big(\begin{pmatrix} 3 \\ 1 \end{pmatrix}\Big).$$

(3) Dans l'xemple 3.3.2(3), on a $r = 2$ et

$$N(q) = \left\{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \,\middle|\, x_1 + 2x_2 - 3x_3 = x_2 - 7x_3 = 0\right\} = \mathrm{vect}\left(\begin{pmatrix} -11 \\ 7 \\ 1 \end{pmatrix}\right).$$

(4) Dans l'xemple 3.3.2(4), on a $r = 3 = \dim(E)$, donc $N(q) = 0$.

3.4 Somme, somme directe

(3.4.1) Définition (somme de sous-espaces). *Soient $F_1, \ldots, F_n$ des sous-espaces vectoriels d'un K-ev E. Leur* **somme** *est le sous-espace vectoriel $F \subset E$ engendré par la réunion $F_1 \cup \cdots \cup F_n$; on la note $F_1 + \cdots + F_n = \sum_{i=1}^n F_i$. On a*

$$F_1 + \cdots + F_n = \mathrm{vect}(F_1 \cup \cdots \cup F_n) = \{x_1 + \cdots + x_n \mid x_i \in F_i\}.$$

(3.4.2) Exemple : Lorsque $F_1, F_2 \subset \mathbf{R}^3$ sont des plans distincts dans $\mathbf{R}^3$ (contenant l'origine), alors $F_1 + F_2 = \mathbf{R}^3$ est l'espace tout entier et $F_1 \cap F_2$ est une droite.
En général, pour des sous-espaces quelconques $F_1, F_2 \subset E$ d'un K-ev E, on a

$$\dim(F_1 \cap F_2) + \dim(F_1 + F_2) = \dim(F_1) + \dim(F_2)$$

(dans l'exemple précédent, on a $1 + 3 = 2 + 2$).

(3.4.3) Définition (somme directe de sous-espaces). *Sous les hypothèses de 3.4.1, on dit que la somme $F_1 + \cdots + F_n$ est* **directe** *(notation : $F_1 \oplus \cdots \oplus F_n = \bigoplus_{i=1}^n F_i$) si tout élément $x \in F_1 + \cdots + F_n$ admet* **l'unique** *décomposition $x = x_1 + \cdots + x_n$ ($x_i \in F_i$). Si c'est le cas, soit $\mathcal{B}_i$ une base de F_i ($i = 1, \ldots, n$); alors la réunion $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_n$ est une base de $F_1 \oplus \cdots \oplus F_n$, donc on a*

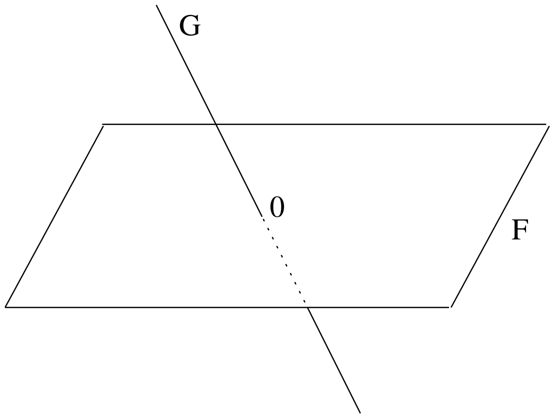
$$\dim(F_1 \oplus \cdots \oplus F_n) = \dim(F_1) + \cdots + \dim(F_n).$$

(3.4.4) Exemples : (1) $K^n = F_1 \oplus + \cdots \oplus F_n$, où les sous-espaces $F_i = \mathrm{vect}(e_i) = Ke_i$ sont les droites engendrées par les vecteurs de la base canonique de K^n.
(2) Si $n = 2$, alors la somme $F_1 + F_2$ est directe $\iff$ $F_1 \cap F_2 = 0$.

(3.4.5) Exercice : Trouver un exemple de trois sous-espaces vectoriels $F_1, F_2, F_3 \subset E$ vérifiant les deux propriétés suivantes : $F_1 \cap F_2 = F_1 \cap F_3 = F_2 \cap F_3 = 0$; la somme $F_1 + F_2 + F_3$ n'est pas directe.

(3.4.6) Définition (sous-espace supplémentaire). *Soient $F, G \subset E$ des sous-espaces vectoriels de E. On dit que G est un* **sous-espace supplémentaire à** *F si l'on a $F \oplus G = E$ ($\iff$ $F \cap G = 0$ et $F + G = E$ $\iff$ $F \cap G = 0$ et $\dim(E) = \dim(F) + \dim(G)$).*

(3.4.7) Exemple:



3.5 Somme orthogonale

Dans les paragraphes §3.5 et §3.6 on note $f : E \times E \longrightarrow K$ une forme bilinéaire symétrique et $q : E \longrightarrow K$ la forme quadratique associée à f $(q(x) = f(x,x))$.

Nous rappelons (voir 3.2.1) que des vecteurs $x, y \in E$ sont **orthogonaux** l'un à l'autre par rapport à f (notation : $x \perp y$ ou $x \perp_f y$ ou $x \perp_q y$) si $f(x,y) = 0$ ($\Longleftrightarrow$ $y \perp x$).

(3.5.1) Définition. *Soit $S \subset E$ un sous-ensemble non vide de E.* **L'orthogonal** *de S est le sous-ensemble*

$$S^{\perp} = \{x \in E \mid \forall y \in S \quad x \perp y\} \subset E.$$

(3.5.2) Remarques: (1) $S^{\perp}$ contient le radical (= le noyau) de q: $S^{\perp} \supset N(q)$.
(2) $\{\overrightarrow{0}\}^{\perp} = E$.
(3) $E^{\perp} = N(q)$.
(4) La propriété bilinéaire de f (voir 2.2.1) entraîne

$$\begin{array}{lcl} x \perp y, \quad x' \perp y & \Longrightarrow & (x + x') \perp y \\ x \perp y, \quad \lambda \in K & \Longrightarrow & (\lambda x) \perp y. \end{array}$$

(5) On déduit de (4) que l'orthogonal $S^{\perp}$ est un **sous-espace vectoriel** de E et que l'on a $S^{\perp} = (\mathrm{vect}(S))^{\perp}$. Autrement dit, on peut supposer que S est un sous-espace vectoriel de E.

(3.5.3) Définition (somme orthogonale). *Soient $F_1, F_2 \subset E$ des sous-espaces vectoriels de E. On dit que F_1* **est orthogonal à** *F_2 (par rapport à f) si l'on a*

$$\forall x \in F_1 \, \forall y \in F_2 \quad x \perp y$$

*(***notation :*** $F_1 \perp F_2$ ou $F_1 \perp_f F_2$ ou $F_1 \perp_q F_2$).*
On dit que E est égal à la **somme orthogonale** *de F_1 et F_2 (***notation :*** $E = F_1 \perp F_2$) si l'on a $E = F_1 \oplus F_2$ et $F_1 \perp F_2$.*

(3.5.4) Exemple (plan euclidien) : Soit $E = \mathbf{R}^2$ $(K = \mathbf{R})$; soit f le produit scalaire euclidien usuel (donc $q(x) = x_1^2 + x_2^2$). Si $F_1, F_2 \subset E$ sont des droites distictes qui passent par l'origine, alors on a $E = F_1 \oplus F_2$. De plus, on a : $E = F_1 \perp F_2 \iff F_1$ est orthogonale à F_2.

(3.5.5) Exemple (plan hyperbolique) : Soient $E = K^2$ et $q(x) = x_1^2 - x_2^2$ (donc $f(x,y) = x_1y_1 - x_2y_2$). Pour tout $a \in K$ on note $D_a = \mathrm{vect}\{\begin{pmatrix} 1 \\ a \end{pmatrix}\}$ la droite de pente a qui passe par l'origine. On note aussi $D_\infty = \mathrm{vect}\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ la droite verticale qui passe par l'origine.

Si $a \neq 0$, alors on a

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in D_a^\perp \iff \begin{pmatrix} 1 \\ a \end{pmatrix} \perp \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \iff y_1 - ay_2 = 0 \iff \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_1 \begin{pmatrix} 1 \\ 1/a \end{pmatrix} \iff \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in D_{1/a}.$$

On a aussi

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in D_0^\perp \iff y_1 = 0 \iff \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in D_\infty, \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in D_\infty^\perp \iff y_2 = 0 \iff \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in D_0.$$

En résumé, on a

$$\forall a \in K \cup \{\infty\} \qquad D_a^\perp = D_{1/a} \tag{3.5.5.1}$$

(où $1/0 = \infty$ and $1/\infty = 0$). En particulier, les droites

$$D_1^\perp = D_1, \qquad D_{-1}^\perp = D_{-1} \tag{3.5.5.2}$$

sont auto-orthogonales (on a $x \perp x$ pour tout $x \in D_1$ or $x \in D_{-1}$)! D'autre part, lorsque $a \neq \pm 1$, on a $D_a \cap D_{1/a} = 0$, donc $E = D_a \oplus D_{1/a}$:

$$\forall a \neq \pm 1 \qquad E = D_a \perp D_{1/a} = D_a \perp D_a^\perp.$$

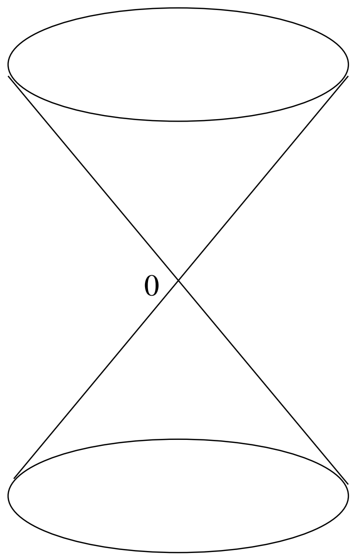
(3.5.6) Définition. *Soit $q : E \longrightarrow K$ une forme quadratique. Un* **vecteur isotrope** *de q est un vecteur non nul $x \in E$ tel que $q(x) = 0$ ($\iff$ $x \perp x$). Une* **droite isotrope** *est un sous-espace vectoriel* vect(x) *de E engendré par un vecteur isotrope. Le* **cône isotrope** *de q est le sous-ensemble de E*

$$C(q) := \{x \in E \mid q(x) = 0\} = \{\vec{0}\} \cup \{\text{vecteurs isotropes de } q\}.$$

Pour tous $x \in C(q)$ et $\lambda \in K$, on a $\lambda x \in C(q)$, car $q(\lambda x) = \lambda^2 q(x)$; il en résulte que $C(q)$ est égal à la réunion de toutes les droites isotropes.

(3.5.7) Exemple (plan hyperbolique): Lorsque $(E, q) = (K^2, x_1^2 - x_2^2)$ est le plan hyperbolique, alors $C(q) = D_1 \cup D_{-1}$ (voir (3.5.5.2)).

(3.5.8) Exemple: Lorsque $E = K^3$ et $q(x) = x_1^2 - x_2^2 - x_3^2$, alors $C(q)$ est le cône "usuel". Soit $D \subset E$ une droite qui passe par l'origine; on a (exercice !) :



C(q)

Si $D \not\subset C(q)$ ($\iff$ D n'est pas isotrope), alors $D^\perp$ est un plan qui ne contient pas D et l'on a $E = D \perp D^\perp$. Si $D \subset C(q)$ ($\iff$ D est isotrope), alors $D^\perp$ est un plan qui contient D (donc $D + D^\perp = D^\perp \neq E$). Plus précisement, $D^\perp$ est le plan tangent à $C(q)$ le long de la droite D.

(3.5.9) Somme orthogonale (formulation matricielle) : On suppose que l'on a $E = F_1 \oplus F_2$ (où $n_i = \dim(F_i) < \infty$). On fixe les bases respectives $\mathcal{B}_i$ de F_i $(i = 1, 2)$; leur réunion $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ est une base de E. La matrice de f dans la base $\mathcal{B}$ se décompose naturallement

$$A = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

où chaque bloc $B_{ij} \in M_{n_i,n_j}(K)$ représent la restriction $f_{ij} : F_i \times F_j \longrightarrow K$ de la forme f à $F_i \times F_j$ (f_{ij} est la fonction $f(x, y)$, où $x \in F_i$ et $y \in F_j$) dans les bases $\mathcal{B}_i$ et $\mathcal{B}_j$ $(i, j = 1, 2)$. La matrice $A = {}^tA$ ètant symétrique, on a

$$B_{ii} = {}^tB_{ii}, \qquad B_{12} = {}^tB_{21}.$$

Par définition, on a

$$E = F_1 \perp F_2 \iff f_{12} = 0,\ f_{21} = 0 \iff B_{12} = 0,\ B_{21} = 0 \iff A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \qquad (3.5.9.1)$$

où l'on a noté $A_i = B_{ii} \in M_{n_i}(K)$ la matrice de la restriction de f à l'espace $F_i \times F_i$ dans la base $\mathcal{B}_i$.

(3.5.10) Définition (somme orthogonale – le cas général). *Soient $F_1, \ldots, F_k \subset E$ des sous-espaces vectoriels de E tels que $E = F_1 \oplus \cdots \oplus F_k$. On dit que E est égal à la* **somme orthogonale** *de $F_1, \ldots, F_k$ si l'on a $F_i \perp F_j$ pour tous $i \neq j$ (***notation :** *$E = F_1 \perp \cdots \perp F_k$).*

(3.5.11) Formulation matricielle : Soit $E = F_1 \oplus \cdots \oplus F_k$; soit $\mathcal{B}_i$ une base de F_i $(i = 1, \ldots, k)$. On note $A \in M_n(K)$ la matrice de f dans la base $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ de E. L'énoncé (3.5.9.1) se généralise de la façon suivante :

$$E = F_1 \perp \cdots \perp F_k \iff A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_k \end{pmatrix}, \qquad (3.5.11.1)$$

où l'on a noté $A_i \in M_{n_i}(K)$ ($n_i = \dim(F_i)$) la matrice de la restriction de f à $F_i \times F_i$ dans la base $\mathcal{B}_i$ $(i = 1, \ldots, k)$.

Exemple : Lorsque $\mathcal{B} = \{e_1, \ldots, e_n\}$ est une base de E, alors on a $E = F_1 \oplus \cdots \oplus F_n$, où $F_i = \mathrm{vect}(e_i)$ est la droite engendrée par le vecteur e_i. De plus, on a

$$E = F_1 \perp \cdots \perp F_n \iff \forall i \neq j \quad e_i \perp e_j \iff A \text{ est une matrice diagonale} \iff$$
$$\iff \mathcal{B} \text{ est une base orthogonale.}$$

3.6 L'orthogonal : le cas non-dégénéré

(3.6.1) Exemple (réduction au cas non-dégénéré) : Soient $E = K^n$, $f(x, y) = x_1y_1 + \cdots + x_ry_r$ $(0 \leq r \leq n)$. La matrice $A \in M_n(K)$ de f dans la base canonique $\mathcal{B} = \{e_1, \ldots, e_n\}$ de K^n se décompose en blocs

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

D'après 3.5.9, il en résulte que la décomposition $E = F_1 \oplus F_2$, où l'on a noté $F_1 = \mathrm{vect}(e_1, \ldots, e_r)$ et $F_2 = \mathrm{vect}(e_{r+1}, \ldots, e_n)$, est **orthogonale :** $E = F_1 \perp F_2$. Plus précisement, $F_2 = N(q)$ n'est rien d'autre que le radical (= le noyau) de q. La matrice $A_1 = I_r \in M_r(K)$ étant inversible, on a

$$E = F_1 \perp N(q), \qquad \text{la restriction de } q \text{ à } F_1 \text{ est non} - \text{dégénéré.}$$

Voici le résultat général :

(3.6.2) Proposition. *Soit $q : E \longrightarrow K$ une forme quadratique, où $n = \dim(E) < \infty$. Soit $F \subset E$ un sous-espace supplémentaire quelconque de $N(q)$; alors on a $E = F \perp N(q)$ et la restriction q_F de q à l'espce F ($q_F : F \longrightarrow K$, $q_F(x) = q(x)$) est une forme quadratique* **non-dégénérée**.

Preuve. Soient $\mathcal{B}_1$ est $\mathcal{B}_2$ des bases respectives de F et $N(q)$. Par définition, les restrictions de f (= de la forme polaire de q) aux espaces $E \times N(q)$ et $N(q) \times E$ sont égales à zéro; il en résulte que la matrice $A \in M_n(K)$ de la forme f dans la base $\mathcal{B}_1 \cup \mathcal{B}_2$ de E est égale à

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix},$$

où $A_1 \in M_r(K)$ ($r = \dim(F)$) est la matrice de q_F dans la base $\mathcal{B}_1$. L'égalité

$$r = \dim(F) = n - \dim(N(q)) \overset{(3.2.2.2)}{=\!=\!=} \mathrm{rg}(q) = \mathrm{rg}(A) = \mathrm{rg}(A_1)$$

entraîne que la matrice A_1 est inversible, donc la forme q_F est non-dégénérée.

(3.6.3) Proposition. *Soit $q : E \longrightarrow K$ une forme quadratique* **non-dégénérée** *sur un espace vectoriel de dimension $n = \dim(E) < \infty$. Pour tout sous-espace vectoriel $F \subset E$ on a*

$$\dim(F^\perp) = n - \dim(F), \qquad (F^\perp)^\perp = F.$$

Preuve. Il suffit de démontrer la relation $\dim(F^\perp) = n - \dim(F)$, pour tout sous-espace vectoriel F de E (elle entraîne $\dim (F^\perp)^\perp = n - \dim(F^\perp) = n - (n - \dim(F)) = \dim(F)$; comme $F \subset (F^\perp)^\perp$ par définition, on a forcément $F = (F^\perp)^\perp$).

Un choix de base de E nous permet d'idéntifier E à K^n et la forme polaire f de q à la fonction $f(x, y) = {}^t x A y$ ($x, y \in K^n$). La matrice $A \in M_n(K)$ est **inversible**, car la forme q est non-dégénérée. Soit $v_1, \ldots, v_r$ ($r = \dim(F)$) une base de $F \subset K^n$; on note

$$C = (v_1 \mid \cdots \mid v_r) \in M_{n,r}(K)$$

la matrice dont les colonnes sont égales à $v_1, \ldots, v_r$. Comme les lignes de la matrice transposée ${}^tC \in M_{r,n}(K)$ sont égales à ${}^tv_1, \ldots, {}^tv_r$, on a

$$F^\perp = \{y \in K^n \mid v_1 \perp y, \ldots, v_r \perp y\} = \{y \in K^n \mid {}^tv_1 A y = \cdots = {}^tv_r A y = 0\} = \{y \in K^n \mid {}^tC A y = 0\}.$$

Les relations

$$\begin{aligned} \dim(F^\perp) &= n - \mathrm{rg}({}^tCA) && \text{d'après la formule du rang (1.2.7.2)} \\ \mathrm{rg}({}^tCA) &= \mathrm{rg}({}^t({}^tCA)) = \mathrm{rg}({}^tAC) = \mathrm{rg}(C) && \text{puisque } {}^tA \text{ est inversible} \\ \mathrm{rg}(C) &= \dim \mathrm{vect}(v_1, \ldots, v_r) = \dim(F) && \text{par définition,} \end{aligned}$$

alors montrent que l'on a $\dim(F^\perp) = n - \dim(F)$.

(3.6.4) Corollaire. *Sous les hypothèses de 3.6.3, si* $F \cap F^{\perp} = \{\vec{0}\}$, *alors* $E = F \perp F^{\perp}$ *est la somme orthogonale de* F *et* $F^{\perp}$.

Preuve. Comme $F \perp F^{\perp}$ par définition, il suffit de démontrer que l'on a $E = F \oplus F^{\perp}$, ce qui résulte de l'hypothèse $F \cap F^{\perp} = \{\vec{0}\}$ et de la relation $\dim(F) + \dim(F^{\perp}) = \dim(E)$.

(3.6.5) Exemples : On suppose que $\dim(F) = 1$ ($\iff$ $F = \mathrm{vect}(x)$, $x \in E$, $x \neq \vec{0}$).
(1) Lorsque x **est isotrope**, alors $x \perp x$, donc $F \subset F^{\perp}$ et $F + F^{\perp} = F^{\perp}$ est un hyperplan dans E (un sous-espace vectoriel de E de dimension $\dim(E) - 1$).
(2) Lorsque x **n'est pas isotrope**, alors $f(x, \lambda x) = \lambda\, f(x,x) = \lambda\, q(x) \neq 0$ pour tout scalaire non nul $\lambda \in K$, donc $F \cap F^{\perp} = \{\vec{0}\}$; il résulte de 3.6.4 que l'on a $E = F \perp F^{\perp}$.

(3.6.6) Exercice (le cas dégénéré). *Soit* $q : E \longrightarrow K$ *une forme quadratique sur un espace vectoriel de dimension finie. Pour tout sous-espace vectoriel* $F \subset E$, *on a* $(F^{\perp})^{\perp} = F + N(q)$.

(3.6.7) Voici une preuve abstraite du Théorème 3.3.1.

(3.6.8) = (3.3.1) Théorème. *Pour toute forme quadratique* $q : E \longrightarrow K$ *(*$n = \dim(E) < \infty$*) il existe une base* $\mathcal{B} = \{e_1, \ldots, e_n\}$ *telle que l'on ait* $E = Ke_1 \perp \cdots \perp Ke_n$ *(* $\iff$ $\mathcal{B}$ *est une base orthogonale de* E, *d'après 3.5.11).*

Preuve. Grâce à 3.6.2, on peut supposer que $E = F$, c'est-à-dire que la forme q est non-dégénérée. Si $n = 1$, on choisit $e_1 \in E$, $e_1 \neq \vec{0}$. Si $n > 1$, alors $q \neq 0$, donc il existe $e_1 \in E$ tel que $q(e_1) \neq 0$. Il résulte de 3.6.5(2) que l'on a $E = Ke_1 \perp E_1$, où $E_1 = (Ke_1)^{\perp}$ est un sous-espace vectoriel de E de dimension $n - 1$. La restriction de q à E_1 étant non-dégénérée (exercice : pourquoi ?), le même argument s'applique : il existe $e_2 \in E_1$ tel que $q(e_2) \neq 0$, donc $E_1 = Ke_2 \perp E_2$, où $E_2 = (Ke_2)^{\perp}$ est un sous-espace vectoriel de E de dimension $n - 2$. Si l'on répète le même argument n-fois, on obtient une décomposition orthogonale $E = Ke_1 \perp \cdots \perp Ke_n$.

(3.6.9) Exercice. *Soient* $q : E \longrightarrow K$ *une forme quadratique et* $F, G \subset E$ *des sous-espaces vectoriels; alors on a* $(F+G)^{\perp} = F^{\perp} \cap G^{\perp}$. *Lorsque* $\dim(E) < \infty$ *et* q *est non-dégénérée, alors on a* $(F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.

3.7 Formes quadratiques sur C et R

(3.7.1) Proposition. *Pour toute forme quadratique* $q : E \longrightarrow \mathbf{C}$ *sur un* $\mathbf{C}$-ev E *de dimension* $n < \infty$ *il existe une base* $e_1, \ldots, e_n$ *de* E *telle que*

$$q(x) = x_1^2 + \cdots + x_r^2 \qquad\qquad (x = \sum_{j=1}^{n} x_j e_j).$$

L'entier $r = \mathrm{rg}(q)$ *(*$0 \leq r \leq n$*) ne dépend que de* q.
Formulation matricielle : *Pour toute matrice symétrique complexe* $A = {}^tA \in M_n(\mathbf{C})$ *il existe une matrice inversible* $P \in M_n(\mathbf{C})$ *telle que* ${}^tPAP = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

Preuve. Soit $r = \mathrm{rg}(q)$. D'après Théorème 3.3.1, il existe une base $e'_1, \ldots, e'_n$ de E et des scalaires non nuls $d_1, \ldots, d_r \in \mathbf{C}$ tels que

$$q(x) = d_1 x_1'^2 + \cdots + d_r x_r'^2 \qquad\qquad (x = \sum_{j=1}^{n} x'_j e'_j).$$

Pour tout $j = 1, \ldots, r$ on choisit $\alpha_j \in \mathbf{C}$ tel que $\alpha_j^2 = d_j$. Comme $d_j x_j'^2 = (\alpha_j x'_j)^2$, le changement de coordonnées

$$x_j = \begin{cases} \alpha_j x'_j, & j = 1, \dots, r \\ x'_j, & j = r+1, \dots, n, \end{cases} \qquad e_j = \begin{cases} \frac{1}{\alpha_j} e'_j, & j = 1, \dots, r \\ e'_j, & j = r+1, \dots, n \end{cases}$$

transforme q à la forme $x_1^2 + \cdots + x_r^2$.

(3.7.2) Théorème (Sylvester). *Pour toute forme quadratique $q : E \longrightarrow \mathbf{R}$ sur un $\mathbf{R}$-ev E de dimension $n < \infty$ il existe une base $e_1, \dots, e_n$ de E telle que*

$$q(x) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2 \qquad (x = \sum_{j=1}^n x_j e_j).$$

Le couple (s, t) ne dépend que de q; on l'appelle **la signature de** *q. Le rang de q est égal á* $\mathrm{rg}(q) = s + t$.

Formulation matricielle : *Pour toute matrice symétrique réelle $A = {}^t\!A \in M_n(\mathbf{R})$ il existe une matrice inversible $P \in M_n(\mathbf{R})$ telle que* ${}^tPAP = \begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Preuve. Soit $r = \mathrm{rg}(q)$. D'après Théorème 3.3.1, il existe une base $e'_1, \dots, e'_n$ de E et des scalaires non nuls $d_1, \dots, d_r \in \mathbf{R}$ tels que

$$q(x) = d_1 x_1'^2 + \cdots + d_r x_r'^2 \qquad (x = \sum_{j=1}^n x'_j e'_j).$$

Quitte à changer l'ordre des variables, on peut supposer que l'on a

$$d_1, \dots, d_s > 0 > d_{s+1}, \dots, d_r;$$

on pose $t = r - s$. Le changement de coordonnées

$$x_j = \begin{cases} \sqrt{d_j}\, x'_j, & j = 1, \dots, s \\ \sqrt{-d_j}\, x'_j, & j = s+1, \dots, s+t = r \\ x'_j, & j = r+1, \dots, n \end{cases}$$

transforme q à la forme quadratique

$$q(x) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2.$$

Il faut encore démontrer que l'entier s (donc l'entier $t = \mathrm{rg}(q) - s$, il aussi) ne dépend que de la forme q. Soit $f_1, \dots, f_n$ une base de E telle que

$$q(x) = y_1^2 + \cdots + y_{s'}^2 - y_{s'+1}^2 - \cdots - y_r^2 \qquad (x = \sum_{j=1}^n y_j f_j).$$

Les sous-espaces vectoriels de E

$$F_1 = \mathrm{vect}(e_{s+1}, \dots, e_n), \qquad F_2 = \mathrm{vect}(f_1, \dots, f_{s'})$$

satisfont aux propriétés suivantes :

$$\left.\begin{array}{rl} \forall x \in F_1 & q(x) \le 0 \\ (\forall x \in F_2,\ x \ne \vec{0}) & q(x) > 0 \end{array}\right\} \implies F_1 \cap F_2 = \{\vec{0}\}.$$

En particulier, on a

$$s' = \dim(F_2) \leq \dim(E) - \dim(F_1) = s.$$

En échangeant les rôles des bases $\{e_i\}$ et $\{f_j\}$, le même argument montre que l'on a $s \leq s'$, d'où $s = s'$.

(3.7.3) Exemples (calcule de la signature) : On sait, grâce à la preuve du théorème de Sylvester, que l'entier s (resp., t) est égal au nombre de termes diagonaux strictement positives (resp., strictement negatives) d'une diagonalisation quelconque de la forme quadratique q.

(1) Dans 3.3.2(3), on a diagonalisé la forme quadratique

$$q(x) = x_1^2 + 4x_1x_2 - 6x_1x_3 + 3x_2^2 + 2x_2x_3 - 40x_3^2$$

de la façon suivante :

$$q(x) = (x_1 + 2x_2 - 3x_3)^2 - (x_2 - 7x_3)^2 = x_1'^2 - x_2'^2 + 0 \cdot x_3'^2,$$

donc

$$\mathrm{rg}(q) = 2, \qquad \mathrm{sign}(q) = (1, 1).$$

(2) Soient

$$E = \{\text{matrices réelles symétriques } X = {}^tX = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \in M_2(\mathbf{R})\}, \qquad q(X) = \det(X) = x_1x_3 - x_2^2.$$

On a

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1' + x_3' \\ x_2' \\ x_1' - x_3' \end{pmatrix}, \qquad q(X) = x_1x_3 - x_2^2 = x_1'^2 - x_2'^2 - x_3'^2,$$

donc

$$\mathrm{rg}(q) = 3, \qquad \mathrm{sign}(q) = (1, 2).$$

(3.7.4) Exercice. *Déterminer le rang et la signature de la forme quadratique* $q : \mathbf{R}^4 \longrightarrow \mathbf{R}$

$$q\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right) = x_1^2 + 2x_1x_2 - 4x_1x_4 + 3x_2^2 + 8x_2x_3 + 6x_3^2 - 2x_4^2.$$

(3.7.5) Exercice. *Soient* $E = M_n(\mathbf{R})$, $q(X) = \mathrm{Tr}(X^2)$ $(X \in E)$.

(1) *Déterminer la forme polaire* $f : E \times E \longrightarrow \mathbf{R}$ *de* q.

(2) *Montrer que l'on a* $E = E_- \perp E_+$, *où*

$$E_\pm = \{X \in M_n(\mathbf{R}) \mid {}^tX = \pm X\}.$$

(3) *Montrer que l'on a, pour tout élément non nul* $X \in E_\pm$, *l'inégalité* $\pm q(X) > 0$.

(4) *Déterminer le rang et la signature de* q.

[Indication : $\mathrm{Tr}(XY) = \mathrm{Tr}(YX)$.*]*

4. Les espaces euclidiens

Un **espace euclidien** est une variante abstraite de l'espace vectoriel réel $\mathbf{R}^n$ muni du produit scalaire euclidien usuel $(x \mid y) = x_1y_1 + \cdots + x_ny_n$. L'espace $\mathbf{R}^n$ admet **une base orthonormée canonique**, à savoir la base canonique, alors qu'un espace euclidien général a beaucoup de bases orthonormées, dont aucune n'est distinguée (exemple: tout sous-espace vectoriel de $\mathbf{R}^n$ est un espace euclidien). Dans le chapitre 4, $K = \mathbf{R}$.

4.1 Notions de base

(4.1.1) Définition. *Soit $q : E \longrightarrow \mathbf{R}$ une forme quadratique sur un espace vectoriel réel E. On dit que la forme q est* **définie positive** *(resp.,* **définie négative***) si l'on a $(\forall x \in E,\ x \neq \overrightarrow{0})\ \ q(x) > 0$ (resp., $q(x) < 0$). On dit que la forme q est* **positive** *(resp.,* **négative***) si l'on a $(\forall x \in E)\ \ q(x) \geq 0$ (resp., $q(x) \leq 0$). On dit que la forme q est* **indéfinie** *s'il existe $x, y \in E$ vérifiant $q(x) > 0 > q(y)$.*

(4.1.2) Exemple: Lorsque $\dim(E) = n < \infty$, alors le Théorème de Sylvester 3.7.2 entraîne qu'il existe une base de E dans laquelle

$$q(x) = x_1^2 + \cdots x_s^2 - x_{s+1}^2 - \cdots - x_{s+t}^2\ (+0 \cdot x_{s+t+1}^2 + \cdots + 0 \cdot x_n^2),$$

où $s, t \geq 0$, $\mathrm{rg}(q) = s + t \leq n$, $\mathrm{sign}(q) = (s, t)$. Il en résulte que l'on a

$$\begin{array}{rcl} q \text{ est positive} & \Longleftrightarrow & t = 0 \\ q \text{ est négative} & \Longleftrightarrow & s = 0 \\ q \text{ est indéfinie} & \Longleftrightarrow & s, t > 0 \\ q \text{ est définie positive} & \Longleftrightarrow & s = n \iff t = 0,\ \mathrm{rg}(q) = n \\ q \text{ est définie négative} & \Longleftrightarrow & t = n \iff s = 0,\ \mathrm{rg}(q) = n \end{array}$$

En particulier, une forme quadratique **définie positive** (resp., **définie négative**) est non-dégénérée (ce qui aussi résulte directement de la définition – exercice !).

(4.1.3) Définition. *Un* **espace préhilbertien** *(réel) est un* $\mathbf{R}$*-ev E muni d'une forme bilinéaire symétrique $f : E \times E \longrightarrow \mathbf{R}$ (le* **produit scalaire***) dont la forme quadratique associée est définie positive. On note $(x \mid y) = f(x, y)$. Un* **espace euclidien** *est un espace préhilbertien (réel) de dimension finie.*

(4.1.4) Exemples d'espaces préhilbertiens réels : (1) $E = \mathbf{R}^n$ muni du produit scalaire euclidien usuel

$$(x \mid y) = x_1y_1 + \cdots + x_ny_n.$$

(2) $E = \{f : [a, b] \longrightarrow \mathbf{R} \mid f \text{ continue}\} = C([a, b]; \mathbf{R})$,

$$(f \mid g) = \int_a^b f(t)g(t)\, dt.$$

(3) Tout sous-espace vectoriel d'un espace euclidien (resp., d'un espace préhilbertien) est un espace euclidien (resp., un espace préhilbertien).

(4.1.5) Bases orthonormées. Le Théorème de Sylvester entraîne (voir 4.1.2) que tout espace euclidien E (de dimension $n = \dim(E)$) admet une base $e_1, \ldots, e_n$ vérifiant

$$(x \mid x) = x_1^2 + \cdots + x_n^2 \qquad\qquad (x = \sum_{i=1}^n x_ie_i);$$

on appelle une telle base une **base orthonormée** de E. Les propriétés suivantes sont équivalentes :

$$\{e_i\} \text{ est une base orthonormée} \iff f(x,y) = (x \mid y) = x_1y_1 + \cdots + x_ny_n \qquad (y = \sum_{i=1}^{n} y_ie_i) \iff$$

$$\iff \text{la matrice du produit scalaire dans la base } \{e_i\} \text{ est égale à } I_n \iff$$

$$\iff (e_i \mid e_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

(4.1.6) Notation: Dans la suite, jusqu'à fin du chapitre 4, on considère un espace préhilbertien (réel) E sur lequel le produit scalaire est noté $(x \mid y)$ $(x, y \in E)$. En utilisant la notation de 3.2.1, on a : $x \perp y \iff (x \mid y) = 0$.

(4.1.7) Définition. *Un* **système orthonormé** *dans E est un sous-ensemble non vide $\{e_i\}$ de E tel que*

$$(e_i \mid e_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

(4.1.8) Proposition. *Si $\{e_i\}$ est un système orthonormé dans E et si $x = \sum_i \lambda_i e_i$ ($\lambda_i \in \mathbf{R}$; la somme est supposée d'être finie), alors on a*

$$\lambda_j = (x \mid e_j).$$

En particulier, les vecteurs $\{e_i\}$ sont **linéarement indépendants** *(on prend $x = \vec{0}$).*

Preuve. On a

$$(x \mid e_j) = (\sum_i \lambda_i e_i \mid e_j) = \sum_i \lambda_i \, (e_i \mid e_j) = \lambda_j.$$

(4.1.9) Plus généralement, un **système orthogonal** dans E est un sous-ensemble non vide $\{f_i\}$ de E tel que $f_i \neq \vec{0}$ et $(f_i \mid f_j) = 0$ ($\iff f_i \perp f_j$) si $i \neq j$. Dans ce cas, on a $(f_i \mid f_i) > 0$, et les vecteurs

$$e_i = \frac{f_i}{\sqrt{(f_i \mid f_i)}}$$

forment un système orthonormé. L'énoncé de 4.1.8 est remplacé par

$$x = \sum_i \lambda_i f_i \implies \lambda_j = \frac{(x \mid f_j)}{(f_j \mid f_j)},$$

car

$$(x \mid f_j) = (\sum_i \lambda_i f_i \mid f_j) = \sum_i \lambda_i \, (f_i \mid f_j) = \lambda_j \, (f_j \mid f_j),$$

ce qui entraîne (voir 4.1.8) que les vecteurs $\{f_i\}$ sont linéarement indépendants.

(4.1.10) En particulier, si $\dim(E) = n < \infty$ et $u_1, \ldots, u_n \in E$, alors on a

$$u_1, \ldots, u_n \text{ est un système orthogonal dans } E \iff u_1, \ldots, u_n \text{ est une base orthogonale de } E$$

$$u_1, \ldots, u_n \text{ est un système orthonormé dans } E \iff u_1, \ldots, u_n \text{ est une base orthonormée de } E.$$

4.2 La norme euclidienne (= la longeur)

(4.2.1) Définition. *La* **norme** *d'un vecteur $x \in E$ est le nombre réel positif $\|x\| := \sqrt{(x \mid x)} \geq 0$. En particulier, si $\dim(E) = n < \infty$ et si l'on écrit $x = \sum_{i=1}^{n} x_i e_i$ dans une base* **orthonormée** *$\{e_i\}$, alors on a $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$.*

(4.2.2) Propriétés de base de la norme : (1) $\|x\| \geq 0$, et $\|x\| = 0 \iff x = \vec{0}$.
(2) Pour tout $\lambda \in \mathbf{R}$, $\|\lambda x\| = |\lambda| \, \|x\|$.
(3) Pour tous $x, y \in E$, on a (voir 3.1.4)

$$\|x+y\|^2 = (x+y \mid x+y) = (x \mid x) + (y \mid y) + (x \mid y) + (y \mid x) = \|x\|^2 + \|y\|^2 + 2(x \mid y),$$

d'où

$$(x \mid y) = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2).$$

(4.2.3) L'inégalité de Cauchy-Schwarz. *Pour tous $x, y \in E$, on a* $|(x \mid y)| \leq \|x\|\|y\|$.

Preuve. Si $x = \vec{0}$, alors l'énoncé ne dit que $0 \leq 0$. On peut supposer, donc, que $x \neq \vec{0}$, ce qui entraîne $(x \mid x) = \|x\|^2 > 0$. Pour tout $t \in \mathbf{R}$, le nombre réel

$$\|tx + y\|^2 = (tx + y \mid tx + y) = t^2 \, (x \mid x) + 2t \, (x \mid y) + (y \mid y) \geq 0$$

est positif. On considère l'égalité

$$t^2 \, (x \mid x) + 2t \, (x \mid y) + (y \mid y) = (x \mid x) \left(t + \frac{(x \mid y)}{(x \mid x)} \right)^2 + \frac{(x \mid x)(y \mid y) - (x \mid y)^2}{(x \mid x)};$$

si l'on pose

$$t = -\frac{(x \mid y)}{(x \mid x)},$$

on obtient

$$\frac{(x \mid x)(y \mid y) - (x \mid y)^2}{(x \mid x)} \geq 0,$$

d'où (en utilisant l'inégalité $(x \mid x) > 0$)

$$(x \mid x)(y \mid y) \geq (x \mid y)^2 \implies \sqrt{(x \mid x)}\sqrt{(y \mid y)} \geq \sqrt{(x \mid y)^2} = |(x \mid y)|.$$

(4.2.4) Exercice. *L'argument précédant montre que l'on a*

$$|(x \mid y)| = \|x\| \cdot \|y\| \iff x, y \text{ sont linéairement dépendants.}$$

(4.2.5) Corollaire ("inégalité triangulaire"). *Pour tous $x, y \in E$, on a* $\|x \pm y\| \leq \|x\| + \|y\|$.

Preuve. On a

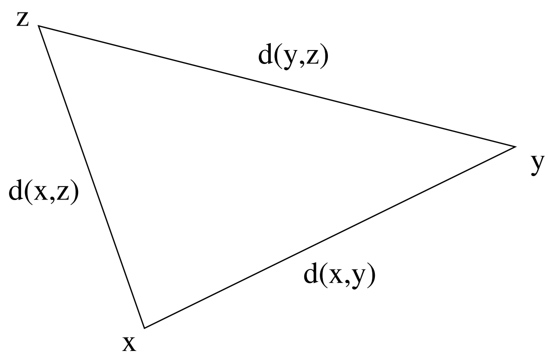
$$(\|x\| + \|y\|)^2 - (\|x \pm y\|)^2 = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 - (\|x\|^2 \pm 2\,(x \mid y) + \|y\|^2) = 2(\|x\|\|y\| \mp (x \mid y)) \geq 0.$$

(4.2.6) Distance. Si l'on définit la **distance** des (extremités de) deux vecteurs $x, y \in E$ par la formule usuelle

$$d(x, y) := \|x - y\|, \tag{4.2.6.1}$$

alors on peut reécrire l'inégalité triangulaire sous la forme plus commode (pour $x, y, z \in E$)

$$d(x,z) = \|x-z\| = \|(x-y)+(y-z)\| \leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z). \tag{4.2.6.2}$$



(4.2.7) En particulier, lorsque $E = C([a,b], \mathbf{R})$ (voir Exemple 4.1.4(2)), la distance de deux fonctions continues $f, g : [a,b] \longrightarrow \mathbf{R}$ est égale à

$$d(f,g) = \sqrt{\int_a^b (f(t)-g(t))^2\, dt}, \tag{4.2.7.1}$$

ce qui signifie que la distance $d(f,g)$ est petite si les valeurs $f(t)$ et $g(t)$ sont proches "en moyenne". Il y a plusieurs façons de mesurer la proximité dans l'espace $C([a,b], \mathbf{R})$; on peut utiliser, par exemple, les distances

$$d_\infty(f,g) = \sup_{t\in[a,b]} |f(t)-g(t)|$$

$$d_p(f,g) = \left(\int_a^b (f(t)-g(t))^p\, dt\right)^{1/p} \qquad (p \geq 1).$$

Il s'avère qu'il y a un lien etroit ("dualité") entre les distances d_p et d_q si

$$\frac{1}{p} + \frac{1}{q} = 1,$$

ce qui explique le rôle particulier joué par la distance $d(f,g) = d_2(f,g)$ "autoduale".

(4.2.8) L'angle. L'angle de deux vecteurs non nuls $x, y \in E$ est l'unique élément $\beta \in [0,\pi]$ tel que

$$\cos(\beta) = \frac{(x \mid y)}{\|x\|\|y\|}. \tag{4.2.8.1}$$

On note que le terme à droite dans (4.2.8.1) appartient bien à l'interval $[-1,1]$, grâce à l'inégalité de Cauchy-Schwarz. De plus, on a

$$\beta = \frac{\pi}{2} \iff \cos(\beta) = 0 \iff (x \mid y) = 0 \iff x \perp y.$$

4.3 Orthogonalité, orthogonalisation

(4.3.1) La méthode d'orthogonalisation de Gram-Schmidt

Soient $u_1, \ldots, u_m \in E$ $(m \geq 1)$ des vecteurs **linéairement indépendants**. On va construire un **système orthogonal** $f_1, \ldots, f_m \in E$ vérifiant

$$\mathrm{vect}(u_1) = \mathrm{vect}(f_1), \qquad \mathrm{vect}(u_1, u_2) = \mathrm{vect}(f_1, f_2), \qquad \mathrm{vect}(u_1, u_2, \ldots, u_m) = \mathrm{vect}(f_1, f_2, \ldots, f_m).$$

Si l'on pose $e_i = f_i/\|f_i\|$, on obtient un **système orthonormé** $e_1, \ldots, e_m \in E$ vérifiant

$$\mathrm{vect}(u_1) = \mathrm{vect}(e_1), \qquad \mathrm{vect}(u_1, u_2) = \mathrm{vect}(e_1, e_2), \qquad \mathrm{vect}(u_1, u_2, \ldots, u_m) = \mathrm{vect}(e_1, e_2, \ldots, e_m).$$

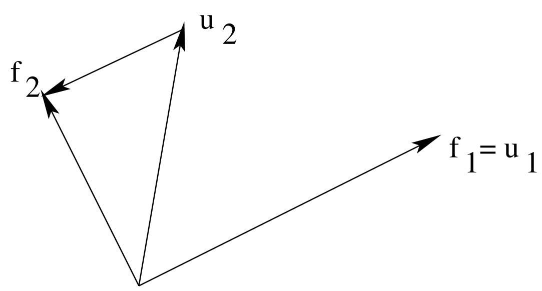
Première étape. On pose $f_1 = u_1$; l'indépendance linéaire des vecteurs $u_1, \ldots, u_m$ entraîne $f_1 \neq 0$, d'où $(f_1 \mid f_1) > 0$.

Deuxième étape. On cherche à trouver un vecteur de la forme

$$f_2 = u_2 + \lambda f_1 \qquad\qquad (\lambda \in \mathbf{R})$$

vérifiant

$$f_1 \perp f_2.$$



Comme

$$f_1 \perp f_2 \iff 0 = (f_1 \mid u_2 + \lambda f_1) = (f_1 \mid u_2) + \lambda (f_1 \mid f_1) \iff \lambda = -\frac{(f_1 \mid u_2)}{(f_1 \mid f_1)}\,,$$

on définit

$$f_2 = u_2 - \frac{(f_1 \mid u_2)}{(f_1 \mid f_1)}\, f_1.$$

La relation $u_2 = f_2 - \lambda f_1$ alors entraîne

$$\mathrm{vect}(u_1, u_2) = \mathrm{vect}(f_1, u_2) = \mathrm{vect}(f_1, f_2).$$

Troisième étape. On cherche à trouver un vecteur de la forme

$$f_3 = u_3 + \lambda_1 f_1 + \lambda_2 f_2 \qquad\qquad (\lambda_i \in \mathbf{R})$$

vérifiant

$$\mathrm{vect}(f_1, f_2) \perp f_3.$$

Le calcul précédant montre (en utilisant la relation $f_1 \perp f_2$) que l'on a

$$f_1 \perp f_3 \iff 0 = (f_1 \mid u_3 + \lambda_1 f_1 + \lambda_2 f_2) = (f_1 \mid u_3) + \lambda_1 (f_1 \mid f_1) \iff \lambda_1 = -\frac{(f_1 \mid u_3)}{(f_1 \mid f_1)}$$
$$f_2 \perp f_3 \iff 0 = (f_2 \mid u_3 + \lambda_1 f_1 + \lambda_2 f_2) = (f_2 \mid u_3) + \lambda_2 (f_2 \mid f_2) \iff \lambda_2 = -\frac{(f_2 \mid u_3)}{(f_2 \mid f_2)}\,,$$

ce qui nous conduit à définir

$$f_3 = u_3 - \frac{(f_1 \mid u_3)}{(f_1 \mid f_1)} f_1 - \frac{(f_2 \mid u_3)}{(f_2 \mid f_2)} f_2.$$

La relation $u_3 = f_3 - \lambda_1 f_1 - \lambda_2 f_2$ entraîne

$$\mathrm{vect}(u_1, u_2, u_3) = \mathrm{vect}(f_1, f_2, u_3) = \mathrm{vect}(f_1, f_2, f_3).$$

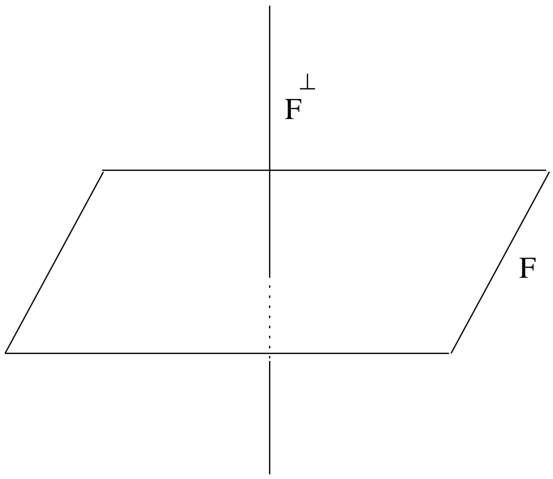
Si l'on répète la même procédure m-fois, on obtient le système orthogonal $f_1, \ldots, f_m$ requis; si l'on pose $e_i = f_i/\|f_i\|$, on obtient le système orthonormé correspondant $e_1, \ldots, e_m$.

(4.3.2) Exercice. *Appliquer la méthode de Gram-Schmidt aux éléments*

(1) $u_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$, $u_3 = \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix}$ *de* $E = \mathbf{R}^3$ *(muni du produit scalaire euclidien usuel);*

(2) $u_1 = 1$, $u_2 = t$, $u_3 = t^2$, $u_4 = t^3$ *de* $E = C([-1,1], \mathbf{R})$ *(on obtient les quatre premiers* ***polynômes de Legendre****).*

(4.3.3) Proposition-Définition("Supplémentaire orthogonal"). *Pour tout sous-espace vectoriel* $F \subset E$ *on a* $F \cap F^\perp = \{\overrightarrow{0}\}$ *(en utilisant la notation de §3.5). Si* $\dim(F) < \infty$*, alors on a* $E = F \oplus F^\perp = F \perp F^\perp$*; on dit que* $F^\perp$ *est* **le supplémentaire orthogonal de** F.



Preuve. Si $x \in F \cap F^\perp$, alors $x \perp x$, d'où $\|x\|^2 = (x \mid x) = 0$, ce qui entraîne $x = \overrightarrow{0}$. Si $\dim(E) < \infty$, alors on peut déduire la décomposition $E = F \oplus F^\perp$ du Corollaire 3.6.4. En général (en supposant que $m = \dim(F) < \infty$), il faut démontrer que l'on a $E \stackrel{?}{=} F + F^\perp$ (voir 3.4.4(2)), c'est-à-dire que tout vecteur $x \in E$ se décompose

$$x \stackrel{?}{=} y + z, \qquad y \in F, \qquad z \in F^\perp \quad (\Longrightarrow y \perp z).$$

On suppose d'abord qu'une telle décomposition existe. On choisit une base orthonormée $e_1, \ldots, e_m$ de F (une telle base existe, d'après 4.1.5). Si l'on écrit

$$y = \sum_{i=1}^m y_i e_i, \qquad z = x - y = x - \sum_{i=1}^m y_i e_i \qquad (y_i \in \mathbf{R}),$$

alors on a

$$z \in F^\perp \iff (\forall i = 1, \ldots, m) \quad e_i \perp z \iff (\forall i = 1, \ldots, m) \quad 0 = (e_i \mid z) = (e_i \mid x) - y_i.$$

Réciproquement, si l'on définit

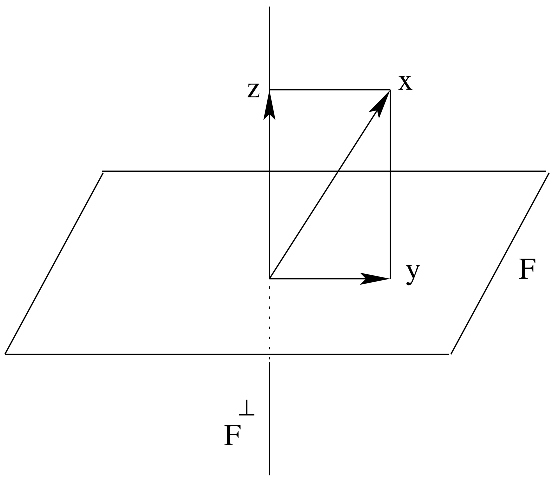
$$y := \sum_{i=1}^{m} (e_i \mid x) e_i \in F, \qquad z = x - y, \tag{4.3.3.1}$$

alors on obtient la décomposition cherchée

$$x = y + z, \qquad y \in F, \qquad z \in F^{\perp}. \tag{4.3.3.2}$$

(4.3.4) Projections orthogonales. Si $F \subset E$ est un sous-espace vectoriel de dimension finie $m = \dim(F) < \infty$, alors la preuve de 4.3.3 montre que tout vecteur $x \in E$ admet l'unique décomposition (4.3.3.2). L'application

$$\begin{aligned} p_F : E &\longrightarrow F \\ x &\mapsto y, \end{aligned}$$



qui associe à x sa "F-composante" y est linéaire; on l'appelle **la projection orthogonale de E sur F.** Pour toute base orthonormée $e_1, \dots, e_m$ de F, on a

$$y = p_F(x) = \sum_{i=1}^{m} (e_i \mid x) e_i, \tag{4.3.4.1}$$

d'après (4.3.3.1). De plus, l'orthogonalité $y \perp z$ entraîne

$$\|x\|^2 = \|y\|^2 + \|z\|^2 + 2(y \mid z) = \|y\|^2 + \|z\|^2 = \sum_{i=1}^{m} (e_i \mid x)^2 + \|z\|^2 \geq \sum_{i=1}^{m} (e_i \mid x)^2,$$

où l'inégalité devient l'égalité $\iff z = 0 \iff x = y \in F$. On vient de démontrer l'énoncé suivant :

(4.3.5) Proposition("L'inégalité de Bessel"). *Pour tout vecteur $x \in E$ et tout système orthonormé $e_1, \dots, e_m$ de E, on a*

$$\sum_{i=1}^{m} (e_i \mid x)^2 \leq \|x\|^2,$$

où l'inégalité devient l'égalité $\iff x \in \mathrm{vect}(e_1, \dots, e_m)$.

(4.3.6) En exemple en dimension infinie*

Soit E l'espace $E = C([-1,1], \mathbf{R})$ muni du produit scalaire

*** Ce paragraphe est facultatif.**

$$(f \mid g) = \int_{-1}^{1} f(t)g(t)\,dt.$$

Les fonctions

$$e_0 = \frac{1}{\sqrt{2}}, \qquad e_n = \cos(\pi n t), \qquad f_n = \sin(\pi n t) \qquad\qquad (n \geq 1)$$

forment un système orthonormé dans E (exercice !). On fixe une fonction $g \in E$ et l'on pose

$$a_0 = (g \mid 1), \qquad a_n = (g \mid \cos(\pi n t)), \qquad b_n = (g \mid \sin(\pi n t)) \qquad\qquad (n \geq 1).$$

Si l'on applique 4.3.5 au système orthonormé $e_0, e_1, \ldots, e_m, f_1, \ldots, f_m$ (où l'on a fixé un entier $m \geq 1$), on obtient

$$\frac{a_0^2}{2} + \sum_{n=1}^{m} (a_n^2 + b_n^2) \leq \|g\|^2.$$

L'entier m étant quelconque (et les termes $a_n^2 + b_n^2$ étant positifs), on en déduit l'inégalité

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \leq \|g\|^2 = \int_{-1}^{1} g(t)^2\,dt. \tag{4.3.6.1}$$

Par exemple, pour $g(t) = t$, on a

$$a_i = 0, \qquad b_n = \int_{-1}^{1} t \sin(\pi n t)\,dt = (-1)^{n-1}\,\frac{2}{\pi n}, \qquad \|t\|^2 = \frac{2}{3}, \tag{4.3.6.2}$$

donc l'inégalité (4.3.6.1) devient

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \leq \frac{\pi^2}{6}. \tag{4.3.6.3}$$

Exercice. *Soit $F = \mathrm{vect}(f_1, f_2, \ldots)$. Montrer que $g(t) = t$ ne se décompose pas sous la forme*

$$g = y + z, \qquad y \in F, \qquad z \in F^{\perp};$$

en particulier, $F + F^{\perp} \neq E$.

Que se passe-t-il ? Il faut qu'on considère des **combinaisons linéaires infinies**. Par exemple, on a

$$\text{“}t = \sum_{n=1}^{\infty} b_n\, f_n\text{”}$$

(où b_n sont les nombres évalués en (4.3.6.2)) pour la **convergence par rapport à la norme** $\|\cdot\|$ **:**

$$\lim_{m \longrightarrow \infty} \|t - \sum_{n=1}^{m} b_n\, f_n\| = 0,$$

ce qui entraîne que l'inégalité (4.3.6.3) devient l'égalité

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Plus généralement, on a, pour toute fonction $g \in C([-1, 1], \mathbf{R})$,

$$\lim_{m \longrightarrow \infty} \|g - \frac{a_0}{2} - \sum_{n=1}^{m} (a_n\, e_n + b_n\, f_n)\| = 0,$$

ce qui entraîne que l'inégalité (4.3.6.1) devient l'égalité (**l'égalité de Parseval**) :

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) = \|g\|^2.$$

4.4 Isométries, matrices orthogonales

Dans ce paragraphe, E est un espace euclidien de dimension $n \geq 1$. On va étudier les isométries de E (= les applications linéaires $E \longrightarrow E$ qui conservent les distances).

(4.4.1) Changement de base orthonormée. Nous rappelons (voir 4.1.5) qu'une base $\mathcal{B} = e_1, \ldots, e_n$ de E est **orthonormée** $\iff$ la matrice A du produit scalaire dans la base $\mathcal{B}$ est égale à la matrice identité I_n. Si c'est le cas, et si $\mathcal{B}' = e'_1, \ldots, e'_n$ est une autre base de E, alors la matrice A' du produit scalaire dans la base $\mathcal{B}'$ est égale à

$$A' = {}^tPAP = {}^tPI_nP = {}^tPP,$$

où $P \in M_n(\mathbf{R})$ est la matrice (inversible) de pasage de $\mathcal{B}$ à $\mathcal{B}'$ (d'après (2.3.6.1)). En particulier, on a

$$\mathcal{B}' \text{ est aussi orthonormée} \quad \iff \quad {}^tPP = I_n. \tag{4.4.1.1}$$

(4.4.2) Définition. *Une matrice $P \in M_n(\mathbf{R})$ est dite* **orthogonale** *si ${}^tPP = I_n$ (ce qui entraîne $\det(P)^2 = 1$, d'où $\det(P) = \pm 1$).*

(4.4.3) Reformulation: (1) Si l'on note $v_1, \ldots, v_n \in \mathbf{R}^n$ les colonnes d'une matrice $P \in M_n(\mathbf{R})$, alors ${}^tv_1, \ldots, {}^tv_n$ sont les lignes de la matrice transposée $P \in M_n(\mathbf{R})$, donc les coefficients de la matrice $Q = {}^tPP$ sont égaux au produits scalaires

$$Q_{ij} = (v_i \mid v_j).$$

En particulier, on a

$$P \text{ est } \textbf{orthogonale} \quad \iff \quad Q = I_n \quad \iff$$

$$\iff \quad \text{les colonnes } v_1, \ldots, v_n \text{ forment un } \textbf{système orthonormé} \text{ dans } \mathbf{R}^n$$

$$\iff \quad \text{les colonnes } v_1, \ldots, v_n \text{ forment une } \textbf{base orthonormée} \text{ de } \mathbf{R}^n.$$

(2) D'après (4.4.1.1), la matrice de passage entre deux bases orthonormées est orthogonale.

(4.4.4) Proposition-Définition. *Une* **isométrie de** *E est une application linéaire $u : E \longrightarrow E$ qui vérifie les conditions équivalentes suivantes :*
(1) $\forall x, y \in E \qquad (u(x) \mid u(y)) = (x \mid y)$.
(2) $\forall x \in E \qquad \|u(x)\| = \|x\|$.
(3) $\forall$ *base orthonormée* $e_1, \ldots, e_n$ *de* E, $u(e_1), \ldots, u(e_n)$ *est aussi une base orthonormée de* E.
(3') $\exists$ *base orthonormée* $e_1, \ldots, e_n$ *de* E *telle que* $u(e_1), \ldots, u(e_n)$ *soit aussi une base orthonormée de* E.
(4) $\forall$ *base orthonormée* $e_1, \ldots, e_n$ *de* E, *la matrice* U *de* u *dans la base* $\{e_i\}$ *est orthogonale:* ${}^tUU = I_n$.
(4') $\exists$ *base orthonormée* $e_1, \ldots, e_n$ *of* E *telle que la matrice* U *de* u *dans la base* $\{e_i\}$ *soit orthogonale:* ${}^tUU = I_n$.
[En particuler, les isométries de $\mathbf{R}^n$ (muni du produit scalaire euclidien usuel) sont les applications linéaires $X \mapsto UX$, où $U \in M_n(\mathbf{R})$ est une matrice orthogonale.]

Preuve. Il s'agit de montrer que les conditions sont équivalentes. Les implications (3) $\Longrightarrow$ (3') et (4) $\Longrightarrow$ (4') sont automatiques.

(1) $\Longrightarrow$ (2): La condition (1) entraîne $\|u(x)\|^2 = (u(x) \mid u(x)) \overset{(1)}{=} (x \mid x) = \|x\|^2$.
(2) $\Longrightarrow$ (1): La condition (2) entraîne

$$\begin{aligned}2(u(x) \mid u(y)) &= \|u(x)+u(y)\|^2 - \|u(x)\|^2 - \|u(y)\|^2 = \|u(x+y)\|^2 - \|u(x)\|^2 - \|u(y)\|^2 \overset{(2)}{=} \\ &= \|x+y\|^2 - \|x\|^2 - \|y\|^2 = 2(x \mid y).\end{aligned}$$

(1) $\Longrightarrow$ (3): La condition (1) entraîne $(u(e_i) \mid u(e_j)) = (e_i \mid e_j) = \delta_{ij}$.
(3') $\Longrightarrow$ (1): La condition (3') entraîne $(e_i \mid e_j) = (u(e_i) \mid u(e_j)) = \delta_{ij}$, pour tous $i, j = 1, \ldots, n$. Si $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{j=1}^n y_j e_j$ sont des vecteurs quelconques de E, alors on a

$$(u(x) \mid u(y)) = \left(\sum_{i=1}^n x_i u(e_i) \,\middle|\, \sum_{j=1}^n y_j u(e_j) \right) = \sum_{i,j=1}^n x_i y_j \, (u(e_i) \mid u(e_j)) = \sum_{i=1}^n x_i y_i = (x \mid y).$$

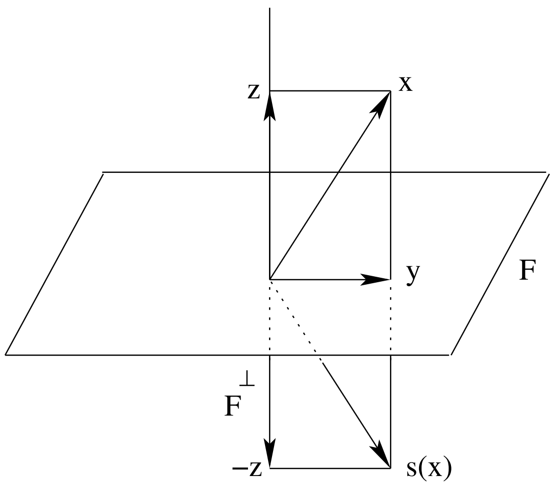
(4') $\Longrightarrow$ (1) $\Longrightarrow$ (4): Soit $\mathcal{B} = e_1, \ldots, e_n$ une base orthonormée de E. Si l'on associe au vecteur $x = \sum_{i=1}^n x_i e_i \in E$ le vecteur colonne $X = {}^t(x_1, \ldots, x_n) \in \mathbf{R}^n$ (et de même pour $y \in E$), alors l'application u (resp., le produit scalaire) s'écrit $X \mapsto UX$ (resp., $(x \mid y) = {}^tXY$), où l'on a noté $U \in M_n(\mathbf{R})$ la matrice de u dans la base $\mathcal{B}$. Les implications cherchées (4') $\Longrightarrow$ (1) $\Longrightarrow$ (4) alors résultent de la formule

$$(u(x) \mid u(y)) = {}^t(UX)\, UY = {}^tX\, {}^tUUY.$$

(4.4.5) Exemple (symétries orthogonales): (1) Pour tout sous-espace vectoriel $F \subset E$, la décomposition orthogonale $E = F \perp F^\perp$ (voir 4.3.3) définit une application linéaire

$$\begin{aligned} s = s_F : E &\longrightarrow E \\ y + z &\mapsto y - z \end{aligned} \qquad (y \in F,\ z \in F^\perp),$$

que l'on appelle **la symétrie orthogonale par rapport à** F.



Cette application est une isométrie, puisqu'on a

$$\begin{aligned}(x \mid x') &= (y + z \mid y' + z') = (y \mid y') + (z \mid z') + (y \mid z') + (z \mid y') = (y \mid y') + (z \mid z') \\ (s(x) \mid s(x')) &= (y - z \mid y' - z') = (y \mid y') + (z \mid z') - (y \mid z') - (z \mid y') = (y \mid y') + (z \mid z')\end{aligned}$$

pour tous $y, y' \in F$, $z, z' \in F^\perp$ et $x = y + z$, $x' = y' + z'$.
(2) Les sous-espaces $F, F^\perp \subset E$ sont égaux aux sous-espaces

$$F = \{x \in E \mid s(x) = x\} = \mathrm{Ker}(s - \mathrm{Id}), \qquad F^\perp = \{x \in E \mid s(x) = -x\} = \mathrm{Ker}(s + \mathrm{Id}).$$

(3) **Lien avec des projections orthogonales :** en utilisant la notation de 4.3.4, on a

$$s_F(x) = y - z = 2y - (y + z) = 2p_F(x) - x = (y + z) - 2z = x - 2p_{F^\perp}(x).$$

(4) **Représentation matricielle :** soit $e_1, \ldots, e_m$ (resp., $e_{m+1}, \ldots, e_n$) une base orthonormée de F (resp., de $F^\perp$); alors $\mathcal{B} = e_1, \ldots, e_n$ est une base orthonormée de E. Comme

$$s(e_i) = \begin{cases} e_i, & i \leq m \\ -e_i, & i > m, \end{cases}$$

la matrice de s dans la base $\mathcal{B}$ est égale à

$$\begin{pmatrix} I_m & 0 \\ 0 & -I_{n-m} \end{pmatrix}.$$

(5) **Réflexions :** si $\dim(F^\perp) = 1$ ($\iff$ $\dim(F) = n - 1$), on dit que s_F est une **réflexion** par rapport à l'hyperplan F. Si $u \in F^\perp$, $u \neq \vec{0}$, alors $e_n := u/\|u\|$ est une base orthonormée de la droite $F^\perp$. La formule (4.3.4.1) (appliquée à l'espace $F^\perp$ au lieu de F) alors s'écrit

$$p_{F^\perp}(x) = (x \mid e_n)e_n = \frac{(x \mid u)u}{(u \mid u)},$$

d'où

$$s_F(x) = x - 2p_{F^\perp}(x) = x - \frac{2(x \mid u)u}{(u \mid u)}.$$

Si $e_1, \ldots, e_{n-1}$ est une base orthonormée de F, alors la matrice de s dans la base (orthonormée) $e_1, \ldots, e_n$ de E est égale à

$$\begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix};$$

en particulier, $\det(s_F) = -1$.

(4.4.6) Exercice. *Décrire géométriquement des symétries orthogonales dans* $\mathbf{R}^3$ *(pour toutes les valeurs possible* $\dim(F) = 0, 1, 2, 3$ *de la dimension de* $F \subset \mathbf{R}^3$*).*

(4.4.7) Isométries - formulation matricielle. Si l'on fixe une base orthonormée quelconque $e_1, \ldots, e_n$ de E, alors E s'identifie à $\mathbf{R}^n$, et le produit scalaire au produit scalaire euclidien usuel. De plus, les isométries de E s'identifient aux isométries de $\mathbf{R}^n$, i.e. aux applications $X \mapsto UX$, où $U \in M_n(\mathbf{R})$ est une matrice orthogonale (d'après 4.4.4).

Notation: Pour tout $n \geq 1$, on note

$$\begin{aligned} O(n) &= \{U \in M_n(\mathbf{R}) \mid {}^tUU = I_n\} = \{\text{isométries de } \mathbf{R}^n\} = O^+(n) \cup O^-(n) \\ SO(n) = O^+(n) &= \{U \in O(n) \mid \det(U) = +1\} \\ O^-(n) &= \{U \in O(n) \mid \det(U) = -1\}. \end{aligned}$$

On appelle $O(n)$ (resp., $SO(n) = O^+(n)$) le **groupe orthogonal** (resp., le **groupe orthogonal spécial**) de $\mathbf{R}^n$.

(4.4.8) Exemples: (1) $O(1) = \{\pm 1\}$, $\quad SO(1) = \{1\}$.

(2) On a (voir 4.5.5-6 ci-dessous)

$$SO(2) = O^+(2) = \{\text{rotations dans } \mathbf{R}^2 \text{ autour d'origine}\}$$
$$O^-(2) = \{\text{réflexions dans } \mathbf{R}^2\}.$$

(3) On a (voir 4.6.5 ci-dessous)

$$SO(3) = O^+(3) = \{\text{rotations dans } \mathbf{R}^3 \text{ autour d'un axe contenant l'origine}\}.$$

4.5 Isométries de $\mathbf{R}^2$

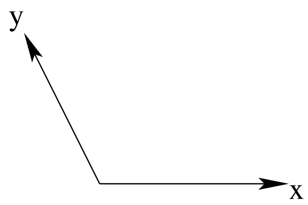
Dans ce paragraphe, $E = \mathbf{R}^2$ est muni du produit scalaire euclidien usuel $(x \mid y) = x_1y_1 + x_2y_2$.

(4.5.1) Les angles orientés dans $\mathbf{R}^2$. Soient $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ des vecteurs non nuls de $\mathbf{R}^2$. On peut raffiner la définition générale (voir (4.2.8.1)) de l'angle non orienté $\beta \in [0, \pi]$ de x et y (qui ne dépend de l'ordre de x et y) de la façon suivante. On pose

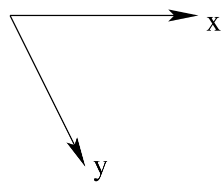
$$A = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

et l'on définit **l'angle orienté entre les vecteurs x et y** (l'ordre de x et y est important) par la formule

$$\widehat{xy} = \gamma = \begin{cases} \beta, & \text{si } A \geq 0 \quad (\gamma \in [0, \pi]) \\ -\beta, & \text{si } A < 0 \quad (\gamma \in (-\pi, 0)) \end{cases}$$



A > 0



A < 0

Il est plus commode de considérer un angle orienté comme une "classe de nombres réels modulo 2π", plutôt qu'un nombre réel qui appartient à l'interval $(-\pi, \pi]$. Ce point de vue signifie que les angles orientés sont représentés par des nombres réels quelconques, et que deux nombres réels déterminent le même angle orienté $\iff$ leur différence est un multiple entier de 2π (par exemple, -5π et π représentent le même angle orienté). Tout angle orienté admet l'unique représentant qui appartient à $(-\pi, \pi]$.

(4.5.2) Description géométrique des isométries de $\mathbf{R}^2$

Rotations (autour d'origine) :

Pour tout angle orienté α, la rotation (dans le sens inverse des aiguilles d'une montre) autour d'origine de l'angle α est l'unique isométrie $r : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ vérifiant

$$(\forall x \in \mathbf{R}^2,\ x \neq \vec{0}) \qquad \text{l'angle orienté entre } x \text{ et } r(x) \text{ est égal à } \alpha.$$

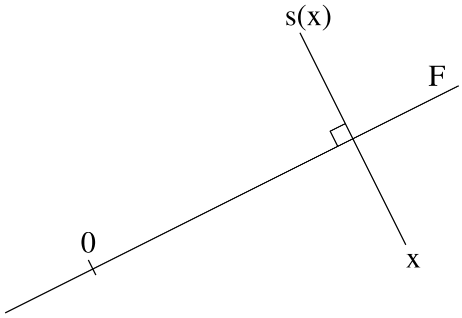
Réflexions :

Pour toute droite F qui passe par l'origine, la réflexion par rapport à F (voir 4.4.5(5)) est une isométrie $s = s_F : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ vérifiant

$$\forall x \in \mathbf{R}^2 \quad s(s(x)) = x, \qquad s \circ s = \mathrm{Id}.$$

La droite F (= **l'axe de** s") est égale à

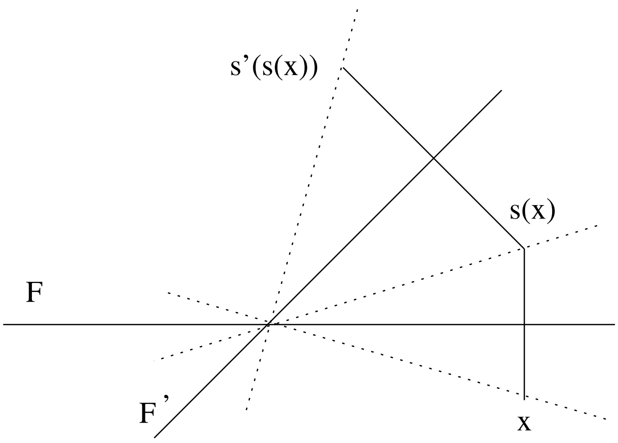
$$F = \{x \in \mathbf{R}^2 \mid s(x) = x\} = \mathrm{Ker}(s - \mathrm{Id}).$$



(4.5.3) Composée de deux réflexions = rotation
Soient $F, F' \subset \mathbf{R}^2$ des droites qui passent par l'origine; on note $s = s_F$, $s' = s_{F'}$ les réflexions correspondantes. On fixe des vecteurs non nuls $u \in F$ et $u' \in F'$. Si α est l'angle orienté entre u et u', alors

$$s' \circ s = r$$

est la rotation autour d'origine d'angle (orienté) 2α.



De même,

$$s \circ s' = s^{-1} \circ (s')^{-1} = (s' \circ s)^{-1} = r^{-1}$$

est la rotation autour d'origine d'angle (orienté) -2α.

(4.5.4) Proposition. *On a*

$$O(2) = \left\{ \begin{pmatrix} a & -\lambda c \\ c & \lambda a \end{pmatrix} \,\middle|\, \lambda = \pm 1,\ a, c \in \mathbf{R},\ a^2 + c^2 = 1 \right\}, \qquad \det \begin{pmatrix} a & -\lambda c \\ c & \lambda a \end{pmatrix} = \lambda.$$

Preuve. Soient $a, b, c, d \in \mathbf{R}$. On a

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2) &\iff \begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix} \text{ est une base orthonormée de } \mathbf{R}^2 \\ &\iff a^2 + c^2 = b^2 + d^2 = 1,\ \begin{pmatrix} b \\ d \end{pmatrix} \in \left(\mathrm{vect}\left(\begin{pmatrix} a \\ c \end{pmatrix}\right)\right)^{\perp} = \mathrm{vect}\left(\begin{pmatrix} -c \\ a \end{pmatrix}\right) \\ &\iff a^2 + c^2 = 1,\ \begin{pmatrix} b \\ d \end{pmatrix} = \lambda \begin{pmatrix} -c \\ a \end{pmatrix},\ \lambda^2 = 1. \end{aligned}$$

(4.5.5) Proposition. (1) *On a*

$$SO(2) = O^+(2) = \left\{ \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \,\middle|\, a, c \in \mathbf{R},\ a^2 + c^2 = 1 \right\} = \left\{ R(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \,\middle|\, \alpha \in \mathbf{R} \right\}.$$

(2) *L'isométrie définie par la matrice $R(\alpha)$ est la rotation autour d'origine d'angle (orienté) α.*

Preuve. (1) La première égalité résulte de 4.5.4; la seconde du fait que les points du cercle $a^2 + c^2 = 1$ s'écrivent $a = \cos(\alpha)$, $c = \sin(\alpha)$ ($\alpha \in \mathbf{R}$).
(2) Comme $\sin(\alpha + 2\pi) = \sin(\alpha)$ (et de même pour cos), la matrice $R(\alpha)$ ne dépend que de "la classe de α modulo 2π" (voir 4.5.1), donc on peut supposer que l'on a $\alpha \in (-\pi, \pi]$.
Pour tout vecteur $x = \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbf{R}^2$ on a $\|R(\alpha)x\| = \|x\| = \sqrt{u^2 + v^2}$ (puisque la multiplication par $R(\alpha)$ est une isométrie). Soit $\beta \in [0, \pi]$ l'angle (non orienté) de x et $R(\alpha)x$ (voir 4.2.8); il résulte de

$$(x \mid R(\alpha)x) = (\begin{pmatrix} u \\ v \end{pmatrix} \,\middle|\, \begin{pmatrix} au - cv \\ cu + av \end{pmatrix}) = u(au - cv) + v(cu + av) = a(u^2 + v^2)$$

que l'on a

$$\cos(\beta) = \frac{(x \mid R(\alpha))}{\|x\| \|R(\alpha)x\|} = a = \cos(\alpha),$$

d'où $\alpha = \pm\beta$. D'autre part, l'angle orienté entre x et $R(\alpha)x$ est égal à

$$\gamma = \begin{cases} \beta, & \text{si } A \geq 0 \\ -\beta, & \text{si } A < 0, \end{cases}$$

où

$$A = \begin{vmatrix} u & au - cv \\ v & cu + av \end{vmatrix} = c(u^2 + v^2) = (u^2 + v^2)\sin(\alpha);$$

on en déduit

$$\begin{aligned} A \geq 0 &\iff \sin(\alpha) \geq 0 \iff \alpha \in [0, \pi] \iff \alpha = \beta \\ A < 0 &\iff \sin(\alpha) < 0 \iff \alpha \in (-\pi, 0) \iff \alpha = -\beta, \end{aligned}$$

d'où $\gamma = \alpha$.

(4.5.6) Proposition. *On a*

$$O^-(2) = \left\{ \begin{pmatrix} a & c \\ c & -a \end{pmatrix} \,\middle|\, a, c \in \mathbf{R},\ a^2 + c^2 = 1 \right\} = \left\{ g \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \,\middle|\, g \in O^+(2) \right\} = \{\text{réflexions dans } \mathbf{R}^2\}.$$

Preuve. Les deux premières égalités résultent de 4.5.4-5. On sait que $\{\text{réflexions dans } \mathbf{R}^2\} \subset O^-(2)$, d'après 4.4.5(5). Réciproquement, l'application linéaire

$$s : \mathbf{R}^2 \longrightarrow \mathbf{R}^2, \qquad s : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} u \\ -v \end{pmatrix},$$

qui est représentée par la matrice

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in O^-(2),$$

est une réflexion par rapport à l'axe horisontal $F = \mathrm{vect}(u)$, $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Pour toute matrice

$$g = \begin{pmatrix} a & c \\ c & -a \end{pmatrix} \in O^-(2) \qquad\qquad (a^2 + c^2 = 1),$$

on a

$$g = \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = R(\alpha)S.$$

Soit $u' \in \mathbf{R}^2$ un vecteur non nul tel que l'angle orienté entre u et u' soit égal à $\alpha/2$. Si l'on note $S' \in O^-(2)$ la matrice qui représente la réflexion de l'axe F', alors on a $S'S = R(\alpha)$, d'après 4.5.3. On en déduit que la matrice

$$g = R(\alpha)S = S'SS = S'$$

représente une réflexion, d'où $\{$réflexions dans $\mathbf{R}^2\} = O^-(2)$.

(4.5.7) Formulaire. Pour simplifier, on identifie toute matrice $g \in O(2)$ à l'isométrie de $\mathbf{R}^2$ représentée par g.

(1) $R(\alpha)R(\beta) = R(\alpha + \beta) = R(\beta)R(\alpha)$, $R(\alpha)^{-1} = R(-\alpha)$.

(2) Pour toute réflexion $s \in O^-(2)$ et toute rotation $r = R(\alpha) \in O^+(2)$, on a $s^2 = I$ ($\iff$ $s^{-1} = s$) et

$$srs^{-1} = sR(\alpha)s = R(-\alpha) = R(\alpha)^{-1} = r^{-1}.$$

Pour démontrer cette formule, on peut faire un dessin (exercice !), ou l'on écrit

$$s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} R(\beta), \qquad sR(\alpha)s^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} R(\beta)R(\alpha)R(-\beta) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} =$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} R(\alpha) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = R(-\alpha) = R(\alpha)^{-1}.$$

(3) Les valeurs propres de la matrice $R(\alpha)$ sont égales à $\cos(\alpha) \pm i\sin(\alpha) = e^{\pm i\alpha}$ (un énoncé plus général est démontré en 4.5.9(5) ci-dessous).

(4) Les valeurs propres d'une réflexion $s \in O^-(2)$ sont égales à ± 1; l'axe de s est égal à $\mathrm{Ker}(s - I)$ (d'après 4.4.5(5) et 4.4.5(2)).

(4.5.8) Formules complexes. Il existe une identification naturelle ($\mathbf{R}$-linéaire) de $\mathbf{R}^2$ à $\mathbf{C}$:

$$\mathbf{R}^2 \longrightarrow \mathbf{C}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x + iy. \tag{4.5.8.1}$$

Pour tout nombre complex $w = u + iv$, la multiplication par w

$$z = x + iy \mapsto wz = (u + iv)(x + iy) = (ux - vy) + i(vx + uy)$$

définit, en utilisant l'identification (4.5.8.1), une application $\mathbf{R}$-linéaire $\mathbf{R}^2 \longrightarrow \mathbf{R}^2$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} ux - vy \\ vx + uy \end{pmatrix} = \begin{pmatrix} u & -v \\ v & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix};$$

on note

$$M(w) = M(u+iv) = \begin{pmatrix} u & -v \\ v & u \end{pmatrix} \in M_2(\mathbf{R}) \tag{4.5.8.2}$$

la matrice représentative de cette multiplication.

(4.5.9) Propriétés des matrices $M(w)$

(1) $M(w+w') = M(w) + M(w')$ (car $(w+w')z = wz + w'z$);
(2) $M(ww') = M(w)M(w')$ (car $(ww')z = w(w'z)$);
(3) $\forall w \in \mathbf{R} \quad M(w) = \begin{pmatrix} w & 0 \\ 0 & w \end{pmatrix}$;
(4) Lorsque $\alpha \in \mathbf{R}$ et $w = e^{i\alpha} = \cos(\alpha) + i\sin(\alpha)$, alors on a $M(e^{i\alpha}) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} = R(\alpha)$;
(5) Les valeurs propres de la matrice $M(u+iv)$ sont les racines du polynôme $\lambda^2 - 2u\lambda + (u^2+v^2) = 0$, à savoir $\lambda_{1,2} = u \pm iv$ (en particulier, les valeurs propres de la matrice $R(\alpha)$ sont égales à $\cos(\alpha) \pm i\sin(\alpha) = e^{\pm i\alpha}$).

(4.5.10) Les nombres complexes et les isométries de $\mathbf{R}^2$. Il résulte de 4.5.8-9 (et de 4.5.5-6) que les isométries de $\mathbf{R}^2$ s'écrivent sous la forme complexe suivante (en utilisant l'identification (4.5.8.1)) :

$$\text{rotations :} \qquad z \mapsto e^{i\alpha}z; \qquad \text{réflexions :} \qquad z \mapsto e^{i\alpha}\overline{z}. \tag{4.5.10.1}$$

On déduit la seconde formule du fait que l'action du conjugué complex $z = x + iy \mapsto \overline{z} = x - iy$ correspond à l'action de la matrice $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(4.5.11) Exercice. (1) *Lesquelles, parmi les matrices suivantes, sont orthogonales ?*

$$\frac{1}{2}\begin{pmatrix} \sqrt{3} & 1 \\ 1 & -\sqrt{3} \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & -\sqrt{3} \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ -1 & -\sqrt{3} \end{pmatrix}, \quad \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ -1 & \sqrt{3} \end{pmatrix}$$

(2) *Pour toute matrice orthogonale de (1), décrire géométriquement l'isométrie correspondante de* $\mathbf{R}^2$.
(3) *Écrire les isométries de (2) sous la forme complexe (4.5.10.1).*

(4.5.12) Exercice. *Déterminer les vecteurs propres de la matrice $M(u+iv)$. Cette matrice est-elle diagonalisable ?*

4.6 Isométries de $\mathbf{R}^n$

(4.6.1) Proposition. *Toute valeur propre $\lambda \in \mathbf{C}$ d'une matrice orthogonale $U \in O(n)$ vérifie $|\lambda| = 1$.*

Preuve. Voir 6.5.6 ci-dessous.

(4.6.2) Théorème. *Pour toute isométrie $u : E \longrightarrow E$ d'un espace euclidien E de dimension $\dim(E) = n$ il existe une base* **orthonormée** *dans laquelle la matrice de u soit égale à*

$$\begin{pmatrix} I_a & 0 & 0 & \cdots & 0 \\ 0 & -I_b & 0 & \cdots & 0 \\ 0 & 0 & R(\alpha_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R(\alpha_m) \end{pmatrix}$$

où $a+b+2m=n$ et α_j n'est pas un multiple entier de π. En particulier, $\det(u)=(-1)^b$.
Formulation matricielle : *Pour toute matrice* **orthogonale** *$U\in O(n)$ il existe une matrice* **orthogonale** *$P\in O(n)$ telle que $P^{-1}UP$ soit de la forme ci-dessus.*

(4.6.3) Théorème. *Toute matrice* **orthogonale** *$U\in O(n)$ s'écrit $U=s_1\cdots s_r$, où $r\le n$ et chaque s_j est une réflexion.*

(4.6.4) Exemple ($n=2$): Si $U\in O^-(2)$, alors $U=s_1$ est une réflexion. Si $U\in O^+(2)$, alors U est une rotation, donc U est égale au produit de deux réflexions $U=s_1s_2$.

(4.6.5) Exemple ($n=3$): Si $U\in O(3)$, alors il existe un changement de base orthonormée (défini par une matrice $P\in O(3)$) tel que $P^{-1}UP$ s'écrive sous la forme suivante :

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},\qquad \begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix},\qquad \begin{pmatrix}1&0&0\\0&\cos(\beta)&-\sin(\beta)\\0&\sin(\beta)&\cos(\beta)\end{pmatrix},\qquad \begin{pmatrix}-1&0&0\\0&\cos(\beta)&-\sin(\beta)\\0&\sin(\beta)&\cos(\beta)\end{pmatrix}$$

(où β n'est pas un multiple entier de 2π). On peut décrire géométriquement l'isométrie de $\mathbf{R}^3$ représentée par U selon les quatre classes de $P^{-1}UP$ ci-dessus :

(0) L'application identité.
(1) Une réflexion par rapport au plan $\mathrm{Ker}(U-I_3)$.
(2) Une rotation d'axe $\mathrm{Ker}(U-I_3)$ et d'angle $\beta\neq 0$ (déterminé à un signe près) vérifiant $\mathrm{Tr}(U)=1+2\cos(\beta)$.
(3) Un produit de trois réflexions (qui n'est pas une réflexion).

De plus, U appartient à la classe $n=0,1,2,3 \iff \mathrm{rg}(U-I_3)=n \iff U$ est un produit de n, mais pas de moins que n, réflexions.

(4.6.6) Exercice. (1) *Montrer que les matrices suivantes sont orthogonales :*

$$\frac{1}{3}\begin{pmatrix}1&-2&-2\\-2&1&-2\\-2&-2&1\end{pmatrix},\qquad \frac{1}{3}\begin{pmatrix}1&-2&2\\-2&1&2\\-2&-2&-1\end{pmatrix},\qquad \frac{1}{3}\begin{pmatrix}1&2&2\\-2&-1&2\\-2&2&-1\end{pmatrix},\qquad \frac{1}{3}\begin{pmatrix}-1&2&2\\2&-1&2\\2&2&-1\end{pmatrix}.$$

(2) *Pour toute matrice de (1), décrire géométriquement l'isométrie de $\mathbf{R}^3$ correspondante (en utilisant 4.6.5).*

(4.6.7) Exercice. *Soit $U\in O(3)$.* (1) *Montrer que $\lambda=\det(U)$ est une valeur propre de U.*
(2) *Soit $x\in\mathbf{R}^3$ un vecteur propre de U de valeur propre λ. Montrer que le plan $x^\perp$ vérifie $U(x^\perp)=x^\perp$.*
(3) *En déduire l'énoncé de 4.6.5.*

5. Diagonalisation "difficile" d'une forme quadratique réelle

5.1 Le résultat principal

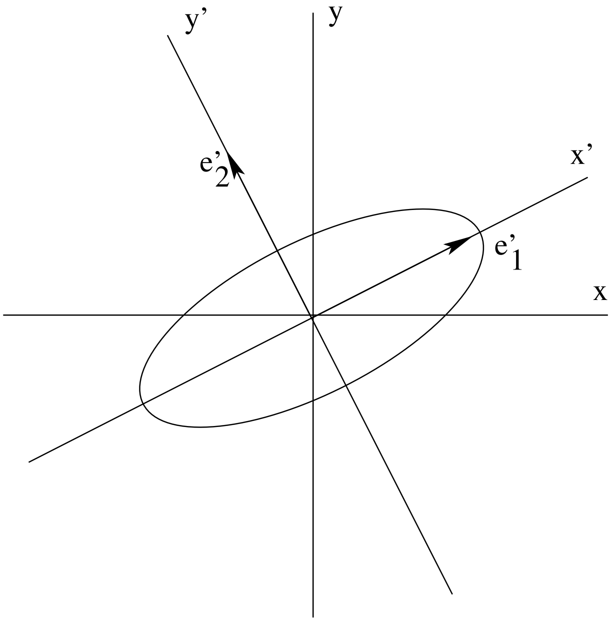
(5.1.1) Exemple: La forme quadratique

$$q : \mathbf{R}^2 \longrightarrow \mathbf{R}, \qquad q(\begin{pmatrix} x \\ y \end{pmatrix}) = 2x^2 - 4xy + 5y^2$$

est définie positive; l'équation

$$q = 2x^2 - 4xy + 5y^2 = 1 \tag{5.1.1.1}$$

définit une **ellipse :**



On cherche à trouver une **rotation** $P \in O^+(2)$ qui diagonalise la forme quadratique q, i.e. qui transforme l'équation (5.1.1.1) à une équation diagonale

$$d_1 x'^2 + d_2 y'^2 = 1, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

La matrice de la forme q dans la base canonique $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ est égale à $A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$. Il s'agit de trouver une matrice $P \in M_2(\mathbf{R})$ vérifiant

$$^tPP = I_2, \qquad \det(P) = 1 \tag{5.1.1.2}$$

$$^tPAP = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} = D. \tag{5.1.1.3}$$

Comme $^tP = P^{-1}$, la condition (5.1.1.3) équivaut à

$$P^{-1}AP = D \iff (Ae'_1 \mid Ae'_2) = AP = PD = (d_1e'_1 \mid d_2e'_2), \tag{5.1.1.4}$$

où l'on a noté e'_1, e'_2 les colonnes de la matrice $P = (e'_1 \mid e'_2)$. Autrement dit, (5.1.1.4) signifie que e'_1, e'_2 sont des **vecteurs propres** de A, et d_1, d_2 sont les **valeurs propres** correspondantes (voir 1.6.6). De plus, la première condition de (5.1.1.2) entraîne que les vecteurs e'_1, e'_2 forment une **base orthonormée** de $\mathbf{R}^2$.

On va déterminer les valeurs propres et les vecteurs propres de A. Les racines du polynôme caractéristique de A

$$P_A(t) = t^2 - \mathrm{Tr}(A)t + \det(A) = t^2 - 7t + 6 = (t-1)(t-6)$$

sont égales à $\lambda_1 = 1$ et $\lambda_2 = 6$. On peut résoudre les équations

$$(A - I)e'_1 = 0, \qquad \|e'_1\| = 1, \qquad (A - 6I)e'_2 = 0, \qquad \|e'_2\| = 1,$$

sans difficulté; on obtient

$$e'_1 = \pm\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad e'_2 = \pm\frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

On a $e'_1 \perp e'_2$ (voir 6.5.6 ci-dessous), donc tous les choix possibles des signes donnent lieu à une matrice **orthogonale** $P \in O(2)$. Il faut choisir les signes pour que l'on ait $\det(P) = 1$; par exemple, on peut prendre

$$P = (e'_1 \mid e'_2) = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix},$$

ce qui définit une **rotation** $P \in O^+(2)$ vérifiant

$${}^tPAP = P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}.$$

On peut aussi faire le calcul à la main :

$$\begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix} = x'e'_1 + y'e'_2, \qquad x = \frac{2x' - y'}{\sqrt{5}}, \qquad y = \frac{x' + 2y'}{\sqrt{5}}$$

$$2x^2 - 4xy + 5y^2 = \frac{2(4x'^2 - 4x'y' + y'^2) - 4(2x' - y')(x' + 2y') + 5(x'^2 + 4x'y' + 4y'^2)}{5} = x'^2 + 6y'^2.$$

Voici le résultat général.

(5.1.2) Théorème. *Toute matrice* **réelle symétrique** $A = {}^tA \in M_n(\mathbf{R})$ *admet n vecteurs propres $e'_1, \ldots, e'_n \in \mathbf{R}^n$ qui forment une* **base orthonormée** *de $\mathbf{R}^n$ (par rapport au produit scalaire euclidien usuel). La matrice $P = (e'_1 \mid \cdots \mid e'_n) \in M_n(\mathbf{R})$ est orthogonale (${}^tPP = I_n$) est l'on a*

$${}^tPAP = P^{-1}AP = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \qquad\qquad (Ae'_j = \lambda_j e'_j,\ \lambda_j \in \mathbf{R}).$$

L'isométrie $X = PX'$ transforme la forme quadratique $q(x) = {}^tXAX$ à une forme quadratique diagonale

$${}^t(PX')A(PX') = {}^tX'\,{}^tPAPX' = \lambda_1 x'^2_1 + \cdots + \lambda_n x'^2_n.$$

(5.1.3) Remarques. Soit $A = {}^tA \in M_n(\mathbf{R})$ une matrice réelle symétrique.

(1) Quitte à remplacer un vecteur propre e'_j par $-e'_j$, on peut trouver P telle que l'on ait $\det(P) = 1$ ($\Longleftrightarrow P \in O^+(n)$).
(2) Toutes les valeurs propres de A sont réelles (voir 6.5.6 ci-dessous).
(3) Les vecteurs propres de A qui correspondent aux valeurs propres distinctes sont orthogonaux par rapport au produit scalaire euclidien usuel (voir 6.5.6 ci-dessous).
(4) Si la matrice A a n valeurs propres distinctes, alors l'énoncé de Théorème 5.1.2 résulte de (2) et (3).
(5) Si la matrice A a des valeurs propres multiples, alors on déduit de (2) et (3) que Théorème 5.1.2 équivaut à la diagonalisabilité de la matrice A (dans ce cas on choisit une base orthonormée de chaque sous-espace propre $\mathrm{Ker}(A - \lambda_i I)$, et ensuite on prend leurs réunion).
(6) Afin de démontrer Théorème 5.1.2, on le traduit à une forme abstraite.

(5.1.4) Exercice. *Trouver une rotation $P \in O^+(3)$ qui diagonalise la forme quadratique*

$$q : \mathbf{R}^3 \longrightarrow \mathbf{R}, \qquad q(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = x^2 - 16xy + 8xz + y^2 + 8yz + 7z^2.$$

5.2 Adjoint

Il faut d'abord trouver uve variante abstraite de la condition de symétrie $A = {}^tA$.

(5.2.1) Exemple: Le produit scalaire euclidien usuel $(x \mid y) = x_1y_1 + \cdots + x_ny_n$ sur $\mathbf{R}^n$ satisfait à la propriété suivante :

$$\forall M \in M_n(\mathbf{R}) \quad \forall x, y \in \mathbf{R}^n \qquad (Mx \mid y) = {}^t(Mx)y = {}^tx\,{}^tMy = (x \mid {}^tMy).$$

Le passage à la matrice transposée admet la variante abstraite suivante.

(5.2.2) Proposition-Définition. *Soit E un espace euclidien. Pour toute application linéaire $f : E \longrightarrow E$ il existe l'unique application linéaire ${}^tf : E \longrightarrow E$ (on l'appelle* **l'adjoint de** *f) vérifiant*

$$\forall x, y \in E \qquad (f(x) \mid y) = (x \mid {}^tf(y)).$$

Pour toute base **orthonormée** *$\mathcal{B}$ de E, on a*

$$(\text{la matrice de } {}^tf \text{ dans la base } \mathcal{B}) = {}^t(\text{la matrice de } f \text{ dans la base } \mathcal{B}).$$

Preuve. On fixe une base orthonormée quelconque de E; soit $M \in M_n(\mathbf{R})$ la matrice de f dans $\mathcal{B}$. Si l'on associe à tout vecteur $x \in E$ le vecteur colonne $X \in \mathbf{R}^n$ des coordonnées de x dans $\mathcal{B}$, on a $(x \mid y) = {}^tXY$. On suppose que l'application ${}^tf$ existe; soit $N \in M_n(\mathbf{R})$ la matrice de ${}^tf$ dans $\mathcal{B}$. Le calcul que l'on a effectué dans 5.2.1 montre que l'on a

$$\forall x, y \in E \quad {}^tX\,{}^tMY = {}^t(MX)Y = (f(x) \mid y) = (x \mid {}^tf(y)) = {}^tX(NY) = {}^tXNY,$$

d'où

$$\forall X, Y \in \mathbf{R}^n \qquad {}^tX\,{}^tMY = {}^tXNY \iff N = {}^tM.$$

En particulier, la matrice N est déterminée par M, ce qui montre l'unicité de ${}^tf$. Réciproquement, si l'on **définit** ${}^tf : E \longrightarrow E$ comme l'application linéaire représentée dans $\mathcal{B}$ par la matrice ${}^tM$, le calcul précédent montre que l'on a

$$\forall x, y \in E \quad (f(x) \mid y) = (x \mid {}^tf(y)).$$

(5.2.3) Définition. *Soit E un espace euclidien. Une application linéaire $f : E \longrightarrow E$ est dite* **symétrique** *si l'on a ${}^t f = f$. Cette condition équivaut à $\forall x, y \in E \quad (f(x) \mid y) = (x \mid f(y)) \iff$ la matrice de f dans une ($\iff$ dans toute) base orthonormée de E est symétrique.*

(5.2.4) Théorème (variante abstraite de 5.2.1). *Soit E un espace euclidien. Pour toute application linéaire* **symétrique** *$f = {}^t f : E \longrightarrow E$ il existe une base* **orthonormée** *$e'_1, \ldots, e'_n$ de E dont les éléments sont des vecteurs propres de f : $f(e'_j) = \lambda_j e'_j \ (\lambda_j \in \mathbf{R})$.*

Preuve. On admet une variante faible de 5.1.3(2), à savoir l'existence d'une valeur propre réelle $\lambda \in \mathbf{R}$ de f; soit $x \in E$, $x \neq \vec{0}$, $f(x) = \lambda x$. Le sous-espace vectoriel $\mathbf{R}x = \mathrm{vect}(x) \subset E$ est une droite; son supplémentaire orthogonal $E_1 = (\mathbf{R}x)^\perp \subset E$ est un espace euclidien de dimension $n-1$. On prétend que $f(E_1) \subseteq E_1$: en effet, si $y \in E_1$, alors $(x \mid y) = 0$, donc

$$(x \mid f(y)) = (f(x) \mid y) = (\lambda x \mid y) = \lambda(x \mid y) = 0,$$

ce qui signifie que $f(y) \in E_1$. Il en résulte que la restriction de f à E_1 définit une application linéaire

$$f_1 : E_1 \longrightarrow E_1, \qquad y \mapsto f(y).$$

L'application f étant symétrique, f_1 est aussi symétrique. On pose $e'_1 = x/\|x\|$ et l'on applique le même argument à l'application f_1; on obtient, par récurrence, la base orthonormée cherchée de E.

6. Formes hermitiennes, espaces unitaires

Les espaces unitaires sont des analogues complexes des espaces euclidiens. Leurs généralisations en dimension infinie (les espaces de Hilbert complexes) jouent un rôle fondamental dans la théorie quantique.

6.1 Notions de base

(6.1.1) La positivité du produit scalaire euclidien usuel se déduit de l'inégalité élémentaire

$$\forall x \in \mathbf{R} \qquad x^2 \geq 0.$$

L'analogue complex de cette inégalité

$$\forall z \in \mathbf{C} \qquad z\overline{z} \geq 0$$

nous conduit à l'exemple suivant.

(6.1.2) Exemple: le **produit scalaire hermitien usuel** sur $\mathbf{C}^n$

$$(x \mid y) = {}^t x\, \overline{y} = x_1\overline{y}_1 + \cdots + x_n\overline{y}_n \qquad\qquad (x, y \in \mathbf{C}^n)$$

est positif au sens suivant

$$(\forall x \in \mathbf{C}^n, x \neq \vec{0}) \qquad (x \mid x) = x_1\overline{x}_1 + \cdots + x_n\overline{x}_n > 0,$$

mais il **n'est pas bilinéaire**; il est linéaire en x

$$(x + x' \mid y) = (x \mid y) + (x' \mid y), \qquad (\lambda x \mid y) = \lambda\,(x \mid y) \qquad\qquad (x, x', y \in \mathbf{C}^n;\ \lambda \in \mathbf{C})$$

et **anti-linéaire** en y:

$$(x \mid y + y') = (x \mid y) + (x \mid y'), \qquad (x \mid \lambda y) = \overline{\lambda}\,(x \mid y) \qquad\qquad (x, y, y' \in \mathbf{C}^n;\ \lambda \in \mathbf{C}).$$

C'est un cas particulier de la notion abstraite suivante.

(6.1.3) Définition. *Soit E un $\mathbf{C}$-ev. Une* **forme hermitienne sur** E *est une application $f : E \times E \longrightarrow \mathbf{C}$ (autrement dit, une fonction $f(x, y)$ en deux variables $x, y \in E$ à valeurs complexes) telle que*
(1) *Si l'on fixe $y \in E$, la fonction $f(x, y)$ est linéaire en x:*

$$f(x + x', y) = f(x, y) + f(x', y), \qquad f(\lambda x, y) = \lambda\, f(x, y) \qquad\qquad (x, x', y \in E;\ \lambda \in \mathbf{C}).$$

(2) *Si l'on fixe $x \in E$, la fonction $f(x, y)$ est anti-linéaire en y:*

$$f(x, y + y') = f(x, y) + f(x, y'), \qquad f(x, \lambda y) = \overline{\lambda}\, f(x, y) \qquad\qquad (x, y, y' \in E;\ \lambda \in \mathbf{C}).$$

(3) $\forall x, y \in E \qquad f(y, x) = \overline{f(x, y)}$.
[Les conditions (1) et (3) entraînent (2).]

6.2 Représentation matricielle des formes hermitiennes

(6.2.1) On suppose que $n = \dim(E) < \infty$. Un choix d'une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de E nous permet d'identifier E à $\mathbf{C}^n$ (voir 1.1.8 ci-dessus). En particulier, soient $x = x_1e_1 + \cdots + x_ne_n$ et $y = y_1e_1 + \cdots + y_ne_n \in E$ des vecteurs de E; on note

$$X = M_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = M_{\mathcal{B}}(y) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbf{C}^n$$

les vecteurs colonnes qui représentent x et y dans la base $\mathcal{B}$.

Si $f : E \times E \longrightarrow \mathbf{C}$ est une forme hermitienne, alors on a

$$f(x,y) = f(\sum_{j=1}^n x_j e_j, \sum_{k=1}^n y_k e_k) = \sum_{j=1}^n x_j\, f(e_j, \sum_{k=1}^n y_k e_k) = \sum_{j=1}^n x_j \sum_{k=1}^n \overline{y}_k\, f(e_j, e_k) = \sum_{j,k=1}^n f(e_j, e_k)\, x_j \overline{y}_k. \tag{6.2.1.1}$$

(6.2.2) Définition. *Sous les hypothèses de 6.2.1, on appelle la matrice carrée $A = (A_{jk}) \in M_n(\mathbf{C})$, où $A_{jk} = f(e_j, e_k)$,* **la matrice de f dans la base $\mathcal{B}$**. *On a*

$$A_{kj} = f(e_k, e_j) = \overline{f(e_j, e_k)} = \overline{A_{jk}}, \qquad {}^t\!A = \overline{A}.$$

(6.2.3) Écriture matricielle. En utilisant la matrice A, on peut exprimer (6.2.1.1) comme le produit matriciel

$$f(x,y) = \sum_{j,k=1}^n A_{jk}\, x_j \overline{y}_k = (x_1 \cdots x_n) \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_n \end{pmatrix} = {}^t\!XA\overline{Y}. \tag{6.2.3.1}$$

(6.2.4) Définition. *Une matrice carrée complexe $B = (B_{jk}) \in M_n(\mathbf{C})$ est dite* **hermitienne** *si l'on a ${}^t\!B = \overline{B}$ ($\Longleftrightarrow$ $B_{kj} = \overline{B}_{jk}$ pour tous j,k $\Longrightarrow$ tous les termes diagonaux $B_{jj} \in \mathbf{R}$ sont réels). En particulier, une matrice carrée réelle est hermitienne $\Longleftrightarrow$ elle est symétrique.*

(6.2.5) D'après 6.2.2, la matrice d'une forme hermitienne dans une base quelconque de E est hermitienne. Réciproquement, pour toute matrice hermitienne $A \in M_n(\mathbf{C})$, la formule (6.2.3.1)

$$f(x,y) = {}^t\!XA\overline{Y}$$

définit une forme hermitienne sur E: les propriétés 6.1.3(1)-(2) sont vérifiées automatiquement; pour démontrer la propriété 6.1.3(3), on suit le calcul de 2.3.5: comme $f(y,x) \in \mathbf{C} = M_1(\mathbf{C})$ est égal à sa transposée, on a, pour tous $x, y \in E$,

$$f(y,x) = {}^t\!YA\overline{X} = {}^t({}^t\!YA\overline{X}) = {}^t\overline{X}\,{}^t\!A\,Y = {}^t\overline{X}\,\overline{A}\,Y = \overline{{}^t\!XA\overline{Y}} = \overline{f(x,y)}.$$

(6.2.6) Changement de base. Soit $\mathcal{B}' = e'_1, \ldots, e'_n$ une autre base de E; on note $P \in M_n(\mathbf{C})$ la matrice de passage (inversible) de $\mathcal{B}$ à $\mathcal{B}'$. En utilisant les formules usuelles

$$x = \sum_{j=1}^n x_j e_j = \sum_{j=1}^n x'_j e'_j, \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}, \qquad X = PX'$$

(où $X, X' \in \mathbf{C}^n$ sont les vecteurs colonnes de coordonnées de $x \in E$ dans les bases respectives $\mathcal{B}, \mathcal{B}'$), on calcule la matrice (hermitienne) $A' \in M_n(\mathbf{C})$ de f dans la base $\mathcal{B}'$:

$$\left.\begin{aligned} f(x,y) &= {}^t\!XA\overline{Y} = {}^t(PX')A\overline{(PY')} = {}^t\!X'({}^t\!PA\overline{P})\overline{Y'} \\ f(x,y) &= {}^t\!X'A'\overline{Y'} \end{aligned}\right\} \Longrightarrow A' = {}^t\!PA\overline{P}.$$

6.3 Les analogies avec la théorie de formes quadratiques

(6.3.1) Soit $f : E \times E \longrightarrow \mathbf{C}$ une forme hermitienne. On appelle **la forme quadratique hermitienne associée à** f la fonction

$$q : E \longrightarrow \mathbf{C}, \qquad q(x) = f(x,x) = {}^tXA\overline{X}.$$

Par exemple, si $E = \mathbf{C}^2$ et si la matrice de f dans la base canonique est égale à

$$A = \begin{pmatrix} 1 & 1+2i \\ 1-2i & 3 \end{pmatrix},$$

alors on a

$$\begin{aligned} f(x,y) &= x_1\overline{y}_1 + (1+2i)x_1\overline{y}_2 + (1-2i)x_2\overline{y}_1 + 3x_2\overline{y}_2 \\ q(x) &= x_1\overline{x}_1 + (1+2i)x_1\overline{x}_2 + (1-2i)x_2\overline{x}_1 + 3x_2\overline{x}_2 \end{aligned} \tag{6.3.1.1}$$

La forme f est déterminée par q, ce qui résulte de la représentation matricielle (voir (6.3.1.1) ci-dessus) ou des formules abstraites suivantes :

$$\begin{gathered} q(x \pm y) = f(x \pm y, x \pm y) = f(x,x) + f(y,y) \pm (f(x,y) + f(y,x)) = q(x) + q(y) \pm (f(x,y) + f(y,x)) \\ q(x \pm iy) = f(x \pm iy, x \pm iy) = f(x,x) + f(y,y) \pm i(-f(x,y) + f(y,x)) = q(x) + q(y) \pm i(-f(x,y) + f(y,x)) \\ 4f(x,y) = q(x+y) - q(x-y) + iq(x+iy) - iq(x-iy). \end{gathered}$$

(6.3.2) La théorie de formes quadratiques (voir le chapitre 3 ci-dessus) admet une variante hermitienne : sous les hypothèses de 6.2.1, on définit

$$\begin{gathered} \mathrm{rg}(f) := \mathrm{rg}(A) \\ f \text{ est non} - \text{dégénérée} \iff \mathrm{rg}(f) = n \iff \det(A) \neq 0 \\ x \perp y \iff f(x,y) = 0 \ (\iff y \perp x) \qquad (x, y \in E) \\ \forall S \subset E \ (S \text{ non vide}) \qquad S^{\perp} := \{y \in E \mid \forall x \in S \quad x \perp y\} = (\mathrm{vect}(S))^{\perp} \ \text{ est un sous} - \text{espace vectoriel de } E \\ N(f) := E^{\perp} \end{gathered}$$

Les résultats du §3.6 sont valables : si f est non-dégénérée et si $F \subset E$ est un sous-espace vectoriel, alors on a $\dim(F^{\perp}) = \dim(E) - \dim(F)$. De plus, si $F \cap F^{\perp} = \{\overrightarrow{0}\}$, alors on a $E = F \oplus F^{\perp} = F \perp F^{\perp}$. Théorème 3.3.1 (diagonalisation "facile" d'une forme quadratique) et Théorème 3.7.2 (théorème de Sylvester) admettent les variantes hermitiennes suivantes (dont les démonstrations sont analogues aux celles du chapitre 3).

(6.3.3) Théorème (diagonalisation facile d'une forme hermitienne). *Pour toute forme hermitienne $f : E \times E \longrightarrow \mathbf{C}$ sur un $\mathbf{C}$-ev E de dimension finie il existe une base $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ de E dans laquelle la matrice de f est diagonale :*

$$f(x,y) = d_1 x'_1 \overline{y'_1} + \cdots + d_n x'_n \overline{y'_n} \qquad \left(x = \sum_{j=1}^{n} x'_j e'_j,\ y = \sum_{j=1}^{n} y'_j e'_j\right)$$

(où $d_1, \ldots, d_n \in \mathbf{R}$).

(6.3.4) Exemples: (1) Pour la forme hermitienne (6.3.1.1), on a

$$\begin{aligned} q(x) &= x_1\overline{x}_1 + (1+2i)x_1\overline{x}_2 + (1-2i)x_2\overline{x}_1 + 3x_2\overline{x}_2 = (x_1 + (1-2i)x_2)(\overline{x}_1 + (1+2i)\overline{x}_2) - 2x_2\overline{x}_2 = \\ &= x'_1\overline{x'_1} - 2x'_2\overline{x'_2}, \qquad x'_1 = x_1 + (1-2i)x_2, \qquad x'_2 = x_2. \end{aligned}$$

(2) Pour la forme hermitienne f dont la matrice (dans la base canonique de $\mathbf{C}^2$) est égale à

$$A = \begin{pmatrix} 0 & a \\ \overline{a} & 0 \end{pmatrix} \qquad\qquad (a \in \mathbf{C},\ a \neq 0),$$

on a

$$q(x) = a x_1 \overline{x}_2 + \overline{a} x_2 \overline{x}_1.$$

Le changement de coordonnées

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} (a x_1 + x_2)/2 \\ (a x_1 - x_2)/2 \end{pmatrix}, \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} (x_1' + x_2')/a \\ x_1' - x_2' \end{pmatrix}$$

transforme q à

$$q(x) = 2(x_1'\overline{x_1'} - x_2'\overline{x_2'}).$$

(6.3.5) Théorème (variante hermitienne du théorème de Sylvester). *Pour toute forme hermitienne $f : E \times E \longrightarrow \mathbf{C}$ sur un $\mathbf{C}$-ev E de dimension finie il existe une base $\mathcal{B} = \{e_1, \ldots, e_n\}$ de E dans laquelle on a*

$$f(x, y) = x_1\overline{y_1} + \cdots + x_s\overline{y_s} - x_{s+1}\overline{y_{s+1}} - \cdots - x_{s+t}\overline{y_{s+t}}.$$

Les entiers $s, t \geq 0$ ne dépendent que de la forme f (on a $s + t = \mathrm{rg}(f)$).

6.4 Espaces unitaires

(6.4.1) Définition. *Une forme hermitienne $f : E \times E \longrightarrow \mathbf{C}$ est dite* **définie positive** *si l'on a $(\forall x \in E,\ x \neq \vec{0})\ f(x, x) > 0$. Si c'est le cas, on dit que E (muni du produit scalaire hermitien $(x \mid y) = f(x, y)$) est un* **espace préhilbertien complex** *(resp., un* **espace unitaire** *si $\dim(E) < \infty$). Une matrice hermitienne $A \in M_n(\mathbf{C})$ est dite* **définie positive** *(notation: $A >> 0$) si la forme hermitienne associée $f(x, y) = {}^tXA\overline{Y}$ l'est : $(\forall X \in \mathbf{C}^n,\ X \neq \vec{0})\ {}^tXA\overline{X} > 0$.*

(6.4.2) Exemples: (1) $E = \mathbf{C}^n$, muni du produit scalaire hermitien usuel $(x \mid y) = x_1\overline{y}_1 + \cdots + x_n\overline{y}_n$, est un espace unitaire.
(2) L'espace des fonctions continues à valeurs complexes

$$E = \{f : [a, b] \longrightarrow \mathbf{C} \mid f \text{ continuous}\} = C([a, b]; \mathbf{C}), \qquad (f \mid g) = \int_a^b f(t)\,\overline{g(t)}\,dt$$

est un espace préhilbertien complex.
(6.4.3) (1) Il résulte du Théorème 6.3.5 que tout espace unitaire E admet une base $\mathcal{B} = e_1, \ldots, e_n$ dans laquelle on a

$$(x \mid y) = x_1\overline{y}_1 + \cdots + x_n\overline{y}_n \qquad\qquad (x = \sum_{j=1}^n x_j e_j,\ y = \sum_{j=1}^n y_j e_j)$$

(voir 6.4.2(1)); on dit que $\mathcal{B}$ est une **base orthonormée** de E.
(2) La définition d'un **système orthonormé** dans E reste inchangée (voir 4.1.7). Par exemple, les fonctions

$$f_m(t) = e^{2\pi i m t} = \cos(2\pi m t) + i\,\sin(2\pi m t) \qquad\qquad (m = 0, \pm 1, \pm 2, \ldots)$$

forme un système orthonormé dans l'espace $C([0, 1]; \mathbf{C})$.

(6.4.4) La norme. La norme d'un vecteur x d'un espace préhilbertien complexe E est définie par la formule usuelle

$$\|x\| := \sqrt{(x \mid x)} \geq 0.$$

L'inégalité de Cauchy-Schwarz (variante hermitienne) :

$$\forall x, y \in E \qquad |(x \mid y)| \leq \|x\|\|y\|. \tag{6.4.4.1}$$

Pour démontrer (6.4.4.1), on modifie légèrement l'argument de 4.2.3 : en supposant que $x \neq \vec{0}$, on a, pour tout $t \in \mathbf{R}$,

$$\begin{aligned} 0 \leq \|tx + y\|^2 &= t^2(x \mid x) + t\left((x \mid y) + \overline{(x \mid y)}\right) + (y \mid y) = \\ &= (x \mid x)\left|t + \frac{(x \mid y)}{(x \mid x)}\right|^2 + \frac{(x \mid x)(y \mid y) - |(x \mid y)|^2}{(x \mid x)}; \end{aligned}$$

si l'on prend $t = -(x \mid y)/(x \mid x)$, on obtient (6.4.4.1) (en utilisant l'inégalité $(x \mid x) > 0$).

6.5 Isométries, matrices unitaires

Soit E un espace unitaire de dimension $n \geq 1$. Les énoncés du § 4.4 admettent les variantes hermitiennes suivantes.

(6.5.1) On fixe une base orthonormée $\mathcal{B}$ de E. Soit $\mathcal{B}'$ une base quelconque de E; on note $P \in M_n(\mathbf{C})$ la matrice (inversible) de passage de $\mathcal{B}$ à $\mathcal{B}'$. Les matrices A, A' du produit scalaire hermitien dans les bases respectives $\mathcal{B}, \mathcal{B}'$ sont égales à

$$A = I_n, \qquad A' = {}^t P A \overline{P} = {}^t P \overline{P},$$

d'où

$$\mathcal{B}' \text{ est une base orthonormée} \quad \Longleftrightarrow \quad A' = I_n \quad \Longleftrightarrow \quad {}^t P \overline{P} = I_n.$$

(6.5.2) Définition. *Une matrice $P \in M_n(\mathbf{C})$ est dite* **unitaire** *si l'on a ${}^t P \overline{P} = I_n$ (ce qui entraine que $|\det(P)| = 1$). On note que, si P est unitaire, alors $\overline{P}$ l'est aussi.*

(6.5.3) Isométries. Une **isométrie de** E est une application linéaire $u : E \longrightarrow E$ vérifiant

$$\forall x \in E \qquad \|u(x)\| = \|x\| \quad \Longleftrightarrow \quad \forall x, y \in E \qquad (u(x) \mid u(y)) = (x \mid y).$$

En écrivant $u(x) = UX$ et $(x \mid y) = {}^t X \overline{Y}$ dans une base orthonormée quelconque de E, le même calcul que l'on a effectué dans 4.4.4

$$(u(x) \mid u(y)) = {}^t(UX)\overline{UY} = {}^t X({}^t U \overline{U})\overline{Y}$$

montre que l'on a

$$\begin{aligned} u \text{ est une isométrie} \quad &\Longleftrightarrow \quad \text{la matrice } U \text{ de } u \text{ dans une base orthonormée de } E \text{ est unitaire} \\ &\Longleftrightarrow \quad \text{la matrice de } u \text{ dans chaque base orthonormée de } E \text{ est unitaire.} \end{aligned}$$

En particulier, les isométries de $\mathbf{C}^n$ (muni du produit scalaire hermitien usuel) sont les applications linéaires

$$u : \mathbf{C}^n \longrightarrow \mathbf{C}^n, \qquad X \mapsto UX, \qquad U \in M_n(\mathbf{C}) \ \textbf{unitaire}.$$

Si l'on utilise la notation suivante

$$U(n) = \{\text{matrices unitaires } U \in M_n(\mathbf{C})\} = \{\text{isométries de } \mathbf{C}^n\}$$
$$SU(n) = \{U \in U(n) \mid \det(U) = 1\},$$

alors on a

$$U(n) \cap M_n(\mathbf{R}) = O(n),$$

et l'on peut identifier

$$U(1) = \{\lambda \in \mathbf{C} \mid \ |\lambda| = 1\}$$

à $SO(2)$, grâce à (4.5.10.1).

(6.5.4) Exercice. *Montrer :*

$$SU(2) = \Big\{ \begin{pmatrix} a & -\overline{c} \\ c & \overline{a} \end{pmatrix} \Big| a, c \in \mathbf{C},\ |a|^2 + |c|^2 = 1 \Big\}.$$

(6.5.5) Proposition (valeurs propres de matrices unitaires et hermitiennes). *Soit $A \in M_n(\mathbf{C})$ une matrice complexe, soient $x, y \in \mathbf{C}^n$ des vecteurs propres de A: $x, y \neq \vec{0}$, $Ax = \lambda x$, $Ay = \mu y$ ($\lambda, \mu \in \mathbf{C}$). On note $(x \mid y) = {}^t x\, \overline{y}$ le produit scalaire hermitien usuel sur $\mathbf{C}^n$.*
(1) *Si la matrice A est unitaire, alors on a $|\lambda| = |\mu| = 1$. Si $\lambda \neq \mu$, alors on a $(x \mid y) = 0$.*
(2) *Si la matrice A est hermitienne, alors on a $\lambda, \mu \in \mathbf{R}$. Si $\lambda \neq \mu$, alors on a $(x \mid y) = 0$.*

Preuve. (1) La matrice A étant unitaire, on a

$$\forall u, v \in \mathbf{C}^n \qquad (Au \mid Av) = {}^t(Au)\overline{Av} = {}^t u\, {}^t A\, \overline{A}\, \overline{v} = {}^t u\, \overline{v} = (u \mid v).$$

Si l'on prend $u = v = x$, on obtient

$$0 \neq (x \mid x) = (Ax \mid Ax) = (\lambda x \mid \lambda x) = \lambda\overline{\lambda}\, (x \mid x) = |\lambda|^2\, (x \mid x),$$

d'où $|\lambda| = 1$ (de même, $|\mu| = 1$). Si l'on prend $u = x$ et $v = y$, on obtient

$$(x \mid y) = (Ax \mid Ay) = (\lambda x \mid \mu y) = \lambda\overline{\mu}\, (x \mid y) = \lambda\mu^{-1}\, (x \mid y);$$

si $\lambda \neq \mu$, on en déduit que $(x \mid y) = 0$.
(2) L'égalité ${}^t A = \overline{A}$ entraîne

$$\forall u, v \in \mathbf{C}^n \qquad (Au \mid v) = {}^t(Au)\overline{v} = {}^t u\, {}^t A\, \overline{v} = {}^t u\, \overline{Av} = (u \mid Av).$$

Si l'on prend $u = v = x$, on obtient

$$\lambda\, (x \mid x) = (\lambda x \mid x) = (Ax \mid x) = (x \mid Ax) = (x \mid \lambda x) = \overline{\lambda}\, (x \mid x);$$

comme $(x \mid x) \neq 0$, il en résulte que $\lambda = \overline{\lambda}$ est réel (de même, μ is réel). Si l'on prend $u = x$ et $v = y$, on obtient

$$\lambda\, (x \mid y) = (\lambda x \mid y) = (Ax \mid y) = (x \mid Ay) = (x \mid \mu y) = \overline{\mu}\, (x \mid y) = \mu\, (x \mid y);$$

si $\lambda \neq \mu$, on en déduit que $(x \mid y) = 0$.

(6.5.6) Corollaire (valeurs propres de matrices orthogonales/matrices symétriques réelles). *Soit $A \in M_n(\mathbf{R})$ une matrice réelle, soient $x, y \in \mathbf{C}^n$ des vecteurs propres complexes de A: $x, y \neq \overrightarrow{0}$, $Ax = \lambda x$, $Ay = \mu y$ ($\lambda, \mu \in \mathbf{C}$). On note $(x \mid y) = {}^t x\, \overline{y}$ le produit scalaire hermitien usuel sur $\mathbf{C}^n$.*
(1) *Si la matrice A est orthogonale, alors on a $|\lambda| = |\mu| = 1$. Si $\lambda \neq \mu$, alors on a $(x \mid y) = 0$.*
(2) *Si la matrice A est symétrique, alors on a $\lambda, \mu \in \mathbf{R}$. Si $\lambda \neq \mu$, alors on a $(x \mid y) = 0$.*

(6.5.7) Théorème (diagonalisation difficile d'une forme hermitienne. *Pour toute forme hermitienne f sur un espace unitaire E il existe une base* **orthonormée** *de E dans laquelle la matrice de f est diagonale (dont les éléments diagonaux sont réels).*
Formulation matricielle : *Pour toute matrice* **hermitienne** *$A \in M_n(\mathbf{C})$ il existe une matrice* **unitaire** *$P \in U(n)$ telle que la matrice ${}^t\overline{P}AP = P^{-1}AP$ est diagonale (dont les éléments diagonaux sont réels). Plus précisement, les colonnes de P sont des vecteurs propres de P qui forment une base orthonormée de $\mathbf{C}^n$ (par rapport au produit scalaire hermitien usuel).*

Preuve. Il faut modifier la preuve de 5.2.4 de la façon suivante. On définit l'adjoint hermitien ${}^h f$ de f vérifiant $(f(x) \mid y) = (x \mid {}^h f(y))$, et l'on montre que les matrices respectives M, N de $f, {}^h f$ (dans une base orthonormée quelconque de E) vérifient $N = {}^t\overline{M}$.

Le même argument permet de démontrer le résultat suivant.

(6.5.8) Théorème (variante hermitienne de 4.6.2). *Pour toute isométrie $u : E \longrightarrow E$ d'un espace unitaire E il existe une base* **orthonormée** *de E dans laquelle la matrice de u est diagonale (dont les éléments diagonaux $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ vérifient $|\lambda_1| = \cdots = |\lambda_n| = 1$).*
Formulation matricielle : *Pour toute matrice* **unitaire** *$U \in U(n)$ il existe une matrice* **unitaire** *$P \in U(n)$ telle que la matrice ${}^t\overline{P}UP = P^{-1}UP \in U(n)$ est diagonale (le modulus de chaque élément diagonal de $P^{-1}UP$ est égal à 1).*

(6.5.9) Voici un petit dictionnaire entre les propriétés de matrices et les propriétés de nombres complexes :

Nombres	Matrices
$z \in \mathbf{C}$	$A \in M_n(\mathbf{C})$
$\overline{z} \in \mathbf{C}$	${}^t\overline{A}$
$\overline{z} = z \in \mathbf{R}$	${}^t\overline{A} = A$ est hermitienne
$\overline{u}u = 1 \iff \lvert u\rvert = 1$	${}^t\overline{U}U = I \iff U \in U(n)$
$\lvert u\rvert = 1 \iff u = e^{it}$, $t \in \mathbf{R}$	$U \in U(n) \iff U = e^{iA}$, $A = {}^t\overline{A}$
$z \in \mathbf{C}$ $z \neq 0$ $\iff$ $z = ru$, $r \in \mathbf{R}_{>0}$, $\lvert u\rvert = 1$	$M \in M_n(\mathbf{C})$ $\det(M) \neq 0$ $\iff$ $M = PU$, $P = {}^t\overline{P} >> 0$, $U \in U(n)$

7. Exemples de groupes finis de provenance géométrique*

Dans ce chapitre on va étudier les groupes d'isométries des objets géométriques divers $C \subset \mathbf{R}^n$, à savoir les groupes

$$G(C) = \{g \in O(n) \mid g(C) = C\}$$
$$G^+(C) = \{g \in O^+(n) \mid g(C) = C\}.$$

7.1 Groupes

(7.1.1) Qu'est-ce qu'un groupe ? On rencontre souvent les **groupes de transformations :** chaque élément d'un groupe G est une transformation d'un objet donné X; on exige que cette transformation soit **inversible** et que son inverse appartienne à G; on exige aussi que la composée de deux transformations de X qui appartiennent à G soit un élément de G.

On peut aussi oublier l'objet X; dans ce cas G sera un groupe abstrait au sens suivant.

(7.1.2) Définition. *Un* **groupe** *est un ensemble G muni d'une* **opération binaire** $*$ *(autrement dit, pour tout couple d'éléments $g, h \in G$ on définit leur "produit" $g * h \in G$) qui satisfait aux propriétés suivantes :*
(1) (Associativité) $\forall g, h, k \in G \qquad g * (h * k) = (g * h) * k$.
(2) (Élément neutre) *Il existe un élément $e = e_G \in G$ tel que* $\forall g \in G \quad g * e = e * g = g$. *Cet élément est unique; on l'appelle* **l'élément neutre** *de G.*
(3) (Inverse) *Pour tout $g \in G$ il existe un élément $h \in G$ tel que $g * h = h * g = e$. L'élément h est unique; on l'appelle* **l'inverse de** *g et l'on le note $h = g^{-1}$.*
Notation : *on écrit parfois $(G, *)$ au lieu de G.*
Un groupe G est dit **commutatif** *(ou abélien) si l'on a* $\forall g, h \in G \quad g * h = h * g$.

(7.1.3) Exemples: (1) $G = \mathbf{R}^n$, $* =$ somme ($e = \overrightarrow{0}$); ce groupe est commutatif.
(2) $G = \mathbf{R}^* = \{r \in \mathbf{R} \mid r \neq 0\}$, $* =$ produit ($e = 1$); ce groupe est commutatif.
(3) $G = GL_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid \det(A) \neq 0\}$, $* =$ produit matriciel ($e = I_n$); si $n > 1$, alors ce groupe n'est pas commutatif.
(4) Pour tout ensemble X, les **permutations de** X

$$G = S_X = \{f : X \longrightarrow X \mid f \text{ est une application bijective } (= \text{ inversible})\}$$

forment un groupe pour l'opération de composition $* = \circ$:

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ & {\scriptstyle g\circ f}\searrow & \downarrow{\scriptstyle g} \\ & & X \end{array} \qquad\qquad (g \circ f)(x) = g(f(x)).$$

L'élément neutre est l'application identité

$$\mathrm{Id} = \mathrm{Id}_X : X \longrightarrow X, \qquad \forall x \in X \quad \mathrm{Id}(x) = x.$$

Si X a au moins 3 éléments, alors le groupe S_X n'est pas commutatif.

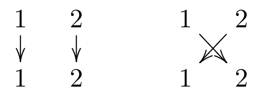
7.2 Les groupes symétriques, les groupes alternés

(7.2.1) Définition. *Soit $n \geq 1$ un entier. Les permutations de l'ensemble $X = \{1, 2, \ldots, n\}$ (= les applications bijectives $\sigma : X \longrightarrow X$) forment le* **groupe symétrique opérant sur** *n* **lettres** *$S_n = S_{\{1,2,\ldots,n\}}$ (pour l'opération de composition).*

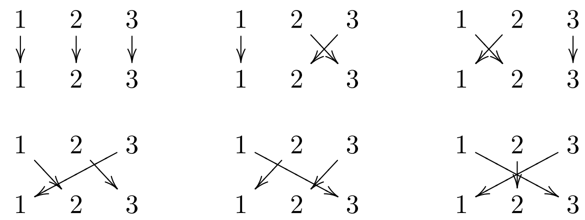
(7.2.2) Exemples: (1) S_1 ne contient que l'application identité

*** Ce chapitre est facultatif.**

(2) S_2 contient deux éléments



(3) S_3 contient six éléments



(4) **Notation :** pour toute permutation $\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$, on écrit

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

Par exemple, les éléments de S_3 s'écrivent

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

(5) **Composition de deux permutations :** par définition, on a $\sigma\tau = \sigma \circ \tau$, donc $(\sigma\tau)(i) = \sigma(\tau(i))$. Par exemple, si

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \qquad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$

alors on calcule

$$\begin{aligned} (\sigma\tau)(1) &= \sigma(\tau(1)) = \sigma(3) = 3 \\ (\sigma\tau)(2) &= \sigma(\tau(2)) = \sigma(1) = 2 \\ (\sigma\tau)(3) &= \sigma(\tau(3)) = \sigma(2) = 1, \end{aligned} \qquad \sigma\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

On peut aussi dessiner le diagram suivant, en utilisant la notation de (3) :

(6) **L'ordre de S_n :** le nombre d'éléments de S_n est égal à

$$|S_n| = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 = n!$$

(pour $\sigma \in S_n$, il y a n valeurs possibles de $\sigma(1)$; si l'on fixe $\sigma(1)$, il y a $n-1$ valeurs possibles de $\sigma(2)$, etc.).
(7) **Notation cyclique :** toute permutation $\sigma \in S_n$ s'écrit comme un produit de **cycles disjoints**. Par exemple, la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 3 & 6 & 2 & 1 \end{pmatrix} \in S_6 \qquad\qquad 1 \longrightarrow 4 \longrightarrow 6, \qquad 2 \longrightarrow 5, \qquad 3$$

est un produit de trois cycles, dont les longuers respectives sont égales à $3, 2, 1$. On écrit

$$\sigma = (146)(25)(3).$$

(7.2.3) Définition. *Une* **transposition** *est une permutation $\sigma \in S_n$ $(n \geq 2)$ qui échange deux éléments de $\{1, \dots, n\}$ et fixe tous les autres. Par exemple, la permutation*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 3 & 4 & 2 & 6 \end{pmatrix} = (25)(1)(3)(4)(6) \in S_6$$

est une transposition.

(7.2.4) Théorème. *Soit $n \geq 2$.* (1) *Toute permutation $\sigma \in S_n$ est un produit de transpositions $\sigma = s_1 \cdots s_r$ (cette écriture n'est pas unique).*
(2) *La parité de r ne dépend que de σ; on définit le* **signe de** *σ comme $\operatorname{sgn}(\sigma) = (-1)^r$.*
(3) *On a $\forall \sigma, \tau \in S_n \qquad \operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$.*
Terminologie : *si $\operatorname{sgn}(\sigma) = +1$ (resp., $\operatorname{sgn}(\sigma) = -1$), on dit que σ est une permutation* **paire** *(resp.,* **impaire***).*

(7.2.5) Exemple: Le groupe S_3 contient trois transpositions (qui vérifient $\operatorname{sgn} = -1$)

$$s_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad s_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad s_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = s_1 s_2 s_1 = s_2 s_1 s_2.$$

Les autres éléments de S_3 (qui vérifient $\operatorname{sgn} = +1$) s'écrivent, par exemple,

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = s_1^2 = s_1 s_1, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = s_1 s_2 = s_2 s_1 s_2 s_1, \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = s_2 s_1 = s_1 s_2 s_1 s_2.$$

(7.2.6) Définition. *Soit $(G, *)$ un groupe. Un* **sous-groupe de** *G est un sous-ensemble $H \subset G$ tel que*
(1) *$e \in H$;*
(2) *$\forall h_1, h_2 \in H \qquad h_1 * h_2 \in H$;*
(3) *$\forall h \in H \qquad h^{-1} \in H$.*
Si c'est le cas, alors $(H, *)$ est aussi un groupe.

(7.2.7) Exemples: (1) Pour tout $n \geq 2$, l'ensemble de toutes les permutations paires est un sous-groupe de S_n

$$A_n = \{\sigma \in S_n \mid \operatorname{sgn}(\sigma) = +1\},$$

que l'on appelle le **groupe alterné opérant sur n lettres**; son ordre (= le nombre d'éléments) est égal à $|A_n| = |S_n|/2 = n!/2$. Par exemple, on a

$$A_3 = \{(1)(2)(3),\ (123),\ (132)\},$$
$$A_4 = \{(1)(2)(3)(4),\ (123)(4),\ (132)(4), \dots, (1)(234),\ (1)(243),\ (12)(34),\ (13)(24),\ (14)(23)\}.$$

(2) Si $G = (S_X, \circ)$ est le groupe des permutations d'un ensemble X et si $Y \subset X$ est un sous-ensemble de X, alors

$$H = \{g \in G \mid g(Y) = Y\}$$

est un sous-groupe de G. Par exemple, si $X = \{1, \dots, n\}$ et $Y = \{n\}$, alors on a $G = S_n$ et $H = S_{n-1}$. Plus généralement, si $Y = \{k+1, \dots, n\}$ (où $1 \leq k \leq n-1$), alors le sous-groupe H devrait être noté $S_k \times S_{n-k}$ (exercice : pourquoi ?).

(3) Soit K un corps, V un K-ev et soient

$$G = (S_V, \circ) = \{f : V \longrightarrow V \mid f \text{ est une application bijective (= inversible)}\}$$
$$H = GL(V) = \{f : V \longrightarrow V \mid f \text{ est une application bijective et linéaire}\};$$

alors $GL(V)$ est un sous-groupe de S_V.

Si $V = K^n$ ($n \geq 1$), alors on identifie l'ensemble des applications linéaires $K^n \longrightarrow K^n$ à $M_n(K)$ (voir (1.2.4.3)), ce qui conduit à l'identification de $GL(K^n)$ au groupe matriciel

$$GL_n(K) = \{A \in M_n(K) \mid A \text{ est inversible}\} = \{A \in M_n(K) \mid \det(A) \neq 0\}.$$

(4) Lorsque $K = \mathbf{R}$, on connaît les sous-groupes suivants de $S_{\mathbf{R}^n}$:

$$SO(n) = O^+(n) \subset O(n) \subset GL_n(\mathbf{R}) \subset S_{\mathbf{R}^n}.$$

7.3 Exemples de sous-groupes finis de $O(n)$

On revient à la question du départ, à savoir à l'étude des sous-groupes de $O(n)$

$$G(Y) = \{g \in O(n) \mid g(Y) = Y\}$$
$$G^+(Y) = \{g \in O^+(n) \mid g(Y) = Y\}$$

(où $Y \subset \mathbf{R}^n$ est un sous-ensemble donné).

(7.3.1) Exercice. *Montrer que, si l'ensemble*

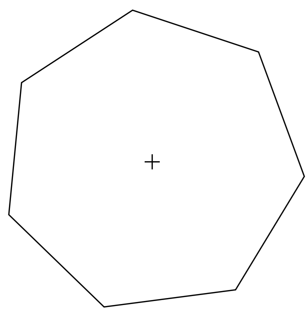
$$G^-(Y) = \{g \in O^-(n) \mid g(Y) = Y\}$$

n'est pas vide, alors on a

$$G^-(Y) = \{sh \mid h \in G^+(Y)\},$$

pour n'importe quel élément $s \in G^-(Y)$.

(7.3.2) Les groupes cycliques, les groupes dièdraux. Soit $Y \subset \mathbf{R}^2$ un polygone régulier à $n \geq 3$ côtés de centre à l'origine.



n=7

(7.3.2.1) Dans ce cas le groupe

$$G^+(Y) = \{\text{rotations autour d'origine d'angles } 2\pi k/n \mid k = 0, \dots, n-1\} = \\ = \{r, r^2, \dots, r^{n-1}, r^n = e\} = C_n$$

est le **groupe cyclique d'ordre** n, engendré par la rotation r (autour d'origine) d'angle $2\pi/n$.

(7.3.2.2) Soit $s \in O^-(2)$ la réflexion par rapport à la droite qui passe par l'origine et par un sommet fixé de Y; alors on a

$$G^-(Y) := \{g \in O^-(2) \mid g(Y) = Y\} = \{s, sr, \dots, sr^{n-1}\}$$

et le groupe

$$G(Y) = \{r, r^2, \dots, r^{n-1}, r^n = e, s, sr, \dots, sr^{n-1}\} = D_n$$

(que l'on note parfois D_{2n}, au lieu de D_n) est le **groupe dièdral d'ordre** $2n$, engendré par deux éléments r, s qui satisfont aux relations suivantes

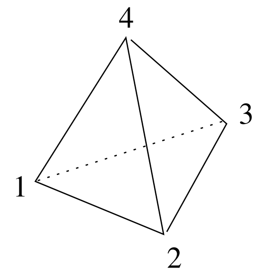
$$r^n = e = s^2, \qquad srs^{-1} = r^{-1}$$

(voir 4.5.7(2)). Tout élément de $G(Y)$ définit une permutation des sommets de Y; dès qu'on numérote les sommets $1, \dots, n$, on obtient une réalisation de D_n en tant qu'un sous-groupe de S_n ($n \geq 3$). Par exemple, pour $n = 3$, on a

$$C_3 = A_3, \qquad D_3 = S_3.$$

(7.3.2.3) Exercice. *Décrire géométriquement tous les éléments de $G^-(Y)$.*
[Distinguer deux cas, selon la parité de n.]

(7.3.3) Le groupe du tetraédre. Soit $T \subset \mathbf{R}^3$ un tetraédre régulier de centre à l'origine; on note $1, \dots, 4$ les sommets de T.



On prétend que

$$G(T) = S_4, \qquad G^+(T) = A_4;$$

en particulier, $|G(T)| = 24$, $|G^+(T)| = 12$.

(7.3.3.1) Plus précisement, tout élément $g \in G(T)$ définit une permutation $\sigma(g) \in S_4$ des sommets de T, et cette correspondance conserve les opérations :

$$\forall g_1, g_2 \in G(T) \qquad \sigma(g_1 \circ g_2) = \sigma(g_1) \circ \sigma(g_2);$$

autrement dit, l'application

$$\sigma : G(T) \longrightarrow S_4, \qquad g \mapsto \sigma(g)$$

est un **homomorphisme de groupes :**

(7.3.3.2) Définition. *Soient* $(G, *)$ *et* $(H, \#)$ *des groupes. Une application* $f : G \longrightarrow H$ *est un* **homomorphisme de groupes** *si elle vérifie la propriété* $\forall g_1, g_2 \in G \quad f(g_1 * g_2) = f(g_1)\#f(g_2)$. *Dans ce cas on a* $f(e_G) = e_H$ *et* $f(g^{-1}) = f(g)^{-1}$ *(pour tout* $g \in G$*).*

(7.3.3.3) On aimerait démontrer que l'application σ (qui conserve les opérations) est bijective, c'est-à-dire que toute permutation $\tau \in S_4$ s'écrit $\tau = \sigma(g)$, pour **l'unique** $g \in G(T)$. Ceci nous permettrait d'identifier $G(T) = (G(T), \circ)$ à $S_4 = (S_4, \circ)$, en utilisant l'application σ.

L'élément g existe, puisque chaque $\tau \in S_4$ est un produit de transpositions $\tau = s_1 \cdots s_r$, et toute transposition

$$s_t = (ij)(k)(l) \in S_4$$

est égale à $s_t = \sigma(g_t)$, où l'on a noté $g_t \in G(T)$ la réflexion par rapport au plan qui passe par k, l et le milieu de ij; on a $\tau = \sigma(g_1) \cdots \sigma(g_r) = \sigma(g)$, où $g = g_1 \cdots g_r$. L'élément g est-il unique? Si l'on a $\tau = \sigma(g) = \sigma(h)$ $(g, h \in G(T))$, on obtient que la permutation $\sigma(gh^{-1}) = \sigma(g)\sigma(h)^{-1} = \tau\tau^{-1} = e = (1)(2)(3)(4)$ fixe tous les sommets de T, donc la matrice $gh^{-1} \in G(T) \subset O(3)$ fixe les vecteurs $\overrightarrow{12}, \overrightarrow{13}, \overrightarrow{14}$; comme ces vecteurs forment une base de $\mathbf{R}^3$, on en déduit que gh^{-1} fixe tous les vecteurs de $\mathbf{R}^3$, d'où $gh^{-1} = I_3$, $g = h$. On vient de montrer que l'application σ est, en effet, un homomorphism de groupes bijectif (= un **isomorphisme de groupes**).

(7.3.3.4) En utilisant la même notation, on a

$$\mathrm{sgn}(s_t) = -1 = \det(g_t)$$

pour tous $t = 1, \ldots, r$, d'où

$$\mathrm{sgn}(\sigma(g)) = \mathrm{sgn}(\sigma(g_1) \cdots \sigma(g_r)) = \mathrm{sgn}(s_1 \cdots s_r) = (-1)^r = \det(g_1 \cdots g_r) = \det(g).$$

Il en résulte que l'on a, pour $g \in G(T)$,

$$g \in G^+(T) \iff \det(g) = +1 \iff \mathrm{sgn}(\sigma(g)) = +1 \iff \sigma(g) \in A_4;$$

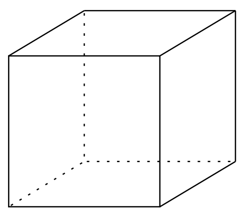
donc l'application σ identifie $G^+(T)$ à A_4.

(7.3.3.5) Exercice. *Décrire géométriquement tous les éléments de* $G^+(T)$.

(7.3.4) Le groupe du cube. Soit $C \subset \mathbf{R}^3$ le cube dont les sommets sont égaux à

$$\begin{pmatrix} \pm 1 \\ \pm 1 \\ \pm 1 \end{pmatrix}$$

(pour toutes les combinaisons possibles des signes).



On prétend que

$$G^+(C) = S_4, \qquad G(C) = S_4 \times \{\pm 1\}$$

(7.3.4.1) Le cube a 4 diagonales, qui passent par les sommets P et $-P$; on note les diagonales $1, \ldots, 4$. Toute isométrie $g \in G(C)$ induit une permutation $\tau(g) \in S_4$ des diagonales, et l'application

$$\tau : G(C) \longrightarrow S_4$$

est un homomorphisme de groupes. Un argument géométrique facile montre que l'on a, pour $g \in G(C)$,

$$\tau(g) = e = (1)(2)(3)(4) \iff g = \pm I_3,$$

ce qui entraîne, pour $g_1, g_2 \in G(C)$,

$$\tau(g_1) = \tau(g_2) \iff \tau(g_1 g_2^{-1}) = \tau(g_1)\tau(g_2)^{-1} = e \iff g_1 g_2^{-1} = \pm I_3 \iff g_1 = \pm g_2.$$

En particulier, un élément $g_1 \in G^+(C)$ est déterminé par $\tau(g_1)$, d'où

$$|G^+(C)| \leq |S_4| = 24, \qquad |G(C)| \leq 2 \cdot |S_4| = 48.$$

(7.3.4.2) D'autre part, on peut écrire directement 48 éléments de $G(C)$, à savoir les matrices suivantes :

$$\begin{pmatrix} \pm1 & 0 & 0 \\ 0 & \pm1 & 0 \\ 0 & 0 & \pm1 \end{pmatrix}, \quad \begin{pmatrix} 0 & \pm1 & 0 \\ \pm1 & 0 & 0 \\ 0 & 0 & \pm1 \end{pmatrix}, \quad \begin{pmatrix} \pm1 & 0 & 0 \\ 0 & 0 & \pm1 \\ 0 & \pm1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 0 & \pm1 & 0 \\ 0 & 0 & \pm1 \\ \pm1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & \pm1 \\ \pm1 & 0 & 0 \\ 0 & \pm1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & \pm1 \\ 0 & \pm1 & 0 \\ \pm1 & 0 & 0 \end{pmatrix}$$

(7.3.4.3) Si l'on met ensemble 7.3.4.1-2, on obtient que

$$|G(C)| = 48, \qquad |G^+(C)| = 24,$$

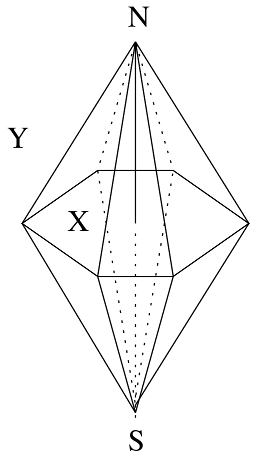
donc τ induit un isomorphisme de groupes entre $G^+(C)$ et S_4. Comme $-I_3 \in G^-(C)$ commute à tous les éléments de $G^+(C)$, on a $G(C) = G^+(C) \times \{\pm I_3\}$.

(7.3.4.4) Exercice. *Trouver une généralisation de 7.3.4.2 en dimension $n \geq 1$ quelconque.*

(7.3.4.5) Exercice. *Y a-t-il un lien entre 7.3.4.4 (pour $n = 2$) et 7.3.2.2 ?*

(7.3.5) Le groupe d'icosaédre. Si $Y \subset \mathbf{R}^3$ est un icosaédre (ou un dodecaédre) régulier de centre à l'origine, alors on a $G^+(Y) = A_5$.

(7.3.6) Les groupes dièdraux (bis). Soit $X \subset \mathbf{R}^2 \subset \mathbf{R}^3$ un polygone régulier à $n \geq 3$ côtés de centre à l'origine. On fixe deux points N, S de la même distance de l'origine sur la droite perpendiculaire au plan $\mathbf{R}^2$; on note $Y \subset \mathbf{R}^3$ l'ensemble obtenu en joignant N et S à tous les points de X :



Dans ce cas on obtient une autre réalisation du groupe dièdral

$$G(X) = D_n \subset O(2),$$

à savoir

$$G^+(Y) = G(X) = D_n \subset O^+(3).$$

En effet, tout élément de $C_n = G^+(X) = D_n \cap O^+(2)$ s'étend à une rotation dans $\mathbf{R}^3$ d'axe NS (et d'angle $2\pi k/n$), tandis que toute réflexion $s_H \in G^-(X)$ (où $H \subset \mathbf{R}^2$ est une droite qui passe par l'origine et soit par un sommet de X, soit par le milieu d'un côté de X) s'étend à une rotation dans $\mathbf{R}^3$ d'axe H et d'angle π.

(7.3.7) Voici une description de tous les sous-groupes finis $G \subset O^+(3)$ (due à Felix Klein) :

(1) $G = C_n$ ou $G = D_n$ (voir 7.3.6; les valeurs dégénérées $n = 1, 2$ sont autorisées).
(2) $G = A_4$; dans ce cas il existe un tetraédre régulier T tel que $G = G^+(T)$.
(3) $G = S_4$; dans ce cas il existe un cube C et un octaédre régulier O tels que $G = G^+(C) = G^+(O)$.
(4) $G = A_5$; dans ce cas il existe un dodecaédre régulier D et un icosaédre régulier I tels que $G = G^+(D) = G^+(I)$.

Cette description entraîne une classification des polyédres réguliers dans $\mathbf{R}^3$.

7.4 Orbites, fixateurs

(7.4.1) Parfois, on peut déterminer l'ordre du groupe $G(Y)$ (où $G^+(Y)$) directement, en utilisant la formule 7.4.3 ci-dessous.

(7.4.2) Définition. *Soit X un ensemble, G un groupe et $\sigma : G \longrightarrow S_X$ un homomorphisme de groupes (par exemple, G peut être un sous-groupe de S_X). Dans ce cas on dit que G* **agit sur** *X. Pour un élément $x \in X$, on définit* **l'orbite de** *x*

$$O(x) = \{\sigma(g)(x) \mid g \in G\}$$

comme l'image de x sous l'action de G; on définit **le fixateur de** *x*

$$G_x = \{g \in G \mid \sigma(g)(x) = x\}$$

comme le sous-groupe de G contenant tous les éléments qui fixent x.

(7.4.3) Proposition. *On suppose qu'un groupe fini G agit sur un ensemble X. Alors on a, pour tout $x \in X$,*

$$|G| = |O(x)| \cdot |G_x|.$$

(7.4.4) Exemples: (1) Lorsque $G = S_n$ $(n \geq 2)$, $X = \{1, \ldots, n\}$ et $x = n$, on a

$$O(x) = \{1, \ldots, n\}, \qquad G_x = S_{n-1},$$

donc

$$|S_n| = n \cdot |S_{n-1}|,$$

ce qui entraîne, par récurrence,

$$|S_n| = n \cdot (n-1) \cdots 2 \cdot 1 = n!$$

(2) Lorsque $C \subset \mathbf{R}^3$ est un cube de centre à l'origine, le groupe $G = G(C)$ agit sur l'ensemble X des côtés de C. Chacun de six côtés $x \in X$ est un carré; on a

$$O(x) = X, \qquad G_x = D_4,$$

d'où $|G(C)| = 6 \cdot 8 = 48$.

(3) Si $C(n) \subset \mathbf{R}^n$ est un cube de dimension n de centre à l'origine, le groupe $G(n) = G(C(n))$ agit sur l'ensemble X des côtés de dimension $n-1$ de $C(n)$. Chaque $x \in X$ est un cube de dimension $n-1$,

$$O(x) = X, \qquad G_x = C(n-1),$$

d'où

$$|C(n)| = 2n \cdot |C(n-1)|.$$

Il en résulte, par récurrence, que l'on a

$$|C(n)| = 2n \cdot (2n-2) \cdots 4 \cdot 2 = 2^n \cdot n!$$

(7.4.5) Exercise. *Y a-t-il un lien entre 7.4.4(3) et 7.3.4.4 ?*