MATHEMATIQUES
4. ALGEBRE

ELIE AZOULAY & JEAN AVIGNANT

ENSEIGNEMENT SUPERIEUR TECHNIQUE

COURS ET
EXERCICES

McGRAW-HILL

ELIE AZOULAY

Ingénieur des Télécommunications
Professeur à l'Ecole Supérieure
d'Ingénieurs en Electrotechnique
et Electronique

JEAN AVIGNANT

Agrégé de l'Université
Professeur à l'I.U.T.
de Ville-d'Avray

MATHEMATIQUES
4. ALGEBRE

COURS ET EXERCICES

McGRAW-HILL

Auckland - Beyrouth - Bogota - Hambourg - Johannesburg
Lisbonne - Londres - Madrid - Mexico - Montréal
New Delhi - New York - Panama - Paris - San Juan
São Paulo - Singapour - Sydney - Tokyo - Toronto

1984

Maquette de couverture : Françoise Rojare

ISBN : 2-7042-1089-6

McGraw-Hill — 28 Rue Beaunier — 75014 Paris

AVANT-PROPOS

L'ouvrage que nous présentons ici s'adresse aux étudiants des écoles supérieures techniques, des I.U.T. et aux élèves de certaines écoles d'ingénieurs — celles en particulier pour lesquelles le passage traditionnel par les classes de mathématiques spéciales n'est pas requis — ainsi qu'aux étudiants du premier cycle universitaire.

L'enseignement des Mathématiques a beaucoup évolué durant les deux dernières décennies. Un formalisme plus rigoureux, des concepts plus généraux et de ce fait plus abstraits se sont imposés. Il demeure cependant un domaine où une attention toute particulière doit être apportée à la pédagogie des mathématiques, c'est celui des enseignements à finalité technique. Dans de tels cas, l'axiomatique rigoureuse, le concept général et abstrait cèdent nécessairement le pas au souci de doter l'élève des moyens utiles à faire rapidement usage d'un certain nombre « d'outils mathématiques » lui permettant d'appréhender, par le biais de modèles appropriés, les notions techniques qui constituent l'essentiel de sa formation. D'ailleurs, nos étudiants sont ceux qui ressentent le plus l'ennui des Mathématiques telles qu'elles sont enseignées aujourd'hui. De celles parfaitement agencées de l'enseignement secondaire aux superbes théories de l'enseignement supérieur, le tout constitue bien peut-être un merveilleux édifice sans faille, à cela près qu'il finit souvent par inspirer morosité et ennui.

Notre choix a été celui d'un formalisme restreint que nous croyons être le plus adéquat et le plus apprécié. Nous avons cependant évité de prendre des libertés quant à la rigueur, et de ce fait, tenu à donner les démonstrations de la plupart des théorèmes et formules, n'omettant délibérément, dans un souci d'allégement, que certaines jugées trop complexes.

Chaque chapitre est suivi d'un résumé et de très nombreux exercices, certains entièrement résolus, d'autres comportant seulement la réponse avec des indications sur la méthode, ceci en particulier à l'usage des autodidactes et des étudiants de formation continue.

Le premier tome reprend les notions de base d'analyse acquises en terminale, qu'il complète et développe concernant les fonctions d'une variable réelle et les fonctions transcendantes usuelles. Il se poursuit par les chapitres consacrés à la notion d'intégrale, aux procédés classiques de recherche de primitives, et enfin à l'extension de la notion d'intégrale. Un dernier chapitre porte sur les fonctions de plusieurs variables dont les résultats seront fréquemment mis à profit dans les tomes suivants.

Le tome 2 complète le précédent pour couvrir l'essentiel de l'analyse du premier cycle. On y trouve les équations différentielles, les intégrales multiples, l'intégrale curviligne, les formules intégrales d'analyse vectorielle ainsi que l'étude des suites et séries (numériques et de fonctions).

Le tome 3 d'analyse aborde les compléments traités au niveau du second cycle, tant dans les écoles d'ingénieurs qu'à l'université, des Mathématiques de la Physique et de la technique : fonctions eulériennes, transformations de Laplace et de Fourier, fonctions de Bessel, fonctions spéciales, équations aux dérivées partielles, fonctions d'une variable complexe, etc.

Le tome 4 rassemble les notions classiques d'algèbre, l'algèbre linéaire ainsi que certaines applications à l'analyse et à la géométrie.

Le tome 5 est celui de la géométrie tant analytique qu'infinitésimale : calculs et opérateurs vectoriels, droites et cercles, coniques et quadriques, fonctions vectorielles d'une variable réelle, courbes planes paramétrées, courbes en polaires et, enfin, courbure et torsion et enveloppes des courbes planes et gauches.

Il est évident que le plan suivi n'est pas nécessairement celui des cours tels qu'ils se déroulent chronologiquement et, de ce fait, de nombreux renvois d'un tome à l'autre ont été indispensables.

Les auteurs remercient par avance le lecteur des critiques ou suggestions qu'il voudra bien leur formuler ainsi que les Editions McGraw-Hill pour le soin qu'elles ont apporté à la réalisation de cet ouvrage.

Jean AVIGNANT
Elie AZOULAY

TABLE DES MATIÈRES

CHAPITRE 1

NOTION D'ENSEMBLE. OPÉRATIONS SUR LES ENSEMBLES

1. ENSEMBLE

Des objets ou éléments présentant une ou plusieurs propriétés communes constituent une collection ou un **ensemble.**

Ces propriétés sont suffisantes pour affirmer qu'un objet appartient ou n'appartient pas à l'ensemble.

Exemples.

Ensemble des nombres entiers.
Ensemble des vecteurs de l'espace à trois dimensions.

2. APPARTENANCE

Si e est l'un des éléments constituant l'ensemble E, on dit que e **appartient à** E et on note

$$e \in E$$

si l'élément e' n'appartient pas à l'ensemble E on écrit $e' \notin E$.

3. DÉTERMINATION D'UN ENSEMBLE

Un ensemble est déterminé si, étant donné un élément quelconque, on peut savoir s'il appartient ou non à cet ensemble.

Dans les cas usuels un ensemble est déterminé

— soit par l'énumération de tous ses éléments, on dit alors que l'ensemble est défini en extension;

— soit par une propriété caractéristique de ses éléments, il s'agit alors d'une définition en compréhension.

Par exemple si on désigne par $\mathbb{N}$ l'ensemble des entiers naturels on aura

$$\mathbb{N} = \{0, 1, 2, 3, \ldots, n, \ldots\}$$

L'ensemble des nombres pairs se notera

$$P = \{x/x = 2n, \forall n \in \mathbb{N}\}$$

qui se lit P est l'ensemble des nombres x tels que x égale $2n$, pour tout n appartenant à $\mathbb{N}$.

4. ÉGALITÉ DE DEUX ENSEMBLES

Deux ensembles sont égaux (ou identiques) si tout élément de l'un est élément de l'autre.

Ainsi les deux ensembles

$$A = \{x, 1, 0, \square\}$$

et

$$B = \{\square, 1, x, 0\}$$

sont égaux et on écrit $A = B$.

5. INCLUSION

On dit que l'ensemble A est **inclus** dans l'ensemble B, ce qui se note

$$A \subset B$$

lorsque tout élément de A appartient à B. On dit que A est inclus dans B ou est un sous-ensemble de B.

Ainsi, si on désigne par $\mathbb{R}$ l'ensemble des nombres réels quelconques, nous aurons

$$\mathbb{N} \subset \mathbb{R}$$

Les inclusions $A \subset B$ et $B \subset C$ entraînent $A \subset C$, on dit que la relation d'inclusion est transitive.

Ainsi, si on désigne par $\mathbb{Z}$ l'ensemble des entiers relatifs (..., -2, -1, 0, 1, 2,...) et par $\mathbb{Q}$ celui des nombres rationnels $\frac{a}{b}$ avec $b \neq 0$ $a \in \mathbb{Z}$ et $b \in \mathbb{Z}$ nous aurons

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

Ensemble vide

L'ensemble **vide** est l'ensemble ne contenant aucun élément. On le note $\varnothing$.

$\varnothing$ est inclus dans tout ensemble. Cette remarque, somme toute triviale, est fréquemment utilisée dans de nombreux raisonnements.

6. VISUALISATION DES ENSEMBLES. DIAGRAMMES

Par commodité et pour rendre plus assimilables certaines notions on utilise différents diagrammes pour représenter les ensembles. Les plus communément utilisés sont obtenus en délimitant par une courbe fermée une zone dans l'espace euclidien de la feuille de papier.

Remarquons que chaque fois que nous définissons un ensemble nous le définissons en fait comme un sous-ensemble d'un autre ensemble. Lorsque nous disons soit l'ensemble A composé des lettres a, b, c nous le définissons comme un sous-ensemble de l'ensemble E des lettres de l'alphabet

$$\{a, b, c, \ldots, y, z\}$$

pris comme référence.

C'est pourquoi toute visualisation d'un ensemble commence par celle de son ensemble référentiel. Celui-ci est généralement représenté par l'ensemble des points, ou croix ou tout autre symbole représentant ses éléments, disposés à l'intérieur d'un rectangle.

Exemples.

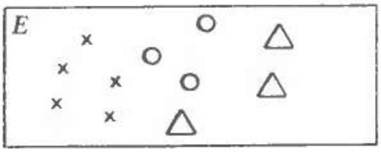


Figure 1.

Tout sous-ensemble de ce référentiel sera délimité par une courbe fermée englobant l'ensemble de ses éléments, les éléments extérieurs à cette courbe appartenant à son complémentaire par rapport au référentiel utilisé.

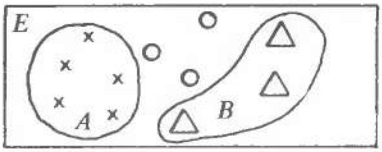


Figure 2.

La figure 3 met en évidence le complémentaire de A par rapport à E.

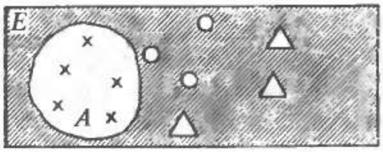


Figure 3.

7. OPÉRATIONS SUR LES ENSEMBLES

1. Réunion

La réunion des deux ensembles A et B est l'ensemble C des éléments qui appartiennent à A ou à B.

On écrit $$C = A \cup B$$
(qui se lit C égale A union B).

L'équivalence suivante traduit la définition

$$\boxed{x \in C = A \cup B \Leftrightarrow x \in A \quad \text{ou} \quad x \in B}$$

Ceci reste vrai si $x \in A$ et $x \in B$.

2. Intersection

L'intersection de deux ensembles A et B est l'ensemble C des éléments qui appartiennent à A et à B.

On écrit alors $$C = A \cap B$$
(qui se lit C égale A inter B).

On a de même

$$\boxed{x \in C = A \cap B \Leftrightarrow x \in A \quad \text{et} \quad x \in B}$$

Les schémas ci-dessous visualisent ces deux opérations.

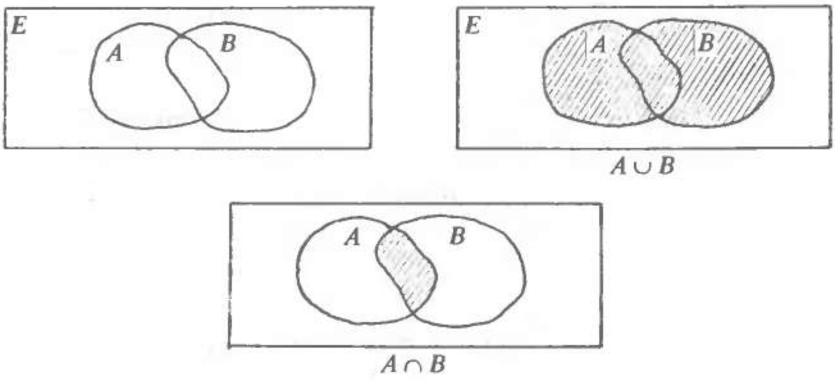


Figure 4.

Si A et B n'ont pas d'élément commun on dit qu'ils sont **disjoints,** alors $A \cap B = \varnothing$. Ainsi l'ensemble des nombres naturels pairs et l'ensemble des nombres naturels impairs ont une intersection vide.

3. Complémentation

Deux sous-ensembles A et B de l'ensemble E sont dits complémentaires si leur réunion est l'ensemble E et leur intersection l'ensemble vide :

$$A \cup B = E$$

$$A \cap B = \varnothing$$

on note alors $$B = \complement_E^A$$

ou encore $$B = A^c$$

ou enfin $$B = \overline{A}$$

ces deux dernières notations sont d'ailleurs à éviter si l'ensemble E par rapport auquel on prend le complément n'est pas évident.

La zone hachurée ci-dessous donne $\overline{A}$.

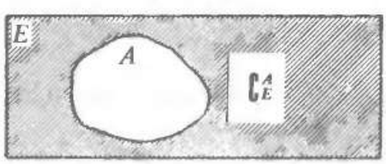


Figure 5.

8. PROPRIÉTÉS DES OPÉRATIONS SUR LES ENSEMBLES

Les propriétés suivantes se démontrent très aisément.

1. Commutativité

L'intersection et la réunion sont des opérations commutatives. Quels que soient les ensembles A et B

$$A \cap B = B \cap A$$

et

$$A \cup B = B \cup A$$

2. Associativité

L'intersection et la réunion sont associatives quels que soient les ensembles A, B et C :

$$(A \cap B) \cap C = A \cap (B \cap C)$$
$$(A \cup B) \cup C = A \cup (B \cup C)$$

3. Distributivité

L'intersection et la réunion sont distributives l'une par rapport à l'autre

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

4. Idempotence

$$A \cup A = A$$
$$A \cap A = A$$

9. LOIS DE MORGAN (*)

Quels que soient les sous-ensembles A et B de E

$$\boxed{(A \cup B)^c = A^c \cap B^c}$$

et

$$\boxed{(A \cap B)^c = A^c \cup B^c}$$

le complémentaire de la réunion de deux ensembles est l'intersection des complémentaires de ces ensembles.

* La démonstration fait l'objet de l'exercice 3.

Le complémentaire de l'intersection de deux ensembles est la réunion des complémentaires de ces ensembles.

On dit qu'il existe une dualité entre les opérations de réunion et d'intersection.

10. DIFFÉRENCE DE DEUX ENSEMBLES

On appelle différence des ensembles A et B, l'ensemble

$$A - B = A \cap B^c$$

et différence symétrique de A et B l'ensemble noté $A \Delta B$ tel que

$$A \Delta B = (A \cap B^c) \cup (A^c \cap B)$$

$A \Delta B$ est l'ensemble des éléments qui appartiennent à un et un seul des ensembles A et B.

11. NOTION DE RECOUVREMENT ET DE PARTITION

Si on désigne par I un ensemble d'indices et une correspondance $f(i) = x_i \in E$ alors $\{x_i\}$, ensemble des x_i, est une partie de E indexée par I. On désigne par $\mathscr{P}(E)$ l'ensemble des parties de E.

On appelle recouvrement de E la réunion $\cup X_i = E$ c'est-à-dire $\forall x \in E \quad \exists X_i$ tel que $x_i \in X_i$

Si de plus $\begin{cases} \forall i \in I & X_i \neq \varnothing \\ X_i \cap X_j = \varnothing & \text{si } i \neq j \end{cases}$

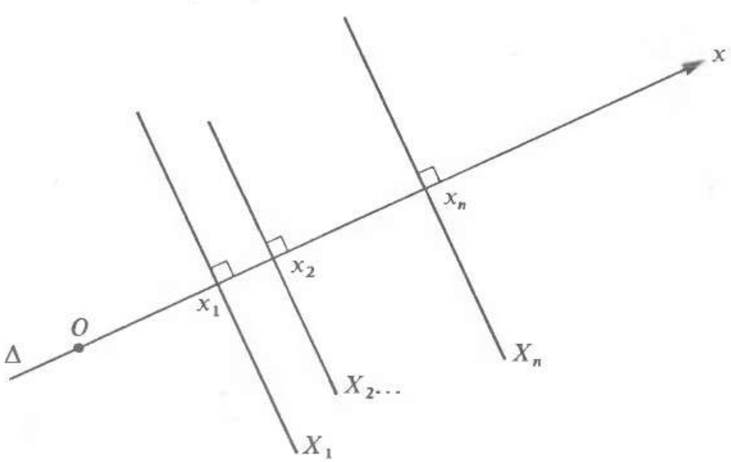


Figure 6.

on dit que les X_i forment une partition de E.

Ainsi dans le cas $E = R^2$

les perpendiculaires à Ox aux points d'abscisses $x_1, x_2, \ldots, x_n$ forment une partition de E.

12. ENSEMBLES USUELS EN MATHÉMATIQUES

On désigne généralement les ensembles les plus usuels par une lettre en gras ou à double barre :

N ou $\mathbb{N}$		l'ensemble des entiers naturels
N* ou $\mathbb{N}^*$		l'ensemble des entiers positifs
Z ou $\mathbb{Z}$		l'ensemble des entiers relatifs (positifs, négatifs ou nuls)
Q ou $\mathbb{Q}$		l'ensemble des nombres rationnels $\left(\frac{p}{q}, p \text{ et } q \in Z\right)$
R ou $\mathbb{R}$		l'ensemble des réels
$\mathbf{R}^+$ ou $\mathbb{R}^+$		l'ensemble des réels positifs
C ou $\mathbb{C}$		l'ensemble des nombres complexes

13. PRODUIT D'ENSEMBLES

Soit deux ensembles A et B et deux éléments a et b avec $a \in A$ et $b \in B$.

L'ensemble des couples (a, b) pris dans cet ordre est appelé ensemble produit cartésien des ensembles A et B.

On le note $A \times B$

— Il faut d'abord remarquer que la précision de l'ordre a puis b n'est pas superflue. Tout couple appartenant à $A \times B$ est constitué d'un élément appartenant à A puis d'un élément appartenant à B. (b, a) n'est pas en général élément de $A \times B$ (ce n'est le cas que si a et $b \in A \cap B$ ou si $a = b$).

— De plus si A et B sont des ensembles finis et si on désigne par

card A le nombre des éléments de A

et card B le nombre des éléments de B

on aura card $(A \times B)$ = card $A \times$ card B. En effet le nombre des couples du type (a, b) avec $a \in A$ et $b \in B$ est obtenu en faisant correspondre à tout élément $a \in A$ tous les éléments de B, soit (card B) éléments. Ceci devant être répété autant de fois qu'il y a d'éléments dans A, le nombre des éléments de $A \times B$ est bien card $A \times$ card B.

— Le produit cartésien $A^2 = A \times A$ est l'ensemble des couples (a_1, a_2) avec $a_1 \in A$ et $a_2 \in A$ et d'une façon générale

$$A^n = A \times A \times \ldots \times A$$

est l'ensemble des familles à n éléments

$$(a_1, a_2, \ldots, a_n)$$

où $a_1 \in A$, $a_2 \in A$, …, $a_n \in A$.

14. RELATION. GRAPHE D'UNE RELATION

Soit une propriété caractéristique des éléments d'un sous-ensemble G de $A \times B$: elle définit une **relation** R entre a et b, a décrivant A et b décrivant B.

G s'appelle le graphe de la relation R.

Exemple.

L'ensemble des nombres réels $\mathbf{R}$ peut être représenté géométriquement par une droite et l'ensemble des couples de nombres réels, c'est-à-dire les éléments de l'ensemble $\mathbf{R}^2$ peut être représenté par un plan : chaque point du plan est caractérisé par son ordonnée et son abscisse dans un repère formé par deux droites orthogonales.

Dans ce repère la droite d'équation $y = 3x$, par exemple, est un sous-ensemble de $\mathbf{R}^2$: cette droite est le graphe de la relation R, $y = 3x$.

Résumé du chapitre 1

- **Appartenance.**

$$a \in A$$

- **Inclusion.**

$$A \subset B \Rightarrow \forall \alpha \in A, \quad \alpha \in B$$

- **Réunion d'ensembles.**

$$x \in A \cup B \Rightarrow x \in A \quad \text{ou (et)} \quad x \in B$$

- **Intersection d'ensembles.**

$$x \in A \cap B \Rightarrow x \in A \quad \text{et} \quad x \in B$$

• **Complémentation.**

$$x \notin A \Rightarrow x \in \complement_E^A \quad \text{ou} \quad x \in \overline{A}$$

• **Propriétés des opérations de réunion et d'intersection.**

— Elles sont commutatives.

$$A \cap B = B \cap A \quad \text{et} \quad A \cup B = B \cup A$$

— Elles sont associatives.

$$A \cup (B \cup C) = (A \cup B) \cup C \quad \text{et} \quad A \cap (B \cap C) = (A \cap B) \cap C$$

— Chacune des lois est distributive par rapport à l'autre.

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

• **Différence de *A* et *B*.**

$$A - B = A \cap B^c$$

• **Différence symétrique de *A* et *B***

$$A \Delta B = (A \cap B^c) \cup (A^c \cap B)$$

• **Lois de Morgan.**

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

• **Recouvrement de *E*.**

$\{X_i\}_{i \in I}$ constituent un recouvrement de E si

$$\cup X_i = E$$

• **Partition de *E*.**

$\{X_i\}_{i \in I}$ constituent une partition de E si

$$\cup X_i = E$$
$$X_i \cap X_j = \varnothing \quad \forall_i, \forall_j \quad i \neq j$$

• **Produit cartésien d'ensembles.**

$$A \times B = \{(a, b) \quad \text{avec} \quad a \in A \quad \text{et} \quad b \in B\}$$

EXERCICES

■ **1.** Soit A, B, C trois sous-ensembles de E. Montrer que si

$$A \cup C \subset A \cup B \text{ et } A \cap C \subset A \cap B,$$

l'ensemble C est inclus dans B.

Soit $x \in C$. Donc $x \in A \cup C$.
$A \cup C \subset A \cup B \Rightarrow x \in A \cup B$ c'est-à-dire x appartient à A ou B:
Si $x \in A$, $x \in A \cap C$ et d'après la relation $A \cap C \subset A \cap B$, $x \in B$.
Donc $x \in B$, pour tout x de C. D'où $C \subset B$.

■ **2.** Montrer que si F et G sont des sous-ensembles de E:

a) $F \subset G \Leftrightarrow F \cup G = G \Leftrightarrow F^c \cup G = E$.
b) En déduire que $F \subset G \Leftrightarrow F \cap G = F \Leftrightarrow F \cap G^c = \varnothing$.

---

a) $F \subset G$ entraîne que tout x appartenant à F appartient à G. $x \in F \cup G$ veut dire x appartient à F ou G, donc à G: $F \cup G = G$.

Si $x \in F^c$ et $F \subset G$, x appartient, soit à G, soit à G^c: donc $G^c \subset F^c$ et comme $G^c \cup G = E$, *a fortiori* $F^c \cup G = E$.

Réciproquement $F^c \cup G = E \Rightarrow G^c \subset F^c \Rightarrow F \subset G$.

b) $F \subset G \Rightarrow$ tout x appartenant à F appartient à G, donc à $F \cap G$. Donc $F = F \cap G$ et réciproquement si $F = F \cap G$, tout x appartenant à $F \cap G$ appartient à F, donc $F \subset G$. Tout x appartenant à F appartient à G, donc est hors de G^c: $F \cap G^c$ est donc l'ensemble vide. Réciproquement si $F \cap G^c = \varnothing$ tout x de F est hors de G^c donc dans G et $F \subset G$.

■ **3.** Soit A et B, deux sous-ensembles de E. Démontrer les lois de Morgan:

a) $A^c \cup B^c = (A \cap B)^c$.
b) $A^c \cap B^c = (A \cup B)^c$.

---

a) Soit x tel que $x \in A^c \cup B^c$: cela signifie que x appartient à A^c ou B^c.

Si $x \in A^c$, $\quad x \notin A \Rightarrow x \notin A \cap B \Rightarrow x \in (A \cap B)^c$
Si $x \in B^c$, $\quad x \notin B \Rightarrow x \notin A \cap B \Rightarrow x \in (A \cap B)^c$

Ceci est vrai pour tout x de $A^c \cup B^c$, donc on en déduit

$$A^c \cup B^c \subset (A \cap B)^c.$$

Soit x tel que $x \in (A \cap B)^c$: x n'est pas un élément commun à A et B.

Si $x \in A$, $x \notin B \Rightarrow x \in B^c \Rightarrow x \in A^c \cup B^c$.
Si $x \in B$, $x \notin A \Rightarrow x \in A^c \Rightarrow x \in A^c \cup B^c$.
Si $x \notin A$ et $x \notin B \Rightarrow x \in A^c$ et $x \in B^c \Rightarrow x \in A^c \cup B^c$.

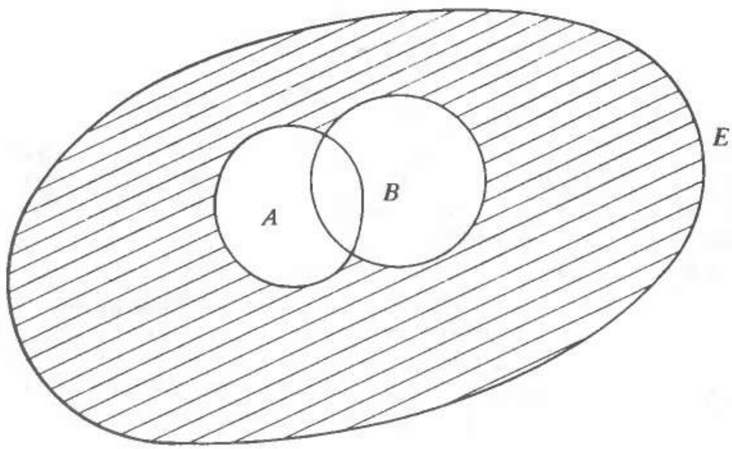
Ceci est vrai pour tout x de $(A \cap B)^c$. On en déduit que

$$(A \cap B)^c \subset A^c \cup B^c$$

On a démontré plus haut que $A^c \cup B^c \subset (A \cap B)^c$. La seule possibilité est

$$(A \cap B)^c = A^c \cup B^c$$

b) Ayant démontré *a*) rigoureusement, nous nous contentons du schéma intuitif d'Euler :



La partie ombrée représente $A^c \cap B^c$. On vérifie alors que

$$A^c \cap B^c = (A \cup B)^c$$

■ **4.** Soit des ensembles E et F. Si A est un sous-ensemble de E et B un sous-ensemble de F, montrer que $A \times B$ est inclus dans $E \times F$.

Soit (x, y) un élément de $A \times B$. Cela signifie que $x \in A$ et $y \in B$. Donc $x \in E$ et $y \in F$ et $(x, y) \in E \times F$. Ceci est vrai pour tout (x, y) de $A \times B$, donc

$$A \times B \subset E \times F$$

■ 5. Nombre de parties d'un ensemble fini.

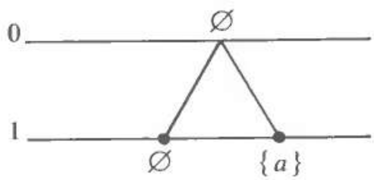
Il est souvent nécessaire de pouvoir dénombrer tous les sous-ensembles que l'on peut former avec les éléments d'un ensemble fini E. Dans ce dénombrement il ne faut pas oublier l'ensemble vide $\varnothing$ ainsi que l'ensemble E qui est une partie de lui-même, celle contenant tous les éléments.

Nous voulons dénombrer les parties d'un ensemble contenant n éléments. Une méthode simple est donnée dans le chapitre « Analyse combinatoire ». Cependant la méthode graphique dite méthode de l'arbre est souvent utilisée pour obtenir le résultat et nous la présentons ici.

On divise le plan (ici la feuille de papier) en niveaux, de haut en bas, que l'on numérote. Chaque niveau représente un ensemble et le numéro de ce niveau est le nombre d'éléments de l'ensemble.

Chaque niveau (représenté par une droite horizontale) comporte un certain nombre de points, chaque point représentant une partie de l'ensemble. Donc si l'on sait construire ces points, on connaîtra le nombre de parties de l'ensemble.

Il est facile de construire le niveau 0 et le niveau 1 :

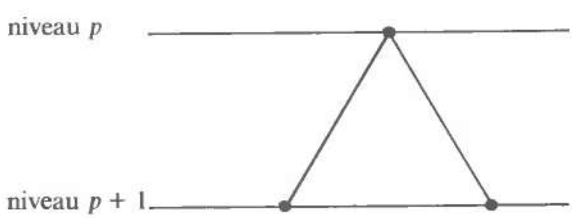


: une partie, $\varnothing$ lui-même

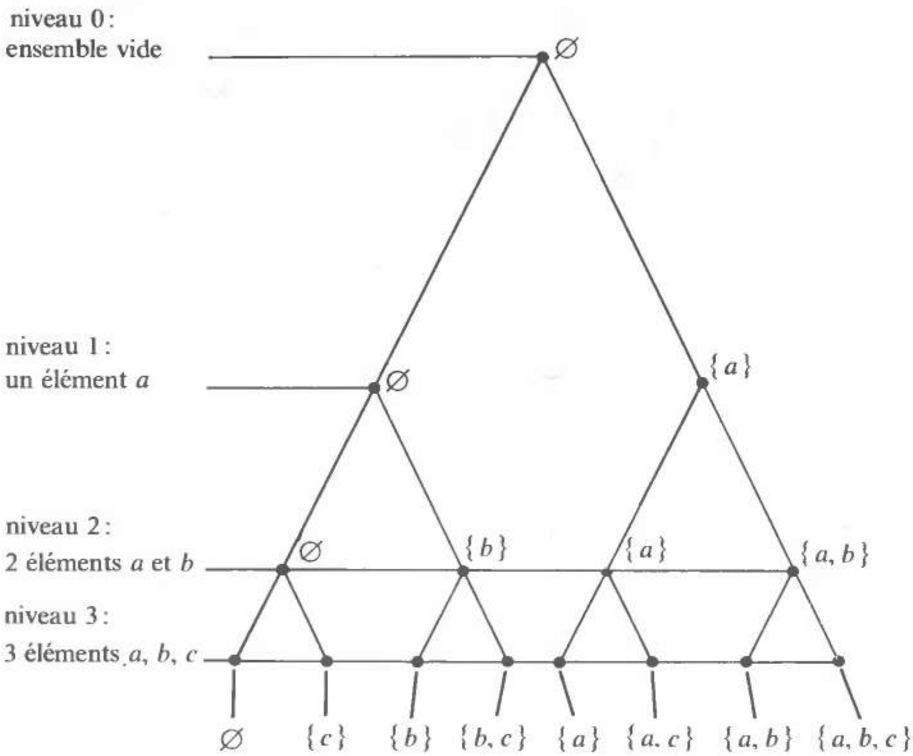
: deux parties, $\varnothing$ et $\{a\}$.

Le problème est alors le suivant : connaissant le nombre de parties d'un ensemble E_p à p éléments, quel est le nombre de parties de l'ensemble E_{p+1} à $p + 1$ éléments, tel que E_p soit inclus dans E_{p+1} ?

E_p est inclus dans E_{p+1} : ceci signifie qu'on obtient E_{p+1} en ajoutant un élément α à E_p. Il s'offre alors une alternative : une partie de E_{p+1} contient α ou ne contient pas α. Les parties de E_{p+1} ne contenant pas α sont les parties de E_p; d'autre part, les parties de E_{p+1} contenant α sont aussi celles de E_p auxquelles on ajoute l'élément α. Une partie de E_p fournit donc deux parties de E_{p+1} ce que l'on traduit par le schéma :



A partir du niveau 0, on peut donc construire l'arbre suivant :



Le niveau 0 possède une partie soit 2^0, le niveau 1 possède deux parties soit 2^1; le niveau 2 possède quatre parties soit 2^2 et le niveau 3 possède huit parties soit 2^3. Le nombre de parties d'un ensemble E_p est donc 2^p pour $p = 0, 1, 2, 3$. La méthode de l'arbre montre que ce résultat est vrai pour tout p.

■ **6.** Soit deux ensembles $A = \{ a_1, a_2, a_3, a_4 \}$ et $B = \{ b_1, b_2, b_3, b_4, b_5 \}$

a) Écrire le produit cartésien $A \times B$.

b) Quel est le nombre de parties de $A \times B$?

$$a)\quad A \times B = \left\{ \begin{matrix} a_1\, b_1, & a_1\, b_2, & a_1\, b_3, & a_1\, b_4, & a_1\, b_5 \\ a_2\, b_1, & a_2\, b_2, & a_2\, b_3, & a_2\, b_4, & a_2\, b_5 \\ a_3\, b_1, & a_3\, b_2, & a_3\, b_3, & a_3\, b_4, & a_3\, b_5 \\ a_4\, b_1, & a_4\, b_2, & a_4\, b_3, & a_4\, b_4, & a_4\, b_5 \end{matrix} \right\}$$

b) card $(A \times B) = 20 =$ card $A \times$ card B.
card $\mathscr{P}\,(A \times B) = 2^{20}$.

■ 7. Soit l'ensemble E possédant n éléments. Quel est le nombre d'éléments de E^p, p entier ? Quel est le nombre de parties de E^p ?

E^p désigne le produit cartésien $\underbrace{E \times E \times \ldots \times E.}_{p \text{ fois}}$

Le nombre d'éléments de E^p est n^p.

Raisonnons par récurrence : soit le produit $E_1 \times E_2$; à chaque élément a de E_1 on peut faire correspondre tous les éléments de E_2 un par un. On forme ainsi n éléments de $E_1 \times E_2$. Or $E_1 \equiv E_2 \equiv E$, c'est-à-dire que E_1 renferme n éléments. $E_1 \times E_2 = E^2$ renferme donc n^2 éléments. Supposons que E^{p-1} renferme n^{p-1} éléments; cherchons le nombre d'éléments de $E^p = E^{p-1} \times E$. A chaque élément de E, on peut faire correspondre n^{p-1} éléments de E^{p-1} et former n^{p-1} éléments de E^p. Il y a n éléments dans E, donc $n^{p-1} \times n = n^p$ éléments dans E^p.

Le nombre de parties de E^p est alors 2^{n^p} (voir exercice 5).

■ 8. Soit l'ensemble E : $\{1, 2, 3, 4, 5, 6\}$. On définit sur E^2 la relation $x : y$ qui signifie y divise x, $y \in E$, $x \in E$.

Quel est le graphe de cette relation ?

Le graphe de la relation : est l'ensemble des couples (x, y) de E^2 tels que y divise x :

$$G = \{(2, 1), (3, 1), (4, 1), (5, 1), (6, 1), (4, 2), (6, 2), (6, 3)\}$$

■ 9. On appelle fonction caractéristique de A, partie de l'ensemble E, une application f de E dans l'ensemble à deux éléments $\{0, 1\}$, telle que

$$f(x) = 0 \quad \text{si} \quad x \notin A$$
$$f(x) = 1 \quad \text{si} \quad x \in A$$

Soit A et B deux parties de E et leurs fonctions caractéristiques f et g. Quels sont les ensembles A_1, A_2, A_3 dont les fonctions caractéristiques sont :

a) $1 - f$;

b) fg;

c) $f + g - fg$.

---

a) Soit A_1 l'ensemble dont la fonction caractéristique est $1 - f$.

Si $x \in A$, $\quad 1 - f(x) = 0 \Rightarrow x \notin A_1$
Si $x \notin A$, $\quad 1 - f(x) = 1 \Rightarrow x \in A_1$

donc $A_1 = A^c$.

b) Soit A_2 l'ensemble dont la fonction caractéristique est fg.

Si $x \in A$	et	$x \notin B$	$fg = 0$
Si $x \notin A$	et	$x \in B$	$fg = 0$
Si $x \in A$	et	$x \in B$	$fg = 1$

donc $A_2 = A \cap B$.

c) Soit A_3 l'ensemble dont la fonction caractéristique est $f + g - fg$.

Si $x \in A$	et	$x \notin B$	$f + g - fg = 1$
Si $x \notin A$	et	$x \in B$	$f + g - fg = 1$
Si $x \in A \cap B$			$f + g - fg = 1$
Si $x \notin A \cup B$			$f + g - fg = 0$

donc $A_3 = A \cup B$.

■ **10.** Soit un ensemble E et deux parties A et B de E.

a) Démontrer que $A \,\Delta\, B = \complement_A (A \cap B) \cup \complement_B (A \cap B)$.

b) Démontrer que quelles que soient les parties A, B, C de E

$$(A \,\Delta\, B) \,\Delta\, C = A \,\Delta\, (B \,\Delta\, C)$$

c) Démontrer qu'il existe une partie unique X de E telle que

i) pour toute partie A de E: $A \,\Delta\, X = X \,\Delta\, A = A$,

ii) il existe une partie unique A' de E telle que

$$A \,\Delta\, A' = A' \,\Delta\, A = X.$$

---

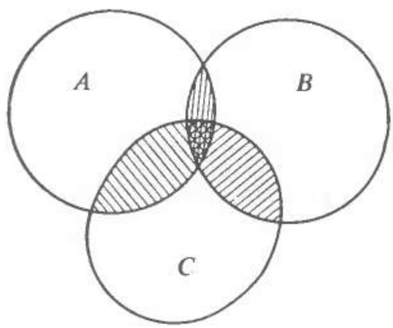
a) C'est évident; cela résulte de la définition de la différence symétrique.

b) $(A \,\Delta\, B) \,\Delta\, C$ est l'ensemble des éléments appartenant à A, B ou C mais n'appartenant pas à $A \cap B$, ni à $A \cap C$, ni à $B \cap C$.

De même $A \,\Delta\, (B \,\Delta\, C)$ est l'ensemble des éléments appartenant à A, B ou C mais n'appartenant pas à $B \cap C$, $A \cap C$ ou $A \cap B$.

Il y a bien identité.

Le diagramme d'Euler donne d'ailleurs un aperçu du résultat.



Les parties en blanc sont $A \Delta B \Delta C$.

c) *i*) $A \Delta X = X \Delta A = A \Leftrightarrow \complement_A (A \cap X) \cup \complement_X (A \cap X) = A$

et $\Rightarrow \complement_X (A \cap X) \subset A.$

Or, d'après la définition du complémentaire d'un ensemble par rapport à un autre ensemble, il vient

$$X = \complement_X (A \cap X) \cup (A \cap X) \quad \text{et} \quad \complement_X (A \cap X) \cap (A \cap X) = \varnothing$$

Or $(A \cap X) \subset A$ et $\complement_X (A \cap X) \subset A$

On en déduit donc $\complement_X (A \cap X) \cup (A \cap X) \subset A$

c'est-à-dire $X \subset A$

D'où $A \cap X = X$ et $\complement_X (A \cap X) = \varnothing$

L'hypothèse s'écrit alors

$$\complement_A (X) \cup \varnothing = \complement_A (X) = A$$

ce qui implique

$$X = \varnothing$$

ii) $\complement_A (A \cap A') \cup \complement_{A'} (A \cap A') = \varnothing \Rightarrow \begin{cases} \complement_A (A \cap A') = \varnothing \\ \complement_{A'} (A \cap A') = \varnothing \end{cases}$

$$\left.\begin{array}{l} \complement_A (A \cap A') = \varnothing \Rightarrow A \cap A' = A \\ \complement_{A'} (A \cap A') = \varnothing \Rightarrow A \cap A' = A' \end{array}\right\} \Rightarrow A' = A.$$

■ 11. Soit $\mathscr{P}(E)$ l'ensemble des parties de E, on désigne par φ une application $\mathscr{P}(E) \to E$. On considère l'ensemble A défini par

$$A = \{x \in E \,/\, \forall\, y \in \mathscr{P}(E),\, x \in y \Rightarrow x \neq \varphi(y)\}$$

Montrer que $\varphi(A) \notin A$

En déduire que φ n'est pas injective.

Que peut-on en déduire pour cardinal de $\mathscr{P}(E)$?

Soit $a = \varphi(A)$; si $a \in A$ on tire de la définition de A que $a \neq \varphi(A)$ donc $a = \varphi(A)$ implique $a \notin A$.

$\exists\, y \in \mathscr{P}(E)$ tel que $a \in y \Rightarrow a = \varphi(y)$

$a \in y$ et $a \notin A \Rightarrow y \neq A$ et φ ne peut donc être injective puisque $a = \varphi(A) = \varphi(y)$

On en déduit que card $\mathscr{P}(E)$ est nécessairement supérieur strictement à card E : il ne peut donc exister d'application de $\mathscr{P}(E)$ dans E qui soit injective.

CHAPITRE 2

APPLICATIONS. RELATIONS D'ORDRE ET D'ÉQUIVALENCE. ANALYSE COMBINATOIRE

1. LOIS DE COMPOSITION D'UN ENSEMBLE

Loi de composition interne

Un ensemble E est muni d'une loi de composition interne si, à certains couples (a, b) d'éléments appartenant à E, correspond un élément bien déterminé c de E.

Exemple.

Si E est l'ensemble des nombres réels, l'addition est une loi de composition interne partout définie : $\forall a$, $\forall b$, c est défini. La division est également une loi interne définie lorsque le diviseur n'est pas nul.

Loi de composition externe

Considérons deux ensembles E et Λ.

Supposons qu'à tout élément $a \in E$ et qu'à tout élément $\lambda \in \Lambda$ on associe un élément $b \in E$. On dit que l'on a défini une loi de composition externe.

A titre d'exemple supposons que E soit l'ensemble des vecteurs libres de l'espace à trois dimensions et Λ l'ensemble des nombres réels; à tout vecteur $V \in E$ on fait correspondre le vecteur λV, $(\lambda \in \Lambda)$ appelé produit du vecteur V par le nombre λ.

λV est un vecteur $\in E$: la multiplication par un scalaire est une loi de composition externe définie dans E.

2. NOTION D'APPLICATION

On appelle **application** d'un ensemble E dans un ensemble F une loi de correspondance f permettant d'associer à tout $x \in E$ un élément $y \in F$.

E est l'**ensemble de départ,** *F* l'**ensemble image.** L'élément y associé à x est l'**image de** x **par** f, et on note

$$x \xrightarrow{f} y$$

Exemples d'applications.

— La projection des points d'une droite D sur un plan P est une application de D dans P.

— $\mathbb{N}$ étant l'ensemble des nombres entiers, l'application

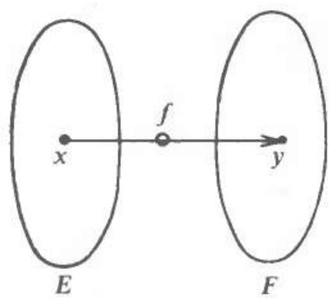


Figure 1.

$$x \xrightarrow{f} y \qquad 2x + 1 \qquad \forall x \in \mathbb{N}$$

est une application de $\mathbb{N}$ dans $\mathbb{N}$.

Surjection

L'image $f(E)$ de E par f est en général une partie de F. Si tout élément de F est l'image par f d'au moins un élément de E, on dit que f est une application de E **sur** F ou encore que f est une application **surjective,** on a alors

$$f(E) = F$$

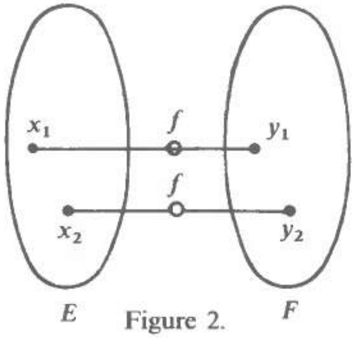


Figure 2.

Injection

Lorsqu'à deux éléments distincts de E correspondent par f deux éléments distincts de F, f réalise une **injection** de E dans F. On dit encore que f est une **application injective.**

On a alors

$$\forall\, x_1, x_2 \in E \quad f(x_1) \neq f(x_2) \Rightarrow x_1 \neq x_2$$

ou ce qui revient au même

$$\forall\, x_1, x_2 \in E \quad f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

Bijection

f est une application **bijective** si elle surjective et injective. f réalise dans ce cas une correspondance **biunivoque** entre E et F : tout élément de F est l'image d'un élément de E et d'un seul.

Dans ce dernier cas, il existe une application inverse f^{-1} de F sur E, puisqu'à tout $y \in F$ on sait associer un élément x bien déterminé de E.

f^{-1} est l'application inverse ou réciproque de f

$$x \xrightarrow{f} y \quad \Leftrightarrow \quad y \xrightarrow{f^{-1}} x$$

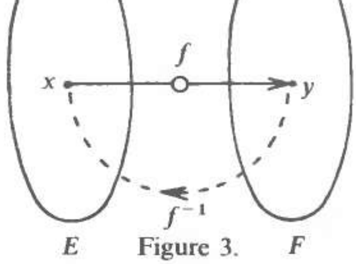


Figure 3.

f^{-1} est elle-même une bijection de F sur E.

Ainsi l'application $f(x) = a + x$ où $x \in \mathbb{Z}$ est une bijection, on peut vérifier aisément qu'elle est bien surjective et injective. Son inverse est

$$f^{-1}(y) = y - a$$

Composition d'applications

Soit trois ensembles E, F, G et deux applications f et g de E dans F et de F dans G. A $x \in E$ correspond dans F un seul y tel que $y = f(x)$. A $y \in F$ correspond dans G un seul z tel que $z = g(y)$.

On définit l'application composée de E dans G par le symbole $g \circ f$ qui à $x \in E$ fait correspondre $z \in G$.

Ainsi si $f(x) = x^2$ et $g(x) = x - 5$ et $E = F = G = \mathbb{Z}$

$$(g \circ f)(x) = x^2 - 5$$

La composée de deux surjections est une surjection. En effet

$$f(E) = F \qquad \text{et} \qquad g(F) = G$$

entraînent que $(g \circ f)(E) = G$

La composée de deux injections est une injection

$$g[f(x_1)] = g[f(x_2)] \Rightarrow f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

et $g \circ f$ est injective.

Il s'ensuit que la composée de deux bijections est une bijection, et en particulier la composition de l'application $f\colon E \to F$ et de sa réciproque $f^{-1}\colon F \to E$ est l'application identique I sur E

$$f^{-1} \circ f = I$$

Involution

Toute bijection f d'un ensemble E sur lui-même, égale à son inverse, est dite involutive

$$\forall\, x \in E \quad f(x) = f^{-1}(x) = x$$

Si f est une bijection involutive de E sur lui-même

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

On dit également que le carré d'une bijection involutive est la transformation identique.

D'une façon générale, toute bijection de E sur lui-même s'appelle permutation de E.

Propriétés des applications

Si $A, B \in \mathscr{P}(E)$ alors $A \subset B \Rightarrow f(A) \subset f(B)$

En effet, si $y \in f(A)$ alors il existe x tel que $y = f(x)$; mais $x \in A \Rightarrow x \in B$ donc $y = f(x) \in f(B)$

On peut démontrer de même que si $A, B \in \mathscr{P}(E)$

alors $$f(A \cup B) = f(A) \cup f(B)$$

et $$f(A \cap B) \subset f(A) \cap f(B)$$

l'égalité n'ayant lieu que si f est injective.

3. NOTION DE RELATION BINAIRE

Soit x un élément d'un ensemble E et y un élément d'un ensemble F. Une relation R entre x et son image y est un lien verbal caractérisant la correspondance entre x et y. Lorsque le couple (x, y) vérifie la relation R on note

$$\boxed{xRy}$$

x est en relation avec y.

Si x et y appartiennent au même ensemble E, la relation R est appelée relation binaire dans E.

Exemples.

Supposons que E et F soient l'ensemble des entiers naturels, x **divise** y est une relation binaire. Elle est vraie pour le couple (2, 8) mais ne l'est pas pour le couple (3, 8).

L'égalité, l'appartenance sont d'autres relations binaires.

Propriétés des relations binaires dans un ensemble

Soit R une relation binaire dans un ensemble E et x, y, z des éléments de E.

- *Réflexivité*

R est réflexive si

$$xRx \qquad \forall x \in E$$

- *Transitivité*

R est transitive si

$$\left.\begin{array}{l} xRy \\ yRz \end{array}\right\} \Rightarrow xRz$$

- *Symétrie*

R est symétrique si

$$xRy \Rightarrow yRx$$

Pour comprendre ces trois notions, il suffit de prendre quelques exemples simples.

L'**égalité** est une relation binaire dans l'ensemble des entiers naturels : on note $x = y$ pour indiquer que le couple (x, y) vérifie la relation binaire d'égalité.

Il est évident que l'égalité est une relation binaire réflexive, transitive et symétrique.

En effet $\quad x = x$ (réflexivité)

$$\left.\begin{array}{l} x = y \\ y = z \end{array}\right\} \quad x = z \text{ (transitivité)}$$

$$x = y \Rightarrow y = x \text{ (symétrie)}$$

— Si R est la **divisibilité** dans l'ensemble des entiers naturels $\mathbb{N}$, R est réflexive puisque tout entier est divisible par lui-même. R est également transitive : si x divise y et si y divise z, x divise z. Par contre, R n'est pas symétrique : si x divise y, y ne divise pas x.

— Enfin, sur l'ensemble des droites de l'espace « D est **orthogonale** à D' » est une relation binaire symétrique, elle n'est cependant ni réflexive, ni transitive.

4. RELATION D'ÉQUIVALENCE

Définition

Une relation définie dans un ensemble E est une relation d'équivalence si elle est :

- réflexive,
- symétrique,
- transitive.

x et y étant deux éléments se correspondant par une relation d'équivalence R, on note

$$\boxed{\begin{array}{c} x \sim y\ (R) \\ x \text{ est équivalent à } y \text{ modulo } R \end{array}}$$

Classe d'équivalence

L'ensemble des éléments y équivalents à un élément donné x constituent une classe d'équivalence de x qu'on note $C(x)$.

$$\underline{\text{Si} \quad x \sim y\ (R) \quad \text{alors} \quad C(x) = C(y).}$$

Si $a \in C(x)$ alors $x \sim a(R)$ donc $a \sim y\,(R)$ et $a \in C(y)$ et $C(x) \subset C\,(y)$ et par symétrie on obtiendrait de même $C(y) \subset C(x)$ d'où

$$C(x) = C(y)$$

De même

$$\underline{\text{si } C(x) = C(y) \text{ alors } x \sim y\,(R)}$$

Si

$$a \in C(x) = C(y)$$

on a $a \sim x\,(R)$ donc $x \sim a\,(R)$ et $a \sim y\,(R)$ d'où

$$x \sim y\,(R)$$

L'ensemble des classes d'équivalence constitue une **partition** de E.
En effet si $x \in E$ $C(x) \neq \varnothing$ puisque R est réflexive et $x \in C(x)$.
Il est par ailleurs évident que $\bigcup C(x) = E$.

Enfin $C(x) \neq C(y) \Rightarrow C(x) \cap C(y) = \varnothing$ en effet, si

$$a \in C(x) \cap C(y) \quad \text{alors} \quad a \sim x\,(R) \quad \text{et} \quad a \sim y\,(R)$$

ce qui entraîne $x \sim y\,(R)$ et donc $C(x) = C(y)$ ce qui est contraire à l'hypothèse.

On peut aussi montrer que réciproquement toute partition de E définit sur E une relation d'équivalence R.

Ensemble quotient

L'ensemble des classes d'équivalence modulo R se nomme ensemble quotient de E par R et se note E/R.

L'application qui à tout $x \in E$ fait correspondre sa classe d'équivalence $C(x)$ est appelée **application canonique**.

5. RELATION D'ORDRE

Relation d'ordre (au sens large)

On appelle relation d'ordre au sens large sur un ensemble E, une relation binaire R qui possède les propriétés suivantes :

- réflexivité : xRx,
- transitivité : xRy et $yRz \Rightarrow xRz$,
- antisymétrie : xRy est incompatible avec yRx, si $x \neq y$

ou, en d'autres termes $\left.\begin{matrix} xRy \\ yRx \end{matrix}\right\} \Rightarrow x = y$

La relation d'**inégalité** ($x \leqslant y$) est l'exemple le plus simple de relation d'ordre au sens large.

$\mathbb{N}$ étant l'ensemble des entiers naturels, la relation « x divise y », x et $y \in \mathbb{N}$, est également une relation d'ordre.

Relation d'ordre (au sens strict)

- elle n'est pas réflexive,
- elle est transitive,
- elle est antisymétrique.

La relation d'inégalité stricte $x < y$ en est l'exemple le plus immédiat.

6. DÉNOMBREMENT DES DIFFÉRENTES APPLICATIONS D'UN ENSEMBLE *E* DANS UN ENSEMBLE *F*

Supposons que les ensembles E et F sont finis et contiennent respectivement p et n éléments

$$E = \{x_1, x_2, \ldots, x_p\}$$
$$F = \{y_1, y_2, \ldots, y_n\}$$

Nombre d'injections de *E* dans *F*. Arrangements

On a vu précédemment que

$$x_1, x_2 \in E \quad \text{et} \quad x_1 \neq x_2 \Rightarrow f_i(x_1) \neq f_i(x_2)$$

pour toute injection f_i si n est au moins égal à p.

Les images $y_1, y_2, \ldots, y_p$ de $x_1, x_2, \ldots, x_p$ par f_i sont un arrangement sans répétition des n éléments de F pris p à p. Le nombre de tels arrangements est désigné par A_n^p. Deux arrangements diffèrent soit par la nature de leurs éléments, soit par l'ordre dans lequel sont rangés ces éléments.

Si $p = 1$, $A_n^1 = n$. En effet E se réduit à $\{x_1\}$. Le nombre d'injections de E dans F est le nombre de correspondances qui, à $x_1 \in E$, conduisent à un des éléments de F, en nombre n.

Si $p > 1$, A_n^{p-1} mesure le nombre d'injections de l'ensemble

$$E' = \{x_1, x_2, \ldots, x_{p-1}\}$$

dans F. Chacune des injections précédentes peut être prolongée par une injection de E dans F faisant correspondre à x_p l'un des $n - (p - 1)$ images restantes. Le nombre de telles injections de x_p a été calculé précédemment et vaut $A_1^{n-p+1} = n - p + 1$.

D'où la relation

$$\begin{aligned} A_n^p &= (n - p + 1)\, A_n^{p-1} \\ &= (n - p + 1)(n - p + 2)\, A_n^{p-2} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &= (n - p + 1)(n - p + 2)\ldots(n - 1)n \end{aligned}$$

Dans le cas particulier où $p = n$, l'application est de plus surjective et donc bijective. A_n^n mesure le nombre de correspondances biunivoques entre E et F :

$$A_n^n = 1 \,.\, 2 \,.\, \ldots \,.\, (n - 1) \,.\, n = n!$$

Si $E = F$ $n!$ mesure le nombre de permutations de E dans lui-même. En convenant que $0! = 1$, l'expression de A_n^p peut être donnée par

$$\boxed{A_n^p = \frac{n!}{(n - p)!}}$$

Nombre d'injections de E dans un sous-ensemble à p éléments de F. Combinaisons

On appelle combinaison à p éléments de F toute partie P de F contenant p éléments. De cette définition découle que deux combinaisons sont distinctes si elles ne sont pas constituées des mêmes éléments. Le tiercé dans le désordre fournit un exemple de combinaison.

Relation entre les arrangements de n objets pris p à p et les combinaisons de n objets pris p à p

Soit E un ensemble fini à n éléments.

Soit A le sous-ensemble de $\mathbb{N}$ à p éléments : $A = \{1, 2, \ldots, p\}$.

Soit $\mathscr{A}$ un arrangement de A dans E, qui à A fait correspondre un sous-ensemble ordonné à p éléments de E. Lorsqu'on change l'ordre de ces p éléments, on obtient les différents arrangements du même sous-ensemble de E, c'est-à-dire de la même combinaison de p éléments de E : pour une combinaison donnée à p éléments, le nombre d'arrangements possibles de ces p éléments est le nombre de manières d'en changer l'ordre, c'est-à-dire le nombre de permutations, $p!$

Appelant A_n^p le nombre total d'arrangements de n objets pris p à p et C_n^p le nombre de combinaisons de n objets pris p à p, on en déduit la relation importante :

$$\boxed{A_n^p = p!\, C_n^p}$$

Nombre de combinaisons de n éléments pris p à p

De la formule ci-dessus on déduit

$$C_n^p = \frac{A_n^p}{p!}$$

Le nombre A_n^p est, d'après le paragraphe précédent

$$A_n^p = \frac{n!}{(n-p)!}$$

On en déduit donc

$$\boxed{C_n^p = \frac{n!}{p!\,(n-p)!}}$$

Remarques

— C_n^p se note aussi $\binom{n}{p}$

— La formule $C_n^p = \dfrac{n!}{p!\,(n-p)!}$ montre que $C_n^p = C_n^{n-p}$.

7. APPLICATIONS : BINÔME DE NEWTON

Formule du binôme

Soit a et b deux éléments de $\mathbb{R}$ et n un entier positif. Nous désirons développer en une somme de facteurs le produit $(a+b)^n$

$$(a+b)^n = \underbrace{\underset{1}{(a+b)}\underset{2}{(a+b)}\ldots\underset{n}{(a+b)}}_{n \text{ facteurs}}$$

$$= a\,\underbrace{(a+b)\ldots(a+b)}_{(n-1)\text{ facteurs}} + b\,\underbrace{(a+b)\ldots(a+b)}_{(n-1)}$$

On continue la décomposition de proche en proche et l'on regroupe :

- termes en a^n : on prend a dans chaque facteur, d'une seule manière;
- termes en $a^{n-1}b$: on prend a dans $(n-1)$ facteurs et b dans le dernier facteur. Il y a n manières de choisir b, donc n termes $a^{n-1}b$;

• terme général $a^{n-p}b^p$: on prend a dans $n - p$ facteurs et b dans les p autres facteurs. On choisit p facteurs contenant b parmi n : c'est une combinaison, en effet l'ordre n'importe pas puisque le produit de nombres réels est commutatif. Il y a C_n^p combinaisons de p facteurs parmi n, donc C_n^p termes $a^{n-p}b^p$.

On continue le dénombrement des termes jusqu'au $n^{\text{ième}}$ terme qui est b^n muni du coefficient $C_n^n = 1$.

Le développement du binôme $(a + b)^n$ s'écrit alors

$$(a + b)^n = C_n^0a^n + C_n^1a^{n-1}b + \dots + C_n^pa^{n-p}b^p + \dots + C_n^nb^n$$

en remarquant que

$$C_n^0 = C_n^n = 1$$

La formule du binôme est

$$\boxed{\begin{aligned}(a + b)^n &= C_n^0a^n + C_n^1a^{n-1}b + \dots + C_n^pa^{n-p}b^p + \dots + C_n^nb^n \\ &= \sum_{p=0}^{n} C_n^pa^{n-p}b^p\end{aligned}}$$

Remarque.

On retrouve souvent cette formule avec $a = 1$.

$$(1 + b)^n = \sum_{p=0}^{n} C_n^pb^p$$

Propriétés des coefficients C_n^p

1) $C_n^p = C_n^{n-p}$ par application de la formule

$$C_n^p = \frac{n!}{(n - p)!\,p!}$$

2) $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$

C_n^p est le nombre de parties à p éléments d'un ensemble B à n éléments. Soit α un élément de B. Le nombre de parties à p éléments de B est égal à la somme du nombre de parties à p éléments contenant α et du nombre de parties à p éléments ne contenant pas α.

a) Le nombre de parties à p éléments ne contenant pas α est le nombre de parties à p éléments du sous-ensemble de B à $n - 1$ éléments ne contenant pas α, c'est-à-dire C_{n-1}^p.

b) Le nombre de parties à p éléments contenant α est égal au nombre de parties obtenues en supprimant α, c'est-à-dire en formant des combinaisons à $p - 1$ éléments pris dans le sous-ensemble de B à $n - 1$ éléments ne contenant pas α. Ce nombre est alors C_{n-1}^{p-1}.

D'où le résultat

$$\boxed{C_n^p = C_{n-1}^{p-1} + C_{n-1}^p}$$

Remarque.

Le calcul d'après la valeur trouvée pour C_n^p redonne ce résultat directement :

$$C_n^p = \frac{n!}{p!(n-p)!} \quad C_{n-1}^{p-1} = \frac{(n-1)!}{(p-1)!(n-p)!} \quad C_{n-1}^p = \frac{(n-1)!}{p!(n-p-1)!}$$

$$\begin{aligned} C_{n-1}^{p-1} + C_{n-1}^p &= \frac{(n-1)!}{(p-1)!(n-p)!} + \frac{(n-1)!}{p!(n-p-1)!} \\ &= (n-1)!\left[\frac{p}{p!(n-p)!} + \frac{n-p}{p!(n-p)!}\right] \\ &= (n-1)!\frac{n}{p!(n-p)!} = \frac{n!}{p!(n-p)!} = C_n^p \end{aligned}$$

Le triangle de Pascal

Le triangle de Pascal sert à calculer les coefficients du développement du binôme en utilisant la relation précédente

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$$

qui permet de passer de la puissance $n - 1$ à la puissance n :

1						$(a+b)^0$
1	1					$(a+b)^1$
1	2	1				$(a+b)^2$
1	3	3	1			$(a+b)^3$
1	4	6	4	1		$(a+b)^4$
1	5	10	10	5	1	$(a+b)^5$

. . . .

L'application directe de la formule donne la règle suivante : le coefficient de la ligne n et de la colonne p s'obtient en ajoutant le coefficient de la ligne $n - 1$ et de la colonne p au coefficient de la ligne $n - 1$ et de la colonne $p - 1$. Par exemple, le coefficient 6, quatrième ligne troisième colonne, est la somme de 3, troisième ligne troisième colonne, et de 3, troisième ligne deuxième colonne.

Nombre de parties d'un ensemble E fini contenant n éléments

Nous avons déjà calculé le nombre de parties d'un ensemble fini par la méthode de l'arbre (voir chap. 1, exercice 5). Voici une démonstration plus simple, application directe de la formule du binôme.

Les parties de l'ensemble E sont

— les parties contenant 0 élément : il y a $C_n^0 = 1$ manière de prendre 0 élément parmi n. On obtient l'ensemble vide $\varnothing$;

— les parties contenant un élément : il y a C_n^1 manières de prendre un élément parmi n, donc C_n^1 parties à un élément;

. . . .

— les parties contenant p éléments : il y a C_n^p manières de prendre p éléments parmi n, donc C_n^p parties à p éléments;

. . . .

— les parties contenant n éléments : il y a $C_n^n = 1$ manière de prendre n éléments parmi n. On obtient l'ensemble E.

Le nombre total de parties est donc

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^p + \dots + C_n^n$$

On remarque que cette somme est le développement du binôme

$$(a + b)^n \quad \text{avec} \quad a = b = 1.$$

La valeur de la somme est donc $(1 + 1)^n = 2^n$.

La distribution multinomiale

Une interprétation possible de la formule des combinaisons est la suivante : on dispose de n objets, que l'on désire ranger dans deux cases en mettant n_1 objets dans la première et $n_2 = n - n_1$ objets dans la seconde.

Combien de façons y a-t-il de les ranger ainsi ?

La réponse est évidemment

$$C_n^{n_1} = \frac{n!}{n_1!\,(n - n_1)!} = \frac{n!}{n_1!\,n_2!}$$

Cette distribution s'appelle distribution binomiale.

On peut généraliser à une distribution multinomiale, c'est-à-dire chercher le nombre N de façons de distribuer n objets dans k cases en mettant n_1 objets dans la première case, n_2 dans la seconde et ainsi de suite. On obtient

$$\boxed{N = \frac{n!}{n_1!\,n_2!\dots n_k!}}$$

On trouvera une démonstration de cette formule dans l'exercice n° 10.

Résumé du chapitre 2

- **Lois de composition.**

Loi de composition interne sur E.

$$a, b \in E \quad (a, b) \xrightarrow{L} c \in E$$

ou encore au couple (a, b) d'éléments de E la loi L fait correspondre un élément c de E.

Loi de composition externe.

$$a \in E \quad \text{et} \quad \lambda \in \Lambda \quad (a, \lambda) \xrightarrow{L} b \in E$$

- **Applications de E dans F.**

Loi qui fait correspondre à tout x de E, un y et un seul de F; x est la variable, y est l'image de x par l'application.

Application injective.

A tout élément y de F, on peut faire correspondre *au plus* un élément x de E

$$x \neq x' \Rightarrow f(x) \neq f(x')$$

Application surjective.

A tout élément y de F on peut faire correspondre *au moins* un élément x de E

$$f(E) = F$$

Application bijective.

Elle est surjective et injective.

- **Relation binaire sur E**

Propriétés possibles d'une relation binaire:

— réflexivité $\quad xRx \quad \forall x \in E$
— symétrie $\quad xRy \Rightarrow yRx$
— transitivité $\quad xRy$ et $yRz \Rightarrow xRz$.

- **Relation d'équivalence sur *E*.**

Elle est réflexive, symétrique et transitive.

Classe d'équivalence.

L'ensemble des éléments équivalents à x constituent la classe d'équivalence $C(x)$ de x

Si $x \sim y\ (R)$ $C(x) = C(y)$

Les classes d'équivalence sont disjointes et constituent une partition de E.

Ensemble quotient E/R

C'est l'ensemble des classes d'équivalence pour la relation d'équivalence R.

- **Relation d'ordre.**

C'est une relation binaire

réflexive,

$$\forall x \quad xRx$$

antisymétrique

$$xRy \ \text{ et } \ yRx \Rightarrow x = y$$

transitive

$$xRy \ \text{ et } \ yRz \Rightarrow xRz$$

- **Analyse combinatoire.**

Arrangements.

Le nombre $A_n^p = \dfrac{n!}{(n-p)!}$ appelé arrangement de n objets pris p à p mesure le nombre d'applications injectives d'un ensemble à p éléments dans un ensemble à n éléments.

Permutations

Si $n = p$, $A_n^n = n!$ est le nombre de permutations de n objets.

Combinaisons.

$$C_n^p = \frac{A_n^p}{p!} = \frac{n!}{(n-p)!\,p!}$$

est le nombre de combinaisons de n objets pris p à p.

$$C_n^p = C_n^{n-p}$$

$$C_n^p = C_{n-1}^{p-1} + C_{n-1}^p$$

- **Développement du binôme de Newton.**

$$(a + b)^n = C_n^0 a^n + C_n^1 a^{n-1} b + \ldots + C_n^p a^{n-p} b^p + \ldots + C_n^n b^n$$
$$= \sum_{p=0}^{n} C_n^p a^{n-p} b^p$$

EXERCICES

■ **1.** Soit X et Y deux ensembles possédant respectivement n et p éléments. Déterminer le nombre $a(n, p)$ d'applications de X dans Y.

Une application de X dans Y fait correspondre à un élément particulier x_0 de X un élément quelconque de Y, soit un des p éléments de Y.

Il existe donc p possibilités de correspondance pour x_0. A chacune d'entre elles on peut faire correspondre les différentes applications de $X - \{x_0\}$ dans Y d'où

$$a(n, p) = p \cdot a(n - 1, p)$$

et ainsi de suite

$$a(n, p) = p^{n-1} a(1, p)$$

Or $$a(1, p) = p$$

D'où

$$a(n, p) = p^n$$

■ **2.** On donne les ensembles $E = \{a, b, c, d\}$ et $F = \{1, 2, 3, 4, 5, 6\}$.

a) Soit une relation R_1 définie sur $E \times F$, donnée par son graphe

$$G_1 = \{(a, 1), (a, 2), (b, 1), (c, 2), (d, 4)\}$$

R_1 est-elle une application?

b) Soit la relation R_2 donnée par son graphe G_2

$$G_2 = \{(a, 1), (b, 2), (c, 4), (d, 6)\}$$

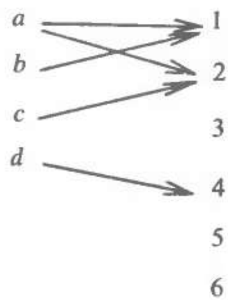
R_2 est-elle une application? Si oui, la caractériser.

c) Peut-on définir une application surjective de E dans F?

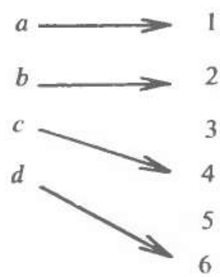
---

a) R_1 n'est pas une application de E dans F : l'ensemble de départ est (a, b, c, d), l'ensemble d'arrivée est $(1, 2, 4)$; à l'élément a, R_1 fait correspondre les éléments 1 et 2, ce qui est contraire à la définition de l'application.

G_1 se représente par le réseau de flèches :



b) R_2 est une application injective, représentée par le réseau de flèches :



c) On ne peut pas définir une application f surjective de E dans F car la relation $f(E) = F$ ne peut être vérifiée : en effet card $F >$ card E.

■ 3. Soit l'ensemble $E = \{1, 2, 3, 4\}$ soit le graphe G de la relation R défini par

$$G = \{(1, 3), (3, 4), (1, 4), (2, 1), (1, 2), (2, 2)\}$$

R est-elle réflexive, symétrique ou transitive ?

R n'est pas réflexive car les couples $(1, 1)$, $(3, 3)$, $(4, 4)$ n'appartiennent pas à G.

R n'est pas symétrique : par exemple $3 R 4$ et non $4 R 3$.

R est transitive $1 R 3$ et $3 R 4 \Rightarrow 1 R 4$

$2 R 1$ et $1 R 2 \Rightarrow 2 R 2$

■ **4.** On examinera si les relations binaires suivantes sont réflexives, symétriques ou transitives :

a) La relation d'orthogonalité pour les droites d'un plan.

b) La relation d'orthogonalité pour les cercles d'un plan.

c) La relation $a^2 + a = b^2 + b$ *pour les entiers relatifs*.

---

a) Soit D et D' deux droites d'un plan.

$D \perp D'$ signifie que D est orthogonale à D'.

$D \perp D$ n'est pas vérifié : une droite n'est pas perpendiculaire à elle-même. $\perp$ n'est pas réflexive.

$D \perp D' \Rightarrow D' \perp D$: D orthogonale à D' entraîne D' orthogonale à D. $\perp$ est symétrique.

$D \perp D'$ et $D' \perp D''$ entraîne que D et D'' sont parallèles et non perpendiculaires. $\perp$ n'est pas transitive.

b) Soit deux cercles C et C', de rayons différents de zéro.

$C \perp C'$ signifie que C et C' se coupent et qu'aux points d'intersection les tangentes sont perpendiculaires.

$C \perp C$ n'est pas vérifié. $\perp$ n'est pas réflexive.

$C \perp C' \Rightarrow C' \perp C$: $\perp$ est symétrique.

$C \perp C'$ et $C' \perp C''$ n'entraîne pas $C \perp C''$ (faire une figure pour s'en rendre compte). $\perp$ n'est pas transitive.

c) La relation $a\rho b : a^2 + a = b^2 + b$ est

— réflexive $a\rho a : a^2 + a = a^2 + a$,

— symétrique $a\rho b \Rightarrow b\rho a$:

$$a^2 + a = b^2 + b \Rightarrow b^2 + b = a^2 + a$$

— transitive

$$\left.\begin{array}{l} a^2 + a = b^2 + b \\ \hline b^2 + b = c^2 + c \end{array}\right\} \Rightarrow a^2 + a = c^2 + c$$

■ **5.** On définit sur $\mathbb{R}^2$ la relation ρ

$$(x, y)\, \rho\, (x', y') \Leftrightarrow x + y = x' + y'$$

1) Montrer que ρ est une relation d'équivalence.

2) Trouver la classe d'équivalence du couple (0, 0).

3) Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f : \quad (x, y) \to x + y$$

Montrer que deux éléments de $\mathbb{R}^2$ équivalents modulo ρ ont même image par f et que deux éléments non équivalents ont des images distinctes.

4) En déduire qu'entre l'ensemble quotient $\mathbb{R}^2/\rho$ et $\mathbb{R}$ il existe une bijection g que l'on précisera.

---

1) ρ est réflexive, symétrique et transitive. C'est bien une relation d'équivalence.

2) La classe d'équivalence du couple (0, 0) est constituée par l'ensemble des couples $(x, y) \in \mathbb{R}^2$ vérifiant $x + y = 0$: c'est l'ensemble des points situés sur la deuxième bissectrice du plan xOy.

3) Si $u, v \in \mathbb{R}^2$ et $u \rho v \Leftrightarrow f(u) = f(v)$.

4) g est l'application qui à une classe fait correspondre la somme des composants d'un représentant quelconque de cette classe.

g est injective d'après 3)

Par ailleurs $\forall \alpha \in \mathbb{R}$ on peut considérer que α est l'image de

$$\left(\frac{\alpha}{2}, \frac{\alpha}{2}\right) \in \mathbb{R}^2$$

donc g est surjective, et g est enfin bijective.

■ **6.** Calculer le minimum du produit $p!\,q!$ sachant que $p + q = n$, n donné.

Rechercher le minimum de $p!\,q!$ équivaut à chercher le minimum de $p!\,(n-p)!$ ou encore le maximum de

$$C_n^p = \frac{n!}{p!\,(n-p)!}$$

En utilisant le triangle de Pascal, on remarque que le maximum de C_n^p est obtenu pour $C_n^{n/2}$ si n est pair, ou $C_n^{(n+1)/2}$ si n est impair.

On en déduit, suivant le cas :

$$p = \frac{n}{2} \qquad \text{et} \qquad q = \frac{n}{2}$$

$$p = \frac{n+1}{2} \qquad \text{et} \qquad q = \frac{n-1}{2}$$

■ **7.** Démontrer la formule

$$1 \times 1! + 2 \times 2! + \ldots + n \times n! = (n+1)! - 1$$

Nous allons démontrer cette formule par récurrence :

$$1 \times 1! = (1+1)! - 1 = 1$$
$$1 \times 1! + 2 \times 2! = (2+1)! - 1 = 5$$
$$\ldots\ldots\ldots$$

Supposons que cette formule soit vraie à l'ordre $n - 1$, montrons qu'elle est vraie à l'ordre n:

$$1 \times 1! + 2 \times 2! + \ldots + (n-1) \times (n-1)! = n! - 1$$

par hypothèse.

$$1 \times 1! + 2 \times 2! + \ldots + (n-1)(n-1)! + n \times n!$$
$$= n! - 1 + n \times n! = n!(n+1) - 1 = (n+1)! - 1$$

d'où le résultat.

■ **8.** Démontrer la formule

$$1 - C_n^1 + C_n^2 - C_n^3 + \ldots + (-1)^p C_n^p = (-1)^p C_{n-1}^p$$

On remplace chaque terme de la somme par la quantité

$$C_n^p = C_{n-1}^p + C_{n-1}^{p-1},$$

formule démontrée dans le cours:

$$1 - C_{n-1}^0 - C_{n-1}^1 + C_{n-1}^1 + C_{n-1}^2 - \ldots + (-1)^{p-1} C_{n-1}^{p-2}$$
$$+ (-1)^{p-1} C_{n-1}^{p-1} + (-1)^p C_{n-1}^{p-1} + (-1)^p C_{n-1}^p$$

Les termes s'éliminent deux à deux à l'exception de

$$(-1)^p C_{n-1}^p, \quad \text{puisque} \quad C_{n-1}^0 = 1$$

D'où le résultat.

■ **9.** Calculer les sommes

$$s_1 = \sum_{i=0}^{n} C_n^i$$
$$s_2 = \sum_{i=0}^{n} (-1)^i C_n^i$$
$$s_3 = \sum_{i=0}^{n} C_n^{2i} \text{ si } n \text{ pair}$$
$$= \sum_{i=0}^{n-1} C_n^{2i} \text{ si } n \text{ impair}$$

On trouve

$$s_1 = 2^n$$
$$s_2 = 0$$
$$s_3 = 2^{n-1}$$

■ **10.** *a*) On dispose de k cases et de n objets. On désire disposer les objets dans les cases de la manière suivante :

$$\left.\begin{array}{l} n_1 \text{ objets dans la première case} \\ n_2 \text{ objets dans la deuxième case} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ n_k \text{ objets dans la } k^{\text{ième}} \text{ case} \end{array}\right\} \quad \text{avec } \sum_{i=1}^{k} n_i = n$$

Quel est le nombre de dispositions possibles ?

b) On désire disposer 8 objets dans 3 cases, le nombre d'objets par case étant indiqué par les schémas :

i)	1	3	4
ii)	3	2	3
iii)	0	6	2

Quel est le nombre de telles dispositions dans chaque cas ?

---

a) On dispose n_1 objets dans la 1re case : il y a $C_n^{n_1}$ dispositions possibles. Ayant choisi une disposition, on prend n_2 objets parmi ceux qui restent et on les place dans la case 2. Il y a $C_{n-n_1}^{n_2}$ dispositions possibles dans la case 2 pour une disposition dans la case 1, soit pour les $C_n^{n_1}$ dispositions de 1, $C_n^{n_1} C_{n-n_1}^{n_2}$ dispositions pour 1 et 2. On continue ainsi jusqu'à la dernière case. Le nombre de dispositions est donc au total :

$$C_n^{n_1}\, C_{n-n_1}^{n_2}\, C_{n-n_1-n_2}^{n_3} \cdots C_{n-n_1-n_2-\ldots-n_{k-1}}^{n_k}$$

$$= \frac{n!}{n_1!(n-n_1)!} \times \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \times \ldots \times \frac{(n-n_1-n_2\ldots-n_{k-1})!}{n_k!\left(n-\sum\limits_{i=1}^{k} n_i\right)!}$$

$$= \frac{n!}{n_1!\, n_2! \ldots n_k!}$$

b) L'application de la formule ci-dessus conduit à

i) $\dfrac{8!}{1!\,3!\,4!} = 280$

ii) $\dfrac{8!}{3!\,2!\,3!} = 560$

iii) $\dfrac{8!}{0!\,6!\,2!} = 28$

CHAPITRE 3

STRUCTURES ALGÉBRIQUES

Les notions qui suivent présentent de l'intérêt tant sur le plan terminologique que structurel avant d'aborder l'étude des espaces vectoriels.

1. GROUPE

Une loi T confère à un ensemble E la structure de groupe si elle possède simultanément les quatre propriétés suivantes : elle est interne; elle est associative; elle définit un élément neutre de E; chaque élément de E possède un symétrique par rapport à T.

T est interne

Soit deux éléments x et y de E, composés par la loi T : $xTy = z$. Nous avons vu que T est interne, si z est aussi un élément de E, quels que soient x et y.

$\mathbb{N}$ étant l'ensemble des entiers positifs, la multiplication, notée $\times$, est une loi interne sur $\mathbb{N}$ et nous avons vu qu'il en était de même de l'addition.

T est associative

x, y, z étant des éléments de E, composés par la loi T, on dit que T est associative si cette loi de composition vérifie la relation suivante

$$(xTy)\ Tz = xT(yTz)$$

T définit un élément neutre

On dit que e est élément neutre de E pour la loi T, si, quel que soit x élément de E, $xTe = eTx = x$.

Exemple.

1 est élément neutre de $\mathbb{N}$ pour la loi $\times$ (multiplication),
0 est élément neutre de $\mathbb{N}$ pour la loi $+$ (addition).

x est symétrisable

Cela signifie que tout élément x de E possède un symétrique x' pour la loi T, défini par $xTx' = x'Tx = e$, élément neutre pour T, de E.

Exemple

Soit $\mathbb{Z}$ l'ensemble des entiers relatifs. Tout élément x est symétrisable pour la loi addition notée +

$$x + (-x) = 0 \qquad -x = x'$$

Les éléments de $\mathbb{N}$ ne sont pas symétrisables pour la loi + ni pour la loi ×

Remarque.

Si la loi T est commutative, on dit que le groupe est **abélien.**

2. ANNEAU

Un ensemble E possède la structure d'anneau s'il est muni de deux lois internes, notées T et L, possédant simultanément les propriétés suivantes :

- T confère à E la structure de groupe abélien.

T est associative, commutative, définit un élément neutre de E et tout élément de E possède un symétrique pour la loi T.

- L est distributive et associative par rapport à T.

On dit que L est **distributive** par rapport à T si, quels que soient x, y, z, éléments de E, L et T vérifient la relation

$$(xTy)\, Lz = (xLz)\, T(yLz)$$

Exemple

Soit **Z** muni des lois addition et multiplication; ces lois, notées + et ×, confèrent à **Z** la structure d'anneau commutatif :

+ et × sont internes : le produit ou la somme de nombres entiers est un nombre entier.

+ et × sont associatives : quels que soient a, b, c appartenant à **Z** les relations suivantes sont vérifiées :

$$(a + b) + c = a + (b + c)$$
$$(a \times b) \times c = a \times (b \times c)$$

+ et × sont commutatives :

$$a \times b = b \times a$$
$$a + b = b + a$$

L'élément neutre de **Z** pour la loi $+$ est 0.
L'élément neutre de **Z** pour la loi $\times$ est 1.
Tout élément a de **Z** admet un symétrique pour la loi $+$: c'est son opposé $-a$.
La loi $+$ confère donc à **Z** la structure de groupe abélien.
La multiplication est distributive par rapport à l'addition.

3. CORPS

On appelle corps un anneau K tel que la deuxième loi confère une structure de groupe à l'ensemble $K - \{e\}$, où e est l'élément neutre de la première loi. Cela signifie que, pour la deuxième loi de l'anneau, il existe un élément neutre et chaque élément est symétrisable.

Exemple

L'ensemble des nombres réels **R** a une structure de corps pour l'addition et la multiplication.
L'addition est interne, associative, commutative, définit l'élément neutre 0. Chaque élément possède un symétrique qui est son opposé. L'addition confère donc à **R** la structure de groupe abélien.
La multiplication est interne, associative et distributive par rapport à l'addition. Cela suffit pour doter **R** d'une structure d'anneau. De plus la multiplication possède un élément neutre 1, et chaque élément, sauf 0, est symétrisable. Le symétrique de a est son inverse $\frac{1}{a}$.

On peut donc munir **R** de la structure de corps.

Remarque.
Si la deuxième loi L est commutative, on dit que le corps est commutatif : ainsi **R** est un corps commutatif.

Résumé du chapitre 3

STRUCTURES ALGÉBRIQUES

- **Groupe.**

E a une structure de groupe relativement à T si cette loi est interne et possède les propriétés :

(A) associativité
(N) existence d'un élément neutre
(S) existence d'un symétrique.

• **Groupe abélien.**

T est, de plus, commutative.

• **Anneau**

A a une structure d'anneau s'il est muni d'une loi T lui conférant la structure de groupe abélien, et d'une loi L associative et distributive par rapport à T.

• **Corps.**

K est un corps si la loi L confère une structure de groupe à l'ensemble $K - \{e\}$ où $e$est l'élément neutre de T.

EXERCICES

■ **1.** N étant l'ensemble des entiers naturels, on considère la loi de composition interne $*$ sur **N** définie par

$$a * b = a^2 + b^2$$

Cette loi est-elle :

1) commutative ?
2) associative ?
3) munie d'un élément neutre ?

---

1) $b * a = b^2 + a^2 = a * b$ la loi est bien commutative.

2) $a, b, c \in \mathbf{N}$

$$(a * b) * c = (a^2 + b^2)^2 + c^2$$
$$a * (b * c) = a^2 + (b^2 + c^2)^2$$

Les deux quantités ci-dessus ne sont généralement pas égales (par exemple, pour $b = c = 0$ et $a = 2$).

3) Si e est un élément neutre

$$a * e = e * a = a$$

c'est-à-dire $$a^2 + e^2 = e^2 + a^2 = a$$

Or l'égalité ci-dessus $a^2 + e^2 = a$ n'est pas possible pour $a = 2$, par exemple, car alors

$$e = \sqrt{2} \notin \mathbf{N}$$

■ **2.** On munit un ensemble de la structure d'anneau non commutatif en prenant la loi addition, comme première loi interne, et comme seconde loi interne la loi $*$ définie par

$$x * y = xy - yx$$

a) Démontrer que la loi $*$ est distributive par rapport à l'addition.

b) Démontrer que $x * y = - y * x$.

c) Démontrer que

$$x * (y * z) + y * (z * x) + z * (x * y) = 0$$
$$x * (y * z) - (x * y) * z = (z * x) * y$$

$$\begin{aligned} a)\ x * (y + z) &= x(y + z) - (y + z)\,x \\ &= xy - yx + xz - zx \\ &= x * y + x * z. \end{aligned}$$

b) $x * y = xy - yx = -(yx - xy) = -(y * x)$

c) $x * (y * z) + y * (z * x) + z * (x * y)$

$$\begin{aligned} &= x\,(yz - zy) - (yz - zy)\,x + y\,(zx - xz) - (zx - xz)\,y \\ &\quad + z\,(xy - yx) - (xy - yx)\,z \\ &= xyz - xzy - yzx + zyx + yzx - yxz - zxy + xzy \\ &\quad + zxy - zyx - xyz + yxz = 0 \end{aligned}$$

$x * (y * z) - (x * y) * z$

$$\begin{aligned} &= x\,(yz - zy) - (yz - zy)\,x - (xy - yx)\,z \\ &\quad + z\,(xy - yx) = xyz - xzy - yzx + zyx - xyz \\ &\quad + yxz + zxy - zyx \\ &= yxz - yzx + zxy - xzy = (zx - xz)\,y - y\,(zx - xz) \\ &= (z * x) * y \end{aligned}$$

■ **3.** Soit les couples de nombres rationnels a et b pris dans cet ordre. Montrer que si on munit l'ensemble de ces couples des lois $+$ et $\times$ on lui confère la structure de corps.

Si α et α' sont deux couples avec $\alpha = (a, b)$ et $\alpha' = (a', b')$,
$+$ est défini par : $\alpha + \alpha' = (a + a', b + b')$
$\times$ est défini par : $\alpha\alpha' = (aa' + 2\,bb', ab' + a'b)$

$+$ et $\times$ sont des lois internes.

En effet $a + a'$, $b + b'$, $aa' + 2\,bb'$, $ab' + a'b$ sont des nombres rationnels puisque les rationnels forment un corps **Q**. Les propriétés de **Q** entraînent les propriétés suivantes :

$$\begin{aligned} +\ \text{est associative} : (\alpha + \alpha') + \alpha'' &= ((a + a') + a'', (b + b') + b'') \\ &= (a + a' + a'', b + b' + b'') \end{aligned}$$

$$= ((a + (a' + a''), b + (b' + b''))$$
$$= \alpha + (\alpha' + \alpha'')$$

+ admet un élément neutre $e = (0, 0)$.

Chaque élément $\alpha = (a, a')$ a un symétrique $\alpha' = (-a, -a')$.

× est associative :

$$\begin{aligned}(\alpha\alpha')(\alpha'') &= (aa' + 2bb', ab' + a'b)(a'', b'') \\ &= ((aa' + 2bb')a'' + 2(ab' + a'b)b'', \\ &\quad (aa' + 2bb')b'' + a''(ab' + a'b)) \\ &= (a(a'a'' + 2b'b'') + 2b(a'b'' + b'a''), \\ &\quad a(a'b'' + b'a'') + b(a'a'' + 2b'b'')) \\ &= \alpha(\alpha'\alpha'').\end{aligned}$$

× est commutative : $\alpha\alpha' = (aa' + 2bb', ab' + a'b)$
$$= (a'a + 2b'b, a'b + ab') = \alpha'\alpha.$$

× possède un élément neutre $\alpha = (1, 0)$.

Chaque élément α possède un symétrique α' :

$$\alpha\alpha' = (aa' + 2bb', ab' + a'b) = (1, 0)$$

$$\begin{cases} aa' + 2bb' = 1 \\ ab' + a'b = 0 \end{cases}$$

on multiplie la première ligne par b

$$aa'b + 2b^2b' = b$$

D'après la deuxième équation $a'b = -ab'$, d'où

$$-a^2b' + 2b^2b' = b \quad \text{et} \quad b' = \frac{b}{2b^2 - a^2}$$

et
$$a' = \left(1 - \frac{2b^2}{2b^2 - a^2}\right)\frac{1}{a} = -\frac{a}{2b^2 - a^2}$$

Le symétrique de (a, b) est

$$\left(-\frac{a}{2b^2 - a^2}, \frac{b}{2b^2 - a^2}\right)$$

a et b sont rationnels, donc $a \neq \sqrt{2}\,b$.

■ **4.** Soit quatre fonctions f_i : $i = 1, 2, 3, 4$ définies par

$$x \to f_1(x) = x$$
$$x \to f_2(x) = 1/x$$
$$x \to f_3(x) = -x$$
$$x \to f_4(x) = -1/x$$

Montrer que ces fonctions forment un groupe pour la loi $*$ définie par

$$f_i * f_i = f_i(f_j)$$

La loi $*$ est interne :

$$\begin{array}{llll}
x \to (f_1 * f_2)(x) = 1/x = f_2(x) & & x \to (f_2 * f_1)(x) = 1/x = f_2(x) \\
x \to (f_1 * f_3)(x) = -x = f_3(x) & & x \to (f_2 * f_3)(x) = -1/x = f_4(x) \\
x \to (f_1 * f_4)(x) = -1/x = f_4(x) & & x \to (f_2 * f_4)(x) = -x = f_3(x) \\
x \to (f_1 * f_1)(x) = x = f_1(x) & & x \to (f_2 * f_2)(x) = x = f_1(x) \\
x \to (f_3 * f_1)(x) = -x = f_3(x) & & x \to (f_4 * f_1)(x) = -1/x = f_4(x) \\
x \to (f_3 * f_2)(x) = -1/x = f_4(x) & & x \to (f_4 * f_2)(x) = -x = f_3(x) \\
x \to (f_3 * f_4)(x) = 1/x = f_2(x) & & x \to (f_4 * f_3)(x) = 1/x = f_2(x) \\
x \to (f_3 * f_3)(x) = x = f_1(x) & & x \to (f_4 * f_4)(x) = x = f_1(x)
\end{array}$$

Le tableau ainsi écrit montre en outre que :

la loi est associative (ce que l'on vérifie aisément);
la loi est commutative;
l'élément neutre est f_1;
chaque fonction est son propre symétrique.

■ **5.** Soit A un anneau tel que $\forall x \in A$ on ait la relation $x^2 = x$

Montrer que pour tout x dans A on a $x + x = 0$ et que A est commutatif.

On a

$$\begin{aligned}(x + x) = (x + x)^2 &= x^2 + x^2 + x^2 + x^2 \\ &= x + x + x + x\end{aligned}$$

d'où on tire $\quad x + x = 0 \quad$ soit $\quad x = -x$.

Soit $\quad x, y \in A$

$$\begin{aligned}(x + y) &= (x + y)^2 \\ &= x^2 + xy + yx + y^2 \\ &= x + xy + yx + y\end{aligned}$$

égalité qui se réduit à

$$xy + yx = 0$$

et puisque $x = -x$

$$xy = yx$$

A est bien commutatif.

■ **6.** On définit sur l'ensemble A des couples d'entiers algébriques les lois $+$ et $\times$ par :

$$(a, b) + (a', b') = (a + a', b + b')$$
$$(a, b) \times (a', b') = (aa' + bb', ab' + ba')$$

l'égalité $(a, b) = (a', b')$ impliquant $a = a'$ et $b = b'$.

1) Montrer que A est un anneau commutatif unitaire.

2) Déterminer les éléments de A qui ont un inverse.

3) Dans quel cas a-t-on

$$(a, b) \times (x, y) = (a, b) \times (u, v) \Rightarrow (x, y) = (u, v)$$

---

1) A est un groupe abélien pour l'addition, son élément neutre étant $(0, 0)$.

La multiplication est distributive pour l'addition du fait de la linéarité, et est commutative.

L'élément unité est le couple $(1, 0)$.

La multiplication est associative. Posons

$$(\alpha, \beta) = [(a, b)(a', b')] \times (a'', b'')$$

$$\alpha = (aa' + bb')\, a'' + (ab' + ba')\, b'' = aa'a'' + bb'a'' + ab'b'' + ba'b''$$
$$\beta = (ab' + ba')\, a'' + (aa' + bb')\, b'' = bb'b'' + ab'a'' + ba'a'' + aa'b''$$

α et β sont des fonctions symétriques de (a, a', a'') et (b, b', b''), d'où

$$[(a, b)(a', b')] \times (a'', b'') = (a, b) \times [(a', b')(a'', b'')]$$

2) Si (a, b) admet un inverse (x, y), on aura

$$(a, b) \times (x, y) = (ax + by, ay + bx) = (1, 0)$$

soit
$$\begin{cases} ax + by = 1 \\ ay + bx = 0 \end{cases} \qquad (1)$$

ou
$$\frac{x}{a} = -\frac{y}{b} = k$$

avec $k(a^2 - b^2) = 1$

Si $a = b$ le système (1) ne peut avoir de solution.

Si $a \neq b$ on obtient alors $x = \dfrac{a}{a^2 - b^2}$, $y = \dfrac{-b}{a^2 - b^2}$ qui n'est acceptable que si $(x, y) \in A$ c'est-à-dire x et y entiers; il en est alors de même pour $x + y = \dfrac{1}{a + b}$ et $x - y = \dfrac{1}{a - b}$.

Alors $a + b = \pm 1$ et $a - b = \pm 1$ qui est bien suffisante puisque alors $a^2 - b^2 = \pm 1$ et $(x, y) \in A$.

Les seuls couples possibles sont donc $(0, 1)$, $(0, -1)$, $(1, 0)$, $(-1, 0)$ chacun d'eux étant son propre inverse.

3) L'égalité $(a, b) \times (x, y) = (a, b) \times (u, v)$ s'écrit encore

$$(a, b) \times (x - u, y - v) = 0$$

Il s'agit donc de voir dans quel cas $(a, b) \times (r, s) = 0$ entraîne $r = s = 0$

soit
$$\begin{cases} ar + bs = 0 \\ br + as = 0 \end{cases} \qquad (2)$$

Si $a^2 - b^2 \neq 0$ le système (2) n'a de solution que pour $r = s = 0$.
Si $a^2 - b^2 = 0$ le système (2) admet une infinité de solution.
Donc la simplification n'est possible que pour $a = \pm b$.

CHAPITRE 4

NOMBRES COMPLEXES

Afin de généraliser la notion de racines d'une équation algébrique, on a introduit les nombres complexes ou imaginaires.

1. DÉFINITION

Si a et b sont deux nombres réels quelconques, le couple ordonné $[a, b]$ est appelé *nombre complexe* ou *imaginaire*.

Représentation

Dans le plan Ox, Oy rapporté à un repère orthonormé, le point M d'abscisse a et d'ordonnée b est l'*image* du nombre complexe $z = [a, b]$. On dit encore que z a pour *affixe* M.

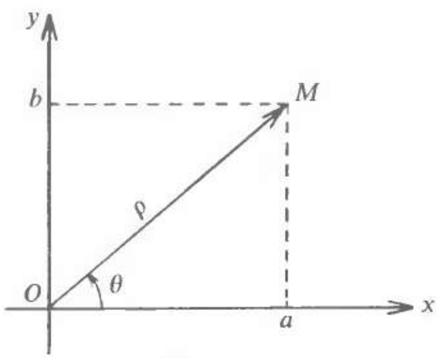


Figure 1.

Module. Argument

L'angle $\theta = (Ox, OM)$ est l'*argument* de z. Il est défini à $2\,K\pi$ près.
Le *module* de z est la longueur du vecteur $\overrightarrow{OM}$. On le désigne par

$$\rho = |z| = |OM|$$

a et b sont liés à ρ et θ par

$$a = \rho \cos \theta,$$
$$b = \rho \sin \theta$$

Réciproquement, ρ et θ peuvent être déterminés en fonction de a et de b:

$$a^2 + b^2 = \rho^2 (\cos^2 \theta + \sin^2 \theta) \quad \text{et} \quad \frac{b}{a} = \frac{\rho \sin \theta}{\rho \cos \rho}$$

soit

$$\rho = \sqrt{a^2 + b^2} \qquad \operatorname{tg} \theta = \frac{b}{a}$$

et le nombre complexe $[a, b]$ se note également (ρ, θ).

2. OPÉRATIONS SUR LES NOMBRES COMPLEXES

Égalité de deux nombres complexes

Les nombres complexes $z_1 = [a_1, b_1]$ et $z_2 = [a_2, b_2]$ sont égaux si et seulement si $a_1 = a_2$ et $b_1 = b_2$, c'est-à-dire si leurs images M_1 et M_2 sont confondues.

Les deux nombres complexes z_1 et z_2 sont *conjugués* si $a_1 = a_2$ et $b_1 = - b_2$. Leurs images sont symétriques par rapport à l'axe Ox.

Le conjugué de z se note $\bar{z}$.

On en conclut que si $z = (\rho, \theta)$ $\bar{z} = (\rho, - \theta)$.

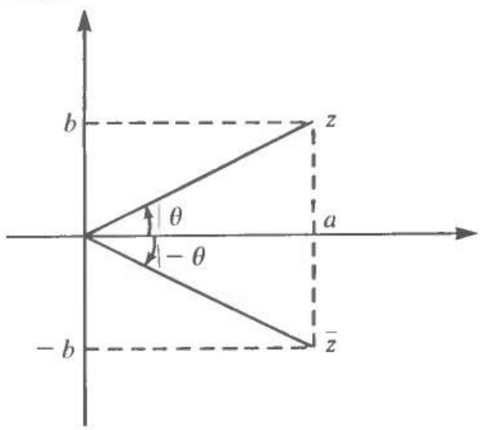


Figure 2.

et de même si $z = a + ib$

$$\bar{z} = a - ib$$

On remarque que l'application $z \to \bar{z}$ est une application bijective de $\mathbb{C}$ sur $\mathbb{C}$ satisfaisant à :

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$$

$$\overline{z_1 z_2} = \overline{z_1} \,.\, \overline{z_2}$$

Elle est de plus involutive puisque

$$\overline{(\overline{z_1})} = z_1 \qquad \forall\, z_1 \in \mathbb{C}$$

Notons enfin que si $z = a + ib$

$$z + \bar{z} = 2\,a \qquad a \in \mathbb{R}$$

$$z\bar{z} = (|\, z \,|)^2 = a^2 + b^2$$

Addition et soustraction de nombres complexes

Par définition la somme des nombres complexes $z_1 = [a_1, b_1]$ et $z_2 = [a_2, b_2]$ est le nombre complexe :

$$\boxed{Z = z_1 + z_2 = [a_1 + a_2,\, b_1 + b_2]}$$

M_1 et M_2 étant les images de z_1 et z_2, l'image M de Z est telle que

$$\overrightarrow{OM} = \overrightarrow{OM_1} + \overrightarrow{OM_2}$$

De l'inégalité triangulaire $OM < OM_1 + OM_2$ on déduit

$$|\, z_1 + z_2 \,| \leqslant |\, z_1 \,| + |\, z_2 \,|$$

qu'on peut immédiatement généraliser

$$|\, z_1 + z_2 + \ldots + z_n \,| \leqslant |\, z_1 \,| + |z_2\,| + \ldots + |\, z_n \,|$$

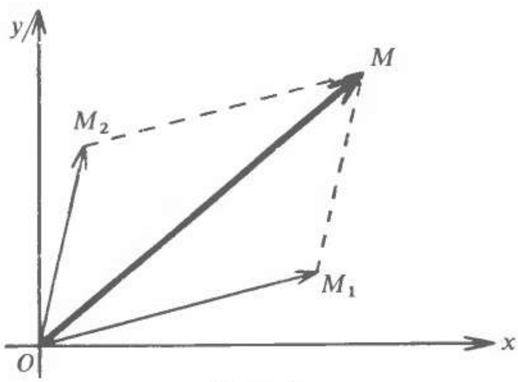


Figure 3.

Le module d'une somme de nombres complexes est inférieur ou égal à la somme des modules.

La différence des nombres complexes z_1 et z_2, soit $Z = z_1 - z_2$, est définie par $Z + z_2 = z_1$ et

$$Z = [a_1 - a_2, b_1 - b_2]$$

Les propriétés de commutativité et d'associativité de l'addition des nombres complexes découlent de la définition de la somme de deux nombres complexes.

Produits de deux nombres complexes

Le produit des nombres complexes $z_1 = (\rho_1, \theta_1)$ et $z_2 = (\rho_2, \theta_2)$ est, par définition, le nombre complexe de module $\rho_1 \cdot \rho_2$ et d'argument $\theta_1 + \theta_2$.

$$\boxed{Z = z_1 \cdot z_2 = (\rho_1 \cdot \rho_2, \theta_1 + \theta_2)}$$

Si $Z = [a, b]$, $z_1 = [a_1, b_1]$ et $z_2 = [a_2, b_2]$ on établit facilement les relations liant a et b à a_1, b_1, a_2 et b_2.

En effet

$$\begin{aligned} a &= \rho_1\rho_2 \cos(\theta_1 + \theta_2) \\ &= \rho_1\rho_2 (\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2) \\ &= (\rho_1 \cos\theta_1) \cdot (\rho_2 \cos\theta_2) - (\rho_1 \sin\theta_1)(\rho_2 \sin\theta_2) \\ &= a_1a_2 - b_1b_2 \end{aligned}$$

De même

$$\begin{aligned} b &= \rho_1\rho_2 \sin(\theta_1 + \theta_2) \\ &= \rho_1\rho_2 (\sin\theta_1 \cos\theta_2 + \sin\theta_2 \cos\theta_1) \\ &= a_1b_2 + a_2b_1 \end{aligned}$$

d'où l'égalité

$$[a_1, b_1]\,[a_2, b_2] = [a_1a_2 - b_1b_2, a_1b_2 + a_2b_1]$$

Le résultat ci-dessus se généralise au cas du produit de n nombres complexes

$$|z_1 \cdot z_2 \cdot z_3 \ldots z_n| = |z_1| \cdot |z_2| \ldots |z_n|$$

et

$$\arg(z_1 \cdot z_2 \cdot z_3 \ldots z_n) = \arg z_1 + \arg z_2 + \ldots + \arg z_n \; (2\,\pi)$$

3. QUOTIENT DE DEUX NOMBRES COMPLEXES

Si $Z = \frac{z_1}{z_2}$ est le quotient de z_1 par z_2 on peut, à partir de la définition de la multiplication, identifier le module R et l'argument Θ de Z.

$z_1 = Zz_2$ conduit à

$$\rho_1 = R\rho_2 \qquad \text{soit} \quad R = \frac{\rho_1}{\rho_2}$$

$$\theta_1 = \Theta + \theta_2 \qquad \text{soit} \quad \Theta = \theta_1 - \theta_2$$

En particulier

$$\frac{1}{z} = \frac{1}{\rho(\cos\theta + i\sin\theta)} = \frac{1}{\rho}(\cos\theta - i\sin\theta) \qquad \text{si } \rho \neq 0$$

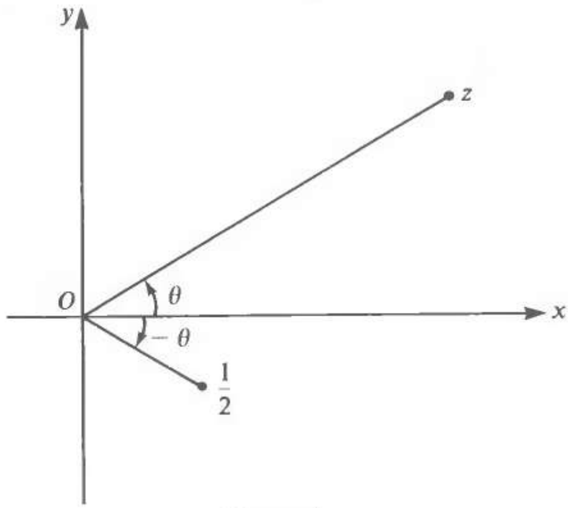


Figure 4.

4. NOMBRES RÉELS. NOMBRES IMAGINAIRES PURS. NOTATION ALGÉBRIQUE ET TRIGONOMÉTRIQUE DES NOMBRES COMPLEXES

Les nombres de la forme $z = [a, 0]$ ont leur image située sur l'axe des x. Nous pouvons confondre l'ensemble des nombres réels et l'ensemble des nombres complexes du type $z = [a, 0]$.

Les nombres z de la forme $[0, b]$ ont leur image située sur l'axe Oy. On dit qu'ils sont **imaginaires purs.** L'argument d'un nombre imaginaire pur est $\frac{\pi}{2}$ à $K\pi$ près.

Le nombre i

Le nombre imaginaire pur [0, 1] de module + 1 et d'argument $+\frac{\pi}{2}$ (à $2K\pi$ près) est désigné par i.

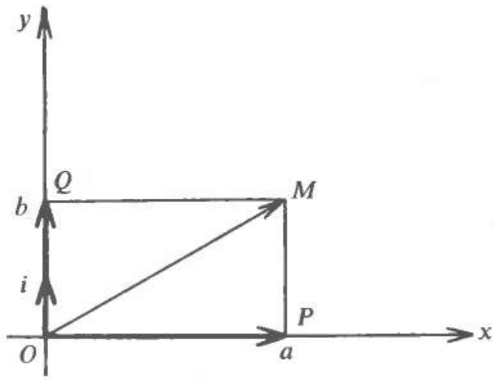


Figure 5.

Le nombre complexe $z = [a, b]$ d'image M est la somme des nombres $[a, 0]$ et $[0, b]$ ayant respectivement P et Q pour image

$$[a, b] = [a, 0] + [0, b]$$

or $[0, b] = b \, . \, i$, et d'après les définitions précédentes

$$\boxed{z = a + ib}$$

$a + ib$ est la forme algébrique de z; a est la partie réelle et b la partie imaginaire de z. On note souvent $a = \mathscr{R}(z)$ et $b = \mathscr{I}(z)$.

Et puisque $a = \rho \cos \theta$ et $b = \rho \sin \theta$ on peut encore écrire

$$\boxed{z = \rho\,(\cos \theta + i \sin \theta)}$$

qui est la forme trigonométrique de z.

Le nombre complexe i ayant pour module 1 et pour argument $+\frac{\pi}{2}$, la multiplication par i correspond donc à la rotation de centre O et d'angle $\frac{\pi}{2}$. En particulier i^2 a pour module 1 et pour argument π, ce qui conduit à $i^2 = 1 \, . \, (\cos \pi + i \sin \pi) = -1$, soit

$$\boxed{i = \sqrt{-1}}$$

Applications.

Si $\quad z_1 = a_1 + ib_1 \quad$ ou $\quad z_1 = \rho_1 (\cos \theta_1 + i \sin \theta_1)$
$\quad\quad z_2 = a_2 + ib_2 \quad\quad\quad z_2 = \rho_2 (\cos \theta_2 + i \sin \theta_2)$

La définition de la somme de deux nombres complexes conduit à

$$\boxed{z_1 + z_2 = a_1 + a_2 + i\,(b_1 + b_2)}$$

celle du produit à

$$z_1 \, . \, z_2 = (a_1a_2 - b_1b_2) + i(a_1b_2 + b_1a_2)$$

ou encore

$$\boxed{z_1z_2 = \rho_1\rho_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]}$$

enfin, la division de z_1 par z_2 conduit à

$$\boxed{\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2} [\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)}$$

5. FORMULE DE MOIVRE

Considérons le nombre complexe de module 1 et d'argument θ:

$$z = \cos \theta + i \sin \theta$$

Des résultats ci-dessus nous déduisons que z^n a pour module 1 et pour argument $n\theta$ soit

$$\boxed{(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta}$$

6. RACINES $n^{\text{ièmes}}$ D'UN NOMBRE COMPLEXE

Soit l'équation

$$z^n = \lambda \tag{1}$$

où λ est le nombre complexe $r(\cos\alpha + i\sin\alpha)$.

Posons $z = \rho(\cos\theta + i\sin\theta)$. Le module de z^n est ρ^n et son argument vaut $n\theta$.

L'égalité (1) équivaut à

$$\rho^n = r \quad \text{et} \quad n\theta = \alpha + 2K\pi$$

soit

$$\rho = \sqrt[n]{r} \quad \text{et} \quad \theta = \frac{\alpha}{n} + \frac{2K\pi}{n}$$

Lorsque K prend les valeurs $0, 1, 2, \ldots, n-1$, θ prend n valeurs distinctes, et on obtient pour l'équation (1) n racines ayant même module.

Exemple.

Racines cubiques de l'unité

$$z^3 = 1$$

$$= \cos 2K\pi + i\sin 2K\pi$$

et

$$z = \cos\frac{2K\pi}{3} + i\sin\frac{2K\pi}{3}$$

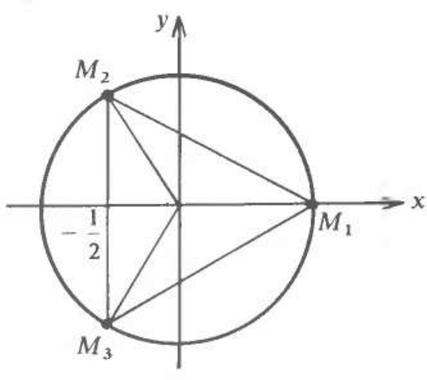


Figure 6.

Pour $K = 0, 1, 2$, on trouve

$$z_1 = 1$$

$$z_2 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = j,$$

$$z_3 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2} = j^2$$

Ces trois racines vérifient la relation

$$1 + j + j^2 = 0$$

et leurs images sont les sommets d'un triangle équilatéral.

7. RÉSOLUTION DE L'ÉQUATION DU SECOND DEGRÉ

Soit l'équation

$$ax^2 + bx + c = 0 \qquad (2)$$

avec $\Delta = b^2 - 4\,ac < 0$ et a, b et c réels. En mettant le trinôme sous sa forme canonique, l'équation (2) s'écrit

$$\left(x + \frac{b}{2\,a}\right)^2 - \frac{b^2 - 4\,ac}{4\,a^2} = 0$$

ou

$$\left(x + \frac{b}{2\,a}\right)^2 = \frac{i^2\,(-\Delta)}{4\,a^2}$$

et

$$x = -\frac{b}{2\,a} \pm i\frac{\sqrt{-\Delta}}{2\,a}$$

L'équation du second degré dont le discriminant est négatif a deux racines complexes conjuguées

$$\frac{-b \pm i\sqrt{-\Delta}}{2\,a}$$

Exemple.

L'équation $x^2 - x + 1 = 0$ admet pour racines les quantités

$$x = \frac{1 \pm i\sqrt{3}}{2}$$

8. TRANSFORMATIONS GÉOMÉTRIQUES SIMPLES

Parmi les nombreuses applications des nombres complexes, nous en signalons dès maintenant certaines fort simples d'ordre géométrique.

Quelques aspects ont été évoqués dans le tome 3 (transformations conformes).

Si $\mathscr{D} \subset \mathbb{C}$, considérons les applications $f: \mathscr{D} \to \mathbb{C}$ ci-dessous et désignons par m et M les affixes de z et $Z = f(z)$

1) $\underline{f(z) = z + a}$ où $a \in \mathbb{C}$

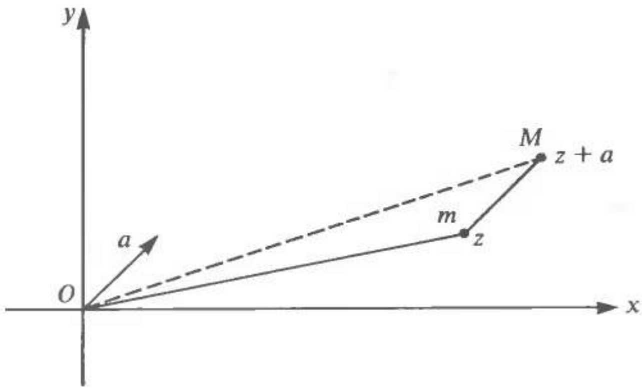


Figure 7.

On voit que la transformation géométrique associée est une translation de vecteur a.

2) $\underline{f(z) = az}$ où $a \in \mathbb{C}$

Si $z = (\rho, \theta)$ et $a = (r, \alpha)$ M se déduit de m par similitude de centre O, de rapport r et d'angle α

$$Z = (r \, . \, \rho, \theta + \alpha)$$

Si $r = |a| = 1$ alors $Z = z\,(\cos\alpha + i \sin\alpha)$ ou, relation inverse équivalente, $z = Z\,(\cos\alpha - i\sin\alpha)$

On en tire

$$\begin{cases} X = x\cos\alpha + y\sin\alpha \\ Y = -\,x\sin\alpha + y\cos\alpha \end{cases}$$

ou

$$\begin{cases} x = X\cos\alpha - Y\sin\alpha \\ y = X\sin\alpha + Y\cos\alpha \end{cases}$$

qui sont les traditionnelles formules de changement d'axes par rotation.

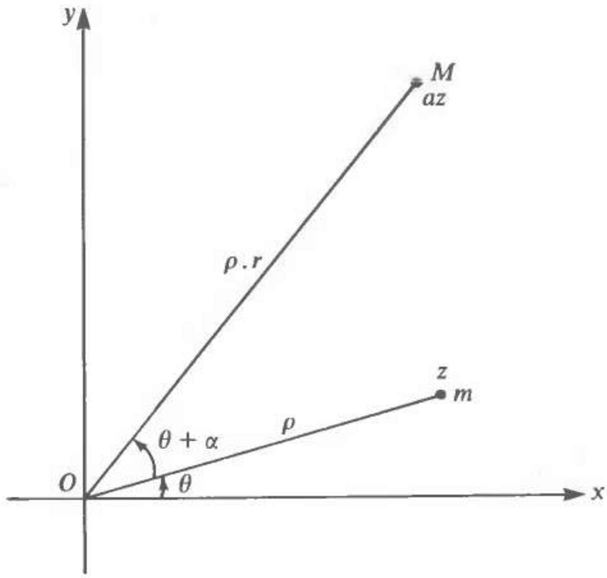


Figure 8.

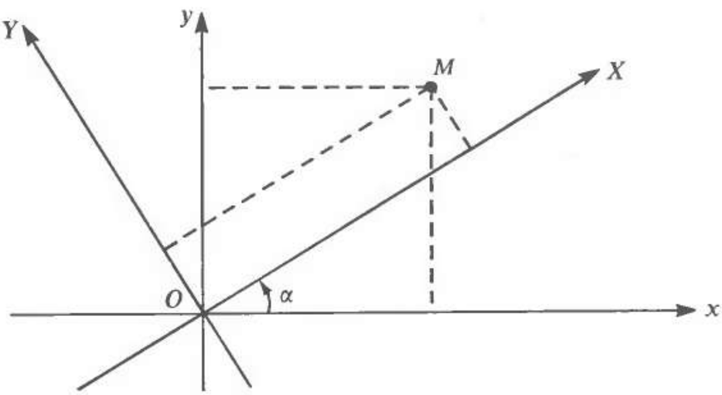


Figure 9.

3) $\underline{f(z) = az + b}$ $\quad a, b \in \mathbb{C}$

Si $a \neq 1$ les points doubles de $Z = az + b$ sont donnés par $z = Z = z_0$ donc

$$z_0 (1 - a) = b \quad \text{et} \quad z_0 = \frac{b}{1 - a}$$

L'expression de $f(z) - z_0 = a(z - z_0)$ montre que c'est une similitude de centre M_0 affixe de z_0, d'angle $\alpha = \arg a$ et de rapport $|a|$; elle se réduit à une rotation pour $|a| = 1$.

4) $f(z) = \dfrac{1}{z}$

On passe de m à M par le produit d'ailleurs commutatif de
— l'inversion de pôle O et puissance 1,
— une symétrie par rapport à l'axe des x.

5) $f(z) = \dfrac{az + b}{cz + d}$ $\qquad a, b, c, d \in \mathbb{C}$

Ce cas, fonction homographique de z, regroupe en fait tous les cas précédents; il suffit en effet d'écrire

$$Z = f(z) = \frac{a}{c} + \frac{ad - bc}{c\left(z + \frac{d}{c}\right)}$$

et il se ramène aux transformations géométriques rencontrées ci-dessus.

Posons
$$\lambda = \frac{a}{c} \in \mathbb{C}$$
$$\mu = \frac{bc - ad}{c^2}$$
$$\nu = \frac{d}{c}$$
$$f(z) = \lambda + \frac{\mu}{z + \nu}$$

d'après ce qui a précédé

$$\left\{\begin{array}{lll} z \longrightarrow z_1 = z + \nu & & \text{est une translation} \\ z_1 \longrightarrow z_2 = \dfrac{1}{z_1} & & \text{est une inversion suivie d'une symétrie} \\ z_2 \longrightarrow z_3 = \mu z_2 & & \text{est une similitude de centre } O \\ z_3 \longrightarrow z_4 = z_3 + \lambda & & \text{est une translation} \end{array}\right.$$

et la transformation homographique est donc la composée de ces différentes transformations élémentaires.

9. e^z POUR z COMPLEXE. FORMULES D'EULER

Soit $z = x + iy$. $(x, y) \in \mathbb{R}^2$

On définit e^z par (*)

$$e^z = e^{x+iy} = e^x (\cos y + i \sin y)$$

Pour $y = 0$, $z = x$ est réel et e^z se réduit à e^x. Dans le cas où z est imaginaire pur on obtient

$$e^{iy} = \cos y + i \sin y$$

et en changeant y en $-y$

$$e^{-iy} = \cos y - i \sin y$$

En ajoutant et retranchant membre à membre ces deux égalités, on obtient les formules d'Euler :

$$\cos y = \frac{e^{iy} + e^{-iy}}{2} \qquad \sin y = \frac{e^{iy} - e^{-iy}}{2i}$$

On définit de même pour $z \in \mathbb{C}$:

$$\operatorname{ch} z = \frac{e^z + e^{-z}}{2} \qquad \operatorname{sh} z = \frac{e^z - e^{-z}}{2}$$

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}$$

$$\operatorname{th} z = \frac{\operatorname{sh} z}{\operatorname{ch} z} \qquad \operatorname{tg} z = \frac{\sin z}{\cos z}$$

les relations liant les lignes hyperboliques et trigonométriques de z ayant été établies au tome 2 chapitre 8, nous les rappelons :

$$\sin iz = i \operatorname{sh} z \quad \cos iz = \operatorname{ch} z$$
$$\operatorname{sh} iz = i \sin z \quad \operatorname{ch} iz = \cos z$$

Nous rappelons aussi les développements en série étudiés au tome 3 :

$$\sin z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$$
$$\cos z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

(*) Une justification plus complète de e^z par les séries entières a été vue dans le tome 3.

$$\operatorname{sh} z = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}$$

$$\operatorname{ch} z = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!}$$

Résumé du chapitre 4

- **Nombres complexes.**

$z = [a, b]$

L'affixe de z est le point de coordonnées a, b.

Module de z $\quad \rho = \sqrt{a^2 + b^2}$

Argument de z $\quad \theta = \operatorname{Arc\,tg} \frac{b}{a}$

$$i = \sqrt{-1}$$

$$z = a + ib = (\rho, \theta)$$

$$= \rho(\cos\theta + i\sin\theta)$$

Produit de $z_1 = (\rho_1, \theta_1)$ par $z_2 = (\rho_2, \theta_2)$

$$z_1 \,.\, z_2 = (\rho_1 \,.\, \rho_2, \theta_1 + \theta_2)$$

- **Formule de Moivre.**

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$$

- **Formules d'Euler.**

$$\cos\alpha = \frac{e^{i\alpha} + e^{-i\alpha}}{2}$$

$$\sin\alpha = \frac{e^{i\alpha} - e^{-i\alpha}}{2i}$$

EXERCICES

■ **1.** Simplifier les expressions

$$\sqrt{-\frac{4}{9}};\ (2 + i) + (7 - 3i);\ \left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right)^3;\ (-1 + i)^3$$

$$\sqrt{-\frac{4}{9}} = \sqrt{\frac{4\,i^2}{9}} = \frac{2\,i}{3}$$

$$(2 + i) + (7 - 3\,i) = 9 - 2\,i$$

$$\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right)^3 = \cos\frac{3\,\pi}{3} + i\sin\frac{3\,\pi}{3} = -1$$

$$\begin{aligned}(-1 + i)^3 &= (-1)^3 + 3\,(-1)^2\,i + 3\,(-1)\,.\,i^2 + i^3\\ &= -1 + 3\,i - 3\,i^2 - i\\ &= 2 + 2\,i.\end{aligned}$$

■ **2.** Calculer

$$A = (2 - 5\,i)(3 + 8\,i)(1 + i)$$
$$B = \frac{(2 + i)^2 + (2 - i)^2}{(2 + i)^2 - (2 - i)^2}$$
$$C = (1 + 2\,i)^4.$$

$$(2 - 5\,i)(3 + 8\,i) = 6 + (16 - 15)\,i - 40\,i^2 = 46 + i$$

$$A = (46 + \mathrm{i})(1 + \mathrm{i}) = 46 + \mathrm{i}\,(46 + 1) + \mathrm{i}^2 = 45 + 47\,i$$
$$(2 + i)^2 = 4 + 4\,i + i^2 = 3 + 4\,i$$
$$(2 - i)^2 = 4 - 4\,i + i^2 = 3 - 4\,i$$
$$B = \frac{3 + 4\,i + 3 - 4\,i}{3 + 4\,i - 3 + 4\,i} = \frac{6}{8\,i} = -\frac{3}{4}\,i$$
$$\begin{aligned}C &= (1 + 2\,i)^4 = [(1 + 2\,i)^2]^2 = (1 + 4\,i + 4\,i^2)^2 = (-3 + 4\,i)^2\\ &= 9 - 24\,i + 16\,i^2 = -7 - 24\,i\end{aligned}$$

■ **3.** Mettre sous la forme $a + ib$ les quantités suivantes

$$\frac{17}{4 - i};\ \frac{1 + i}{1 - i};\ \frac{2 - 3\,i}{4 - 5\,i};\ \frac{1}{5 - 3\,i} - \frac{1}{5 + 3\,i};\ \left(\frac{4\,i^{11} - i}{1 + 2\,i}\right)^2$$

$$\frac{17}{4 - i} = \frac{17\,(4 + i)}{(4 - i)(4 + i)} = \frac{17\,(4 + i)}{4^2 - i^2} = \frac{17\,(4 + i)}{16 + 1} = 4 + i$$
$$\frac{1 + i}{1 - i} = \frac{(1 + i)^2}{(1 - i)(1 + i)} = \frac{1 + 2\,i + i^2}{1 - i^2} = \frac{2\,i}{2} = i$$
$$\frac{2 - 3\,i}{4 - 5\,i} = \frac{(2 - 3\,i)(4 + 5\,i)}{(4 - 5\,i)(4 + 5\,i)} = \frac{8 + i\,(10 - 12) - 15\,i^2}{16 - 25\,i^2} = \frac{23 - 2\,i}{41}$$
$$\frac{1}{5 - 3\,i} - \frac{1}{5 + 3\,i} = \frac{5 + 3\,i - 5 + 3\,i}{25 - 9\,i^2} = \frac{3}{17}\,i$$
$$i^{11} = i^{10}\,.\,i = (i^2)^5\,.\,i = (-1)^5\,.\,i = -i$$

et

$$\left(\frac{4\,i^{11} - i}{1 + 2\,i}\right)^2 = \frac{25\,i^2}{1 + 4\,i + 4\,i^2} = \frac{-\,25}{-\,3 + 4\,i}$$

$$= -\,\frac{25\,(-\,3 - 4\,i)}{(-\,3 + 4\,i)\,(-\,3 - 4\,i)} = 3 + 4\,i$$

■ **4.** Représenter et mettre sous forme trigonométrique les nombres complexes

$$2 + 2\,i;\ -\,1 - i\,\sqrt{3};\ \sqrt{3} - i;\ -\,\frac{5}{2}\,i$$

Module $\qquad \rho = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\,\sqrt{2}$

Argument $\qquad \operatorname{tg}\theta = \dfrac{2}{2} = 1, \quad \theta = \dfrac{\pi}{4}\ (2\,K\pi)$

$$2 + 2\,i = 2\,\sqrt{2}\left(\cos\frac{\pi}{4} + i\,\sin\frac{\pi}{4}\right)$$

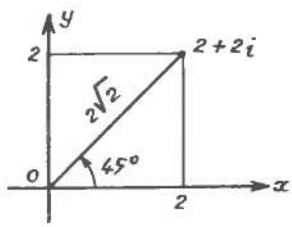


Figure 10.

Module $\qquad \rho = \sqrt{1 + 3} = 2$

Argument $\qquad \operatorname{tg}\theta = \sqrt{3} \qquad \theta = \pi + \dfrac{\pi}{3}\,(2\,K\pi)$

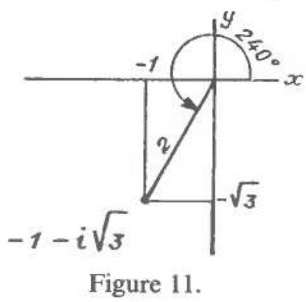


Figure 11.

Module $\rho = \sqrt{1 + 3} = 2$

Argument $\operatorname{tg} \theta = -\dfrac{1}{\sqrt{3}}$ $\theta = \dfrac{11\pi}{6} (2K\pi)$

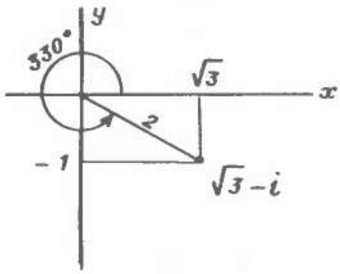


Figure 12.

Module $\rho = \dfrac{5}{2}$

Argument $\theta = \dfrac{3\pi}{2} (2K\pi)$

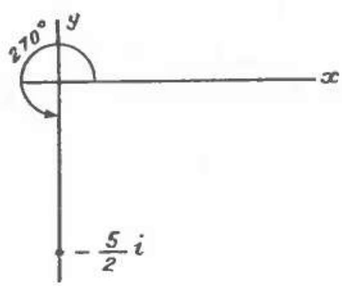


Figure 13.

■ 5. Utiliser la formule de Moivre pour calculer $\cos 3\theta$ et $\sin 3\theta$.

$$\begin{aligned}\cos 3\theta + i \sin 3\theta &= (\cos\theta + i\sin\theta)^3 \\ &= \cos^3\theta + 3\cos^2\theta \,.\, i\sin\theta \\ &\quad + 3\cos\theta\, i^2 \sin^2\theta + i^3\sin^3\theta.\end{aligned}$$

En égalisant séparément partie réelle et partie imaginaire, nous obtenons

$$\begin{aligned}\cos 3\theta &= \cos^3\theta - 3\cos\theta\sin^2\theta\\ &= \cos^3\theta - 3\cos\theta(1-\cos^2\theta)\\ &= 4\cos^3\theta - 3\cos\theta,\\ \sin 3\theta &= 3\cos^2\theta\sin\theta - \sin^3\theta\\ &= 3\sin\theta - 4\sin^3\theta.\end{aligned}$$

■ **6.** Quelles conditions doivent vérifier les images M_1 et M_2 des nombres complexes z_1 et z_2 pour que $\frac{z_1}{z_2}$ soit réel, ou imaginaire pur.

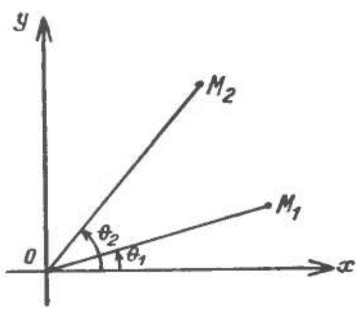


Figure 14.

L'argument de $\frac{z_2}{z_1}$ est

$$\widehat{M_1OM_2} = \theta_2 - \theta_1$$

$\frac{z_2}{z_1}$ sera réel si $\widehat{M_1OM_2}$ est multiple de π, c'est-à-dire si M_1 et M_2 sont alignés avec O; $\frac{z_2}{z_1}$ est imaginaire pur pour $\widehat{M_1OM_2} = \frac{\pi}{2}$.

■ **7.** Quelles sont les racines carrées de i.

i a pour module 1 et pour argument $\frac{\pi}{2} + 2K\pi$

$$\sqrt{i} = \cos\left(\frac{\pi}{4} + K\pi\right) + i \sin\left(\frac{\pi}{4} + K\pi\right)$$

Pour $K = 0$ on trouve $\frac{1}{\sqrt{2}}(1 + i)$

Pour $K = 1$ on trouve $-\frac{1}{\sqrt{2}}(1 + i)$

■ **8.** Trouver les trois racines cubiques de $-2 + 2i$.

$$-2 + 2i = 2(-1 + i) = 2\sqrt{2}(\cos 135^\circ + i \sin 135^\circ).$$

Les racines cubiques de cette quantité ont pour module

$$\sqrt[3]{2^{3/2}} = \sqrt{2}$$

et pour argument

$$\frac{135^\circ + K\,360^\circ}{3}$$

ce sont
$$\sqrt{2}(\cos 45^\circ + i \sin 45^\circ) = 1 + i$$
$$\sqrt{2}(\cos 165^\circ + i \sin 165^\circ) = -1{,}366 + i\,.\,0{,}366$$
$$\sqrt{2}(\cos 285^\circ + i \sin 285^\circ) = 0{,}366 - i\,.\,1{,}366$$

■ **9.** Mettre sous forme $a + ib$ les quantités

$$e^{i\pi},\ e^{i\pi/2},\ 5\,e^{i\pi/4},\ e^{7i\pi},\ e^{7i\pi/2}.$$

$$e^{i\pi} = \cos \pi + i \sin \pi = -1$$

$$e^{i\pi/2} = \cos\frac{\pi}{2} + i \sin\frac{\pi}{2} = i$$

$$5\,e^{i\pi/4} = 5\left(\cos\frac{\pi}{4} + i \sin\frac{\pi}{4}\right)$$
$$= \frac{5}{\sqrt{2}}(1 + i)$$

$$e^{7i\pi} = \cos 7\pi + i \sin 7\pi$$
$$= -1$$

$$e^{7i\pi/2} = \cos\frac{7\pi}{2} + i\sin\frac{7\pi}{2}$$
$$= -i.$$

■ **10.** Module et argument de $e^{1+i\sqrt{3}}$.

$$e^{1+i\sqrt{3}} = e(\cos\sqrt{3} + i\sin\sqrt{3})$$

Le module est e et l'argument $\sqrt{3}$.

■ **11.** Calculer les deux sommes

$$U_n = 1 + a\cos\theta + a^2\cos 2\theta + \dots + a_n\cos n\theta$$
$$V_n = a\sin\theta + a^2\sin 2\theta + \dots + a_n\sin n\theta$$

et les limites U et V lorsque $n \to \infty$

Pour utiliser la formule de Moivre, formons $U_n + iV_n$:

$$U_n + iV_n = 1 + a(\cos\theta + i\sin\theta) + a^2(\cos\theta + i\sin\theta)^2 + \dots + a^n(\cos\theta + i\sin\theta)^n$$
$$= \sum_{k=1}^{n} a^k e^{ik\theta}$$

progression géométrique de raison $ae^{i\theta}$

d'où
$$U_n + iV_n = \frac{1 - a^{n+1}[\cos(n+1)\theta + i\sin(n+1)\theta]}{1 - a(\cos\theta + i\sin\theta)}$$

Pour obtenir U_n et V_n il suffit de séparer parties réelle et imaginaire, en multipliant par la quantité conjuguée du dénominateur, on trouve

$$U_n = \frac{[1 - a^{n+1}\cos(n+1)\theta](1 - a\cos\theta) + a^{n+2}\sin\theta\sin(n+1)\theta}{1 - 2a\cos\theta + a^2}$$

$$V_n = \frac{a\sin\theta[1 - a^{n+1}\cos(n+1)\theta] - a^{n+1}(1 - a\cos\theta)\sin(n+1)\theta}{1 - 2a\cos\theta + a^2}$$

Les limites obtenues pour n tendant vers l'infini n'existent que si $|a| < 1$

On a alors
$$U + iV = \frac{1}{1 - a(\cos\theta + i\sin\theta)}$$

soit

$$U = \frac{1 - a\cos\theta}{1 - 2a\cos\theta + a^2} \qquad V = \frac{a\sin\theta}{1 - 2a\cos\theta + a^2}$$

■ **12.** Quelle condition faut-il imposer à z pour que $|z + 5| = |z - i|$?

La condition $|z + 5| = |z - i|$ est équivalente à

$$(z + 5)^2 = (z - i)^2$$

et pour $z = x + iy \quad x, y \in \mathbb{R}$ on a

$$(x + 5)^2 + y^2 = x^2 + (1 - y)^2$$

soit encore

$$y = -5x - 12$$

équation d'une droite

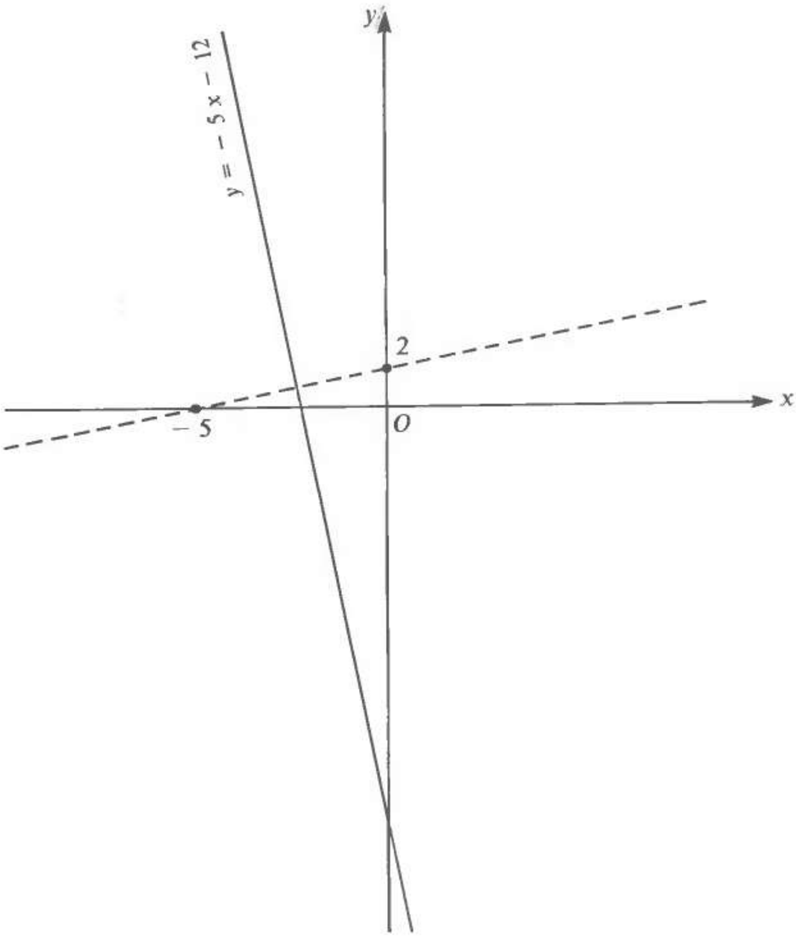


Figure 15.

Ce résultat était d'ailleurs prévisible puisque $|z + 5| = |z - i|$ traduit que l'affixe de M est équidistant des points A d'affixe -5 soit A $(-5, 0)$ et B d'affixe i soit $B\,(0, 1)$ et l'ensemble cherché n'est autre que la médiatrice de AB.

■ **13.** L'image de $z = x + iy$ décrivant le cercle

$$x^2 + y^2 - 2\,a\,x = 0$$

qu'en est-il de celle de $Z = \dfrac{z-1}{z+1}$?

Posons

$$Z = X + i\,Y$$
$$z\,\bar{z} = x^2 + y^2$$
$$z + \bar{z} = 2\,x$$

et l'équation du cercle est également donnée par la relation

$$z\bar{z} - a\,(z + \bar{z}) = 0 \tag{1}$$

L'expression de z en fonction de Z est déduite de

$$zZ + Z - z + 1 = 0$$

soit

$$z = -\frac{Z+1}{Z-1}$$

et en portant dans (1)

$$\frac{(Z+1)}{(Z-1)}\cdot\frac{(\bar{Z}+1)}{(\bar{Z}-1)} + a\left[\frac{(Z+1)}{(Z-1)} + \frac{(\bar{Z}+1)}{(\bar{Z}-1)}\right] = 0$$

on obtient après réduction

$$Z\bar{Z} + Z + \bar{Z} + 1 + a\,(Z\bar{Z} + \bar{Z} - Z - 1 + Z\bar{Z} - \bar{Z} + Z - 1) = 0$$

soit encore

$$Z\bar{Z} + Z + \bar{Z} + 1 + 2\,a\,(Z\bar{Z} - 1) = 0$$

et remplaçant $Z\bar{Z}$ par $X^2 + Y^2$, $Z + \bar{Z}$ par $2\,X$, il reste

$$(1 + 2\,a)\,(X^2 + Y^2) + 2\,X + 1 - 2\,a = 0$$

Si $a \neq -\dfrac{1}{2}$ Z décrit donc un cercle.

Si $a = -\dfrac{1}{2}$ Z décrit la droite d'abscisse -1.

■ **14.** z décrivant un cercle centré à l'origine, que décrit Z sachant que

$$Z^2 = z + 1$$

Posons $$z = \rho\, e^{i\theta}$$
$$Z^2 = (X + iY)^2 = 1 + \rho\, e^{i\theta}$$
soit $$X^2 - Y^2 + 2\, iXY = 1 + \rho \cos\theta + i\rho \sin\theta$$
séparant parties réelles et imaginaires, on a
$$X^2 - Y^2 = 1 + \rho \cos\theta$$
$$2\, XY = \rho \sin\theta$$
et après élimination de θ
$$(X^2 - Y^2 - 1)^2 + 4\, X^2Y^2 = \rho^2$$
ou encore
$$(X^2 + Y^2)^2 - 2\,(X^2 - Y^2) = \rho^2 - 1$$
quartique qui, pour $\rho = 1$, se réduit à une lemniscate (voir t. 5).

■ **15.** A z d'affixe m on associe $Z = \dfrac{i}{z + i}$ d'affixe M.

Que décrit M quand m décrit une droite ou un cercle?

Posons $z_1 = z + i$; m_1 image de z_1 se déduit de m par une translation parallèle à Oy et d'amplitude 1.

$z_2 = \dfrac{1}{z_1}$ a pour image m_2 qui se déduit de m_1 par la composition, dans un ordre indifférent, d'une inversion de pôle O et de puissance 1 et d'une symétrie par rapport à Ox.

Enfin M se déduit par une rotation d'angle $+\dfrac{\pi}{2}$ et de centre O puisque $Z = iz_2$.

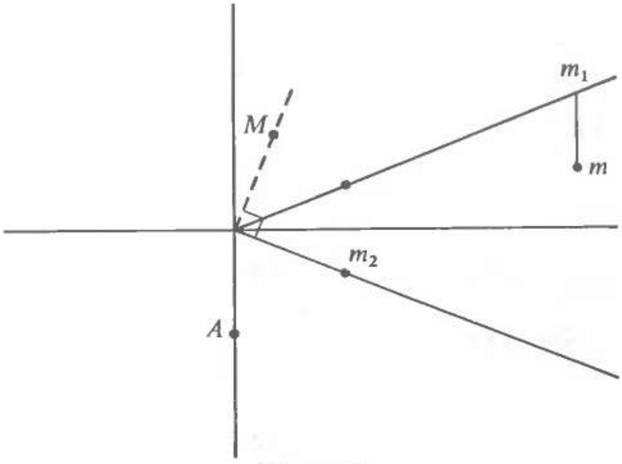


Figure 16.

Si m décrit une droite Δ ne passant pas par le point A image de $-i$, m_1 décrit une droite parallèle à Δ ne passant pas par l'origine; l'inversion de pôle O et de puissance 1 transforme cette droite en un cercle et M décrit donc aussi un cercle; si m décrit une droite passant par A, M se déplace sur une droite issue de O.

Si m décrit un cercle ne passant pas par A, alors M décrit également un cercle.

Si m décrit un cercle passant par A, M décrit une droite.

■ **16.** On considère l'équation

$$z^6 - 2z^3\cos\theta + 1 = 0 \qquad \theta \in \mathbb{R}$$

en cherchant les racines, décomposer le trinôme $z^6 - 2z^3\cos\theta + 1$ en produit de trois trinômes du second degré à coefficients réels. Pour quelles valeurs de θ les images des racines de l'équation sont-elles les sommets d'un hexagone régulier ou d'un triangle équilatéral ?

Posons $Z = z^3$.

L'équation $Z^2 - 2Z\cos\theta + 1 = 0$ a pour discriminant

$$\delta' = \cos^2\theta - 1 = -\sin^2\theta$$

d'où les racines

$$Z_1 = \cos\theta + i\sin\theta = e^{i\theta} \quad \text{et} \quad Z_2 = \cos\theta - i\sin\theta = e^{-i\theta}$$

D'où pour

$$z_j = e^{\pm i\left(\frac{\theta}{3} + \frac{2K\pi}{3}\right)} \qquad 1 \leqslant j \leqslant 6 \qquad 1 \leqslant k \leqslant 3$$

$$\begin{cases} z_1 = \cos\dfrac{\theta}{3} + i\sin\dfrac{\theta}{3} \\ z_2 = \cos\left(\dfrac{\theta}{3} + \dfrac{2\pi}{3}\right) + i\sin\left(\dfrac{\theta}{3} + \dfrac{2\pi}{3}\right) \\ z_3 = \cos\left(\dfrac{\theta}{3} + \dfrac{4\pi}{3}\right) + i\sin\left(\dfrac{\theta}{3} + \dfrac{4\pi}{3}\right) \\ z_4 = \cos\dfrac{\theta}{3} - i\dfrac{\theta}{3} = \overline{z_1} \\ z_5 = \cos\left(\dfrac{\theta}{3} + \dfrac{2K\pi}{3}\right) - i\sin\left(\dfrac{\theta}{3} + \dfrac{2K\pi}{3}\right) = \overline{z_2} \\ z_6 = \cos\left(\dfrac{\theta}{3} + \dfrac{4K\pi}{3}\right) - i\sin\left(\dfrac{\theta}{3} + \dfrac{4K\pi}{3}\right) = \overline{z_3} \end{cases}$$

L'équation proposée peut donc s'écrire

$$(z - z_1)(z - \overline{z_1})(z - z_2)(z - \overline{z_2})(z - z_3)(z - \overline{z_3}) = 0$$

Or $\quad z + \bar{z} = 2\,\mathcal{R}(z)$

$$z \,.\, \bar{z} = |\,z\,|^2$$

et l'égalité ci-dessus conduit à la décomposition demandée

$$\left(z^2 - 2\,z\cos\frac{\theta}{3} + 1\right)\left[z^2 - 2\,z\cos\left(\frac{\theta}{3} + \frac{2\,\pi}{3}\right) + 1\right]$$

$$\left[z^2 - 2\,z\cos\left(\frac{\theta}{3} + \frac{4\,\pi}{3}\right)\right]$$

Les expressions des racines montrent que z_1, z_2, z_3 sont les sommets d'un triangle équilatéral inscrits dans le cercle de rayon 1 centré en O, z_4, z_5, z_6 celles d'un triangle symétrique du précédent par rapport à Ox.

Les six points figurant les racines seront sommets d'un hexagone régulier pour $\dfrac{\theta}{3} = \dfrac{\pi}{6}$ soit $\theta = \dfrac{\pi}{2}$.

Ils seront sommets d'un triangle équilatéral si les deux triangles précédents sont confondus, soit si $\dfrac{\theta}{3} = 0$ ou $\theta = 0$.

CHAPITRE 5

POLYNÔMES. FRACTIONS RATIONNELLES

1. DÉFINITION

Un polynôme de degré n a la forme

$$A = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$$

où a_0 au moins est différent de 0.

Ainsi écrit, on dit que A est ordonné suivant les puissances décroissantes de la variable $x \in \mathbb{R}$; nous nous limiterons à ce cas en remarquant cependant que x peut appartenir à un ensemble quelconque non nécessairement sous-ensemble de $\mathbb{R}$. Ainsi x peut représenter une matrice carrée pour laquelle nous verrons que sont parfaitement définies les opérations d'addition, de multiplication par un scalaire et de produit. Dans un tel cas $A(x)$ est alors lui-même une matrice.

L'ensemble des considérations et résultats qui vont suivre pourraient donc s'étendre au cas plus général où l'**indéterminée** x représente des êtres mathématiques assez divers, et donner une plus large portée à la théorie élémentaire.

Notation

La suite ordonnée des coefficients de A: $a_0, a_1, \dots, a_n, 0, 0 \dots$ a au plus n termes non nuls.

On convient de noter

$$A = [a_0, a_1, \dots, a_n]$$

Polynôme nul

Le polynôme nul est le polynôme dont tous les coefficients sont nuls.

On peut le noter

$$[0] = [0, 0, 0, \dots]$$

2. ADDITION DES POLYNÔMES

Par définition la somme $A + B$ des polynômes

$$A = [a_0, a_1, a_2, \ldots, a_p]$$
$$B = [b_0, b_1, b_2, \ldots, b_q]$$

est le polynôme

$$A + B = [a_0 + b_0, a_1 + b_1, \ldots, a_m + b_m, \ldots]$$

On constate que cette définition implique que

$$d^0 (A + B) \leqslant \text{Sup } [d^0 (A), d^0 (B)]$$

en désignant par $d^0 (P)$ le degré du polynôme $P(x)$.

De la définition de la somme de A et B il est simple de déduire que l'addition des polynômes est une loi de composition interne associative et commutative. Elle possède un élément neutre le polynôme $[0] = [0, 0, \ldots]$.

De plus le polynôme $A = [a_0, a_1, \ldots, a_n]$ admet un opposé, le polynôme

$$-A = [-a_0, -a_1, \ldots, -a_n]$$

Ainsi l'addition confère à l'ensemble des polynômes une structure de groupe abélien.

3. MULTIPLICATION DES POLYNÔMES

On appelle produit des polynômes

$$A = [a_0, a_1, \ldots, a_p]$$
$$B = [b_0, b_1, \ldots, b_q]$$

le polynôme

$$C = A \,.\, B = [c_0, c_1, \ldots, c_m]$$

tel que $c_0 = a_0 b_0$, $c_1 = a_1 b_0 + b_1 a_0$ et d'une façon générale

$$c_k = \sum_{i+j=k} a_i \,.\, b_j$$

En particulier le coefficient du terme de plus haut degré est

$$c_m = a_p b_q$$

ou
$$d^0 (A \,.\, B) = d^0 (A) + d^0 (B)$$

On peut ainsi vérifier que la multiplication des polynômes est associative et distributive par rapport à l'addition.

Muni des deux lois de composition que sont l'addition et la multiplication des polynômes, **l'ensemble des polynômes a une structure d'anneau commutatif.**

4. MULTIPLICATION PAR UN SCALAIRE

Si $A = [a_0, a_1, \ldots, a_n]$ et $\lambda \in \mathbb{C}$ le produit λA est le polynôme $[\lambda a_0, \lambda a_1, \ldots, \lambda a_n]$. La loi de composition externe ci-dessus définie possède les propriétés :

$$\left\{ \begin{aligned} \lambda(\mu A) &= (\lambda\mu) A \\ (\lambda + \mu) A &= \lambda A + \mu A \\ 1 . A &= A \\ \lambda . (A + B) &= \lambda A + \lambda B \end{aligned} \right.$$

On en déduit que l'ensemble des polynômes relativement à l'addition et à la multiplication par un scalaire a une structure d'espace vectoriel (voir chap. 6). On voit d'ailleurs que les $(n + 1)$ monômes

$$1, x, x^2, \ldots, x^n$$

sont linéairement indépendants et constituent une base de dimension $n + 1$ pour l'espace vectoriel des polynômes de degré au plus égal à n.

5. POLYNÔME DÉRIVÉ

Définition

On appelle polynôme dérivé de

$$P(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n$$

le polynôme

$$P'(x) = n a_0 x^{n-1} + (n - 1) a_1 x^{n-2} + \ldots + a_{n-1}$$

cette définition découlant de la notion de dérivée de l'application $x \to P(x)$, fonction de la variable réelle (ou complexe x).

L'application $P(x) \to P'(x)$ vérifie les propriétés évidentes suivantes :

1) $(P + Q)' = P' + Q'$ (dérivée d'une somme)

2) $\forall\, k \in \mathbb{C}, \quad (kP)' = kP'$

3) $(PQ)' = P'Q + Q'P$

cette dernière propriété est simple à établir.

Si P et Q sont des monômes

$$P = ax^n \qquad Q = bx^p$$

$$(P . Q)' = ab\,[x^{n+p}]' = (n + p)\, abx^{n+p-1}$$

$$P'Q + Q'P = nabx^{n+p-1} + pabx^{n+p-1} = (PQ)'$$

Si P et Q sont des polynômes, leur produit est une somme de monômes, et on est ramené au cas précédent.

4) Le polynôme dérivé de P^n est $nP^{n-1}P'$.

Si $n = 2$, ce résultat découle de 3) et se généralise par récurrence sur n.

6. FORMULES DE MAC LAURIN ET DE TAYLOR

Le polynôme $P(x)$ ordonné suivant les puissances croissantes permet d'écrire

$$P(x) = \sum_{i=0}^{n} \alpha_i x^i$$

d'où

$$P'(x) = \sum_{i=1}^{n-1} i\alpha_i x^{i-1}$$

$$P''(x) = \sum_{i=2}^{n-2} i\,(i-1)\,x^{i-2}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$P^{(n)}(x) = \alpha_n$$

De ces égalités on tire

$$P(0) = \alpha_0$$

$$P'(0) = \alpha_1$$

$$P''(0) = 2\,\alpha_2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$P^{(n)}(0) = n\,(n-1)\,(n-2)\ldots 1\,\alpha_n = n!\,\alpha_n$$

d'où

$$\boxed{P(x) = P(0) + x\frac{P'(0)}{1!} + x^2\frac{P''(0)}{2!} + \ldots + x^n\frac{P^{(n)}(0)}{n!}} \qquad (1)$$

qui est la formule de Mac Laurin pour un polynôme.

Pour obtenir la formule de Taylor de $P(x)$ il suffit de poser

$$Q(x) = P(x+h).$$

La formule précédente de Mac Laurin pour $Q(x)$ donne

$$Q(x) = Q(0) + x\frac{Q'(0)}{1!} + x^2\frac{Q''(0)}{2!} + \ldots + x^n\frac{Q^{(n)}(0)}{n!} \qquad (2)$$

et comme

$$Q(x) = P(x+h) = \alpha_0 + \alpha_1(h+x) + \ldots + \alpha_n(h+x)^n$$

on tire

$$Q(0) = P(h)$$
$$Q'(x) = \alpha_1 + 2\,\alpha_2\,(h + x) + \ldots + n\alpha_n\,(h + x)^{n-1}$$
$$Q'(0) = P'(h)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

enfin

$$Q^{(n)}(x) = n!\,\alpha_n \Rightarrow Q^{(n)}(0) = P^{(n)}(h)$$

soit en portant dans (2)

$$\boxed{P(x + h) = P(h) + \frac{x\,P'(h)}{1!} + x^2\,\frac{P''(h)}{2!} + \ldots + x^n\,\frac{P^{(n)}(h)}{n!}} \quad (3)$$

qui est la formule de Taylor pour un polynôme.

Applications.

• *Formule du binôme*

La formule de Taylor appliquée au polynôme $P(x) = x^n$ conduit à

$$P'(x) = nx^{n-1}, \qquad P''(x) = n\,(n-1)\,x^{n-2}$$
$$P^{(p)}(x) = n\,(n-1)\ldots(n-p+1)\,x^{n-p}$$
$$P^{(n)}(x) = n!$$

Il en résulte pour $Q(x) = P(x + h) = (x + h)^n$

$$(x + h)^n = h^n + nh^{n-1}x + \frac{n\,(n-1)}{2!}\,x^2h^{n-2} + \ldots$$
$$+ \frac{n\,(n-1)\ldots(n-p+1)}{p!}\,x^ph^{n-p} + \ldots + x^n$$

qui, pour $x = a$ et $h = b$, redonne la formule classique du binôme de Newton :

$$(a + b)^n = a^n + na^{n-1}\,b + \frac{n\,(n-1)}{2!}\,a^{n-2}b^2 + \ldots$$
$$+ C_n^p\,a^{n-p}b^p + \ldots\; b^n$$

• *Ordre de multiplicité d'une racine.*

L'ordre de multiplicité d'une racine x_0 de $P(x)$ est le plus grand entier p tel que P soit divisible par $(x - x_0)^p$ $(p \leqslant n)$ (voir § 11).

La formule de Taylor donne le développement de $P(x)$

$$P(x_0) + P'(x_0)\,(x - x_0) + \ldots + \frac{P^{(p-1)}}{(p-1)!}\,(x - x_0)^{p-1}$$
$$+ (x - x_0)^p\left[\frac{P^{(p)}(x_0)}{p!} + \frac{P^{(p+1)}(x_0)}{(p+1)!}\,(x - x_0) + \ldots + \frac{P^{(n)}(x_0)}{n!}\,(x - x_0)^{n-p}\right]$$

Désignons par $r(x)$ la première ligne de ce développement de $P(x)$

$$r(x) = P(x_0) + P'(x_0)(x - x_0) + \dots + \frac{P^{(p-1)}(x_0)}{(p-1)!}(x - x_0)^{p-1}$$

Si x_0 est racine d'ordre p de $P(x)$, $r(x)$ est identiquement nul :

$$r(x_0) = 0 \Rightarrow P(x_0) = 0$$

$$\frac{r(x_0)}{x - x_0} = 0 \Rightarrow P'(x_0) = 0$$

$$\frac{r(x_0)}{(x - x_0)^2} = 0 \Rightarrow P''(x_0) = 0$$

. . . .

$$\frac{r(x_0)}{(x - x_0)^{p-1}} = 0 \Rightarrow P^{(p-1)}(x_0) = 0$$

Réciproquement, si

$$P(x_0) = P'(x_0) = \dots P^{(p-1)}(x_0) = 0$$

mais $P^{(p)}(x_0) \neq 0$ d'après l'écriture ci-dessus de $P(x)$, $P(x)$ est divisible par $(x - x_0)^p$ et admet x_0 comme racine d'ordre p.

7. DIVISION DES POLYNÔMES SUIVANT LES PUISSANCES DÉCROISSANTES OU DIVISION EUCLIDIENNE

Étant donné deux polynômes de degrés respectifs n et p

$$A = a_0x^n + a_1x^{n-1} + \dots + a_n$$
$$B = b_0x^p + b_1x^{p-1} + \dots + b_p$$

avec $p \leqslant n$, on appelle quotient et reste de la division de A par B deux autres polynômes Q et R tels que l'on ait l'identité

$$A = BQ + R \tag{4}$$

le degré de R étant inférieur à celui de B; A est le dividende et B le diviseur; on dit que A est divisible par B si $R = 0$.

Si Q et R existent ils sont **uniques.**

En effet, s'il existait deux autres polynômes Q' et R' vérifiant

$$A = BQ' + R' \tag{5}$$

en rapprochant (4) et (5) on obtiendrait

$$BQ + R = BQ' + R'$$

soit

$$B(Q - Q') = R' - R \tag{6}$$

les polynômes R' et R étant par hypothèse de degrés inférieurs à celui de B, il en est de même du polynôme $R' - R$.

Or le polynôme $B(Q - Q')$ est de degré au moins égal à celui de B et l'égalité (3) est impossible à moins que, $Q - Q' = 0$ et $R - R' = 0$ soit $Q = Q'$ et $R = R'$.

8. RECHERCHE DU QUOTIENT ET DU RESTE

Nous allons dégager la règle qui permet de déterminer le quotient Q et le reste R de la division de A par B.

Considérons les monômes de plus haut degré de A et B : a_0x^net b_0x^p.

Il existe un monôme q_1 tel que

$$a_0x^n = b_0x^p \cdot q_1$$

q_1 a pour degré $n - p$ (et se réduit à une constante si $n = p$).

Le polynôme

$$A_1 = A - Bq_1 \tag{7}$$

est de degré inférieur à celui de A puisque les termes de degré n se détruisent dans A_1. Si le degré de A_1 est de plus inférieur à celui de B, alors $A = Bq_1 + A_1$ est de la forme $A = BQ + R$: le quotient est q_1 et le reste A_1. Si au contraire le degré de A_1 est supérieur à celui de B, on recommence l'opération : on divise le terme de plus haut degré de A_1 par celui de B. Soit q_2 le quotient. On obtient un nouveau polynôme A_2.

$$A_2 = A_1 - Bq_2$$

Si le degré de q_2 est inférieur à celui de B, l'opération est terminée, sinon on poursuit jusqu'à l'ordre K :

$$A_k = A_{k-1} - Bq_k$$

où le degré K est inférieur à p, degré du polynôme B. Alors

$$A = Bq_1 + A_1$$
$$A_1 = Bq_2 + A_2$$
$$\dots\dots\dots\dots\dots\dots$$
$$A_{k-1} = Bq_K + A_k$$

ce qui permet d'écrire en additionnant :

$$A = B(q_1 + q_2 + \dots + q_K) + A_K$$

Le quotient est le polynôme

$$Q = q_1 + q_2 + \dots + q_K$$

et le reste $R = A_K$. C'est sur ce principe que repose la recherche du quotient et du reste.

Pratiquement, on dispose le dividende et le diviseur comme dans une division artihmétique. On effectue la division du terme de degré n du polynôme A par celui de degré p du polynôme B. On obtient ainsi le premier terme q_1 du quotient. On multiplie le diviseur par q_1 et on retranche ce produit du dividende. La différence ainsi obtenue est le polynôme A_1 sur lequel on recommence l'opération jusqu'à ce qu'on arrive à une différence de degré inférieur à celui du diviseur. C'est cette dernière différence qui est le reste.

L'exemple qui suit concrétise les différentes phases:

Diviser $x^5 - 7x^3 + x - 6$ par $x^3 + x^2 - x - 1$

$$
\begin{array}{lr|l}
 & x^5 \qquad - 7x^3 \qquad + x - 6 & x^3 + x^2 - x - 1 \\ \cline{3-3}
 & -x^5 - x^4 + x^3 + x^2 \qquad\qquad & x^2 - x - 5 \\ \hline
\mathbf{A_1} & -x^4 - 6x^3 + x^2 + x - 6 & \\
 & +x^4 + x^3 - x^2 - x \quad & \\ \hline
\mathbf{A_2} & -5x^3 \qquad\qquad - 6 & \\
 & +5x^3 + 5x^2 - 5x - 5 & \\ \hline
\mathbf{Reste} & 5x^2 - 5x - 11 &
\end{array}
$$

Le quotient est $x^2 - x - 5$ et le reste $5\,x^2 - 5\,x - 11$ puisque de degré inférieur à celui du diviseur.

9. DIVISION SUIVANT LES PUISSANCES CROISSANTES

En ordonnant les polynômes suivant les puissances décroissantes on a obtenu une première division des polynômes.

D'une façon strictement identique, on procède à la division des polynômes suivant les puissances croissantes.

Si
$$A = a_0 + a_1x + \dots + a_nx^n$$
$$B = b_0 + b_1x + \dots + b_px^p$$

avec $b_0 \neq 0$, il existe un couple unique Q et R de polynômes tels que

$$A = BQ + x^{m+1}R \qquad \textbf{avec degré de} \quad Q \leqslant m$$

Q et R étant toujours le quotient et le reste de la division suivant les puissances croissantes de A et B à l'ordre m.

La définition précédente appelle une justification de possibilité.

Posons $Q = c_0 + c_1x + \dots + c_mx^m$ et montrons qu'on peut déterminer $(m + 1)$ coefficients de ce quotient de telle manière que les termes de degré $0, 1, 2, \dots, m$ dans la différence $A - BQ$ soient nuls. Nous obtenons les équations:

$$a_0 - b_0c_0 = 0$$
$$a_1 - b_1c_0 - b_0c_1 = 0$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

$$a_m - b_m c_0 - \dots - b_0 c_m = 0$$

Moyennant la condition $b_0 \neq 0$, on peut calculer de proche en proche les coefficients de $Q(x)$ comme dans le cas de la division suivant les puissances décroissantes.

Sur le plan opératoire la méthode précédemment développée conduit au résultat, les calculs étant arrêtés lorsque le reste contient x^{m+1} en facteur. Les calculs se disposent de la même façon.

Notons qu'en cas de divisibilité le quotient obtenu est le même que celui obtenu suivant les puissances décroissantes.

Exemples.

La division à l'ordre 2 de $1 + x$ par $1 + x^2$ conduit à

$$\begin{array}{r|l}
1 + x \qquad\qquad\qquad & 1 + x^2 \\
\hline
+ x - x^2 \qquad\qquad & 1 + x - x^2 \\
- x^2 - x^3 \qquad & \\
- x^3 + x^4 &
\end{array}$$

et nous avons l'identité

$$1 + x = (1 + x^2)(1 + x - x^2) + x^3(-1 + x)$$

montrant que le quotient et le reste sont

$$Q(x) = 1 + x - x^2$$
$$R(x) = x - 1$$

Ainsi la division de

$$A = 2 + 3x - 3x^3 - 10x^4 + 8x^5$$

par $B = 1 + x - 2x^2$ à l'ordre 3 conduit à

$$\begin{array}{r|l}
2 + 3x \qquad\qquad - 3x^3 - 10x^4 + 8x^5 & 1 + x - 2x^2 \\
\hline
- 2 - 2x + 4x^2 \qquad\qquad\qquad\qquad\qquad & 2 + x + 3x^2 - 4x^3 \\
\hdashline
x + 4x^2 - 3x^3 \qquad\qquad\qquad & \\
- x - x^2 + 2x^3 \qquad\qquad\qquad & \\
\hdashline
3x^2 - x^3 - 10x^4 \qquad & \\
- 3x^2 - 3x^3 + 6x^4 \qquad & \\
\hdashline
- 4x^3 - 4x^4 + 8x^5 & \\
+ 4x^3 + 4x^4 - 8x^5 & \\
\hdashline
0 \qquad\qquad &
\end{array}$$

Ici $R = 0$ A est divisible par B.

10. PGCD DE POLYNÔMES

Notion d'idéal

Nous avons établi la structure d'anneau (§ 3) pour l'ensemble $K(x)$ des polynômes.

Si $\mathscr{I}$ est un sous-ensemble de $K(x)$ vérifiant les propriétés suivantes :

a) $A(x)$ et $B(x) \in \mathscr{I} \Rightarrow A(x) - B(x) \in \mathscr{I}$

b) $A(x) \in \mathscr{I}$ et $P(x) \in K(x) \Rightarrow A(x) \cdot P(x) \in \mathscr{I}$.

$\mathscr{I}$ est appelé idéal de $K(x)$.

Conséquences de cette définition

• Le polynôme nul noté [0] est un idéal. L'anneau $K(x)$ est un idéal.

• Tout idéal de $K(x)$ non vide contient [0] car s'il contient $A(x)$ il contient $A(x) - A(x) = [0]$ d'après *a*).

• Si $A(x), B(x) \in \mathscr{I}$ alors $\forall\, P(x), Q(x) \in K(x)$

$$A \cdot P + B \cdot Q \in \mathscr{I}$$

En effet d'après *b*) il contient $A \cdot P$ et $B \cdot (-Q)$ donc

$$A \cdot P - B \cdot (-Q) = A \cdot P + B \cdot Q$$

Notion de polynôme unitaire ou normalisé

On appelle **polynôme unitaire ou normalisé** un polynôme dont le coefficient du terme de plus haut degré est l'unité.

Tout polynôme non identiquement nul est normalisé si on le divise par le coefficient de son terme de plus haut degré.

Si deux polynômes sont normalisés, leur produit l'est nécessairement (cela découle de la règle donnant le terme de plus haut degré du produit).

Théorème fondamental

Tout idéal $\mathscr{I}$ de $K(x)$ se compose

— ou bien du seul polynôme [0]

— ou bien de tous les multiples d'un polynôme A non nul.

Parmi les polynômes de $\mathscr{I}$, soit A l'un quelconque de degré minimum (*) et B l'un des polynômes de $\mathscr{I}$.

La division de B par A conduit à

$$B = AQ + R \qquad \text{avec} \quad d^0(R) < d^0(A)$$

(*) Ou représentant de la classe normalisée qu'ils constituent.

Or $R \in \mathscr{I}$ puisque $R = B - AQ$ et $B, A \in \mathscr{I}$ et R est nécessairement nul puisque, si tel n'était pas le cas, $\mathscr{I}$ contiendrait un polynôme de degré inférieur à celui de A, ce qui est contraire aux hypothèses d'où

$$B = AQ \qquad \forall B \in \mathscr{I}$$

Idéal engendré par deux polynômes A et B

Tout polynôme D diviseur commun de A et B est diviseur de $AU + BV$, où U, $V \in K(x)$. $\mathscr{I}$ensemble des combinaisons linéaires $AU + BV$ est un idéal, comme on peut le vérifier rapidement.

Si $A \cdot U + B \cdot V \in \mathscr{I}$ et $A \cdot U' + B \cdot V' \in \mathscr{I}$ alors

$(A \cdot U + B \cdot V) - (A \cdot U' + B \cdot V') = (U - U') \cdot A + (V - V') \cdot B$

est une combinaison linéaire de A et B et appartient à $\mathscr{I}$.

Par ailleurs si

$$P \in K(x), \qquad P \cdot (A \cdot U + B \cdot V) = A \cdot (P \cdot U) + B \cdot (P \cdot V)$$

appartient à $\mathscr{I}$comme autre combinaison linéaire de A et B.

D'après le théorème fondamental précédent, si A et B ne sont pas tous deux nuls, $\mathscr{I}$ contient un polynôme $D \neq [0]$ de degré minimum dont tous les autres sont des multiples.

D divise en particulier A et B puisque

$$A = A \cdot 1 + 0 \cdot B \Rightarrow A \in \mathscr{I}$$
$$B = 0 \cdot A + 1 \cdot B \Rightarrow B \in \mathscr{I}$$

Si on pose $D = U_0A + V_0B$ tout diviseur D' de A et B divise D et le degré de D' est au plus égal à celui de D.

On en conclut :

Deux polynômes A et B admettent un plus grand commun diviseur (PGCD), D, défini à un facteur unitaire près. D est une combinaison linéaire de A et B. Tout diviseur commun à A et B divise D et réciproquement.

Polynômes premiers entre eux

Deux polynômes sont premiers entre eux si D est de degré 0 : leur PGCD normalisé est 1.

Théorème de Bezout

Si A et B sont premiers entre eux le polynôme normalisé 1 appartient à l'idéal engendré par A et B. Réciproquement si le polynôme normalisé 1 appartient à l'idéal engendré par A et B, il est nécessairement le polynôme de degré minimal de cet idéal.

Il existe dans $\mathscr{I}$ deux polynômes U_0 et V_0 tels que

$$\boxed{AU_0 + BV_0 = 1}$$

avec $d^0(U_0) < d^0(B)$ et $d^0(V_0) < d^0(A)$.

Propriété des polynômes premiers entre eux

Si $A(x)$ est un polynôme premier avec les polynômes $B(x)$ et $C(x)$ il est premier avec leur produit BC.

En effet, il existe dans $K(x)$ des polynômes U_1, V_1, U_2, V_2 tels que

$$\left.\begin{aligned} AU_1 + BV_1 = 1 \\ AU_2 + CV_2 = 1 \end{aligned}\right\} \Rightarrow A\,(AU_1U_2 + CU_1V_2 + BU_2V_1) + BC(V_1V_2) = 1$$

égalité qui montre que A et BC sont premiers entre eux.

On en déduit par récurrence que si A est premier avec les polynômes $B_1, B_2, \ldots, B_n$ il est premier avec $B_1 \cdot B_2 \ldots B_n$ et donc, quels que soient les entiers $\alpha_1, \alpha_2, \ldots, \alpha_n$, il est premier avec $B_1^{\alpha_1} \cdot B_2^{\alpha_2} \cdot \ldots \cdot B_n^{\alpha_n}$.

Théorème de Gauss

Si A est un polynôme premier avec le polynôme B et s'il divise le produit BC, A divise C.

Donc la condition nécessaire et suffisante pour que les polynômes $A(x), B(x) \in K(x)$ soient premiers entre eux est qu'il existe $U_0(x), V_0(x) \in K(x)$ tels que

$$AU_0 + BV_0 = 1$$

Plus précisément, on démontre que $A(x)$ et $B(x)$ étant donnés, s'ils sont premiers entre eux, il existe une infinité de polynômes $U(x)$ et $V(x)$ satisfaisant $A(x) \cdot U(x) + B(x) \cdot V(x) = 1$ mais un seul de ces couples vérifie simultanément l'identité de Bezout et la condition supplémentaire

$$\begin{cases} d^0\,[U(x)] < d^0\,[B(x)] \\ d^0\,[V(x)] < d^0\,[A(x)] \end{cases}$$

Il existe donc dans $K(x)$ des polynômes $U(x)$, $V(x)$ et $Q(x)$ tels que

$$AU + BV = 1 \qquad \text{et} \qquad BC = AQ$$

on en tire
$$AC \cdot U + BC \cdot V = C$$

d'où

$$A\,(CU + QV) = C$$

qui montre bien que A divise C.

Calcul pratique du PGCD

L'algorithme d'Euclide permettant de trouver le PGCD de deux nombres entiers est celui que nous généralisons ici au cas des polynômes.

Supposons que le degré de B, noté $d^0(B)$ vérifie $d^0(B) \geqslant d^0(A)$ on a

$$\begin{aligned} B &= Aq + A_1 && \text{avec} \quad d^0(A_1) < d^0(A) \\ A &= q_1 A_1 + A_2 && \text{avec} \quad d^0(A_2) < d^0(A_1) \end{aligned}$$

....

$$A_{n-1} = q_n A_n + A_{n+1} \qquad \text{avec} \quad d^0(A_{n+1}) < d^0(A_n)$$

....

Les diviseurs communs à A et B divisent donc la suite des restes A_1, $A_2, \ldots, A_{n+1}$ et réciproquement tout diviseur de deux termes consécutifs de cette suite divise tous les autres.

Deux cas peuvent se présenter :

1) l'un des restes est nul : le PGCD est donc le dernier reste non nul;

2) on arrive à un reste de degré nul, soit une constante, les deux polynômes n'ont pas de PGCD étant premiers entre eux.

11. DIVISION PAR $x - a$

Si un polynôme $P(x)$ est divisible par $x - a$ le reste de la division est $P(a)$.

En effet, la division de $P(x)$ par $x - a$ s'écrit

$$P(x) = Q \cdot (x - a) + R \tag{8}$$

Si on fait $x = a$ dans l'égalité (8) il vient

$$P(a) = R$$

qui est bien une constante puisque le degré de R doit être inférieur à celui de $x - a$, c'est-à-dire de degré 0.

On peut donc dire que pour qu'un polynôme $P(x)$ soit divisible par $x - a$ il faut que $P(a) = 0$. Et réciproquement, si $P(a) = 0$, $P(x)$ est divisible par $x - a$. Ainsi $P(x) = x^m - a^m$ est bien divisible par $x - a$ puisque $P(a) = 0$.

Pour $m = 3$, $x^3 - a^3 = (x - a)(x^2 + ax + a^2)$

Pour $m = 4$, $x^4 - a^4 = (x - a)(x^3 + ax^2 + a^2x + a^3)$

D'une façon générale, on appelle racines d'un polynôme $P(x)$ les valeurs $x_1, x_2, \ldots, x_K$ pour lesquelles $P(x) = 0$.

P(x) se met donc sous la forme

$$P(x) = (x - x_1)(x - x_2) \dots (x - x_k)\, h(x) \tag{9}$$

si $x_1, x_2, \dots, x_k$ sont racines simples de $P(x) = 0$.

12. POLYNÔMES ET ÉQUATIONS SUR LE CORPS $\mathbb{C}$ DES NOMBRES COMPLEXES

Un polynôme non identiquement nul admet dans le corps $\mathbb{C}$ des nombres complexes un nombre de racines distinctes au plus égal à n.

Pour $n = 1$, c'est évident $az + b = 0 \Rightarrow z = -\dfrac{b}{a}$. Admettons que cette propriété reste vraie à l'ordre $p - 1$, et soit $P(z)$ un polynôme de degré p dont z_0 est une racine.

On a donc

$$P(z) = (z - z_0)\, Q(z)$$

et $Q(z)$ est un polynôme de degré $p - 1$. Une racine quelconque de $P(z)$ annule le produit $(z - z_0)\, Q(z)$ et conduit soit à $z = z_0$ soit à $Q(z_0) = 0$. Étant de degré $p - 1$, $Q(z)$ a au plus $p - 1$ racines distinctes, donc $P(z)$ a au plus $1 + (p - 1) = p$ racines distinctes, ce qui établit par récurrence la propriété.

Conséquences

- Si le nombre de racines de $P(z)$ est supérieur au degré de $P(z)$, $P(z)$ est identiquement nul.
- Si $P(z)$ et $Q(z)$ sont égaux pour un nombre de valeurs de z supérieur à leurs degrés, ils sont nécessairement identiques. En effet, leur différence est un polynôme qui s'annulera pour un nombre de valeurs de z supérieur à son degré et donc identiquement nul : les polynômes sont donc égaux.

13. THÉORÈME DE D'ALEMBERT

On admettra le théorème suivant :

Tout polynôme de degré supérieur ou égal à 1 admet au moins une racine dans $\mathbb{C}$.

Ceci conduit à :

Tout polynôme de degré $n \geqslant 1$ admet n racines réelles ou complexes distinctes ou confondues.

Si $P(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n$ avec $a_0 \neq 0$, $P(z)$ admet au moins une racine $z_1 \in \mathbb{C}$ donc $P(z) = (z - z_1)\, P_1(z)$ et il suffit de poursuivre pour $P_1(z)$.

14. FACTORISATION D'UN POLYNÔME

L'écriture de $P(x)$ donnée par (9) constitue la factorisation d'un polynôme et reste valable quelle que soit la nature réelle ou complexe des racines. Si $x_1, x_2, \ldots, x_n$ sont les racines de $P(x)$ on aura

$$P(x) = k\,(x - x_1)\,(x - x_2) \ldots (x - x_n) \tag{10}$$

et la valeur de k s'obtient en comparant (10) à

$$P(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n$$

Les termes de plus haut degré étant identiques $k = a_0$ et

$$P(x) = a_0\,(x - x_1)\,(x - x_2) \ldots (x - x_n) \tag{11}$$

Pour plus de commodité, désignons par x la racine réelle et z la racine complexe. Les coefficients $a_0, a_1, \ldots, a_n$ étant réels et $\bar{z}$ désignant le conjugué de z, on vérifie immédiatement, puisque

$$P(\bar{z}) = a_0 \bar{z}^n + a_1 \bar{z}^{n-1} + \ldots + a_n$$

que $P(\bar{z})$ est le conjugué de $P(z)$

$$P(\bar{z}) = \overline{P(z)}$$

Il suffit pour s'en assurer de poser $z = p + iq$, alors $\bar{z} = p - iq$, et dans le développement de $P(\bar{z})$ comparé à celui de $P(z)$, seuls changent de signe les termes imaginaires purs.

On en conclut que si $P(z)$ admet z comme racine d'ordre β il admet également $\bar{z}$ comme racine d'ordre β.

En regroupant dans (11) les parenthèses identiques et tenant compte de la propriété ci-dessus concernant les racines complexes, on peut écrire

$$P(x) = a_0\,(x - x_1)^{\alpha_1}\,(x - x_2)^{\alpha_2} \ldots (x - x_i)^{\alpha_i}$$
$$\times\ [(x - z_1)(x - \bar{z}_1)]^{\beta_1}\,[(x - z_2)\,(x - \bar{z}_2)]^{\beta_2} \ldots [(x - z_2)\,(x - \bar{z}_2)]^{\beta_j} \tag{12}$$

$x_1, x_2 \ldots x_i$ désignant les racines réelles d'ordre respectif $\alpha_1, \alpha_2 \ldots \alpha_i$.
$z_1, z_2 \ldots z_j$ désignant les racines complexes d'ordre respectif $\beta_1, \beta_2 \ldots \beta_j$.

En écrivant $z = p + iq$, on explicite la partie réelle et la partie imaginaire de la racine complexe z; on peut regrouper les termes contenant z et son conjugué $\bar{z}$

$$\begin{aligned}(x - z)\,(x - \bar{z}) &= (x - p - iq)\,(x - p + iq) \\ &= (x - p)^2 + q^2\end{aligned}$$

et (12) peut se mettre sous la forme

$$P(x) = a_0 (x - x_1)^{\alpha_1} (x - x_2)^{\alpha_2} \dots (x - x_i)^{\alpha_i}$$
$$\times [(x - p_1)^2 + q_1^2]^{\beta_1} \cdot [(x - p_2)^2 + q_2^2]^{\beta_2} \dots [(x - p_j)^2 + q_j^2]^{\beta_j}$$

avec $$\alpha_1 + \alpha_2 + \dots + 2(\beta_1 + \beta_2 \dots) = n$$

(d'après le théorème de D'Alembert) les termes entre crochets étant des trinômes du second degré.

Remarques.

L'expression (12) ci-dessus montre que les racines complexes d'un polynôme à coefficients réels sont toujours en nombre pair.

L'ordre de multiplicité de deux racines conjuguées est le même.

Donc tout polynôme de degré impair a au moins une racine réelle.

15. RELATIONS ENTRE COEFFICIENTS ET RACINES

L'expression (12) ci-dessus permet de mettre $P(x)$ sous la forme d'un produit de binômes normalisés du premier degré, aux racines réelles ou complexes $r_1, r_2, \dots, r_n$ non nécessairement distinctes :

$$P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$$
$$= a_0 (x - r_1)(x - r_2)(x - r_3) \dots (x - r_n) \tag{13}$$

Les fonctions symétriques élémentaires des racines sont obtenues en identifiant les coefficients de même degré dans (13) :

$$S_1 = \sum_{i=1}^{n} r_i = r_1 + r_2 + \dots + r_n = -\frac{a_1}{a_0}$$
$$S_2 = \sum_{1 \leq j_1 < j_2 \leq n} r_{j_1} \cdot r_{j_2} = r_1 r_2 + r_1 r_3 + \dots + r_{n-1} r_n = \frac{a_2}{a_0}$$

. . . .

de même S_p désignant la somme de C_n^p produits p à p des n racines

$$S_p = \sum_{1 \leq j_1 < j_2 \dots < j_k \leq n} r_{j_1} r_{j_2} \dots r_{j_k} = (-1)^p \frac{a_p}{a_0}$$

et le produit des n racines

$$S_n = r_1 r_2 r_3 \dots r_n = (-1)^n \frac{a_n}{a_0}$$

16. FRACTIONS RATIONNELLES

Une fraction rationnelle, ainsi qu'on a appris à les intégrer dans le tome 1 (chap. 9) se présente sous la forme $\dfrac{P(x)}{Q(x)}$, avec degré de $Q(x)$

supérieur à degré de $P(x)$. Si le degré de $P(x)$ est supérieur à celui de $Q(x)$ on peut effectuer la division de $P(x)$ par $Q(x)$ suivant les puissances décroissantes pour faire apparaître une partie entière, qui est un polynôme en x et une nouvelle fraction rationnelle $\frac{P_1}{Q_1}$ où le degré de Q_1 est supérieur à celui de P_1. Nous allons donc supposer être dans un tel cas et étudier la décomposition de $\frac{P(x)}{Q(x)}$ avec degré de $Q(x) >$ degré de $P(x)$, les coefficients de $P(x)$ et $Q(x)$ étant réels.

Les valeurs qui annulent $Q(x)$, ou pôles de la fraction rationnelle, sont réelles ou complexes.

Pôles réels : décomposition en éléments simples de première espèce

Si a est racine de $Q(x)$, $Q(a) = 0$ et $Q(x)$ pourra se mettre sous la forme $Q(x) = (x - a)^m Q_1(x)$, m étant l'ordre de multiplicité de la racine a.

Remplaçons $Q(x)$ par $(x - a)^m Q_1(x)$ dans $\frac{P(x)}{Q(x)}$.

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x - a)^m Q_1(x)} \qquad \text{avec} \qquad Q_1(a) \neq 0$$

$P(a) \neq 0$, sinon $x - a$ serait facteur commun à $P(x)$ et $Q(x)$ ce qui est contraire à l'hypothèse $\frac{P(x)}{Q(x)}$ irréductible.

Partie principale relative à un pôle a d'ordre m

Si a est pôle d'ordre m de la fraction irréductible, et en posant $x - a = X$ ou $x = a + X$ on a

$$\frac{P(x)}{Q(x)} = \frac{P(a + X)}{X^m Q_1 (a + X)}$$

On peut ordonner les polynômes $P(a + X)$ et $Q_1(X)$ suivant les puissances croissantes de X au moyen de la formule de Taylor ou, ce qui est généralement le plus utilisé, par un calcul direct

$$P(a + X) = (A_1 + A_2X + \dots + A_mX^{m-1})\, Q_1 (a + X) + X^mR(X) \tag{14}$$

et en divisant par $X^mQ_1(a + X)$

$$\frac{P(a + X)}{X^m Q_1(a + X)} = \frac{A_1}{X^m} + \frac{A_2}{X^{m-1}} + \dots + \frac{A_m}{X} + \frac{R(X)}{Q_1(a + X)}$$

et en revenant à x

$$\frac{P(x)}{Q(x)} = \frac{A_1}{(x-a)^m} + \frac{A_2}{(x-a)^{m-1}} + \ldots + \frac{A_m}{x-a} + \frac{R(x-a)}{Q_1(x)} \quad (15)$$

$\dfrac{P(x)}{Q(x)}$ se présente sous la forme de la somme :

— d'un polynôme en $\dfrac{1}{x-a}$ de degré m sans terme constant, appelé partie principale relative au pôle a;

— d'une fraction rationnelle qui n'admet plus a comme pôle.

Réciproquement, supposons que $\dfrac{P(x)}{Q(x)}$ soit mise sous la forme (15) ci-dessus; en posant $x - a = X$ et en multipliant par X^m on obtient l'identité (14). Or cette dernière est l'expression de la division d'ordre $m - 1$ suivant les puissances croissantes de $P(a + X)$ par $Q_1(a + X)$, division dont le quotient est unique. Il y a donc unicité de la forme (15).

La détermination des coefficients $A_1, A_2, \ldots, A_m$ peut se faire de proche en proche

$$\frac{P(x)}{(x-a)^m Q_1(x)} = \frac{A_1}{(x-a)^m} + \frac{P_1(x)}{(x-a)^{m-1} Q_1(x)}$$

et en multipliant par $(x-a)^m$

$$\frac{P(x)}{Q_1(x)} = A_1 + \frac{P_1(x)(x-a)}{Q_1(x)}$$

et pour $x = a \quad A_1 = \dfrac{P(a)}{Q_1(a)}$

$$\frac{P_1(x)}{(x-a)^{m-1} Q_1(x)} = \frac{A_2}{(x-a)^{m-1}} + \frac{P_2(x)}{(x-a)^{m-2}}$$

et $A_2 = \dfrac{P_1(a)}{Q_1(a)}$ (obtenu après avoir multiplié par $(x-a)^{m-1}$ et en faisant $x = a$) et ainsi de suite...

De même aux racines simples b et c réelles d'ordre respectif p et q correspondront les suites

$$\frac{B_1}{(x-b)^p}, \frac{B_2}{(x-b)^{p-1}}, \ldots, \frac{B_p}{x-b}$$

ainsi que

$$\frac{C_1}{(x-c)^q}, \frac{C_2}{(x-c)^{q-1}}, \ldots, \frac{C_q}{x-c}$$

rentrant dans la décomposition de $\dfrac{P(x)}{Q(x)}$ et dénommées parties principales de $\dfrac{P}{Q}$ par rapport aux pôles b et c.

La forme la plus générale

$$\frac{P(x)}{Q(x)} = \left(\frac{A_1}{(x-a)^m} + \ldots + \frac{A_m}{x-a}\right) + \left(\frac{B_1}{(x-b)^p} + \ldots + \frac{B_p}{x-b}\right) + \ldots + \left(\frac{L_1}{(x-l)^r} + \ldots + \frac{L_r}{x-l}\right)$$

obtenue pour les différentes racines $a, b, c, \ldots, l$ d'ordre $m, p, q, \ldots, r$ s'appelle décomposition en éléments simples, et elle est unique comme on a pu le vérifier.

Pôles complexes : décomposition en éléments de seconde espèce

Si $Q(x)$ admet des racines complexes, elle sont conjuguées 2 à 2.

Deux racines conjuguées sont racines d'un trinôme du second degré $x^2 + px + q = 0$ avec $p^2 - 4q < 0$, alors $Q(x)$ est décomposable en produit de facteurs de la forme

$$(x^2 + px + q)^m Q_1(x)$$

m étant l'ordre de multiplicité des deux racines considérées, et

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x^2 + px + q)^m Q_1(x)}$$

On peut trouver deux constantes M et N de sorte que

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x^2 + px + q)^m Q_1(x)} = \frac{M_1 x + N_1}{(x_2 + px + q)^m} + \frac{P_1(x)}{Q_1(x)}$$

L'existence de M_1 et N_1 et leur calcul découle d'une simple identification.

Les coefficients M_1 et N_1 s'obtiennent en multipliant les membres de la dernière égalité par $(x^2 + px + q)^m$ et en faisant $x^2 + px + q = 0$ et ainsi de suite pour $\dfrac{P_1}{Q_1}$ égal à la différence

$$\frac{P(x)}{Q(x)} - \frac{M_1 x + N_1}{(x^2 + px + q)^m}$$

Le procédé opératoire reste le même que celui utilisé pour les racines réelles. Ainsi $\dfrac{P(x)}{Q(x)}$ donnera la décomposition suivante

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x^2+px+q)\,Q_1(x)} = \frac{M_1x+N_1}{(x^2+px+q)^m} + \frac{M_2x+N_2}{(x^2+px+q)^{m-1}} + \ldots + \frac{M_mx+N_m}{x^2+px+q} + \frac{P_m(x)}{Q_1(x)}$$

D'une façon générale, si $Q(x)$ admet des racines réelles $a, b, c \ldots$ d'ordre $\alpha, \beta, \gamma \ldots$ et des racines complexes d'ordre $\delta, \lambda \ldots$ on aura

$$Q(x) = A_0(x-a)^\alpha (x-b)^\beta (x-c)^\gamma \ldots (x^2+px+q)^\delta (x^2+p'x+q')^\lambda \ldots$$

et la décomposition de $\dfrac{P(x)}{Q(x)}$ qui est **unique** prendra la forme de la somme la plus générale

$$\frac{P(x)}{Q(x)} = \frac{A_1}{(x-a)^\alpha} + \ldots + \frac{A_\alpha}{x-a} + \frac{B_1}{(x-b)^\beta} + \ldots + \frac{B_\beta}{x-b} + \frac{C_1}{(x-c)^\gamma} + \ldots + \frac{C_\gamma}{x-c} + \frac{M_1x+N_1}{(x^2+px+q)^\delta} + \ldots + \frac{M_\delta x+N_\delta}{x^2+px+q} + \frac{M'_1x+N'_1}{(x^2+p'x+q')^\lambda} + \ldots + \frac{M'_\lambda x+N'_\lambda}{x^2+p'x+q'}$$

Ces résultats ont été utilisés lors de l'intégration des fractions rationnelles. A cette occasion de nombreuses décompositions ont été effectuées à titre d'exercice. On en retrouvera quelques autres à la fin de ce chapitre.

Résumé du chapitre 5

- **Définition.**

  Le polynôme $A = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ est identifié par la suite ordonnée $[a_0, a_1, \ldots, a_n]$ de ses coefficients.

- **Égalité.**

  Deux polynômes A et A' sont égaux si $a_i = a'_i \quad \forall i$

- **Somme.**

$$A + A' = [a_i + a'_i]$$

- **Produit par un scalaire.**

$$\lambda A = [\lambda a_i]$$

• **Produit de deux polynômes.**

$$AA' = [c_i] \qquad \text{où} \qquad c_i = \sum_{k=0}^{i} a_k a'_{i-k}$$

$K(x)$, ensemble des polynômes, a une structure d'anneau commutatif.

• **Division de polynômes.**

Si A est un polynôme de degré n et B un polynôme de degré inférieur à n l'égalité

$$A = BQ + R$$

définit d'une façon unique le quotient Q et le reste R de la division de A par B (avec degré de $R <$ degré de B).

Si $R = 0$ on dit que B divise A.
Les polynômes Q et R sont uniques.

• **Division par $x - a$ d'un polynôme $P(x)$.**

$P(x)$ est divisible par $(x - a) \Leftrightarrow P(a) = 0$.

• **Division suivant les puissances croissantes.**

$$A, B \in K(x) \quad \text{et} \quad B(0) \neq 0$$

$\forall k \in \mathbb{N}$, il existe un et un seul couple de polynômes Q et R, quotient et reste de la division de A par B suivant les puissances croissantes à l'ordre k tel que

$$A = BQ + x^{k+1}R \qquad \text{avec} \qquad d^0(Q) \leqslant k$$

• **Notion d'idéal.**

$\mathcal{I}$ sous-ensemble de $K(x)$ vérifiant

$$A(x), B(x) \in \mathcal{I} \Rightarrow A(x) - B(x) \in \mathcal{I}$$
$$A(x) \in \mathcal{I} \quad \text{et} \quad P(x) \in K(x) \Rightarrow A(x) \,.\, P(x) \in \mathcal{I}$$

$\mathcal{I}$ est appelé idéal.

Théorème 1 : *Idéal engendré par un polynôme.*

Tout idéal $\mathcal{I}$ de l'anneau $K(x)$ est engendré par un polynôme $A(x)$, unique à un facteur multiplicatif près : $\forall B(x) \in K(x)$ B est divisible par $A(x)$.

Théorème 2 : *Idéal engendré par deux polynômes.*

Si $A(x), B(x) \in K(x)$ et ne sont pas tous deux nuls, l'ensemble

$$\{ A(x)\, U(x) + B(x)\, V(x) \}$$

où $U(x)$, $V(x) \in K(x)$ est un idéal de $K(x)$ donc engendré par un polynôme $D(x)$ unique à un facteur multiplicatif près.

Tout diviseur de $A(x)$ et $B(x)$ divise $D(x)$ et réciproquement. $D(x)$ est le PGCD de $A(x)$ et $B(x)$.

Si $D(x)$ est de degré 0, A et B sont premiers entre eux.

- **Identité de Bezout.**

Si $A(x)$ et $B(x)$ sont premiers entre eux, il existe U_0, $V_0 \in K(x)$ tels que

$$AU_0 + BV_0 = 1$$

- **Calcul pratique du PGCD.**

Il se fait grâce à l'algorithme d'Euclide $B = Aq + A_1$ avec $d^0(A_1) < d^0(A)$.

Si $A_1 = 0$ A est le PGCD, sinon on divise A par A_1 et on recommence jusqu'à obtenir un reste nul.

- **Racines d'un polynôme. Décomposition.**

Si $A(x)$ est de degré $n \in \mathbb{N}^*$ il admet au moins une racine sur $\mathbb{C}$. En comptant chaque racine multiple avec son ordre de multiplicité, il possède exactement n racines distinctes ou confondues.

Tout polynôme de $K(x)$ est décomposable en un produit de facteurs uniques à un facteur multiplicatif près.

- **Factorisation d'un polynôme.**

Si $z = p + jq$ est racine de $P(x) = 0$

$\bar{z} = p - jq$ est également racine de $P(x) = 0$,

avec le même ordre de multiplicité que z.

Si x_i est racine réelle d'ordre α_i et si z_i est racine complexe d'ordre β_i, $P(x)$ se met sous la forme :

$$K(x - x_1)^{\alpha_1}(x - x_2)^{\alpha_2} \dots [(x - p_1)^2 + q_1]^{\beta_1} \cdot [(x - p_2)^2 + q_2]^{\beta_2} \dots$$

- **Relation entre les coefficients et les racines.**

Si

$$\begin{aligned} P(x) &= a_0x^n + a_1x^{n-1} + \dots + a_n \\ &= a_0(x - r_1)(x - r_2)\dots(x - r_n) \end{aligned}$$

où $r_1, r_2, \dots, r_n$ sont les n racines de $P(x)$ comptées avec leur ordre de multiplicité, on a

$$S_1 = \sum_{i=1}^{n} r_i = -\frac{a_1}{a_0}$$

$$S_2 = \sum_{i<j} r_i r_j = \frac{a_2}{a_0}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$S_n = r_1 . r_2 . \ldots . r_n = (-1)^n \frac{a_n}{a_0}$$

- **Décomposition d'une fraction rationnelle.**

Si $d^0(Q) > d^0(P)$, et α_i étant l'ordre de multiplicité de la racine réelle x_i, β_i celui des deux racines complexes conjuguées de $x^2 + px + q = 0$ avec $p^2 - 4q < 0$ alors :

$$\frac{P(x)}{Q(x)} = \sum \frac{A_i}{(x - x_i)^{\alpha_i}} + \sum \frac{M_i x + N_i}{[(x^2 + p_i x + q_i]^{\beta_i}}$$

les sommes étant prises pour chacune des racines réelles et complexes.

EXERCICES

1. Décomposer les fractions rationnelles.

a) $\dfrac{1}{x^4 - x}$

b) $\dfrac{1}{x(x^2 + x + 1)^2}$

c) $\dfrac{x^5}{(x^2 - 5x + 6)(x - 1)^2}$

---

a) $\dfrac{1}{x^4 - x} = \dfrac{1}{x(x-1)(x^2 + x + 1)} = \dfrac{A}{x} + \dfrac{B}{x - 1} + \dfrac{Cx + D}{x^2 + x + 1}.$

En multipliant les deux membres par x et en faisant $x = 0$, on obtient $A = -1$.

En opérant de même pour $x - 1$, on obtient $B = \dfrac{1}{3}$ lorsqu'on donne à x la valeur 1.

En multipliant par x et en faisant tendre x vers $+\infty$ on obtient

$$A + B + C = 0 \quad \text{soit} \quad C = \frac{2}{3}$$

Enfin pour $x = -1$, on a

$$\frac{1}{2} = -A - \frac{B}{2} - C + D$$

soit
$$D = \frac{1}{3}$$

et
$$\frac{1}{x^4 - x} = -\frac{1}{x} - \frac{1}{3(x-1)} + \frac{1}{3}\,\frac{2x+1}{x^2+x+1}$$

b) La décomposition de $\dfrac{1}{x(x^2+x+1)^2}$ est de la forme

$$\frac{A}{x} + \frac{Bx+C}{x^2+x+1} + \frac{Dx+E}{(x^2+x+1)^2} = \frac{1}{x(x^2+x+1)^2}$$

La méthode vue dans *a*) donne A immédiatement : $A = 1$; on peut multiplier les deux membres par $(x^2 + x + 1)^2$ et faire $x = j$. Le plus simple reste ici la réduction au même dénominateur et l'identification :

$$A(x^2+x+1)^2 + (Bx+C)\,x(x^2+x+1) + x(Dx+E) = 1$$

On en tire alors

$$\begin{array}{rlll} & A = 1 & & \\ & A + B = 0 & \text{d'où} & B = -1 \\ & -E + C = 0 & \text{d'où} & C = -1 \quad \text{et} \quad E = -1 \\ & B - C = 0 & & \\ \text{enfin} & D - E = 0 & \text{et} & D = E = -1 \end{array}$$

et
$$\frac{1}{x(x^2+x+1)^2} = \frac{1}{x} - \frac{x+1}{x^2+x+1} - \frac{x+1}{(x^2+x+1)^2}$$

c) $x^2 - 5x + 6$ s'annule pour $x = 3$ et $x = 2$ d'où

$$x^2 - 5x + 6 = (x-3)(x-2)$$

On peut par ailleurs effectuer la division de x^5 par

$$(x^2 - 5x + 6)(x-1)^2.$$

La partie entière trouvée est $x + 7$. D'où

$$\frac{x^5}{(x^2 - 5x + 6)(x - 1)^2} = x + 7 + \frac{A}{x-3} + \frac{B}{x-2} + \frac{C}{x-1} + \frac{D}{(x-1)^2}$$

avec

$$A = \frac{3^5}{4} = \frac{213}{4} \qquad B = \frac{2^5}{-1} = -32 \qquad D = \frac{1}{(-2)(-1)} = \frac{1}{2}$$

valeurs calculées suivant la méthode exposée en *a*)

Pour C prenons comme valeur particulière $x = 0$ par exemple, on trouve

$$0 = 7 - \frac{3^4}{4} + \frac{32}{2} - C + \frac{1}{2}$$

d'où $C = -\dfrac{30}{4}$ et

$$\frac{x^5}{(x^2 - 5x + 6)(x - 1)^2} = x + 7 + \frac{213}{4(x-3)} - \frac{32}{x-2} - \frac{30}{4(x-1)} + \frac{1}{2(x-1)^2}$$

■ 2. On considère la famille de polynômes définie par récurrence par

$$P_n(x) = (1 + nx)P_{n-1} + x(1-x)P'_{n-1}(x) \quad \forall n \in \mathbb{N}^*$$

et $\quad P_0(x) = 1$

1) Calculer P_1, P_2, P_3.

2) Quel est le terme de plus haut degré de P_n?

3) Calculer $P_n(1)$ et $P_n(0)$.

4) Montrer que $P_n(x) = x^n P_n\left(\dfrac{1}{x}\right)$. En déduire que les racines de $P_n(x) = 0$ sont deux à deux inverses. Donner une racine commune à tous les P_n pour n impair.

---

1) $P_1(x) = 1 + x$

$P_2(x) = (1 + 2x)(1 + x) + x(1 - x) = x^2 + 4x + 1$

$P_3(x) = (1 + 3x)(x^2 + 4x + 1) + x(1 - x)(2x + 4)$

$\quad = x^3 + 11x^2 + 11x + 1$

2) On voit que le terme de plus haut degré de P_n est x^n pour $0 \leqslant i \leqslant 3$.

Supposons cette propriété vraie à l'ordre $n - 1$, soit que le polynôme P_{n-1} a pour premier terme x^{n-1}, alors de l'expression de P_n

$$P_n = (1 + nx)P_{n-1} + x(1-x)P'_{n-1}(x)$$

on tire que P_n aura pour terme de plus haut degré

$$nx \cdot x^{n-1} - x^2 \cdot (n-1)x^{n-2}$$

soit x^n.

3) $$P_n(1) = (1 + n)\,P_{n-1}(1) = (1 + n)\,n\,P_{n-2}(1)\ldots$$

or $$P_1(1) = 2$$

d'où $$P_n(1) = (n + 1)!$$

$$P_n(0) = P_{n-1}(0) = \ldots = P_2(0) = 1$$

Dans la composition ci-dessus P' ressemble à P.

4) $$P_1(x) = x\,P_1\left(\frac{1}{x}\right)$$

on peut aussi calculer $x^2 P_2\left(\frac{1}{x}\right)$ pour vérifier la propriété également à l'ordre 2.

Supposons que $x^{n-1}P_{n-1}\left(\frac{1}{x}\right) = P_{n-1}(x)$.

On aura alors

$$P'_{n-1}(x) = (n - 1)\,x^{n-2}P_{n-1}\left(\frac{1}{x}\right) - \frac{1}{x^2}\,.\,x^{n-1}P'_{n-1}\left(\frac{1}{x}\right)$$

$$= x^n\left[\left(\frac{1}{x} + n\right)P_{n-1}\left(\frac{1}{x}\right) + \left[\left(\frac{1}{x} - 1\right)(n + 1)\right.\right.$$
$$\left.P_{n-1}\left(\frac{1}{x}\right) - \left(\frac{1}{x} - 1\right)P'_{n-1}\left(\frac{1}{x}\right)\right]$$

$$= x^n\,P_n\left(\frac{1}{x}\right)$$

5) Si x_0 est racine de $P_n(x) = 0$ alors

$$x_0^n\,P_n\left(\frac{1}{x_0}\right) = 0$$

Or x_0 ne peut être égal à 0 puisque $P_n(0) = 1$ donc $\frac{1}{x_0}$ est également racine.

Pour $n = 2\,p + 1$ on voit immédiatement que

$$P_{2p+1}(-1) = (-1)^{2p+1}\,P_{2p+1}(-1)$$

d'où

$$P_{2p+1}(-1) = 0 \qquad \forall\, p$$

et -1 est la racine commune.

■ 3. $P(x)$ étant un polynôme de degré n ayant toutes ses racines réelles distinctes, montrer qu'il en est de même de P'. a étant différent de 0, montrer que $P^2 + a^2$ a toutes ses racines complexes simples.

Soit $x_1, x_2, \ldots, x_n$ les n racines réelles et distinctes de $P(x)$ avec $x_p > x_{p-1}$.

D'après le théorème de Rolle (t. 1, chap. 7) sur chacun des intervalles $]x_{p-1}, x_p[$, $P'(x)$ a au moins une racine. Or ces intervalles étant au nombre de $n-1$ et $P'(x)$ de degré $n-1$ cela implique que chacun des intervalles ne contient qu'une seule des racines de $P'(x)$. Donc $P'(x)$ a bien $(n-1)$ racines réelles et distinctes.

Par ailleurs $P^2 + a^2 \geqslant a^2 > 0$ donc toutes les racines de $P^2 + a^2 = 0$ sont complexes. Or une racine double complexe serait aussi racine de $\dfrac{\mathrm{d}}{\mathrm{d}x}(P^2 + a^2) = 0$ soit de $PP' = 0$ et, d'après ce qui précède, il y a impossibilité, les racines de P et de P' étant réelles.

Donc $P^2(x) + a^2 = 0$ a toutes ses racines simples.

■ **4.** Montrer que si x_0 est racine commune à $A(x)$ et $B(x)$ elle l'est également de leur PGCD et réciproquement.

En déduire les racines multiples de

$$A(x) = x^5 + x^3 - 4x^2 - 3x - 2$$

Si D est le PGCD de A et B, on a $A = DA_1$ et $B = DA_2$.

Si $D(x_0) = 0$ $A(x_0) = D(x_0) \,.\, A_1(x_0) = 0$ et de même $B(x_0) = D(x_0) \,.\, B_1(x_0) = 0$.

Réciproquement si x_0 est racine de $A(x)$ et $B(x)$, d'après le théorème de Bezout il existe deux polynômes U et V tels que

$$D = AU + BV$$

et
$$D(x_0) = A(x_0)\,U(x_0) + B(x_0)\,V(x_0) = 0$$

Si x_0 est racine multiple de $A(x) = x^5 + x^3 - 4x^2 - 3x - 2$ elle est racine de $A'(x) = 5x^4 + 3x^2 - 8x - 3$.

Cherchons le PGCD de A et A' par l'algorithme d'Euclide.

$$A = A' \,.\, \frac{x}{5} - \frac{2}{5}(x^3 + 6x^2 + 6x + 5)$$

$$A' = (x^3 + 6x^2 + 6x + 5)(5x - 30) + 147(x^2 + x + 1)$$

$$x^3 + 6x^2 + 6x + 5 = (x^2 + x + 1)(x + 5)$$

le reste étant nul le PGCD est $x^2 + x + 1$.

Le polynôme A admet comme racines multiples celles de $x^2 + x + 1$ c'est-à-dire

$$j = \frac{-1 + i\sqrt{3}}{2} \quad \text{et} \quad j^2 = \frac{-1 - i\sqrt{3}}{2}$$

Ces racines sont d'ordre au moins égal à 2 donc A est divisible par $(x^2 + x + 1)^2$ qui est de degré 4; A étant de degré 5 il ne peut y en avoir d'autres.

■5. $\alpha, \beta, \gamma, \delta$, étant des entiers naturels on considère les deux polynômes

$$P(x) = x^{4\alpha+3} + x^{4\beta+2} + x^{4\gamma+1} + x^{4\delta}$$

et

$$Q(x) = x^3 + x^2 + x + 1$$

Montrer que P est divisible par Q.

Calculons $P - Q$:

$$P - Q = x^3 (x^{4\alpha} - 1) + x^2 (x^{4\beta} - 1) + x (x^{4\gamma} - 1) + (x^{4\delta} - 1)$$

$x^4 - 1$ divise tout polynôme de la forme $x^{4m} - 1$ car

$$x^{4m} - 1 = (x^4 - 1) [x^{4(m-1)} + x^{4(m-2)} + \ldots + 1]$$

On en conclut que $x^4 - 1$ divise $P - Q$. Or

$$x^4 - 1 = (x - 1)(x^3 + x^2 + x + 1)$$
$$= (x - 1)\ Q(x)$$

Donc $Q(x)$ divise $x^4 - 1$ et, de ce fait, divise $P - Q$, donc divise P.

■6. Montrer que le polynôme

$$P(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}$$

n'admet pas de racines multiples.

Si $P(x)$ admettait une racine multiple x_0 on aurait $P(x_0) = 0$ et de même $P'(x_0) = 0$.

Or $P(x_0) - P'(x_0) = \dfrac{x_0^n}{n!}$ ne s'annule que pour $x_0 = 0$, valeur qui n'annule pas $P(x)$: donc il n'y a pas de racine multiple.

■7. Soit $x^n + \lambda_1 x^{n-1} + \lambda_2 x^{n-2} + \ldots + \lambda_n = 0$ où λ_i est entier $\forall\ i$.

Montrer que si cette équation a une racine rationnelle elle est nécessairement entière et divise λ_n.

En déduire les racines rationnelles de

$$x^4 - 4\,x^3 - 8\,x^2 + 13\,x + 10 = 0$$

Soit $\dfrac{a}{b}$ la racine rationnelle de l'équation

$$\frac{a^n}{b^n} + \lambda_1 \frac{a^{n-1}}{b^{n-1}} + \ldots + \frac{a}{b} \lambda_{n-1} + \lambda_n = 0 \qquad (1)$$

soit

$$\frac{a^n}{b} = -\lambda_1 a^{n-1} - \lambda_2 a^{n-2} b - \dots - \lambda_{n-1} a b^{n-2} - \lambda_n b^{n-1}$$

a^n ainsi que le second membre étant entiers, l'égalité ci-dessus implique $b = \pm 1$; si on porte dans (1) la valeur ci-dessus, on obtient :

$$a^n + \lambda_1 a^{n-1} + \dots + \lambda_{n-1} a = -\lambda_n$$

le premier membre étant divisible par $\pm a$ il en est de même de λ_n.

Les racines rationnelles de l'équation proposée sont les diviseurs de 10, soit au signe près 1, 2 et 5. Une simple vérification directe montre que seuls conviennent -2 et $+5$.

La factorisationn qui s'ensuit est

$$x^4 - 4x^3 - 8x^2 + 13x + 10 = (x + 2)(x - 5)(x^2 - x + 1)$$

et le trinôme du second membre admet pour racines $\dfrac{1 \pm \sqrt{5}}{2}$ qui sont réelles et irrationnelles.

8. Décomposer en éléments simples la fraction rationnelle

$$\frac{x^3 + x^2 + 2}{(x + 1)(x^2 + 1)}$$

On peut procéder par identification à partir de

$$\frac{x^3 + x^2 + 2}{(x + 1)(x^2 + 1)} = E + \frac{A}{x + 1} + \frac{Bx + C}{x^2 + 1}$$

La partie entière est égale à 1.

A s'obtient en multipliant par $x + 1$ et en faisant $x = -1$: on trouve $A = 1$.

De même en multipliant par $x^2 + 1$ et en faisant $x = i$

$$\frac{i^3 + i^2 + 2}{i + 1} = Bi + C \quad \text{d'où} \quad B = -1 \quad \text{et} \quad C = 0$$

et
$$\frac{x^3 + x^2 + 2}{(x + 1)(x^2 + 1)} = 1 + \frac{1}{x + 1} - \frac{x}{x^2 + 1}$$

9. On pose $A_\alpha(x) = x^2 + \alpha \qquad \alpha \in \mathbb{R}$

Montrer qu'il existe :

a) un seul polynôme B_α de degré $n = 4$ tel que

$$A_\alpha \circ B_\alpha = B_\alpha \circ A_\alpha$$

où $A_\alpha \circ B_\alpha = A_\alpha[B_\alpha(x)]$. On mettra $B_\alpha(x)$ sous la forme $H(x) + xK(x)$, H et K étant des polynômes pairs.

b) Montrer que pour deux valeurs différentes de α il existe deux polynômes $C_\alpha(x)$ de degré 3 tels que

$$A_\alpha(x) \circ C_\alpha(x) = C_\alpha(x) \circ A_\alpha(x)$$

---

a) Pour $B_\alpha(x) = H(x) + xK(x)$, H et K polynômes pairs, $H(x)$ ne peut être nul puisque $B_\alpha(x)$ est de degré 4. D'autre part

$$A_\alpha(x) = A_\alpha(-x),$$

d'où

$$B_\alpha[A_\alpha(x)] = B_\alpha[A_\alpha(-x)]$$

L'hypothèse $A_\alpha \circ B_\alpha = B_\alpha \circ A_\alpha$ conduit donc à

$$A_\alpha[B_\alpha(x)] - A_\alpha[B_\alpha(-x)] = 0$$

soit $$[H(x) + xK(x)]^2 + \alpha^2 - [H(x) - xK(x)]^2 - \alpha^2 = 0$$

ou $$H(x) \, . \, K(x) = 0$$

De ce qui précède on conclut

$$K(x) = 0 \qquad \text{et} \qquad B_\alpha(x) = a_0x^4 + a_1x^2 + a_2$$

avec $a_0 \neq 0$.

En identifiant à partir de

$$a_0\,(x^2 + \alpha)^4 + a_1\,(x^2 + \alpha)^2 + a_2 = (a_0x^4 + a_1x^2 + a_2)^2 + \alpha$$

on trouve $$a_0 = 1$$

$$a_1 = 2\,\alpha$$

$$a_2 = \alpha\,(\alpha + 1)$$

et $$B_\alpha(x) = x^4 + 2\,\alpha x^2 + \alpha\,(\alpha + 1)$$

b) Posons $C_\alpha(x) = H(x) + xK(x)$ comme précédemment avec $K \neq 0$ et H et K impairs.

Puisque $C_\alpha(x) = -C_\alpha(-x)$ $H(x)$ est nécessairement nul. Le polynôme cherché est de la forme

$$C_\alpha = a_0x^3 + a_1x \qquad \text{avec} \quad a_0 \neq 0$$

L'identification (analogue à celle de *a*) conduit à $a_0 = 1$ et $a_1 = \dfrac{3\,\alpha}{2}$ avec $\alpha\,(\alpha + 2) = 0$

$$C_\alpha(x) = x^3 + \frac{3\,\alpha}{2}x$$

soit $C_0 = x^3$ et $C_{-2} = x^3 - 3\,x$ auxquels correspondent $A_0 = x^2$ et $A_{-2} = x^2 - 2$.

10. 1) $P(x)$ et $Q(x)$ étant deux polynômes à coefficients réels liés par la relation $(1 - x^2)\,Q^2(x) = 1 - P^2(x)$, trouver les couples P et Q vérifiant la relation précédente lorsque le degré p de $P(x)$ est inférieur à 3.

2) Pour $p > 1$ montrer que les racines de $Q(x)$ sont également racines de $P'(x)$.

3) Montrer que $P'(x) = \pm pQ(x)$.

4) Montrer qu'il existe un intervalle $]\alpha, \beta[\subset]-1, 1[$ dans lequel

$$|P(x)| < 1 \quad \text{et} \quad |P(\alpha)| = |P(\beta)| = 1$$

Pour $p = 0$ on a nécessairement $Q(x) = 0$ d'où $P(x) = \pm 1$.

Pour $p = 1$, posons $P(x) = ax + b$ avec $a \neq 0$.

De la relation liant P et Q on déduit pour le degré de $Q(x)$, $d^0(Q) = d^0(P) - 1$, donc $Q(x) = \text{constante} = c$

$$1 - (ax + b)^2 = c^2(1 - x^2)$$

d'où $\qquad a^2 = c^2 \quad 2ab = 0 \quad \text{et} \quad 1 - b^2 = c^2$

soit $\qquad b = 0 \quad a = \pm 1 \quad c = \pm 1$

Pour $p = 2$ $P(x) = ax^2 + bx + c$ avec $a \neq 0$

$$Q(x) = dx + f$$

L'identification conduit à

$$1 - (ax^2 + bx + c)^2 = (1 - x^2)(dx + f)^2$$

soit encore

$$1 - a^2x^4 - 2abx^3 - x^2(b^2 + 2ac) - 2bcx - c^2 = -d^2x^4 - 2dfx^3 + x^2(d^2 - f^2) + 2dfx + f^2$$

d'où on déduit

$d^2 = a^2$; $ab = df$; $b^2 + 2ac = f^2 - d^2$;

$bc = -df$ et $1 - c^2 = f^2$

On en tire successivement

$$b = f = 0 \quad a = -2 \quad c = \pm 2,$$

soit

$$P(x) = \pm(1 - 2x^2) \quad Q(x) = \pm 2x$$

2) La relation $(1 - x^2)Q^2(x) = 1 - P^2(x)$ donne après dérivation

$$-2xQ^2(x) + 2(1 - x^2)Q(x)Q'(x) = -2P(x)P'(x)$$

Toute racine x_0 de $Q(x)$ conduit à $P(x_0) = 1$ et $P(x_0)P'(x_0) = 0$ d'où $P'(x_0) = 0$.

3) Si $x_1, x_2, \ldots, x_k$ sont les racines distinctes de $Q(x)$ avec les ordres de multiplicité $n_1, n_2, \ldots, n_k$ on a

$$n_1 + n_2 + \ldots + n_k = d^0(Q) = p - 1 \tag{1}$$

x_j est racine d'ordre $n_j - 1$ de $Q'(x)$
donc d'ordre $\geqslant n_j + (n_j - 1) = 2n_j - 1 > 0$ de $Q(x) \cdot Q'(x)$,
donc de P' d'après 2).

On en conclut que

$$(2n_1 - 1) + (2n_2 - 1) + \ldots + (2n_k - 1) \leqslant d^0(P') = p - 1$$

soit encore

$$2(n_1 + n_2 + \ldots + n_k) - k \leqslant p - 1$$

Or

$$k \leqslant n_1 + n_2 + \ldots + n_k = p - 1 \tag{2}$$

et (1) et (2) impliquent $k = p - 1$.

Par conséquent $n_1 = n_2 = \ldots = n_k = 1$: les racines de $Q(x)$ sont donc nécessairement simples et, comme elles sont également racines de $P'(x)$ de même degré que $Q(x)$, on a $P'(x) = aQ(x)$ où a est une constante.

Or si le terme de plus haut degré de $P(x)$ est a_0x^p celui de $Q(x)$ de degré $p - 1$ sera b_0x^{p-1}. L'égalité des termes de plus haut degré qui découle de $1 - P^2(x) = (1 - x^2)\,Q^2(x)$ conduit alors à $a_0^2 = b_0^2$.

La même égalité des termes de plus haut degré déduite de $P' = aQ(x)$ conduit à $pa_0 = ab_0$ soit $a = p\dfrac{a_0}{b_0}$ ou encore $a = p\,.\,(\pm 1)$ et

$$P'(x) = \pm\, p\, Q(x)$$

4) De $(1 - x^2)\,Q(x) = 1 - P^2(x)$ on conclut que

$$1 - P^2(1) = 1 - P^2(-1) = 0.$$

1 et -1 étant les zéros de $1 - P^2(x)$ qui a un nombre de racines fini, on peut donc trouver un intervalle $]\alpha, \beta[\; \subset\;]-1, +1[$ dans lequel $1 - P^2(x) \neq 0$ et $1 - P^2(\alpha) = 1 - P^2(\beta) = 0$; or pour $x \in\,]\alpha, \beta[$ $1 - x^2 > 0$ donc $1 - P^2(x) = (1 - x^2)\,Q^2(x) > 0$ soit $|P(x)| < 1$.

■ **11.** 1) $P(x)$ étant un polynôme vérifiant

$$P(x + 1) + P(x) = 2x^n$$

donner le système permettant d'en calculer les coefficients.

2) En montrant que $Q(x + 1) + Q(x) = 0$ où $Q(x)$ est un polynôme ne peut être satisfaite que par $Q(x)$ identiquement nul, en déduire que si on désigne par $E_k(x)$ le polynôme vérifiant

$$E_k(x + 1) + E_k(x) = 2x^k \qquad k \in \mathbb{N}^*$$

on a

$$E'_n(x + 1) = nE_{n-1}(x) \qquad \text{et} \qquad E_n(x) = (-1)^n E_n(1 - x)$$

---

1) Si p est le degré du polynôme $P(x)$, on a nécessairement

$$2a_px^p = 2x^n$$

d'où $p = n$ et $a_p = 1$

Si $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$

$P(x + 1) = a_n(x + 1)^n + a_{n-1}(x + 1)^n + \ldots + a_1(x + 1) + a_0$

et l'égalité $P(x) + P(x + 1) = 2\,x^n$ conduit à

$$
\begin{aligned}
2\,a_n &= 2\\
C_n^1 a_n + 2\,a_{n-1} &= 0\\
C_n^2 a_n + C_{n-1}^1 a_{n-1} + 2\,a_{n-2} &= 0\\
&\ldots\ldots\ldots\ldots\\
C_n^j a_n + C_{n-1}^{j-1} a_{n-1} + C_{n-2}^{j-2} a_{n-2} + \ldots + 2\,a_{n-j} &= 0\\
&\ldots\ldots\ldots\ldots\\
a_n + a_{n-1} + a_{n-2} + \ldots + a_0 &= 0
\end{aligned}
$$

système qui peut être résolu de proche en proche.

Remarquons que le système linéaire précédent a une solution unique puisque le déterminant associé au système

$$
\begin{vmatrix}
2 & 0 & 0 & \ldots & \\
C_n^1 & 2 & 0 & \ldots & \\
C_n^2 & C_{n-1}^1 & 2 & \ldots & \\
\ldots & \ldots & \ldots & \ldots & \ldots\\
1 & 1 & 1 & \ldots & 2
\end{vmatrix}
$$

a pour valeur évidente 2^{n+1} (il suffit de développer suivant la première ligne), et le système est de Cramer.

2) L'égalité $Q(x) + Q(x + 1) = 0$ conduirait à un système d'équation homogène n'admettant que la solution banale

$$a_n = a_{n-1} = a_{n-2} = \ldots = a_0 = 0$$

on obtient par ailleurs en dérivant

$$E_n'(x + 1) + E_n'(x) = 2\,nx^{n-1}$$

or $$E_{n-1}\,(x + 1) + E_{n-1}\,(x) = 2\,x^{n-1}$$

L'élimination de $2\,x^{n-1}$ conduit à

$$[E_n'(x + 1) - nE_{n-1}(x)] + [E_n'(x) - nE_{n-1}(x)] = 0$$

et en appliquant à $E' - nE_{n-1}$ les résultats trouvés pour $Q(x)$ on a nécessairement

$$E_n'(x) = nE_{n-1}(x)$$

Par ailleurs en changeant x en $-\,x$ dans

$$E_n(x) + E_n(x + 1) = 2\,x^n$$

il vient $$E_n(-\,x) + E_n(1 - x) = (-\,1)^n\,.\,2\,x^n$$

et en éliminant $2x^n$ entre ces deux égalités :

$$[E_n(x) - (-\,1)^n\,E_n(1 - x)] + [E_n(x + 1) - (-\,1)^n\,E_n(-\,x)] = 0$$

qui est de la forme $Q(x) + Q(x+1) = 0$ d'où

$$E_n(x) = -(-1)^n E_n(1-x)$$

■ 12. Décomposer en éléments simples la fraction rationnelle

$$P(x) = \frac{x^3 - x^2 + x}{(1-x^2)(2x-1)}$$

La partie entière s'obtient en divisant suivant les puissances décroissantes, c'est $-\frac{1}{2}$

$$P(x) = -\frac{1}{2} + \frac{A}{1+x} + \frac{B}{1-x} + \frac{C}{2x-1}$$

$$A = \lim_{x \to -1} (1+x)\,P(x) = \frac{1}{2}$$

$$B = \lim_{x \to 1} (1-x)\,P(x) = \frac{1}{2}$$

$$C = \lim_{x \to \frac{1}{2}} (2x-1)\,P(x) = \frac{1}{2}$$

■ 13. Décomposer en éléments simples la fraction rationnelle

$$\frac{1}{x^{10} + x^6}$$

On a l'égalité

$$\frac{1}{x^{10}+x^6} = \frac{1}{x^6(x^4+1)} = \frac{1}{x^6(x^2+1+x\sqrt{2})(x^2+1-x\sqrt{2})}$$

dont la décomposition est de la forme

$$\frac{A_1}{x} + \frac{A_2}{x^2} + \frac{A_3}{x^3} + \frac{A_4}{x^4} + \frac{A_5}{x^5} + \frac{A_6}{x^6} + \frac{Bx+C}{x^2+1+x\sqrt{2}} + \frac{Dx+E}{x^2+1-x\sqrt{2}}$$

$\frac{1}{x^{10}+x^6}$ étant paire, il s'ensuit que $A_1 = A_3 = A_5 = 0$ et de plus $B = -D$ et $C = E = 0$.

Pour obtenir les coefficients restants effectuons la division suivant les puissances croissantes de 1 par $1 + x^4$

$$\frac{1}{1+x^4} = 1 - x^4 + \frac{x^8}{1+x^4}$$

d'où $A_6 = 1$, $A_2 = -1$ et $A_4 = 0$.

L'égalité

$$\frac{x^2}{1+x^4} = \frac{Bx}{x^2+1+x\sqrt{2}} - \frac{Bx}{x^2+1-x\sqrt{2}}$$

permet pour une valeur quelconque de x de tirer B. Par exemple $x = 1$ conduit à $B = -\dfrac{1}{2\sqrt{2}}$.

On en conclut

$$\frac{1}{x^{10}+x^6} = -\frac{1}{x^2} + \frac{1}{x^6} - \frac{\frac{1}{2\sqrt{2}}x}{x^2+1+x\sqrt{2}} + \frac{\frac{1}{2\sqrt{2}}x}{x^2+1-x\sqrt{2}}$$

■ **14.** Décomposer en éléments simples les fractions rationnelles :

$$\frac{x^6}{(x^2+1)^2(x+1)^2}$$ *R.* : $1 - \dfrac{1}{x+1} + \dfrac{1}{4(x+1)^2} + \dfrac{x}{2(x^2+1)^2} - \dfrac{4x+1}{4(x^2+1)}$

$$\frac{x^5}{(x^2-5x+6)(x-1)^2}$$ *R.* : $x + 7 + \dfrac{213}{4(x-3)} - \dfrac{32}{x-2} - \dfrac{30}{4(x-1)} + \dfrac{1}{2(x-1)^2}$

$$\frac{1}{x(x^3-1)}$$ *R.* : $-\dfrac{1}{x} + \dfrac{1}{3(x-1)} + \dfrac{2x+1}{3(x^2+x+1)}$

$$\frac{1-x+x^3}{(x-1)^2(2x-1)}$$ *R.* : $\dfrac{5}{1-2x} + \dfrac{3}{x-1} + \dfrac{1}{(x-1)^3}$

$$\frac{3x^2+2x+1}{(x-1)^2(x+1)^2}$$ *R.* : $\dfrac{7}{4(x-1)^2} + \dfrac{1}{2(x-1)} + \dfrac{1}{2(x+1)^2} - \dfrac{1}{x+1}$

$$\frac{1}{(x^2-1)^4(x^2+1)}$$ *R.* : $-\dfrac{35}{x+1} + \dfrac{12}{(x+1)^2} - \dfrac{4}{(x+1)^3} + \dfrac{1}{(x+1)^4} + \dfrac{1}{16(x^2+1)} + \dfrac{35}{x-1} + \dfrac{12}{(x-1)^2} + \dfrac{4}{(x-1)^3} + \dfrac{1}{(x-1)^4}$

$$\frac{n!}{x(x+1)(x+2)\dots(x+n)} \quad R.: \frac{1}{x} - \frac{x}{x+1} + \frac{n(n-1)}{x+2} + \dots + (-1)^p \frac{C_n^p}{n+p} + \dots + -(1)^n \frac{C_n^n}{x+n}$$

$$\frac{x^2}{x^4 - 2x^2 \cos 2a + 1} \qquad R.: \frac{1}{4\cos a}\left[\frac{x}{x^2 - 2x\cos a + 1} - \frac{x}{x^2 + 2x\cos a + 1}\right]$$

CHAPITRE 6

ESPACES VECTORIELS. APPLICATIONS LINÉAIRES. MATRICES

A. ESPACES VECTORIELS

1. Introduction

Dans les tomes 1, 2, 3, il a été souvent question de vecteurs, au sens de la géométrie élémentaire, c'est-à-dire tels qu'on les rencontre en physique élémentaire. De tels vecteurs dépendent de trois paramètres réels indépendants, le plus souvent leurs coordonnées cartésiennes, ou composantes scalaires, dans un repère constitué de trois vecteurs non coplanaires. Sur ces vecteurs, on a pu définir deux types d'opérations : l'addition et la multiplication par un réel, qui n'exigent pas le choix d'une unité de longueur, ou d'une **norme;** les multiplications, scalaires ou vectorielles, qui impliquent un tel choix.

Enfin, suivant des expressions tombées actuellement en désuétude, les vecteurs apparaissent, à ce niveau, soit comme « libres », c'est-à-dire indépendants de leur origine et définis à une translation près, soit comme « liés », c'est-à-dire d'origines et d'extrémités fixées : la notion de point et d'ensemble de points (courbe, surface) apparaît alors comme essentielle. Toute cette géométrie classique, où un vecteur et un point dépendent de trois paramètres, est « de dimension 3 ».

Dans les chapitres qui suivent, nous allons généraliser ces notions en leur donnant une portée plus abstraite, donc plus vaste.

2. Définition d'un espace vectoriel

On dit qu'un ensemble E est un espace vectoriel (ou possède une structure d'espace vectoriel) sur un corps commutatif K (le plus souvent $\mathbb{R}$ ou $\mathbb{C}$) si on peut définir sur les éléments de E (appelés **vecteurs**) deux opérations, ou lois de composition :

a) *une opération interne, l'addition* (notée $+$), associative, commutative, possédant un élément neutre (noté $\overrightarrow{0_E}$, ou simplement $\vec{0}$), et par rapport à laquelle chaque élément $\vec{u}$ de E possède un symétrique $\vec{u'}$, ou opposé : $\vec{u} + \vec{u'} = \vec{0}$;

b) une opération externe, la multiplication par un élément de K (souvent appelé **scalaire** par opposition aux vecteurs), cette loi externe (produit noté $\alpha\vec{u}$) possédant les propriétés suivantes : $\forall\ \alpha,\ \beta \in K$, $\forall\ \vec{u}, \vec{v} \in E$,

$$\alpha(\vec{u} + \vec{v}) = \alpha\vec{u} + \alpha\vec{v} \quad \text{(distributivité sur } E\text{);}$$
$$(\alpha + \beta)\vec{u} = \alpha\vec{u} + \beta\vec{u} \quad \text{(distributivité sur } K\text{);}$$
$$\alpha(\beta\vec{u}) = (\alpha\beta)\vec{u}$$
$$1 . \vec{u} = \vec{u} \quad \text{(1 étant l'élément unité de } K\text{),}$$

Ces relations constituent les axiomes de la structure d'espace vectoriel. Les opérations (*a*) et (*b*) sont les opérations linéaires sur les vecteurs.

Remarques.

Pour l'addition, E est un groupe commutatif.
Comme

$$[1 + (-1)]\vec{u} = \vec{u} + (-1)\vec{u} = 0 . \vec{u} = \vec{0}$$

l'opposé de $\vec{u}$ est $(-1)\vec{u}$ que nous noterons $-\vec{u}$. D'une façon générale

$$\vec{u} + \vec{v} = \vec{w} \Leftrightarrow \vec{v} = \vec{w} + (-1)\vec{u} = \vec{w} - \vec{u}$$

Si K' est un sous-corps de K, un espace vectoriel sur K est aussi un espace vectoriel sur K' (exemple : $K = \mathbb{C}$, $K' = \mathbb{R}$)

Il importe de bien distinguer 0, élément nul de K ($\mathbb{R}$ ou $\mathbb{C}$ en particulier) et $\vec{0}$, vecteur nul.

3. Exemples

Outre l'ensemble des vecteurs de la géométrie élémentaire, le lecteur pourra vérifier que les ensembles suivants possèdent des structures d'espaces vectoriels sur $\mathbb{R}$ (éventuellement $\mathbb{C}$) :

l'ensemble des polynômes à une variable, de degré inférieur ou égal à n;

l'ensemble des polynômes à une variable, sans limitation de degré;

l'ensemble des solutions d'une équation différentielle linéaire sans second membre;

l'ensemble des fonctions intégrables sur un intervalle $[a, b]$;

l'ensemble des séries numériques convergentes;

l'ensemble des séries entières de rayon de convergence infini;

le corps des complexes.

4. Propriétés immédiates des opérations dans un espace vectoriel

Des axiomes de définition, il résulte :

a) $0 \,.\, \vec{u} = \vec{0}, \quad \forall\, \vec{u} \in E$ en effet $(\alpha + 0)\,\vec{u} = \alpha\vec{u} + 0 \,.\, \vec{u}$,

$$\alpha\vec{u} = \alpha\vec{u} + 0 \,.\, \vec{u}$$

or tout élément d'un groupe est régulier, donc $0 \,.\, \vec{u} = \vec{0}$.

b) De même $\alpha \,.\, \vec{0} = \vec{0}, \quad \forall\, \alpha \in K$.

c) $\alpha\vec{u} = \vec{0} \Rightarrow \alpha = 0$ ou $\vec{u} = \vec{0}$: en effet $\alpha\vec{u} = \vec{0}$ avec $\alpha \neq 0$ implique, α possédant un inverse,

$$\frac{1}{\alpha} \,.\, \alpha\vec{u} = \frac{1}{\alpha}\vec{0} = \vec{0},$$

donc $\vec{u} = \vec{0}$.

5. Familles de vecteurs d'un espace vectoriel. Dépendance, indépendance

p éléments $\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p}$ d'un espace vectoriel E constituent une famille (ou système) de vecteurs de E. Sur cette famille on peut effectuer des **combinaisons linéaires** $\vec{u} = \sum_{i=1}^{i=p} \alpha_i \overrightarrow{u_i}$ en combinant, précisément, les deux opérations linéaires de la définition.

a) Par définition, on dit que les p vecteurs $\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p}$ sont indépendants, ou encore que

la famille $(\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p})$ est libre, si
$\sum_1^p \alpha_i\overrightarrow{u_i} = \vec{0} \Rightarrow \alpha_i = 0, \quad \forall\, i = 1, 2, 3, \ldots, p$

b) Dans le cas contraire, autrement dit **s'il existe p scalaires $\alpha_1, \alpha_2, \ldots, \alpha_p$, non tous nuls, tels que $\sum_1^p \alpha_i \overrightarrow{u_i} = \vec{0}$, on dit que les p vecteurs $\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p}$ sont dépendants, ou encore qu'ils constituent une famille liée.** Une telle relation est une relation de dépendance entre les $\overrightarrow{u_i}$.

Dans ce dernier cas, l'un au moins des α_i, par exemple α_p, n'étant pas nul, on peut écrire $\alpha_p \overrightarrow{u_p} = -\alpha_1 \overrightarrow{u_1} - \alpha_2 \overrightarrow{u_2} - \ldots - \alpha_{p-1} \overrightarrow{u_{p-1}}$, et en posant $\lambda_i = -\dfrac{\alpha_i}{\alpha_p}$ on a

$$\overrightarrow{u_p} = \sum_1^{p-1} \lambda_i \overrightarrow{u_i}$$

un des vecteurs de la famille est combinaison linéaire des $p-1$ autres. La réciproque est immédiate.

De ces définitions résultent aussitôt deux propriétés.

- Si une famille $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p})$ est libre, toute famille extraite de celle-ci (sous-famille ou sous-système) est libre : en effet une relation de dépendance entre q de ces vecteurs serait contradictoire avec l'hypothèse.
- Si une famille $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p})$ est liée, toute famille comportant ces p vecteurs (sur-famille ou sur-système) est liée, puisqu'une relation de dépendance sur $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}$ est aussi une relation de dépendance sur la sur-famille en question.

6. Sous-espaces vectoriels

On appelle sous-espace vectoriel d'un espace vectoriel E, sur un corps K, toute partie de E qui possède la structure d'espace vectoriel sur K.

Soit F un sous-espace de E; étant donné deux vecteurs $\vec{u}$ et $\vec{v}$ de F, toute combinaison $\alpha\vec{u} + \beta\vec{v}$ est un élément de F.

Réciproquement si F est une partie de E telle que

$$\vec{u} \in F, \quad \vec{v} \in F \Rightarrow \alpha\vec{u} + \beta\vec{v} \in F, \quad \forall\, \alpha, \beta \in K,$$

on peut définir sur F les deux opérations (a) et (b) de la définition du paragraphe 2, en prenant $\alpha = \beta = 1$, puis $\beta = 0$ et α quelconque; la première loi est une loi de groupe car $\alpha = 1$, $\beta = -1$ montre que $\forall\, \vec{u}, \vec{v}$ de F, $\vec{u} - \vec{v} \in F$, ce qui est une condition pour que F soit un sous-groupe de E. F est bien un sous-espace vectoriel de E.

Nous retiendrons donc cette caractérisation des sous-espaces :

Pour qu'une partie F d'un espace vectoriel E soit un sous-espace de E, il faut et il suffit que toute combinaison de deux vecteurs de F soit un vecteur de F (ce que l'on exprime en disant que F est stable par combinaison linéaire).

Exemple.

Dans l'espace vectoriel des vecteurs libres de la géométrie élémentaire, les vecteurs parallèles à un plan donné constituent un sous-espace, que l'on appelle **plan vectoriel.**

Le vecteur nul de E est à lui seul un sous-espace de E, et est commun à tous les sous-espaces de E.

7. Intersection de deux sous-espaces

Soit F et G deux sous-espaces d'un espace vectoriel E, et soit $F \cap G = H$. Si $\vec{u}, \vec{v}$ sont deux vecteurs de H, ils appartiennent à la fois à F et G, il en est de même de $\alpha\vec{u} + \beta\vec{v}$, quels que soient les scalaires α,

β: donc $\alpha\vec{u} + \beta\vec{v} \in H$, H est un sous-espace de E, et aussi de chacun des sous-espaces F et G.

Exemple géométrique.

Deux plans vectoriels non confondus se coupent suivant une **droite vectorielle** qui est un sous-espace de l'espace de la géométrie élémentaire.

8. Sommes de sous-espaces. Somme directe. Sous-espaces supplémentaires

a) Si F_1 et F_2 sont deux sous-espaces de l'espace E, l'ensemble F des vecteurs $\vec{u} = \overrightarrow{u_1} + \overrightarrow{u_2}$, où $\overrightarrow{u_1} \in F_1$ et $\overrightarrow{u_2} \in F_2$, est un sous-espace de E, car toute combinaison

$$\alpha(\overrightarrow{u_1} + \overrightarrow{u_2}) + \beta(\overrightarrow{v_1} + \overrightarrow{v_2}) = (\alpha\overrightarrow{u_1} + \beta\overrightarrow{v_1}) + (\alpha\overrightarrow{u_2} + \beta\overrightarrow{v_2}),$$

où $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ sont éléments de F_1 et F_2 respectivement, est élément de F.

Ce sous-espace est appelé somme de F_1 et F_2, on l'écrit $F = F_1 + F_2$. Cette addition s'étend aisément à p sous-espaces de E, elle est commutative et associative, comme l'addition des vecteurs.

On démontre (voir exercice 2 de ce chapitre) que $F_1 + F_2$ est le plus petit sous-espace de E contenant $F_1 \cup F_2$.

b) On dit que la somme $F_1 + F_2$ est directe, ou encore que F_1 et F_2 sont supplémentaires vis-à-vis de F, si la décomposition $\vec{u} = \overrightarrow{u_1} + \overrightarrow{u_2}$ d'un élément quelconque de F en somme de deux éléments de F_1 et F_2 est unique. La somme $F = F_1 + F_2$ se note alors $F_1 \oplus F_2$.

Soit alors $G = F_1 \cap F_2$. Si $\vec{v}$ est un vecteur de G, $\vec{v}$ est aussi un vecteur de F_1, de F_2 et de $F = F_1 + F_2$. Or $\vec{v} = \vec{0} + \vec{v}$, où $\vec{0} \in F_1$ et $\vec{v} \in F_2$; en outre $\vec{v} = \dfrac{\vec{v}}{2} + \dfrac{\vec{v}}{2}$, où $\dfrac{\vec{v}}{2} \in F_1$ et $\dfrac{\vec{v}}{2} \in F_2$.

Si $\vec{v} \neq \vec{0}$, la décomposition $\vec{v} = \overrightarrow{u_1} + \overrightarrow{u_2}$ n'est pas unique, la somme n'est pas directe. Il faut donc, pour $F_1 + F_2 = F_1 \oplus F_2$, que $F_1 \cap F_2 = \{\vec{0}\}$.

Réciproquement, soit $F_1 \cap F_2 = \{\vec{0}\}$. Supposons qu'un vecteur de F s'écrive $\vec{u} = \overrightarrow{u_1} + \overrightarrow{u_2} = \overrightarrow{u'_1} + \overrightarrow{u'_2}$, $\overrightarrow{u_1}$ et $\overrightarrow{u'_1}$ étant des éléments de F_1, $\overrightarrow{u_2}$ et $\overrightarrow{u'_2}$ des éléments de F_2. Alors

$$\overrightarrow{u_1} - \overrightarrow{u'_1} = \overrightarrow{u'_2} - \overrightarrow{u_2} = \vec{v} \in F_1 \cap F_2$$

donc $\vec{v} = \vec{0}$ et la décomposition $\vec{u} = \overrightarrow{u_1} + \overrightarrow{u_2}$ est unique.

En conclusion, si F_1 et F_2 sont deux sous-espaces de E, leur somme est directe si et seulement si leur intersection se réduit au vecteur nul; en d'autres termes

$$\boxed{F_1 + F_2 = F_1 \oplus F_2 \Leftrightarrow F_1 \cap F_2 = \{\vec{0}\}}$$

En particulier, F_1 et F_2 sont deux sous-espaces supplémentaires de E si $F_1 \oplus F_2 = E$. Par exemple, un plan et une droite non parallèles sont supplémentaires dans l'espace de la géométrie élémentaire.

La notion de somme directe s'étend aisément à plusieurs sous-espaces de E.

9. Générateurs d'un espace vectoriel. Base. Coordonnées

a) Soit une famille de p vecteurs $(\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p})$ d'un espace E. L'ensemble des combinaisons linéaires de ces p vecteurs est un sous-espace F de E, comme on le vérifie immédiatement. On dit que F est engendré par les $\overrightarrow{u_i}$ $(i = 1, 2, \ldots, p)$ et que les $\overrightarrow{u_i}$ sont une famille ou un système de générateurs de F. On écrit aussi $F = \text{Vect}\,(\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p})$.

Le lecteur établira sans peine, à titre d'exercice, que F est le plus petit sous-espace de E contenant $\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_p}$.

b) **On appelle base d'un espace vectoriel E une famille libre de générateurs de E** (on dit aussi : une partie libre et génératrice de E).

Par exemple, deux vecteurs indépendants, c'est-à-dire non colinéaires, d'un plan constituent une base de ce plan; les fonctions $y_1 = \sin x$, $y_2 = \cos x$ constituent une base de l'espace vectoriel des solutions de l'équation différentielle $y'' + y = 0$; les polynômes $(1, x, x^2, \ldots, x^n)$, en nombre égal à $n + 1$, constituent une base de l'espace vectoriel des polynômes $P(x)$ de degré $\leqslant n$: en effet $\sum_0^n a_p x^p = 0$, $\forall x$, entraîne que les a_p sont nuls, donc la famille $(1, x, x^2, \ldots, x^n)$ est libre, et elle est évidemment génératrice d'après la définition même d'un polynôme.

Revenons au cas général. Soit B une base de E comportant un nombre fini n de vecteurs (c'est-à-dire de cardinal égal à n); ces vecteurs étant notés $\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_n}$, on sait que

$$\sum_1^n \alpha_i \overrightarrow{e_1} = \vec{0} \Rightarrow \alpha_i = 0, \qquad \forall i = 1, 2, \ldots, n; \qquad (B \text{ est libre})$$

$$\forall \vec{u} \in E, \quad \exists\,(x_1, x_2, \ldots, x_n) \quad \text{tels que } \vec{u} = \sum_1^n x_i \overrightarrow{e_i} \ (B \text{ est génératrice})$$

Si pour ce même vecteur $\vec{u}$, on a $\vec{u} = \sum_1^n x'_i . \overrightarrow{e_i}$, il en résulte

$$\sum_1^n (x_i - x'_i)\, \overrightarrow{e_i} = \vec{0}, \quad \text{donc} \quad x'_i = x_i, \quad \forall i = 1, 2, \ldots, n$$

La décomposition de $\vec{u}$ suivant la base B est donc unique.

Réciproquement, supposons qu'il existe dans E n vecteurs $(\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_n})$ tels que tout vecteur de E puisse s'écrire d'une manière unique sous la forme $\vec{u} = \sum_1^n x_i \, \vec{e_i}$: cette famille est génératrice de E, en particulier le vecteur nul $\vec{0}$ se décompose d'une manière unique. Or on a manifestement la décomposition $\vec{0} = \sum_1^n 0 \, . \, \vec{e_i}$, il en résulte que

$$\sum_1^n \alpha_i \, \vec{e_i} = \vec{0} \Rightarrow \alpha_i = 0$$

donc que la famille $(\vec{e_1}, \ldots, \vec{e_n})$ est libre.

Nous conclurons par le théorème suivant :

Pour qu'une famille de n vecteurs $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ d'un espace vectoriel E soit une base de E, il faut et il suffit que tout vecteur de E admette une décomposition unique suivant cette famille, de la forme $\vec{u} = \sum_1^n x_i \, \vec{e_i}$.

Les scalaires $x_1, x_2, \ldots, x_n$, complètement définis par la donnée du vecteur $\vec{u}$, s'appellent les **coordonnées** (ou composantes scalaires) de $\vec{u}$ suivant la base $B = (\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$.

10. Dimension finie

Dans ce qui précède, nous avons supposé l'existence d'une base d'un espace vectoriel E comportant un nombre fini n de vecteurs. Restons dans ce cas (qui n'est pas toujours réalisé). Nous allons montrer qu'alors toute base de E comporte aussi n vecteurs.

a) Tout d'abord, B étant la base $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$, montrons qu'il existe d'autres bases de E : soit $\vec{u_1}$ un vecteur non nul de E, on a

$$\vec{u_1} = \sum_1^n x_i \, \vec{e_i}$$

avec des x_i non tous nuls; si par exemple $x_1 \neq 0$,

$$\vec{e_1} = \frac{1}{x_1} \, \vec{u_1} - \frac{x_2}{x_1} \, \vec{e_2} - \ldots - \frac{x_n}{x_1} \, \vec{e_n}$$

donc toute combinaison linéaire des $\vec{e_i}$ est aussi combinaison linéaire de $\vec{u_1}, \vec{e_2}, \ldots, \vec{e_n}$ qui forment donc une famille génératrice de E; est-elle libre ? Si on pose

$$\alpha_1 \vec{u_1} + \alpha_2 \vec{e_2} + \ldots + \alpha_n \vec{e_n} = \vec{0},$$

$\alpha_1 \neq 0$ serait contradictoire avec l'hypothèse $x_1 \neq 0$ puisque alors $\vec{u_1}$ ne dépendrait que de $\vec{e_2}, \vec{e_3}, \ldots, \vec{e_n}$; donc il faut $\alpha_1 = 0$; mais alors

$$\sum_2^n \alpha_i \, \vec{e_i} = \vec{0} \Rightarrow \alpha_2 = 0, \alpha_3 = 0, \ldots, \alpha_n = 0$$

puisque $\vec{e_2}, \vec{e_3}, \ldots, \vec{e_n}$ sont indépendants. Donc $(\vec{u_1}, \vec{e_2}, \ldots, \vec{e_n})$ forment une base de E. On peut, dès lors, raisonner par récurrence sur p pour établir que si $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}$ $(p \leqslant n)$ sont p vecteurs indépendants de E, on peut former une base de E contenant ces p vecteurs. Supposons, ce qui est établi pour $p - 1 = 1$, que $\vec{u_1}, \vec{u_2}, \ldots, \overrightarrow{u_{p-1}}$, puissent être complétés par $n - p + 1$ des vecteurs de la base B, par exemple $\vec{e_p}, \overrightarrow{e_{p+1}}, \ldots, \vec{e_n}$ pour former une base de E : on a alors

$$\vec{u_p} = x_1 \vec{u_1} + x_2 \vec{u_2} + \ldots + x_{p-1} \overrightarrow{u_{p-1}} + x_p \vec{e_p} + x_{p+1} \overrightarrow{e_{p+1}} + \ldots + x_n \vec{e_n} \qquad (1)$$

Comme $\vec{u_p}$ n'est pas combinaison linéaire de $\vec{u_1}, \vec{u_2}, \ldots, \overrightarrow{u_{p-1}}$, l'un au moins des $x_p, x_{p+1}, \ldots, x_n$ n'est pas nul, on peut supposer $\vec{e_p}, \ldots, \vec{e_n}$ placés dans un ordre tel que ce soit par exemple x_p : alors $\vec{e_p}$ est combinaison linéaire de $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}, \overrightarrow{e_{p+1}}, \ldots, \vec{e_n}$, et il en est de même de toute combinaison linéaire des $\vec{e_i}$, donc de tout vecteur $\vec{u}$ de E. De plus

$$\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_p \vec{u_p} + \alpha_{p+1} \overrightarrow{e_{p+1}} + \ldots + \alpha_n \vec{e_n} = \vec{0}$$

avec $\alpha_p \neq 0$, serait contradictoire avec (1) puisque $\vec{u_p}$ ne dépendrait pas de $\vec{e_p}$, alors que x_p a été supposé $\neq 0$. Il faut donc $\alpha_p = 0$. Mais alors

$$\alpha_1 \vec{u_1} + \alpha_2 \vec{u_2} + \ldots + \alpha_{p-1} \overrightarrow{u_{p-1}} + \alpha_{p+1} \overrightarrow{e_{p+1}} + \ldots + \alpha_n \vec{e_n} = \vec{0}$$

implique que tous les coefficients sont nuls car

$$\vec{u_1}, \ldots, \overrightarrow{u_{p-1}}, \vec{e_p}, \overrightarrow{e_{p+1}}, \ldots, \vec{e_n}$$

sont indépendants. Donc $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}, \overrightarrow{e_{p+1}}, \ldots, \vec{e_n})$ est une base de E. Il en résulte le **théorème de la base incomplète :**

Si E est un espace vectoriel possédant une base B, et si $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}$ $(p \leqslant n)$ sont p vecteurs de E, indépendants, on peut former une base de E en complétant $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}$ par $n - p$ vecteurs de B, convenablement choisis.

b) En appliquant ce théorème au cas où $p = n$, on conclut aussitôt que tout système libre de n vecteurs constitue une base de l'espace vectoriel E. On démontrera facilement que tout système libre de E comporte au plus n vecteurs, et que toute base de E comporte n vecteurs. D'où la définition de la **dimension** d'un espace vectoriel :

S'il existe un entier naturel n tel que l'espace vectoriel E possède une base constituée de n vecteurs, toute autre base de E est elle aussi constituée de n vecteurs, et cet entier n s'appelle dimension de E. On écrit $n = \dim E$.

Exemple.

L'espace vectoriel des solutions d'une équation différentielle linéaire du second ordre, sans second membre, est de dimension 2. L'espace des vecteurs de la géométrie élémentaire est de dimension 3. L'espace des polynômes $P(x)$ de degré $\leqslant n$ est de dimension $n + 1$.

Par convention, l'espace réduit au vecteur nul, noté $\{\vec{0}\}$, est de dimension 0.

Dans le cas général, il est souvent commode, dans un espace E de dimension n, de privilégier une base dite **canonique** dans laquelle un vecteur $\vec{u}$ est représenté par le n-uplet $(x_1, x_2, \ldots, x_n)$, les vecteurs de la base étant alors

$$(1, 0, 0, \ldots, 0), (0, 1, 0, \ldots, 0), (0, 0, 1, \ldots, 0), \ldots (0, 0, 0, \ldots, 1);$$

on peut même assimiler alors E à $\mathbb{R}^n$, produit cartésien

$$\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$$

du moins lorsque E est un espace vectoriel sur $\mathbb{R}$. C'est pourquoi on parle souvent de « géométrie dans $\mathbb{R}^3$ » pour la géométrie élémentaire.

11. Sous-espaces vectoriels en dimension finie. Rang d'un système de vecteurs

a) Soit F un sous-espace de E, espace vectoriel de dimension n. Un système libre de vecteurs de F est aussi un système de vecteurs de E, donc comporte au maximum n vecteurs, donc $\dim F \leqslant n$. Ainsi

$$\boxed{F \subset E \Rightarrow \dim F \leqslant \dim E}$$

Notons que $F \subset E$ et $\dim F = \dim E \Rightarrow F = E$ (immédiat). En particulier, par généralisation de notions géométriques, un sous-espace de dimension 1 est une droite vectorielle (sa base est un vecteur non nul); un sous-espace de dimension 2 est un plan vectoriel; un sous-espace de dimension $n - 1$ est un hyperplan, il se confond avec un plan si $n = 3$.

Les coordonnées d'un vecteur $\vec{u}$ d'un sous-espace F de dimension p dépendent linéairement de p paramètres qui sont les coordonnées de u dans une base de F. Ainsi, dans un espace E de dimension 5, les trois vecteurs indépendants

$$\overrightarrow{u_1}\,(1, 2, -1, 4, 2), \quad \overrightarrow{u_2}\,(0, 1, 2, -1, 0), \quad \overrightarrow{u_3}\,(-1, 3, 1, 1, -1)$$

engendrent un sous-espace F de dimension 3, et les coordonnées d'un vecteur $\vec{u}$ de F sont de la forme :

$$x_1 = \alpha - \gamma, \quad x_2 = 2\alpha + \beta + 3\gamma, \quad x_3 = -\alpha + 2\beta + \gamma,$$
$$x_4 = 4\alpha - \beta + \gamma, \quad x_5 = 2\alpha - \gamma$$

Un hyperplan, en dimension n, peut être représenté par une équation de la forme $\sum_1^n a_i x_i = 0$ où les a_i ne sont pas tous nuls.

b) **On appelle rang d'un système de p vecteurs $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p})$ de E, avec dim $E = n$, la dimension r du sous-espace Vect $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p})$.** En d'autres termes, r est le nombre maximum de vecteurs que peut comporter un système libre extrait du système donné. Il va de soi que $r \leqslant n$ et $r \leqslant p$, donc $r \leqslant \inf(p, n)$.

c) Si F et G sont deux sous-espaces de E, de dimension finie, soit $H = F \cap G$, $\dim F = p$, $\dim G = q$, $\dim H = s$. Cherchons quelle est la dimension de $F + G$. Une base de H, soit $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_s})$ peut être, d'après le théorème de la base incomplète, complétée dans F par $p - s$ vecteurs indépendants $\vec{f_1}, \vec{f_2}, \ldots, \overrightarrow{f_{p-s}}$, pour donner une base de F, et complétée dans G par $q - s$ vecteurs indépendants $\vec{g_1}, \vec{g_2}, \ldots, \overrightarrow{g_{q-s}}$, pour donner une base de G. La forme générale d'un vecteur de $F + G$ est

$$\begin{aligned}\vec{u} = x_1\vec{f_1} + x_2\vec{f_2} + \ldots + x_{p-s}\overrightarrow{f_{p-s}} + x_{p-s+1}\vec{e_1} + \ldots + x_p\vec{e_s} \\ + y_1\vec{g_1} + y_2\vec{g_2} + \ldots + y_{q-s}\overrightarrow{g_{q-s}} + y_{q-s+1}\vec{e_1} + \ldots + y_q\vec{e_s}\end{aligned}$$

soit en posant $x_{p-s+i} + y_{q-s+i} = z_i$,

$$\begin{aligned}\vec{u} = x_1\vec{f_1} + \ldots + x_{p-s}\overrightarrow{f_{p-s}} + y_1\vec{g_1} \\ + \ldots + y_{q-s}\overrightarrow{g_{q-s}} + z_1\vec{e_1} + \ldots + z_s\vec{e_s}\end{aligned}$$

Ces $(p - s) + (q - s) + s = p + q - s$ vecteurs $\vec{f_1}, \ldots, \vec{e_s}$ sont manifestement indépendants, et $F + G$, engendré par ces vecteurs, a pour dimension $p + q - s$. D'où la formule

$$\boxed{\dim(F + G) = \dim F + \dim G - \dim(F \cap G)}$$

Dans le cas d'une somme directe, $F \cap G = \{\vec{0}\}$, donc

$$\boxed{\dim(F \oplus G) = \dim F + \dim G}$$

Ce résultat s'étend facilement à plusieurs sous-espaces sous la forme

$$\dim \bigoplus_{i=1}^{i=p} F_i = \sum_{i=1}^{i=p} \dim F_i$$

Enfin si F et G sont supplémentaires vis-à-vis de E, c'est-à-dire si $E = F \oplus G$, on a $\dim F + \dim G = n$. F étant donné de dimension p, on peut adjoindre à une base de F $n - p$ vecteurs de E pour obtenir une base de E, ces $n - p$ vecteurs ne sont donc pas dans F puisqu'ils sont indépendants des vecteurs de base de F, et ils engendrent un sous-espace G qui est un supplémentaire de F. On voit que tout sous-espace de E admet, en dimension finie, au moins un supplémentaire dans E. Notons que E admet l'unique supplémentaire $\{\vec{0}\}$.

Rappelons qu'il ne faut pas confondre les supplémentaires de F dans E, qui sont en nombre infini, en général, avec le sous-ensemble complémentaire de F dans E, défini comme l'ensemble des vecteurs $(\vec{u} \in E, \vec{u} \notin F)$, et qui n'est pas un espace vectoriel. Un exemple géométrique illustre ce qui précède : dans $\mathbb{R}^3$, un plan vectoriel P (dimension $p = 2$) a pour supplémentaire une droite vectorielle quelconque Δ non incluse dans P (dimension $q = 1$), et pour complémentaire l'ensemble des vecteurs non situés dans P, qui n'est autre que l'union de tous les supplémentaires de P.

12. Espaces de dimension infinie

Nous venons de voir sous diverses formes que la dimension n, si elle existe, d'un espace vectoriel E, est le nombre maximum de vecteurs que peut renfermer un système libre extrait de E. Un tel maximum n'existe pas forcément : si le nombre des éléments d'un système libre de E n'est pas majoré, on dit que E est de dimension infinie. Une base de E est alors une suite infinie d'éléments de E, indépendants. L'exemple le plus simple est celui de l'espace vectoriel des polynômes $P(x)$, sans limitation de degré, dont une base est évidemment le système $(1, x, x^2, \ldots, x^n, \ldots)$, souvent appelé **base canonique** de cet espace.

Dans les chapitres qui vont suivre, nous envisagerons presque exclusivement des espaces de dimensions finies.

B. APPLICATIONS LINÉAIRES. MATRICES

1. Définition

E et F étant deux espaces vectoriels sur un même corps K, on appelle application linéaire de E dans F une application φ de E dans F telle que

$$\forall \vec{u}, \vec{v} \in E, \quad \varphi(\vec{u} + \vec{v}) = \varphi(\vec{u}) + \varphi(\vec{v}) \tag{2}$$

$$\forall \vec{u} \in E, \quad \forall \lambda \in K, \quad \varphi(\lambda\vec{u}) = \lambda\varphi(\vec{u}) \tag{3}$$

Ces deux axiomes de **linéarité** peuvent d'ailleurs être réunis en un seul

$$\boxed{\varphi(\lambda\vec{u} + \mu\vec{v}) = \lambda\varphi(\vec{u}) + \mu\varphi(\vec{v})} \tag{4}$$

Notons que $\varphi(\vec{u})$, $\varphi(\vec{v})$ sont des vecteurs de F. Ainsi la linéarité implique $\varphi(\vec{0_E}) = \vec{0_F}$. Bien qu'il y ait lieu, théoriquement, de distinguer le vecteur nul de E et celui de F, il n'y a pas d'inconvénient pratique à noter $\vec{0}$ tous les vecteurs nuls. K est le plus souvent $\mathbb{R}$, parfois $\mathbb{C}$. Les calculs auxquels conduit cette notion, bien que souvent intéressants seulement dans $\mathbb{R}$, gardent un sens dans $\mathbb{C}$ dans la plupart des cas.

Une application linéaire respecte la combinaison linéaire de plusieurs vecteurs, c'est-à-dire

$$\varphi\left(\sum_{k=1}^{k=p} \alpha^k \overrightarrow{u_k}\right) = \sum_1^p a_k\, \varphi(\overrightarrow{u_k})$$

d'après (2) et (3).

Exemples.

En géométrie vectorielle de dimension 2 ou 3, les symétries, rotations, homothéties, projections sont des applications linéaires de E dans E ($E = \mathbb{R}^2$ ou $\mathbb{R}^3$); en analyse, la dérivation, l'intégration, les transformations intégrales telles que celles de Fourier et Laplace sont des applications linéaires, dont on précisera facilement l'espace de départ E et l'espace d'arrivée F.

2. Noyau. Image

Soit encore φ une application linéaire de E dans F, espaces vectoriels donnés.

a) On appelle noyau de φ l'ensemble des vecteurs $\vec{u}$ de E dont l'image par φ est le vecteur nul de F. Ce noyau se note ker φ (parfois $N(\varphi)$), notation provenant du mot allemand kern (noyau). Ainsi

$$\boxed{\vec{u} \in \ker \varphi \Leftrightarrow \varphi(\vec{u}) = \vec{0}}$$

Comme

$$\varphi(\vec{u}) = \vec{0} \quad \text{et} \quad \varphi(\vec{v}) = \vec{0} \Rightarrow \varphi\,(\lambda\vec{u} + \mu\vec{v}) = \vec{0}$$

ker φ est un sous-espace vectoriel de E, et $\forall\ \varphi, \vec{0} \in \ker \varphi$.

Exemples.

Dans la projection φ sur un plan P parallèlement à une droite Δ, le noyau est l'ensemble des vecteurs parallèles à Δ, c'est-à-dire la droite vectorielle Δ.

Dans l'homothétie de rapport $k \neq 0$, $\varphi(\vec{u}) = k\vec{u} \Leftrightarrow \vec{u} = \vec{0}$, le noyau est réduit au vecteur nul.

Dans l'application linéaire $y(x) \to (x^2 + 1)\, y' + xy$, ker φ est l'espace vectoriel des solutions de l'équation différentielle

$$(x^2 + 1)\, y' + xy = 0$$

c'est-à-dire l'ensemble des fonctions $y = \lambda\,(x^2 + 1)^{-\frac{1}{2}}$ (dimension 1).

b) On appelle image de φ, et on note Im φ, l'ensemble des vecteurs $\varphi(\vec{u})$ tels que $\vec{u} \in E$. C'est une partie de F. Ainsi

$$\boxed{\vec{v} \in \operatorname{Im} \varphi \Leftrightarrow \exists\ \vec{u} \in E \quad \text{tel que} \quad \varphi(\vec{u}) = (\vec{v})}$$

Si $\overrightarrow{v_1}$ et $\overrightarrow{v_2} \in \operatorname{Im} \varphi$, $\exists \overrightarrow{u_1}$ et $\overrightarrow{u_2}$ dans E tels que

$$\varphi(\overrightarrow{u_1}) = \overrightarrow{v_1}, \ \varphi(\overrightarrow{u_2}) = \overrightarrow{v_2},$$

alors $\lambda\overrightarrow{v_1} + \mu\overrightarrow{v_2} = \varphi(\lambda\overrightarrow{u_1} + \mu\overrightarrow{u_2})$, donc $\lambda\overrightarrow{v_1} + \mu\overrightarrow{v_2} \in \operatorname{Im} \varphi$, on voit que Im φ est un sous-espace de F.

On utilise aussi pour Im φ la notation $\varphi(E)$. Plus généralement, si E' est un sous-espace de E, son image par φ est un sous-espace de Im φ, noté $\varphi(E')$.

Pour le noyau, on utilise parfois la notation $\varphi^{-1}(\vec{0})$, qui peut prêter à confusion puisque l'application φ^{-1} n'existe que si φ est bijective, cas que nous examinerons plus loin.

Dans l'exemple, cité dans (*a*), de la projection sur un plan P parallèlement à une droite Δ, l'image est évidemment le plan P.

3. Application linéaire injective

Rappelons qu'une application, linéaire ou non, est injective si $\varphi(\vec{u}) = \varphi(\vec{v}) \Rightarrow \vec{u} = \vec{v}$.

Si φ est linéaire et injective, $\varphi(\vec{u}) = \vec{0} = \varphi(\vec{0}) \Rightarrow \vec{u} = \vec{0}$, donc

$$\ker \varphi = \{\vec{0}\}$$

Réciproquement, si

$$\ker \varphi = \{\vec{0}\}, \quad \varphi(\vec{u}) = \varphi(\vec{v}) \Rightarrow \varphi(\vec{u} - \vec{v}) = \vec{0}$$
$$\Rightarrow \vec{u} - \vec{v} \in \ker \varphi \Rightarrow \vec{u} - \vec{v} = \vec{0} \Rightarrow \vec{u} = \vec{v}$$

Ainsi **pour qu'une application linéaire φ de E dans F soit injective, il faut et il suffit que ker $\varphi = \{\vec{0}\}$.**

4. Image d'une famille génératrice. Cas où E est de dimension finie

a) Soit $S = (\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_k})$ un système de vecteurs de E, engendrant le sous-espace $E' = \operatorname{Vect}(\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_k})$, d'image $F' = \varphi(E')$ dans F. Tout vecteur de F' est de la forme

$$\varphi\left(\sum_{i=1}^{k} x_i \overrightarrow{u_i}\right) = \sum_{i=1}^{k} x_i \varphi(\overrightarrow{u_i})$$

donc est une combinaison linéaire des vecteurs de $\varphi(S)$. $\varphi(S)$ est donc une famille génératrice de $F' = \varphi(E')$.

L'image d'une famille de générateurs d'un sous-espace de E est une famille de générateurs du sous-espace image dans F.

En particulier si φ est injective et si S est un système libre

$$\sum_1^k x_i \varphi(\overrightarrow{u_i}) = \vec{0} \Rightarrow \varphi\left(\sum_1^k x_i \overrightarrow{u_i}\right) = \vec{0} \Rightarrow \sum_1^k x_i \overrightarrow{u_i} = \vec{0} \Rightarrow x_i = 0$$

puisque $\ker \varphi = \{\vec{0}\}$.

$\varphi(S)$ est donc un système libre.

Donc si φ est injective, l'image d'une base de E' (en particulier d'une base de E, si E est de dimension finie) est une base de $F' = \varphi(E')$ (en particulier de $\operatorname{Im} \varphi$). On a $\dim \varphi(E') = \dim E'$.

Dans le cas général où φ n'est pas nécessairement injective, on a évidemment

$$\boxed{\dim \varphi(E') \leqslant \dim E'}$$

ainsi une application linéaire ne peut accroître la dimension.

b) Supposons $\dim E = n$, alors $\dim \operatorname{Im} \varphi = p \leqslant n$. Comme en outre $\ker \varphi \subset E$, $\dim \ker \varphi = q \leqslant n$.

Nous allons établir une relation entre n, p et q. Pour cela choisissons dans $\ker \varphi$ une base $B = (\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_q})$, dans $\operatorname{Im} \varphi$ une base $C = (\overrightarrow{f_1}, \overrightarrow{f_2}, \ldots, \overrightarrow{f_p})$.

Comme $\vec{v} \in \operatorname{Im} \varphi \Leftrightarrow \exists \vec{u} \in E$ tel que $\varphi(\vec{u}) = \vec{v}$, on peut définir des vecteurs $\overrightarrow{e_{q+1}}, \overrightarrow{e_{q+2}}, \ldots, \overrightarrow{e_{q+p}}$ de E par $\varphi(\overrightarrow{e_{q+i}}) = \overrightarrow{f_i}$, $i = 1, 2, \ldots, p$. Soit $B' = (\overrightarrow{e_{q+1}}, \overrightarrow{e_{q+2}}, \ldots, \overrightarrow{e_{q+p}})$.

Montrons que $B \cup B'$ engendre E. Soit $\vec{u} \in E$, alors $\varphi(\vec{u}) \in \operatorname{Im} \varphi$ et $\varphi(\vec{u}) = \alpha_1 \overrightarrow{f_1} + \alpha_2 \overrightarrow{f_2} + \ldots + \alpha_p \overrightarrow{f_p}$, cette décomposition étant unique; donc

$$\varphi(\vec{u}) = \varphi\,(\alpha_1 \overrightarrow{e_{q+1}} + \alpha_2 \overrightarrow{e_{q+2}} + \ldots + \alpha_p \overrightarrow{e_{q+p}})$$

et

$$\vec{u} = \alpha_1 \overrightarrow{e_{q+1}} + \alpha_2 \overrightarrow{e_{q+2}} + \ldots + \alpha_p \overrightarrow{e_{q+p}} + \vec{z}$$

avec $\vec{z} \in \ker \varphi$.

$\vec{z}$ admet une décomposition unique dans B

$$\vec{z} = \beta_1 \overrightarrow{e_1} + \beta_2 \overrightarrow{e_2} + \ldots + \beta_q \overrightarrow{e_q}$$

de sorte que

$$\vec{u} = \beta_1 \overrightarrow{e_1} + \ldots + \beta_q \overrightarrow{e_q} + \alpha_1 \overrightarrow{e_{q+1}} + \ldots + \alpha_p \overrightarrow{e_{q+p}}$$

$B \cup B'$ est donc un système de générateurs de E, et $n \leqslant p + q$. Voyons si $B \cup B'$ est une base de E, c'est-à-dire si les $\overrightarrow{e_i}$ $(1 \leqslant i \leqslant p + q)$ sont indépendants : posons

$$\vec{u} = \lambda_1 \overrightarrow{e_1} + \lambda_2 \overrightarrow{e_2} + \ldots + \lambda_q \overrightarrow{e_q} + \mu_1 \overrightarrow{e_{q+1}} + \ldots + \mu_p \overrightarrow{e_{q+p}} = \vec{0}$$

Comme les $\overrightarrow{e_i}$ où $1 \leqslant i \leqslant q$ sont des vecteurs du noyau

$$\varphi(\vec{u}) = \mu_1 \overrightarrow{e_{q+1}} + \ldots + \mu_p \overrightarrow{e_{q+p}} = \vec{0} \quad \text{puisque} \quad \varphi(\overrightarrow{e_{q+i}}) = \overrightarrow{f_i}$$

Les $\overrightarrow{f_i}$ étant indépendants, les μ_i sont nuls, et

$$\vec{u} = \vec{0} = \lambda_1 \overrightarrow{e_1} + \lambda_2 \overrightarrow{e_2} + \ldots + \lambda_q \overrightarrow{e_q}$$

mais les $\overrightarrow{e_j}$ sont aussi indépendants $(1 \leqslant j \leqslant q)$, dont les λ_j sont nuls.

$B \cup B'$ est donc une base de E, et $n = p + q$.

On en déduit le théorème fondamental :

Si E et F sont deux espaces vectoriels sur le même corps K, et si E est de dimension finie, on a, pour toute application linéaire de E dans F,

$$\boxed{\dim \ker \varphi + \dim \operatorname{Im} \varphi = \dim E} \qquad (5)$$

c) Définition : **on appelle rang d'une application linéaire φ de E dans F la dimension de Im φ,** lorsque cette dimension est finie. C'est le cas, en particulier, lorsque E est de dimension finie n, et on a rang $(\varphi) \leqslant n$.

D'après la relation (5), rang $(\varphi) = n \Leftrightarrow \ker \varphi = \{0\} \Leftrightarrow \varphi$ injective.

5. Cas où E et F sont de dimensions finies. Matrices associées à une application linéaire

Soit $\dim E = n$, $\dim F = p$. Choisissons dans E une base $\mathscr{B}_E$, dans F une base $\mathscr{B}_F$ définies par

$$\mathscr{B}_E = (\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}),$$
$$\mathscr{B}_F = (\vec{f_1}, \vec{f_2}, \ldots, \vec{f_p}).$$

Soit φ une application linéaire de E dans F, elle est complètement définie si on se donne les images dans F des vecteurs de $\mathscr{B}_E$, $\varphi(\vec{e_1}), \varphi(\vec{e_2}), \ldots, \varphi(\vec{e_p})$: en effet $\vec{u} = \sum_1^n x_j\vec{e_j}$ étant un vecteur quelconque de E, de coordonnées $x_1, \ldots, x_n$, son image est $\varphi(\vec{u}) = \sum_1^n x_j\, \varphi(\vec{e_j})$ donc bien déterminée.

Pour connaître dans la base $\mathscr{B}_E$ de F les coordonnées $y_1, y_2, \ldots, y_p$ de $\varphi(\vec{u})$, il suffit de se donner celles des $\varphi(\vec{e_j})$. Ces coordonnées étant au nombre de p, il est nécessaire d'introduire ici une notation à double indice : nous adopterons la plus usitée (mais non la seule) :

$$\varphi(\vec{e_1}) = a_{11}\vec{f_1} + a_{21}\vec{f_2} + \ldots + a_{p1}\vec{f_p}$$

d'une façon générale

$$\varphi(\vec{e_j}) = a_{1j}\vec{f_1} + a_{2j}\vec{f_2} + \ldots + a_{ij}\vec{f_i} + \ldots + a_{pj}\vec{f_p}$$

ou en abrégé

$$\varphi(\vec{e_j}) = \sum_{i=1}^{i=p} a_{ij}\vec{f_i}$$

Ainsi a_{ij} désigne la $i^{\text{ième}}$ coordonnée de $\varphi(\vec{e_j})$, dans la base $(\vec{f_1}, \vec{f_2}, \ldots, \vec{f_p})$.

L'application de la linéarité donne alors

$$\varphi(\vec{u}) = \sum_{j=1}^{j=n} \left(x_j \sum_{i=1}^{i=p} a_{ij}\vec{f_i} \right)$$

de sorte que la coordonnée y_i de $\varphi(\vec{u})$, c'est-à-dire sa composante scalaire suivant $\vec{f_i}$, est

$$y_i = \sum_{j=1}^{j=n} a_{ij}x_j$$

En détaillant cette écriture

$$\varphi(\vec{u}) \begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dotfill \\ y_p = a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pn}x_n \end{cases}$$

On condense ces relations en une seule par l'écriture matricielle. En effet on constate que, dans les bases choisies $\mathscr{B}_E$ et $\mathscr{B}_F$, l'application φ est parfaitement connue si on se donne le tableau rectangulaire, appelé matrice de φ dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \cdots & a_{pj} & \cdots & a_{pn} \\ \varphi(\vec{e_1}) & \varphi(\vec{e_2}) & & \varphi(\vec{e_j}) & & \varphi(\vec{e_n}) \end{pmatrix}$$

Cette matrice comporte p lignes d'indices $1, 2, \dots, i, \dots, p$, (rangées horizontales), n colonnes d'indices $1, 2, \dots, j, \dots, n$ (rangées verticales). La colonne, ou « vecteur-colonne », de rang j est constituée par les coordonnées dans la base $\mathscr{B}_F$ de F de l'image $\varphi(\vec{e_j})$ du $j^{\text{ième}}$ vecteur de la base $\mathscr{B}_E$ de E. La ligne, ou « vecteur-ligne », de rang i est constituée par les coefficients à affecter aux coordonnées de $\vec{u}$ $(x_1, x_2, \dots, x_n)$ pour obtenir la $i^{\text{ième}}$ coordonnée de $\varphi(\vec{u})$. a_{ij} est l'élément général, ou terme général, de la matrice A, il est à l'intersection de la ligne i et de la colonne j ; i est son indice de ligne, j son indice de colonne.

Il est alors logique de représenter, dans la base $\mathscr{B}_E$ de E, le vecteur $\vec{u}$ par une matrice unicolonne, ou matrice-colonne, et d'opérer de même pour $\varphi(\vec{u}) = \vec{v}$ dans la base $\mathscr{B}_F$ de F, et ainsi d'écrire

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{p1} & a_{p2} \dots a_{pn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

On dit que la matrice $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}$ est le produit par A de la matrice $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et on écrit

$$Y = AX \tag{6}$$

(6) est l'écriture matricielle de la relation $\vec{v} = \varphi(\vec{u})$ dans les bases $\mathscr{B}_E$ et $\mathscr{B}_F$. Il va de soi que si l'on change de base dans E, ou dans F, ou dans les deux espaces, cette écriture matricielle est modifiée. Cependant la matrice d'une application linéaire d'un espace de dimension n dans un espace de dimension p sera toujours une matrice à p lignes et n colonnes, en abrégé une matrice (p, n).

Enfin, nous avons introduit la notion de matrice comme représentant une application linéaire, mais il est clair que la notation matricielle peut être utilisée dans bien d'autres circonstances : pour caractériser dans une base un système de vecteurs d'un espace de dimension finie, pour écrire un système de relations ou d'équations linéaires, etc.

Exemples.

a) Cherchons, dans $\mathbb{R}^2$, la matrice représentant, dans le repère orthonormé $(\vec{i}, \vec{j})$, la rotation d'angle θ : il suffit de remarquer que

$$\varphi(\vec{i}) = \vec{i} \cos\theta + \vec{j} \sin\theta$$

$$\varphi(\vec{j}) = \vec{i} \cos\left(\theta + \frac{\pi}{2}\right) + \vec{j} \sin\left(\theta + \frac{\pi}{2}\right) = -\vec{i} \sin\theta + \vec{j} \cos\theta$$

Cette matrice est donc $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$.

La rotation pour un vecteur (x, y) s'écrit

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

soit $x_1 = x\cos\theta - y\sin\theta$, $y_1 = x\sin\theta + y\cos\theta$, ce que l'on peut retrouver en écrivant $x_1 + iy_1 = (x + iy)(\cos\theta + i\sin\theta)$.

b) Le système d'équations du premier degré

$$\begin{cases} 2x + 3y - z = 1 \\ 3x + 4y + 2z = 4 \end{cases}$$

s'écrit sous forme matricielle

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

c) Les formules de dérivation d'une fonction composée

$$f\,[u(x, y), v(x, y)]$$

(t. 1, chap. 11) s'écrivent sous forme matricielle

$$\begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial u}{\partial y} & \dfrac{\partial v}{\partial y} \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial u} \\ \dfrac{\partial f}{\partial v} \end{pmatrix}$$

ce qui donne naissance à la matrice jacobienne

$$\begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial u}{\partial y} & \dfrac{\partial v}{\partial y} \end{pmatrix}$$

6. Matrice nulle. Égalité. Transposition

Nous appellerons application nulle l'application φ telle que $\forall \vec{u} \in E$, $\varphi(\vec{u}) = \vec{0}$. Sa matrice associée, quelles que soient les bases utilisées, a tous ses termes nuls, c'est-à-dire que $a_{ij} = 0, \quad \forall (i, j)$; c'est, par définition, la matrice nulle de dimension (p, n).

Deux matrices sont égales lorsqu'elles représentent, par rapport à des bases données, la même application linéaire. L'égalité $A = B$ équivaut donc, pour les éléments de A et B, aux relations

$$a_{ij} = b_{ij}, \quad i = 1, 2, \ldots, p; \quad j = 1, 2, \ldots, n$$

On appelle transposée d'une matrice A, et on note ${}^tA$, la matrice dont les lignes sont les colonnes de A et dont les colonnes sont les lignes de A. Ainsi

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \\ -1 & 4 \end{pmatrix} \Rightarrow {}^tA = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 2 & 4 \end{pmatrix}$$

On passe de A à ${}^tA$ en échangeant les lignes et les colonnes; la transposée d'une matrice (p, n) est une matrice (n, p).

La transposition est une opération **unaire** (portant sur un seul élément) involutive, il est en effet évident que ${}^t({}^tA) = A$.

On peut de plus démontrer, et nous admettrons, qu'un changement de base dans E ou dans F respecte la relation $B = {}^tA$ entre la matrice A d'une application linéaire φ de E dans F, et la matrice B d'une application linéaire ψ de F dans E. On dit que ψ est l'application transposée de φ, elle se note ${}^t\varphi$.

7. Opérations linéaires sur les applications linéaires et les matrices

Les notations restant les mêmes : dim $E = n$, dim $F = p$, les bases étant respectivement $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ et $(\vec{f_1}, \vec{f_2}, \ldots, \vec{f_p})$, soit φ et ψ deux applications linéaires de E dans F, de matrices associées A et B. $\vec{u}$ étant un vecteur quelconque de E, on peut définir l'application h, de E dans F, telle que $h(\vec{u}) = \varphi(\vec{u}) + \psi(\vec{u})$. Cette application est linéaire, car

$$\begin{aligned} h(\lambda\vec{u_1} + \mu\vec{u_2}) &= \varphi(\lambda\vec{u_1} + \mu\vec{u_2}) + \psi(\lambda\vec{u_1} + \mu\vec{u_2}) \\ &= \lambda\,[\varphi(\vec{u_1}) + \psi(\vec{u_1})] + \mu\,[\varphi(\vec{u_2}) + \psi(\vec{u_2})] \\ &= \lambda\, h(\vec{u_1}) + \mu\, h(\vec{u_2}). \end{aligned}$$

Par définition $h = \varphi + \psi$. La matrice C de h s'appelle la somme des deux matrices A et B (nécessairement de mêmes dimensions). Si a_{ij}, b_{ij}, c_{ij} sont les éléments de mêmes rangs dans A, B, C, on a $c_{ij} = a_{ij} + b_{ij}$. L'addition des applications linéaires et celle des matrices sont donc commutatives et associatives. Leurs éléments neutres sont respectivement l'application nulle et la matrice nulle. Tout élément, φ ou A, de chacun de ces deux ensembles possède un symétrique :

$$\varphi + \psi = 0 \Leftrightarrow \psi = -\varphi \quad \text{application opposée de } \varphi$$

$$A + B = 0 \Leftrightarrow B = -A \qquad (\text{donc } b_{ij} = -a_{ij})$$

matrice opposée de A

On peut de même définir la multiplication d'une application linéaire, donc d'une matrice, par un scalaire, l'application $\lambda\varphi$ étant celle où $\vec{u}$ a pour image $\lambda\varphi(\vec{u})$, la matrice associée étant λA, d'éléments λa_{ij}. En particulier $0 \,.\, A = 0$, $1 \,.\, A = A$, $(-1)A = -A$.

En résumé $\varphi + \psi$ est définie par $(\varphi + \psi)(\vec{u}) = \varphi(\vec{u}) + \psi(\vec{u})$

$\lambda\varphi$ est définie par $(\lambda\varphi)(\vec{u}) = \lambda \,.\, \varphi(\vec{u}) \quad \forall \vec{u} \in E$

$$\begin{pmatrix} & \vdots & \\ \ldots & a_{ij} & \ldots \\ & \vdots & \end{pmatrix} + \begin{pmatrix} & \vdots & \\ \ldots & b_{ij} & \ldots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \ldots & a_{ij} + b_{ij} & \ldots \\ & \vdots & \end{pmatrix}$$

$$\lambda \begin{pmatrix} & \vdots & \\ \ldots & a_{ij} & \ldots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \ldots & \lambda a_{ij} & \ldots \\ & \vdots & \end{pmatrix}$$

Il en résulte évidemment pour l'ensemble, noté $\mathscr{L}(E, F)$, des applications linéaires de E dans F, et pour l'ensemble, noté $\mathscr{M}_{p,n}$, des matrices à p lignes et n colonnes, une structure d'espace vectoriel. Une base de l'espace vectoriel $\mathscr{M}_{p,n}$ est constituée des matrices A_{ij} dont seul l'élément situé à l'intersection de la ligne i et de la colonne j est non nul, et vaut 1, en effet toute matrice A se décompose d'une façon unique sous la forme

$A = \sum\limits_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant n}} a_{ij} A_{ij}$. Les deux espaces vectoriels en question ont donc pour dimension le produit np.

– D'autre part il est évident que

$$\boxed{\begin{array}{c} {}^t(A + B) = {}^tA + {}^tB \\ {}^t(\lambda A) = \lambda\, {}^tA \end{array}}$$

la transposition est donc une application linéaire de $\mathscr{M}_{p,n}$ dans $\mathscr{M}_{n,p}$.

– Remarquons encore que les définitions des opérations linéaires sur $\mathscr{L}(E, F)$ restent valables même si E et F n'ont pas de dimension finie.

8. Composition des applications linéaires. Multiplication des matrices

a) Soit trois espaces vectoriels E, F, G. Soit φ une application linéaire de E dans F, ψ une application linéaire de F dans G : $\psi \circ \varphi = h$ est une application de E dans G

$$\vec{u} \in E \xrightarrow[\varphi]{} \vec{v} = \varphi(\vec{u}) \in F \xrightarrow[\psi]{} \vec{w} = \psi(\vec{v}) = h(\vec{u}) \in G$$

$\forall\ \vec{u_1}, \vec{u_2}$ de E et $\forall\ \lambda_1, \lambda_2$ de K,

$$\begin{aligned} h(\lambda_1\vec{u_1} + \lambda_2\vec{u_2}) &= \psi\,[\varphi(\lambda_1\vec{u_1} + \lambda_2\vec{u_2})] \\ &= \psi\,[\lambda_1\varphi(\vec{u_1}) + \lambda_2\varphi(\vec{u_2})] \\ &= \lambda_1\psi\,[\varphi(\vec{u_1})] + \lambda_2\psi\,[\varphi(\vec{u_2})] = \lambda_1 h(\vec{u_1}) + \lambda_2 h(\vec{u_2}), \end{aligned}$$

h est donc une application linéaire, composée de φ par ψ.

On peut de même définir la composée de plusieurs applications linéaires de E dans F, de F dans G, de G dans H, etc. Comme toute composition d'applications, cette composition est associative. Elle n'est pas commutative, sauf dans des cas particuliers : d'ailleurs en général $\varphi \circ \psi$ n'a pas de sens. Elle est, du fait de la linéarité, distributive par rapport à la combinaison linéaire :

$$\psi \circ (\lambda_1\varphi_1 + \lambda_2\varphi_2) = \lambda_1\,(\psi \circ \varphi_1) + \lambda_2\,(\psi \circ \varphi_2)$$

si λ_1 et λ_2 sont deux scalaires quelconques.

b) En dimensions finies et dans des bases données, $\psi \circ \varphi$ a, par définition, pour matrice associée la matrice $C = BA$, produit de la matrice A associée à φ par la matrice B associée à ψ. Nous allons établir la règle de calcul de ce produit BA.

Soit $\dim E = n$, base de $E = (\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$
$\dim F = p$, base de $F = (\vec{f_1}, \vec{f_2}, \ldots, \vec{f_p})$
$\dim G = q$, base de $G = (\vec{g_1}, \vec{g_2}, \ldots, \vec{g_q})$

Soit A la matrice de φ : $A = \begin{pmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \\ a_{p1} & a_{p2} \dots a_{pn} \end{pmatrix}$

Soit B la matrice de ψ : $B = \begin{pmatrix} b_{11} & b_{12} \dots b_{1p} \\ b_{21} & b_{22} \dots b_{2p} \\ \dots & \dots \\ b_{q1} & b_{q2} \dots b_{qp} \end{pmatrix}$

Soit C la matrice de $h = \psi \circ \varphi$: $C = \begin{pmatrix} c_{11} & c_{12} \dots c_{1n} \\ c_{21} & c_{22} \dots c_{2n} \\ \dots & \dots \\ c_{q1} & c_{q2} \dots c_{qn} \end{pmatrix}$

On sait que les colonnes de C sont les matrices-colonnes des vecteurs $h(\vec{e_1}), h(\vec{e_2}), \dots, h(\vec{e_n})$.

Or $\varphi(\vec{e_1}) = a_{11}\vec{f_1} + a_{21}\vec{f_2} + \dots + a_{p1}\vec{f_p}$.

Donc
$$\begin{aligned} h(\vec{e_1}) = \psi\,[\varphi(\vec{e_1})] &= a_{11}\psi(\vec{f_1}) + a_{21}\psi(\vec{f_2}) + \dots + a_{p1}\psi(\vec{f_p}) \\ &= a_{11}\,(b_{11}\vec{g_1} + b_{21}\vec{g_2} + \dots + b_{q1}\vec{g_q}) \\ &\quad + a_{21}\,(b_{12}\vec{g_1} + b_{22}\vec{g_2} + \dots + b_{q2}\vec{g_q}) + \dots + \\ &\quad + a_{p1}\,(b_{1p}\vec{g_1} + b_{2p}\vec{g_2} + \dots + b_{qp}\vec{g_q}) \\ &= c_{11}\vec{g_1} + c_{21}\vec{g_2} + \dots + c_{q1}\vec{g_q} \end{aligned}$$

Par identification, la décomposition de $h(\vec{e_1})$ dans une base de G étant unique, il vient
$$\begin{aligned} c_{11} &= b_{11}a_{11} + b_{12}a_{21} + \dots + b_{1p}a_{p1} \\ c_{21} &= b_{21}a_{11} + b_{22}a_{21} + \dots + b_{2p}a_{p1} \\ &\dots\dots\dots\dots \\ c_{q1} &= b_{q1}a_{11} + b_{q2}a_{21} + \dots + b_{qp}a_{p1} \end{aligned}$$

On voit que le premier élément de la première colonne de C est la somme des produits des éléments de la première ligne de B par ceux, de même rang, de la première colonne de A, et ainsi de suite pour le second, ..., le $q^{\text{ième}}$ élément. Cette loi se généralise facilement : c_{ij} est la $i^{\text{ième}}$ coordonnée de $h(\vec{e_j})$, or

$$\varphi(\vec{e_j}) = \sum_{k=1}^{k=p} a_{kj}\vec{f_k}$$

$$h(\vec{e_j}) = \sum_{k=1}^{k=p} a_{kj}\,\psi(\vec{f_h}) = \sum_{k=1}^{k=p} a_{kj}\left(\sum_{i=1}^{i=q} b_{ik}\vec{g_i}\right) = \sum_{i=1}^{i=q} c_{ij}\vec{g_i}$$

Donc
$$\boxed{c_{ij} = \sum_{k=1}^{k=p} b_{ik} a_{kj}}$$

Cette formule traduit bien le fait que, dans la matrice BA, produit de A par B, l'élément c_{ij} intersection de la $i^{\text{ième}}$ ligne et de la $j^{\text{ième}}$ colonne est la somme des produits des éléments de la $i^{\text{ième}}$ ligne de B (en nombre égal à p) par ceux de même rang de la $j^{\text{ième}}$ colonne de A (en nombre aussi égal à p).

Plus brièvement, on dit que **le produit de deux matrices s'effectue, de gauche à droite, ligne par colonne,** selon le schéma suivant

$$\begin{array}{c} \text{ligne } i \to \begin{pmatrix} b_{i1} & b_{i2} & \ldots\ldots & b_{ip} \end{pmatrix} \\ \\ \end{array} \begin{array}{c} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{pj} \end{pmatrix} \\ \uparrow \\ \text{colonne } j \end{array} \begin{array}{c} = \\ \\ \end{array} \begin{array}{c} \begin{pmatrix} \ldots & c_{ij} & \ldots \end{pmatrix} \leftarrow \text{ligne } i \\ \uparrow \\ \text{colonne } j \end{array}$$

La relation $Y = AX$, où X et Y sont des vecteurs-colonnes, apparaît bien comme un cas particulier de la multiplication de deux matrices.

Remarquons que, du point de vue des dimensions, l'opération $BA = C$ obéit au schéma suivant : $(q, p) \times (p, n) = (q, n)$.

Ainsi, le plus souvent, si BA existe, AB n'existe pas. Même si ces deux produits existent, ce qui suppose $q = n$, BA et AB n'ont mêmes dimensions que si de plus $p = n$. Même dans ce dernier cas, on a en général $BA \neq AB$. La multiplication matricielle n'est pas commutative. En revanche, comme la composition des applications linéaires, elle est associative et distributive par rapport à la combinaison linéaire :

$$(AB)\,C = A(BC)$$
$$(\lambda A + \mu B)\,C = \lambda\, AC + \mu\, BC$$
$$C\,(\lambda A + \mu B) = \lambda\, CA + \mu\, CB$$

(pourvu évidemment que ces opérations aient un sens).

Enfin la matrice nulle 0, de dimensions convenables, multipliée par ou multipliant une autre matrice de dimensions convenables donne un produit nul, mais $AB = 0$ n'exige ni $A = 0$ ni $B = 0$.

Exemple.

$$\begin{pmatrix} 1 & 2 \\ -2 & 4 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 & 2 \\ 0 & -1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 6 & 8 \\ -2 & -4 & 20 & 8 \\ 3 & 1 & -10 & 3 \end{pmatrix}$$

— Le lecteur démontrera aisément que

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

C. ENDOMORPHISMES. ALGÈBRE DES MATRICES CARRÉES D'ORDRE *n*

1. Endomorphismes d'un espace vectoriel de dimension finie. Matrices carrées d'ordre *n*

a) Dans le cas général des espaces de dimensions quelconques, finies ou non, seules les opérations linéaires sur $\mathscr{L}(E, F)$ conduisent à des éléments de $\mathscr{L}(E, F)$, autrement dit sont des opérations internes. Dans ce paragraphe, nous supposerons $\dim F = \dim E = n$, ce qui pratiquement revient à supposer $F = E$ pour le calcul matriciel associé.

Une application linéaire de E dans lui-même porte le nom d'endomorphisme (on dit parfois **endomorphie**), ce qui signifie « transformation interne » (du grec *endon* : au-dedans; *morphé* : forme). Il va de soi qu'alors la transposition et la composition sont aussi des opérations internes, on peut définir le composé de φ par lui-même, $\varphi \circ \varphi = \varphi^2$, puis $\varphi \circ \varphi^2 = \varphi^3$, ..., et de proche en proche φ^p. Dans une base donnée de E, la matrice de φ est une matrice carrée (n, n), ou matrice carrée d'ordre n. Elle contient n^2 éléments a_{ij}.

L'ensemble des endomorphismes de E (dans E, par définition même), soit $\mathscr{L}(E, E)$, est noté simplement $\mathscr{L}(E)$. De même l'ensemble des matrices carrées $\mathscr{M}_{n,n}$ est noté $\mathscr{M}_n$, éventuellement $\mathscr{M}_n(K)$, si l'on tient à préciser le corps sur lequel on opère : ainsi $\mathscr{M}_n(\mathbb{R})$ désigne les matrices carrées réelles, $\mathscr{M}_n(\mathbb{C})$ les matrices carrées complexes.

b) Sur $\mathscr{L}(E)$ et sur $\mathscr{M}_n$, on peut définir quelques éléments remarquables.

• L'application neutre, ou application unité, ou coïncidence, notée souvent e, est telle que $e(\vec{u}) = \vec{u}$, $\forall \vec{u} \in E$. C'est l'élément neutre de la composition des endomorphismes, car $e \circ \varphi = \varphi \circ e = \varphi$, $\forall \varphi$. La matrice associée à e, dans une base quelconque $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$, a pour vecteurs-colonnes $\varphi(\vec{e_1}) = \vec{e_1}$, $\varphi(\vec{e_2}) = \vec{e_2}$, ..., $\varphi(\vec{e_n}) = \vec{e_n}$, elle s'écrit donc :

$$I = \begin{pmatrix} 1 & 0 & 0 \ldots 0 \\ 0 & 1 & 0 \ldots 0 \\ 0 & 0 & 1 \ldots 0 \\ \vdots & \vdots & \vdots \ddots \vdots \\ 0 & 0 & 0 \ldots 1 \end{pmatrix}$$

(ou I_n, lorsqu'il importe de préciser la dimension). On a $IA = AI = A$, $\forall\, A \in \mathcal{M}_n$. C'est la **matrice-unité.** Comme elle est évidemment invariante dans un changement de base, il arrive que par abus de langage on la confonde avec e, endomorphisme unité. On écrit aussi $I = \begin{pmatrix} & \vdots & \\ \dots & \delta_{ij} & \dots \\ & \vdots & \end{pmatrix}$ en utilisant le **symbole de Kronecker** $\delta_{ij} = 1$ si $i = j$

$= 0$ si $i \neq j$

• Une matrice de la forme λI, c'est-à-dire où $a_{ij} = 0$ si $i \neq j$, $= \lambda$ si $i = j$, est appelée **matrice scalaire**; l'endomorphisme associé est une homothétie de rapport λ.

• Une **matrice** est dite **diagonale** si $i \neq j \Rightarrow a_{ij} = 0$: ainsi seule la « diagonale », partant de a_{11}, du carré que constitue la matrice, peut contenir des éléments non nuls.

Par exemple $\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$ est une matrice diagonale d'ordre 4.

• Une **matrice** A est **triangulaire** si, d'un côté de la diagonale, elle ne renferme que des zéros. Elle est triangulaire supérieure si les éléments éventuellement non nuls sont au-dessus de la diagonale (c'est-à-dire si $i > j \Rightarrow a_{ij} = 0$), triangulaire inférieure dans le cas contraire.

• Une **matrice** A est **symétrique** si elle est égale à sa transposée ${}^tA$, c'est-à-dire si $a_{ji} = a_{ij}$, $\forall\,(i, j)$; elle est **antisymétrique** si ${}^tA = -A$, c'est-à-dire si $a_{ji} = -a_{ij}$, $\forall\,(i, j)$, en particulier alors $a_{ii} = 0$.

2. Anneau des matrices carrées d'ordre n

Les opérations linéaires et la multiplication confèrent à l'ensemble $\mathcal{M}_n$ des matrices carrées d'ordre n une structure d'anneau non commutatif (c'est-à-dire où la multiplication n'est pas commutative). Le calcul algébrique sur les polynômes dans $\mathbb{R}$ ou $\mathbb{C}$ s'applique aux matrices pour tout ce qui ne fait pas intervenir la commutativité, un polynôme matriciel $P(X)$ résultant du polynôme $P(x)$, où x est un nombre, en remplaçant x^k par X^k pour $k \geqslant 0$, avec la convention $X^0 = I$. Ainsi

$$P(x) = x^2 + 2\,x + 3 \Rightarrow P(X) = X^2 + 2\,X + 3\,I$$

X étant une matrice carrée quelconque et I la matrice unité, d'ordre n.

Les « identités remarquables » de l'algèbre dans $\mathbb{R}$ ou $\mathbb{C}$ sont à remanier, ainsi

$$(A + B)(A - B) = A^2 + BA - AB - B^2 \neq A^2 - B^2$$

$$(A + B)^2 = A^2 + AB + BA + B^2 \neq A^2 + 2\,AB + B^2$$

sauf bien entendu si $AB = BA$: on dit alors que A commute avec B.

– Observons que I commute avec toute matrice A et que l'on peut alors appliquer, par exemple, la formule du binôme

$$(A + I)^p = \sum_{k=0}^{k=p} C_p^k A^k I^{p-k} = \sum_{k=0}^{k=p} C_p^k A^k$$

– On peut facilement vérifier que l'ensemble des matrices diagonales et l'ensemble des matrices triangulaires supérieures (ou inférieures) sont des sous-anneaux de l'anneau des matrices carrées d'ordre n.

– Enfin rappelons que $AB = 0$ n'exige ni $A = 0$, ni $B = 0$.

Par exemple $$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0$$

Il en résulte que, même si $A \neq 0$, $AB = AC \not\Rightarrow B = C$.

Nous reviendrons sur cette question de simplification ou **régularité.** Une matrice A non nulle telle qu'il en existe au moins une autre B, non nulle elle aussi, vérifiant $AB = 0$, est un **diviseur de zéro.**

3. Endomorphisme bijectif. Inverse d'une matrice carrée

a) Revenons momentanément au cas général d'une application linéaire φ de E dans F, les dimensions étant n et p. Nous avons vu que φ est injective si et seulement si $\ker \varphi = \{\vec{0}\}$. Cherchons à quelle condition φ est en outre surjective : ceci équivaut, par définition, à $\operatorname{Im} \varphi = F$. Or on sait que $\dim \operatorname{Im} \varphi + \dim \ker \varphi = n$. Il faut donc $p = \dim F = n$, puisque $\dim \ker \varphi = 0$. Il en résulte qu'une application linéaire de E dans F ne peut être bijective que si $\dim F = \dim E$, et le problème se limite au cas des endomorphismes. $\ker \varphi = \{\vec{0}\}$ est alors suffisante pour que φ soit surjective.

Ainsi, φ étant un endomorphisme de E, espace de dimension n, nous obtenons dès à présent les équivalences suivantes :

$$\boxed{\varphi \text{ bijectif} \Leftrightarrow \ker \varphi = \{\vec{0}\} \Leftrightarrow \operatorname{Im} \varphi = E}$$

D'après la définition générale d'une application bijective, tout vecteur $\vec{v}$ de E admettra alors un antécédent $\vec{u}$ et un seul dans E, autrement dit l'application réciproque ou inverse existera

$$\vec{u} = \varphi^{-1}(\vec{v}) \Leftrightarrow \vec{v} = \varphi(\vec{u}) \quad \text{et} \quad \varphi^{-1} \circ \varphi = e$$

L'existence de φ^{-1} équivaut à la bijectivité. Un endomorphisme bijectif, ou inversible, de E est un **isomorphisme** de E. On dit encore qu'un endomorphisme bijectif est **régulier**, ce qui signifie qu'il autorise la simplification dans un produit

$$\varphi \circ \psi_1 = \varphi \circ \psi_2 \Rightarrow \varphi^{-1} \circ \varphi \circ \psi_1 = \varphi^{-1} \circ \varphi \circ \psi_2 \Rightarrow \psi_1 = \psi_2$$

Un endomorphisme non régulier est dit singulier.

b) Ce qui précède a des conséquences immédiates concernant les matrices. Tout d'abord seule une matrice carrée pourra être éventuellement inversible. Nous appellerons évidemment inverse de A la matrice B, si elle existe, telle que $AB = BA = I$. B est alors notée A^{-1} et est la matrice de φ^{-1}, endomorphisme réciproque de celui défini par A dans une base de E, espace vectoriel de dimension n que l'on peut toujours associer aux matrices carrées d'ordre n. Par abus de langage, on note souvent $\ker A$ et $\operatorname{Im} A$ le noyau et l'image de φ, on peut alors écrire

$$\boxed{A \text{ inversible} \Leftrightarrow \ker A = \{\vec{0}\}}$$

Une matrice inversible est aussi régulière pour la multiplication, car

$$AB = AC \Rightarrow A^{-1}AB = A^{-1}AC \Rightarrow IB = IC \Rightarrow B = C$$

Une matrice non inversible est singulière. En particulier une matrice, diviseur de zéro, est singulière.

Il existe plusieurs méthodes permettant de voir si une matrice est inversible et éventuellement de calculer son inverse. Nous avons vu que $\operatorname{Im} \varphi$ est le sous-espace F engendré par $\varphi(\vec{e_1})$, $\varphi(\vec{e_2})$, ..., $\varphi(\vec{e_n})$. La matrice A associée à φ est donc inversible si et seulement si $F = E$, donc si les vecteurs-colonnes de A, formés respectivement des coordonnées de $\varphi(\vec{e_1})$, $\varphi(\vec{e_2})$, ..., $\varphi(\vec{e_n})$, sont indépendants (c'est-à-dire constituent une base de E). Ainsi

$$\boxed{A \text{ inversible} \Leftrightarrow \text{les vecteurs-colonnes de } A \text{ sont indépendants}}$$

Sans autre théorie, il va de soi que l'on peut, dans des cas simples, « inverser » une matrice A soit en résolvant le système $AX = Y$ sous la forme $X = A^{-1}Y$, soit en posant $A^{-1} = \begin{pmatrix} & \vdots & \\ \dots & b_{ij} & \dots \\ & \vdots & \end{pmatrix}$ et en calculant les b_{ij} par les équations résultant de $AA^{-1} = I$.

Exemple.

Soit à calculer l'inverse de $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 5 \\ -1 & 1 & 4 \end{pmatrix}$

Cherchons à résoudre $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$

$$\text{c'est-à-dire} \begin{cases} x + z = a & (1) \\ 2x + 3y + 5z = b & (2) \\ -x + y + 4z = c & (3) \end{cases}$$

$$(1) \Rightarrow x = a - z, \quad (1) + (3) \Rightarrow y = a + c - 5z$$

$$(2) \Rightarrow 2(a - z) + 3(a + c - 5z) + 5z = b$$

d'où

$$z = \frac{5a - b + 3c}{12}$$

$$x = \frac{7a + b - 3c}{12}$$

$$y = \frac{-13a + 5b - 3c}{12}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 7 & 1 & -3 \\ -13 & 5 & -3 \\ 5 & -1 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad A^{-1} = \frac{1}{12} \begin{pmatrix} 7 & 1 & -3 \\ -13 & 5 & -3 \\ 5 & -1 & 3 \end{pmatrix}$$

c) D'une façon générale, on peut établir que la composée de deux applications bijectives est une application bijective. Dans le cas de deux matrices inversibles A et B, associées à des endomorphismes φ et ψ, l'image d'une base de E par φ est une base de E, dont l'image par ψ est encore une base de E, qui est évidemment l'image de la première par $\psi \circ \varphi$, qui est donc inversible. On constate aisément que

$$\boxed{(BA)^{-1} = A^{-1}B^{-1}}$$

De plus $\qquad A \,.\, A^{-1} = I \Rightarrow {}^t(A^{-1})\,{}^tA = {}^tI = I$

donc $\qquad \boxed{({}^tA)^{-1} = {}^t(A^{-1})}$

A inversible $\Leftrightarrow$ ${}^tA$ inversible $\Leftrightarrow$ les vecteurs-colonnes de ${}^tA$ sont indépendants $\Leftrightarrow$ les vecteurs-lignes de A sont indépendants.

Notons encore que les matrices carrées régulières constituent un groupe multiplicatif.

4. Inversion progressive d'une matrice par prémultiplications successives

Le principe de cette méthode, souvent utilisée, est le suivant.

Si on peut trouver une suite (finie) de matrices $M_1, M_2, \ldots, M_k$ telles que $M_k M_{k-1} \ldots M_2 M_1 A = I$, on a évidemment, par construction même, $M_k M_{k-1} \ldots M_2 M_1 I = A^{-1}$.

Or on peut déterminer ces matrices successives M_i telles que chaque multiplication de A par M_1, de $M_1 A$ par M_2, etc., ait pour effet le remplacement d'un vecteur-ligne de la matrice précédente par une combinaison linéaire de lignes de cette même matrice, les produits successifs $M_1 A$, $M_2 M_1 A$, ..., évoluant vers la matrice-unité. Si on arrive à cette matrice-unité au bout de k opérations, les mêmes transformations menées parallèlement sur I, c'est-à-dire la formation de $M_1 I$, $M_2 M_1 I$, ..., nous amènerons au bout des mêmes k opérations à la matrice A^{-1}.

Le point essentiel est de démontrer que si nous écrivons A sous forme de n vecteurs-lignes à chacun n coordonnées, remplacer X_1 par exemple par une combinaison linéaire des n vecteurs-lignes (ou matrices unilignes) $X_1, X_2, \ldots, X_n$, revient à multiplier A par une matrice : en effet

$$A_1 = \begin{pmatrix} \lambda_1 X_1 + \lambda_2 X_2 + \ldots + \lambda_n X_n \\ X_2 \\ X_3 \\ \ldots \\ X_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ 0 & 1 & 0 \ldots & 0 \\ 0 & 0 & 1 \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 \ldots & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \ldots \\ X_n \end{pmatrix} = MA$$

puisque le vecteur $\sum_1^n \lambda_i X_i$ a pour coordonnées $\sum_1^n \lambda_i a_{i1}$, $\sum_1^n \lambda_i a_{i2}$, ..., $\sum_1^n \lambda_i a_{in}$ produits de la première ligne de M par les colonnes successives de A.

Exemples.

a) Cherchons par cette méthode et en détaillant les opérations l'inverse de $A = \begin{pmatrix} 3 & -2 \\ 5 & 7 \end{pmatrix}$ qui existe puisque $X_1 = (3, -2)$ et $X_2 = (5, 7)$ ne sont pas colinéaires.

Pour faire « évoluer » A vers $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ faisons apparaître un 0 à la place de -2 en remplaçant $X_1 = (3, -2)$ par $7X_1 + 2X_2$ d'où

$$M_1 A = \begin{pmatrix} 7 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} 31 & 0 \\ 5 & 7 \end{pmatrix} = A_1$$

Parallèlement $M_1 I = \begin{pmatrix} 7 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ 0 & 1 \end{pmatrix} = J_1$

Dans A_1 faisons apparaître un 0 à la place de 5 en multipliant la seconde ligne par 31 et lui ajoutant la première, multipliée par -5.

$$M_2A_1 = \begin{pmatrix} 1 & 0 \\ -5 & 31 \end{pmatrix}\begin{pmatrix} 31 & 0 \\ 5 & 7 \end{pmatrix} = \begin{pmatrix} 31 & 0 \\ 0 & 217 \end{pmatrix} = A_2$$

et
$$M_2J_1 = \begin{pmatrix} 1 & 0 \\ -5 & 31 \end{pmatrix}\begin{pmatrix} 7 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ -35 & 21 \end{pmatrix} = J_2$$

Pour passer de A_2 à I, il suffit de diviser les lignes par 31 et 217 respectivement, ce qui revient à effectuer

$$M_3A_2 = \begin{pmatrix} \frac{1}{31} & 0 \\ 0 & \frac{1}{217} \end{pmatrix}\begin{pmatrix} 31 & 0 \\ 0 & 217 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A_3 = I$$

et $$M_3J_2 = \begin{pmatrix} \frac{1}{31} & 0 \\ 0 & \frac{1}{217} \end{pmatrix}\begin{pmatrix} 7 & 2 \\ -35 & 21 \end{pmatrix} = \begin{pmatrix} \frac{7}{31} & \frac{2}{31} \\ -\frac{5}{31} & \frac{3}{31} \end{pmatrix} = J_3 = A^{-1}$$

Il est immédiat de vérifier que J_3 est bien l'inverse de A. Pour une matrice d'ordre 2 cette méthode n'est évidemment pas la plus rapide, d'ailleurs un calcul simple montre que

$$\boxed{\begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}} \quad \text{si } ad - bc \neq 0$$

autrement dit : on échange les termes diagonaux,
on change de signe les deux autres,
on divise par le déterminant $ad - bc$.

Cette formule sera généralisée ultérieurement.

b) Dans la pratique il n'est pas nécessaire d'indiquer les multiplicateurs successifs M_1, M_2, M_3 . . ., et on peut disposer les opérations comme suit.

Soit à déterminer la matrice A^{-1} inverse de

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 2 \\ -1 & 1 & 1 \end{pmatrix}$$

Les combinaisons linéaires de lignes appliquées simultanément à A et I qui transforment A en I, transformant I en A^{-1}.

$L1 \Rightarrow L1 \times 1 - L3 \times 3$ (ligne 1 devient ligne 1 − 3 fois ligne 3)	2	1	0	1	0	− 1
	0	2	2	0	1	0
	− 1	1	1	0	0	1
	2	1	0	1	0	− 1
$L3 \Rightarrow L3 \times 2 + L1$	0	2	2	0	1	0
	0	3	2	1	0	1
	2	1	0	1	0	1
$L3 \Rightarrow 2\,L3 - 3\,L2$	0	2	2	0	1	0
	0	0	− 2	2	− 3	2
	2	1	0	1	0	1
$L2 \Rightarrow L2 + L3$	0	2	0	2	− 2	2
	0	0	− 2	2	− 3	2
	2	0	0	0	1	− 2
$L1 \Rightarrow 2\,L1 - \frac{1}{2}L2$	0	2	0	2	− 2	2
	0	0	− 2	2	− 3	2

En divisant la ligne 1 par $\frac{1}{2}$,

En divisant la ligne 2 par $\frac{1}{2}$,

En divisant la ligne 3 par $-\frac{1}{2}$,

On a

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

5. Changement de base. Matrices semblables

Soit, dans un espace vectoriel E de dimension n, deux bases

$$B = (\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) \qquad B' = (\vec{e'_1}, \vec{e'_2}, \ldots, \vec{e'_n})$$

Nous allons chercher des relations, d'abord entre les coordonnées d'un vecteur $\vec{u}$ de E dans la base B, que nous appellerons conventionnellement l'ancienne base, et dans la base B', nouvelle base; ensuite, entre les matrices d'un même endomorphisme φ dans l'ancienne et la nouvelle base.

Pour cela, définissons B', par rapport à B, par les coordonnées des $\vec{e'_j}$ dans B, soit $\vec{e'_j} = \sum_{i=1}^{i=n} \alpha_{ij}\vec{e_i}$. Ainsi la matrice P du système de vecteurs $(\vec{e'_j})$, dans B, est

$$P = \begin{pmatrix} \alpha_{11} & \alpha_{12} \ldots \alpha_{1n} \\ \alpha_{21} & \alpha_{22} \ldots \alpha_{2n} \\ \ldots & \ldots \\ \alpha_{n1} & \alpha_{n2} \ldots \alpha_{nn} \end{pmatrix}$$

Cette matrice P définit complètement le changement de base $B \to B'$, on l'appelle matrice de changement de base, ou plus couramment **matrice de passage** (de B à B'). C'est une matrice régulière puisque le système des $\vec{e'_j}$ est libre.

a) Changement de base pour un vecteur $\vec{u}$.

Soit $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$ les matrices de $\vec{u}$ dans B et B'.

Par définition

$$\begin{aligned} \vec{u} &= x_1\vec{e_1} + x_2\vec{e_2} + \ldots + x_n\vec{e_n} = x'_1\vec{e'_1} + x'_2\vec{e'_2} + \ldots + x'_n\vec{e'_n} \\ &= x'_1(\alpha_{11}\vec{e_1} + \alpha_{21}\vec{e_2} + \ldots + \alpha_{n1}\vec{e_n}) \\ &\quad + x'_2(\alpha_{12}\vec{e_1} + \alpha_{22}\vec{e_2} + \ldots + \alpha_{n2}\vec{e_n}) + \ldots \\ &\quad + x'_n(\alpha_{1n}\vec{e_1} + \alpha_{2n}\vec{e_2} + \ldots + \alpha_{nn}\vec{e_n}) \end{aligned}$$

d'où

$$\begin{cases} x_1 = \alpha_{11}x'_1 + \alpha_{12}x'_2 + \ldots + \alpha_{1n}x'_n \\ x_2 = \alpha_{21}x'_1 + \alpha_{22}x'_2 + \ldots + \alpha_{2n}x'_n \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = \alpha_{n1}x'_1 + \alpha_{n2}x'_2 + \ldots + \alpha_{nn}x'_n \end{cases}$$

Il en résulte la formule de changement de base pour un vecteur :

$$\boxed{X = PX'} \qquad \text{(ou } X' = P^{-1}X\text{)}$$

b) Changement de base pour un endomorphisme φ

Soit A la matrice de φ dans B. Soit encore $\vec{u}$ un vecteur de E, de matrices X et X' dans B et B', et $\vec{v} = \varphi(\vec{u})$, de matrices Y et Y' dans B et B'.

On a donc $Y = AX, \quad X = PX', \quad Y = PY'$.

Donc

$$PY' = APX', \quad P^{-1}PY' = P^{-1}APX' \quad Y' = P^{-1}APX' = A'X'$$

La matrice A' de φ dans la base B' est donc donnée par la formule de changement de base pour un endomorphisme :

$$\boxed{A' = P^{-1}AP} \qquad \text{(ou } A = PA'P^{-1}\text{)} \tag{7}$$

c) D'une façon générale, **deux matrices A et A' liées par une relation de la forme $A' = P^{-1}AP$ sont dites semblables,** par l'intermédiaire de la matrice de passage P. La similitude de deux matrices peut être vérifiée sans calcul d'inverse puisque $PA' = AP$, où P est régulière, équivaut à (7). Deux matrices semblables expriment le même endomorphisme dans deux bases différentes. La similitude se note $A' \sim A$.

On peut établir que **la similitude est une relation d'équivalence** sur l'anneau des matrices carrées d'ordre n, en effet :

elle est réflexive : $A' = A \Rightarrow A' = I^{-1}AI$, donc $A' \sim A$

elle est symétrique : $A' \sim A \Leftrightarrow A' = P^{-1}AP \Leftrightarrow A = Q^{-1}A'Q$, avec $Q = P^{-1}$, donc $A \sim A'$

elle est transitive : $A' \sim A \Leftrightarrow A' = P^{-1}AP$

$A'' \sim A' \Leftrightarrow A'' = R^{-1}A'R$

donc $$A'' = R^{-1}P^{-1}APR = (PR)^{-1}\, A(PR)$$

et $$\left.\begin{array}{l} A' \sim A \\ A'' \sim A' \end{array}\right| \Rightarrow A'' \sim A$$

Autres propriétés :

$$A' = P^{-1}AP \Rightarrow A'^2 = P^{-1}APP^{-1}AP = P^{-1}A^2P$$

puisque $PP^{-1} = I$.

De même, de proche en proche, $A'^k = P^{-1}A^kP, \quad \forall k \in \mathbb{N}$.

Si A est inversible

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

et en posant $A^k = (A^{-1})^{-k}$ pour k entier négatif on voit que

$$A'^k = P^{-1}A^kP$$

reste valable $\forall k \in \mathbb{Z}$.

Enfin, du fait de la distributivité de la multiplication, si f est un polynôme quelconque, on a $f(A') = P^{-1}f(A)P$.

6. Projecteurs

a) Deux sous-espaces F et G d'un espace vectoriel E sont, nous l'avons vu, supplémentaires si $E = F \oplus G$. Alors, $\forall \vec{u} \in E$, il existe une décomposition unique

$$\vec{u} = \vec{v} + \vec{w}, \qquad \vec{v} \in F, \qquad \vec{w} \in G$$

Par analogie avec la géométrie élémentaire, on dit que $\vec{v}$ est la projection de $\vec{u}$ sur F, parallèlement à G, et que $\vec{w}$ est la projection de $\vec{u}$ sur G, parallèlement à F.

L'application p qui, à $\vec{u}$ quelconque dans E, fait correspondre sa projection $\vec{v}$ sur F, parallèlement à G, s'appelle projecteur de E sur F, parallèlement à G : $p(\vec{u}) = \vec{v}$. On définit de même le projecteur q de E sur G, parallèlement à F, par $q(\vec{u}) = \vec{w}$.

On peut observer que p associe à tout $\vec{u}$ de E l'unique vecteur $\vec{v} = p(\vec{u})$ tel que $p(\vec{u}) \in F$, $p(\vec{u}) - \vec{u} \in G$.

b) Propriétés des projecteurs.

– Un projecteur est une application linéaire.

En effet

$$p(\lambda_1\vec{u_1} + \lambda_2\vec{u_2}) - (\lambda_1\vec{u_1} + \lambda_2\vec{u_2}) \in G$$

d'autre part $p(\vec{u_1}) - \vec{u_1} \in G$, $p(\vec{u_2}) - \vec{u_2} \in G$ (sous-espace de E)

$$\Rightarrow \lambda_1[p(\vec{u_1}) - \vec{u_1}] + \lambda_2[p(\vec{u_2}) - \vec{u_2}] \in G$$

$$\lambda_1 p(\vec{u_1}) + \lambda_2 p(\vec{u_2}) - (\lambda_1\vec{u_1} + \lambda_2\vec{u_2}) \in G$$

donc $$\vec{z} = p(\lambda_1\vec{u_1} + \lambda_2\vec{u_2}) - \lambda_1 p(\vec{u_1}) - \lambda_2 p(\vec{u_2}) \in G$$

Or ce vecteur $\vec{z}$, par définition de p, appartient aussi à F, et comme $F \cap G = \{\vec{0}\}$ on a $\vec{z} = \vec{0}$ et $p(\lambda_1\vec{u_1} + \lambda_2\vec{u_2}) = \lambda_1 p(\vec{u_1}) + \lambda_2 p(\vec{u_2})$

– Cherchons le noyau de p, projecteur de E sur F, parallèlement à G : $p(\vec{u}) = \vec{0} \Rightarrow p(\vec{u}) - \vec{u} = -\vec{u} \in G$ donc $\vec{u} \in G$ et $\ker p \in G$; mais si $\vec{u} \in G$, $\vec{u} = \vec{u} + \vec{0}$, décomposition unique où $\vec{u} \in G$ et $\vec{0} \in F$, montre que $\vec{0} = p(\vec{u})$ donc $\ker p = G$. De même $\ker q = F$.

En outre si E est de dimension finie n,

$\dim F + \dim G = n$, $\dim \operatorname{Im} p + \dim G = n$, donc $\dim \operatorname{Im} p = \dim F$, et comme par définition $\operatorname{Im} p \subset F$, on a $\operatorname{Im} p = F$. De même $\operatorname{Im} q = G$.

Retenons donc qu'en dimension finie, si p est le projecteur de E sur F, sous-espace de E, parallèlement à un supplémentaire G de F, on a $\ker p = G$, $\operatorname{Im} p = F$.

– Enfin considérons

$$p^2 = p \circ p : \quad \forall \vec{u} \in E, \quad p(\vec{u}) - \vec{u} \in G, \text{ or } \quad G = \ker p,$$

donc

$$p\,[p(\vec{u}) - \vec{u}] = \vec{0}, \quad p^2(\vec{u}) = p(\vec{u}).$$

Ceci étant vrai $\forall \vec{u} \in E$, on a $p^2 = p$.

Dans une base quelconque de E (supposée ici de dimension finie), la matrice A d'un projecteur vérifie la relation $A^2 = A$ (qui entraîne $A^k = A \quad \forall\, k \in \mathbb{N}^*$). On dit qu'une telle matrice est idempotente. Un projecteur est un endomorphisme idempotent.

Réciproquement, raisonnant encore en dimension finie n, soit A une matrice telle que $A^2 = A$, c'est-à-dire $A(A - I) = 0$. Alors ou bien $A = 0$, ou bien $A = I$, solutions « triviales »; ou bien A et $A - I$ sont singulières, et alors on peut poser $\ker \varphi = G$, sous-espace de E de dimension s telle que $1 \leqslant s \leqslant n - 1$, φ étant l'endomorphisme de matrice A.

$$\forall \vec{u} \in E, \quad \varphi^2(\vec{u}) = \varphi(\vec{u}), \quad \text{donc} \quad \varphi\,[\vec{u} - \varphi(\vec{u})] = \vec{0}$$

et $\qquad \vec{u} - \varphi(\vec{u}) \in G$, donc $\vec{u} = \varphi(\vec{u}) + \vec{w}$, $\vec{w} \in G$

Mais $\vec{v} = \varphi(\vec{u}) \in \operatorname{Im} \varphi = F$.

Donc tout $\vec{u}$ de E est décomposable sous la forme $\vec{u} = \vec{v} + \vec{w}$, où

$$\vec{v} \in F = \operatorname{Im} \varphi, \ \vec{w} \in G = \ker \varphi, \text{ et } \vec{u} \in F + G, \text{ avec } \dim F + \dim G = n$$

Ainsi $E \subset F + G$, et $n = \dim E \leqslant \dim (F + G)$.

Mais

$$\dim (F + G) = \dim F + \dim G - \dim (F \cap G) = n - \dim (F \cap G)$$

donc $\dim (F \cap G) = 0$, $F \cap G = \{\vec{0}\}$, $E = F \oplus G$, et φ est un projecteur.

En remarquant que la matrice 0 définit le projecteur sur $\{\vec{0}\}$ parallèlement à E, et que la matrice I définit le projecteur sur E parallèlement à $\{\vec{0}\}$, nous pouvons conclure par le théorème :

Pour qu'un endomorphisme de dimension finie soit un projecteur, il faut et il suffit qu'il soit idempotent : $\varphi^2 = \varphi$.

– Remarquons que, par contre, la relation $E = \ker \varphi \oplus \operatorname{Im} \varphi$ n'est pas caractéristique des projecteurs.

D. NOTIONS SUR LA DUALITÉ

1. Forme linéaire

Soit E un espace vectoriel sur un corps K de scalaires. Il est clair que K est un espace vectoriel sur lui-même, de dimension 1. Une application linéaire φ de E dans K associe à tout vecteur $\vec{u}$ de E un élément $\lambda = \varphi(\vec{u})$ de K. Une telle application φ – et, par un abus de langage classique, l'image par φ d'un vecteur de E – se nomme forme linéaire de E dans K.

En dimension finie, B étant une base $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ de E, l'image d'un vecteur $\vec{u}$ de coordonnées $x_1, x_2, \ldots, x_n$ est de la forme

$$\boxed{\varphi(\vec{u}) = \sum_{i=1}^{i=n} a_i x_i} \qquad \text{(forme linéaire à } n \text{ variables)}$$

La base prise pour K est évidemment l'élément unité de K, noté 1. La matrice de φ est de dimensions $(1, n)$, donc elle a une seule ligne, c'est $A = (a_1\ a_2\ \ldots\ a_n)$, on constate bien d'ailleurs que

$$\varphi(\vec{u}) = (a_1\ a_2\ \ldots\ a_n)\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$$

2. Espace dual d'un espace vectoriel. Base duale d'une base de E

L'ensemble des formes linéaires de E dans K est un espace vectoriel de dimension égale à $\dim E \times \dim K = n$, si E est de dimension finie. On l'appelle dual de E et on le note E^*. Ainsi, par définition

$$\boxed{E^* = \mathscr{L}(E, K)}$$

Soit $B = (\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ une base de E. Une forme linéaire f est définie par la donnée de $f(\vec{e_1}), f(\vec{e_2}), \ldots, f(\vec{e_n})$. Puisque $\dim E^* = n$, il suffit de définir n formes indépendantes pour obtenir une base de E^*. Nous poserons

$$\boxed{f_i(\vec{e_j}) = \delta_{ij}}$$

c'est-à-dire

$$\begin{array}{lll} f_1(\vec{e_1}) = 1, & f_1(\vec{e_2}) = 0, \ldots, & f_1(\vec{e_n}) = 0 \\ f_2(\vec{e_1}) = 0, & f_2(\vec{e_2}) = 1, \ldots, & f_2(\vec{e_n}) = 0 \\ \cdots & \cdots & \cdots \\ f_n(\vec{e_1}) = 0, & f_n(\vec{e_2}) = 0, \ldots, & f_n(\vec{e_n}) = 1 \end{array}$$

Ces n formes linéaires sont indépendantes, en effet

$$\sum_1^n \alpha_i f_i(\vec{u}) = 0 \qquad \forall \vec{u} \in E$$

exige
$$\sum_{i=1}^{i=n} \alpha_i f_i(\vec{e_j}) = 0 \quad \forall j = 1, 2, \ldots, n$$

ou $\sum_{i=1}^{i=n} \alpha_i \delta_{ij} = 0$ donc $\alpha_j = 0$.

La base de E^* ainsi définie s'appelle base duale de la base B de E.

Le vecteur $\vec{u}$ de coordonnées $x_1, x_2, \ldots, x_n$ a pour image

$$f(\vec{u}) = f\left(\sum_1^n x_i \vec{e_i}\right) = \sum_1^n x_i f(\vec{e_i}) = \sum_1^n a_i x_i$$

si l'on pose $f(\vec{e_i}) = a_i$; en particulier $f_j(\vec{u}) = \sum_{i=1}^{i=n} x_i \delta_{ij} = x_j$, on voit donc que les $f_j(\vec{u})$ sont les coordonnées de $\vec{u}$ dans la base B.

Résumé du chapitre 6

- **Définition d'un espace vectoriel.**

Ensemble E muni de deux lois de composition :

l'addition $(+)$ pour laquelle E possède une structure de groupe;

la multiplication par un scalaire (notée $\alpha\vec{u}$) possédant les propriétés :

$$\begin{aligned} \alpha(\vec{u} + \vec{v}) &= \alpha\vec{u} + \alpha\vec{v} \\ (\alpha + \beta)\vec{u} &= \alpha\vec{u} + \beta\vec{u} \\ \alpha(\beta\vec{u}) &= (\alpha\beta)\vec{u} \\ 1 . \vec{u} &= \vec{u} \end{aligned}$$

- p vecteurs de E constituent un

système libre si $\sum_1^p \alpha_i \vec{u_i} = \vec{0} \Rightarrow \alpha_i = 0, \quad \forall i = 1, 2, \ldots, p$

système lié si $\exists$ des α_i non tous nuls tels que $\sum_1^p \alpha_i \vec{u_i} = \vec{0}$.

• Une partie F d'un espace vectoriel E est un **sous-espace vectoriel** de E si

$$\forall \vec{u}, \vec{v} \text{ dans } F. \quad \forall \alpha, \beta \text{ scalaires}, \quad \alpha\vec{u} + \beta\vec{v} \in F$$

• **L'intersection de deux sous-espaces** de E est un sous-espace.

• La **somme** $F_1 + F_2$ **de deux sous-espaces,** définie comme ensemble des vecteurs $\vec{u_1} + \vec{u_2}$, où $\vec{u_1} \in F_1$ et $\vec{u_2} \in F_2$, est un sous-espace.

La somme est directe si $F_1 \cap F_2 = \{\vec{0}\}$ on note

$$F_1 + F_2 = F_1 \oplus F_2$$

et un vecteur de $F_1 \oplus F_2$ se décompose d'une façon unique en $\vec{u_1} + \vec{u_2}$, $\vec{u_1} \in F_1$, $\vec{u_2} \in F_2$. On dit que F_1 et F_2 sont supplémentaires si $E = F_1 \oplus F_2$.

• On appelle **base d'un espace vectoriel** E une famille libre de générateurs; tout vecteur $\vec{u}$ de E admet une décomposition unique suivant une base, $\vec{u} = \sum_1^n x_i \vec{e_i}$, les x_i sont les coordonnées de $\vec{u}$.

• Si E possède une base de n vecteurs, toute autre base contient n vecteurs, et n est le nombre maximum d'éléments d'un système libre de E : c'est la **dimension** de E.

• **Base incomplète.**

Si $\dim E = n$, p vecteurs indépendants de E ($p \leqslant n$) peuvent être complétés par $n - p$ autres vecteurs pour former une base de E.

• **En dimension finie.**

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G)$$
$$\dim(F \oplus G) = \dim F + \dim G$$

• **Une application** φ **de** E **dans** F (espaces vectoriels) **est linéaire si :**

$$\forall \vec{u}, \vec{v} \text{ de } E, \quad \forall \lambda, \mu \text{ scalaires}, \quad \varphi(\lambda\vec{u} + \mu\vec{v}) = \lambda\varphi(\vec{u}) + \mu\varphi(\vec{v})$$

- **Noyau de φ, ou ker φ.**

$$\vec{u} \in \ker \varphi \Leftrightarrow \varphi(\vec{u}) = \vec{0} \quad \text{(sous-espace de } E\text{)}$$

- **Image de φ, ou Im φ.**

$$\vec{v} \in \operatorname{Im} \varphi \Leftrightarrow \exists\, \vec{u} \in E \quad \text{tel que} \quad \varphi(\vec{u}) = \vec{v} \quad \text{(sous-espace de } F\text{)}$$

- **φ injective $\Leftrightarrow$ ker $\varphi = \{ \vec{0} \}$.**
- Si E est de dimension finie $\quad \dim \operatorname{Im} \varphi + \dim \ker \varphi = \dim E$

dim Im φ s'appelle le **rang de** φ

- Dans deux bases données $(\vec{e_1}, \ldots, \vec{e_n})$ dans E et $(\vec{f_1}, \ldots, \vec{f_p})$ dans F (dimensions finies), **φ est définie par une matrice A à p lignes et n colonnes,** où la colonne de rang j est constituée des coordonnées de $\varphi(\vec{e_j})$; a_{ij} est l'élément de A commun à la $i^{\text{ième}}$ ligne et à la $j^{\text{ième}}$ colonne.

- **Transposition.**

$^tA = B$ s'obtient en échangeant lignes et colonnes.

- **Opérations linéaires.**

  – Si $\forall\, \vec{u} \in E, \quad f(\vec{u}) = \varphi(\vec{u}) + \psi(\vec{u}), \quad f = \varphi + \psi$
  sa matrice est $\quad A + B \quad$ d'éléments $\quad a_{ij} + b_{ij}$

  – Si $\forall\, \vec{u} \in E, \quad f(\vec{u}) = \lambda\varphi(\vec{u}), \quad f = \lambda\varphi$
  sa matrice est λA d'éléments λa_{ij}.

- **Produit de deux matrices A et B associées à deux applications linéaires φ et ψ.**

C'est la matrice BA associée à $\psi \circ \varphi$, ses éléments sont donnés par

$$c_{ij} = \sum_{k=1}^{k=p} b_{ik}a_{kj} \qquad \text{(multiplication lignes par colonnes)}$$

où $\quad p =$ nombre de lignes de A = nombre de colonnes de B.

- **Un endomorphisme d'un espace vectoriel E est une application linéaire de E dans E.**

- **Sa matrice est carrée d'ordre n.**

Matrice-unité : $\quad a_{ii} = 1, \quad a_{ij} = 0 \quad$ si $i \neq j$.
Matrice diagonale : $\quad a_{ij} = 0 \quad$ si $i \neq j$.

Matrice triangulaire supérieure : $a_{ij} = 0$ si $i > j$.
Matrice symétrique : ${}^tA = A$.
Matrice antisymétrique : ${}^tA = -A$.

• L'ensemble des matrices carrées d'ordre n est un **anneau non commutatif.**

• **φ bijectif,**
matrice A associée inversible $\Leftrightarrow \ker \varphi = \{\vec{0}\}$
$\Leftrightarrow \operatorname{Im} \varphi = E$
$\Leftrightarrow$ vecteurs-colonnes indépendants
$\Leftrightarrow$ vecteurs-lignes indépendants.

Alors φ est un isomorphisme, A est régulière.

$$(BA)^{-1} = A^{-1}B^{-1}, \qquad ({}^tA)^{-1} = {}^t(A^{-1})$$

• **Inversion progressive d'une matrice** (voir cours et exemples).

• **Changement de base.**

Si P = matrice de passage = matrice de la nouvelle base par rapport à l'ancienne, on a

$X = PX'$ (changement de base pour un vecteur)
$A' = P^{-1}AP$ (changement de base pour un endomorphisme).

A et A' sont semblables ($A' \sim A$) si $\exists P$ inversible telle que $A' = P^{-1}AP$.
La similitude est une relation d'équivalence.

• **Projecteurs.**

Si $E = F \oplus G$, $\vec{u} \in E$ détermine $\vec{v} = p(\vec{u})$ tel que

$$p(\vec{u}) \in F, \qquad p(\vec{u}) - \vec{u} \in G$$

$p(\vec{u})$ est la projection de $\vec{u}$ sur F parallèlement à G, p est le projecteur de E sur F parallèlement à G.

C'est une application linéaire, où $\ker p = G$, $\operatorname{Im} p = F$.

• **p = projecteur $\Leftrightarrow p \circ p = p$ (p est idempotent).**

• Une forme linéaire de E est une application linéaire de E dans le corps K de scalaires associé à E. L'ensemble de ces formes linéaires est appelé **espace dual de E**, il est noté **E^***.

On a $\dim E^* = \dim E$ et si les $\vec{e_j}$ sont une base de E, la **base duale** dans E^* est formée des f_i telles que

$$\begin{aligned} f_i(\vec{e_j}) &= \delta_{ij} \quad \text{(symbole de Kronecker)} \\ &= 1 \quad \text{si } i = j \\ &= 0 \quad \text{si } i \neq j \end{aligned}$$

EXERCICES

■ Les ensembles suivants sont-ils des espaces vectoriels sur $\mathbb{C}$, ou éventuellement sur un sous-corps de $\mathbb{C}$:

a) les réels de la forme $x = a + b\sqrt{2}$ où a et b sont des réels rationnels;

b) les séries divergentes;

c) les séries entières de rayon de convergence donné R.

---

a) Oui, sur $\mathbb{Q}$.

b) Non : la somme de deux séries divergentes peut donner une série convergente, par exemple $\sum \frac{1}{n}$ et $\sum \frac{1-n}{n^2}$.

c) Non : par exemple $\sum_0^{+\infty} z^n$ et $\sum_0^{+\infty} \frac{1-n!}{n!} z^n$ ont même rayon de convergence $R = 1$, leur somme $\sum_0^{+\infty} \frac{z^n}{n!}$ a un rayon de convergence infini. Par contre les séries entières de rayon de convergence supérieur ou égal à R constituent un espace vectoriel sur $\mathbb{C}$.

■ **2.** Soit F_1 et F_2 deux sous-espaces d'un espace vectoriel E donné.

a) A quelle condition $F_1 \cup F_2$ est-elle un sous-espace de E ?

b) Dans le cas général, démontrer que $F_1 + F_2$ est le plus petit sous-espace de E contenant $F_1 \cup F_2$.

---

a) Soit $\vec{u_1} \in F_1$, $\vec{u_2} \in \vec{F_2}$, donc $\vec{u_1} \in F$ et $\vec{u_2} \in F$, si

$$F = F_1 \cup F_2$$

Si F est un sous-espace, $\vec{v} = \vec{u_1} + \vec{u_2} \in F$, donc $\vec{v} \in F_1$ ou $\vec{v} \in F_2$.

Si $\vec{v} \in \vec{F_1}$, comme de plus $\vec{u_1} \in F_1$, on a $\vec{u_2} = \vec{v} - \vec{u_1} \in F_1$; ceci est vrai $\forall \vec{u_2} \in F_2$, donc $F_2 \subset F_1$; de même $\vec{v} \in F_2 \Rightarrow F_1 \subset F_2$.

En conclusion, $F_1 \cup F_2$ ne peut être un sous-espace de E que si l'un des deux sous-espaces F_1 et F_2 est inclus dans l'autre.

b) Soit $G = F_1 + F_2$, G est l'ensemble des $\vec{x} = \vec{u_1} + \vec{u_2}$, avec $\vec{u_1} \in F_1$, $\vec{u_2} \in F_2$. Un vecteur $\vec{v}$ de $F = F_1 \cup F_2$ est tel que $\vec{v} \in F_1$ ou $v \in F_2$ or $\vec{v} = \vec{v} + \vec{0}$ et $\vec{0} \in F_1$ et F_2, donc $\vec{v} \in G$ et $F \subset G$. Mais un sous-espace H contenant $F = F_1 \cup F_2$ doit contenir tout $\vec{u_1}$ de F_1 et tout $\vec{u_2}$ de F_2, donc contient $\vec{u_1} + \vec{u_2}$, il en résulte $G \subset H$. G, étant un sous-espace commun à tous les sous-espaces contenant F, est donc le plus petit sous-espace contenant $F_1 \cup F_2$.

■ **3.** Dans l'espace vectoriel E des fonctions $f(x)$ d'une variable réelle définies sur un intervalle $[-a, +a]$, que peut-on dire de l'ensemble P des fonctions paires et de l'ensemble I des fonctions impaires ?

P et I sont deux sous-espaces de E : la vérification (par le critère $f \in P$, $g \in P \Rightarrow \lambda F + \mu g \in P$) est immédiate. En écrivant

$$f(x) = \frac{1}{2}[f(x) + f(-x)] + \frac{1}{2}[f(x) - f(-x)]$$

on voit que $E = P + I$. $P \cap I = \{\varphi(x)\}$, telles que
$\varphi(-x) = \varphi(x) = -\varphi(x)$, donc $\varphi = 0$.

donc P et I sont supplémentaires dans E.

■ **4.** Soit E l'espace vectoriel des polynômes $P(x)$ à une variable, de degré $\leqslant k$. Montrer que les $k + 1$ polynômes

$$Q_0(x) = 1, \quad Q_1(x) = x + R_1(x), \quad Q_2(x) = x^2 + R_2(x), \ \ldots, \quad Q_k(x) = x^k + R_k(x)$$

où les $R_i(x)$ sont de degrés $\leqslant i - 1$, forment une base de E.

b) Application : en prenant $Q_0(x) = 1$, $Q_1(x) = x$, $Q_2(x) = x(x-1), \ldots, Q_k(x) = x(x-1)(x-2)\ldots(x-k+1)$ montrer que, si $P(x)$ est un polynôme de degré k, la fonction $f(z)$ définie par la série entière $f(z) = \sum_0^{+\infty} a_n P(n) z^n$ peut s'exprimer à l'aide des fonctions $\varphi(z)$, $\varphi'(z)$, $\varphi''(z)$, $\ldots$, $\varphi^{(k)}(z)$, avec $\varphi(z) = \sum_0^{+\infty} a_n z^n$, série de rayon de convergence $R > 0$.

Calculer explicitement $f(z) = \sum_0^{+\infty} \frac{(n+1)^3}{n!} z^n$.

c) $P(x)$ étant un polynôme donné de degré k, montrer que le système $P(x), P'(x), P''(x), \ldots, P^{(k)}(x)$ constitue une base de l'espace E, et déterminer les composantes, dans cette base, du polynôme $P(x + a)$, a donné.

---

a) Comme $\dim E = k + 1$, il suffit d'établir que les Q_i sont indépendants. Une relation de dépendance $\sum_0^k \alpha_i Q_i = 0$ s'écrit plus explicitement

$$\alpha_0 + \alpha_1 (x + R_1(x)) + \alpha_2 (x^2 + R_2(x)) + \ldots + \alpha_{k-1} (x^{k-1} + R_{k-1}(x)) \\ + \alpha_k (x^k + R_k(x)) = 0$$

En annulant le terme en x^k, il vient $\alpha_k = 0$. En annulant ensuite le terme en x^{k-1}, il vient $\alpha_{k-1} = 0$, et ainsi de suite. Tous les α_i sont donc nuls et $(Q_0, Q_1, \ldots, Q_k)$ est bien une base de E.

On peut aussi remarquer que, par rapport à la base canonique $(1, x, x^2, \ldots, x_k)$, la matrice de $(Q_0, Q_1, \ldots, Q_k)$ est triangulaire supérieure, de la forme

$$\begin{pmatrix} 1 & a_{12} & a_{13} & & a_{1n} \\ 0 & 1 & a_{23} & & a_{2n} \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Son déterminant est égal à 1, donc non nul, et les Q_i sont bien indépendants.

b) D'après (*a*), les polynômes

$$1, x, x(x-1), \ldots, x(x-1)\ldots(x-k+1)$$

forment une base de E, dans laquelle $P(x)$ peut se décomposer d'une façon unique :

$$P(x) = b_0 + b_1 x + b_2 x(x-1) + \ldots + b_k x(x-1)\ldots(x-k+1)$$

Donc

$$\begin{aligned} f(z) &= \sum_0^{+\infty} a_n [b_0 + b_1 n + b_2 n(n-1) + \ldots + b_k n(n-1) \\ &\qquad\qquad \ldots (n-k-1)] z^n \\ &= b_0 \sum_0^{+\infty} a_n z^n + b_1 \sum_1^{+\infty} n a_n z^n + b_2 \sum_2^{+\infty} n(n-1) a_n z^n \\ &\qquad + \ldots + b_k \sum_k^{+\infty} n(n-1)\ldots(n-k+1) a_n z^n \end{aligned}$$

Sur le disque ouvert $|z| < R$, $\varphi(z)$ est dérivable à un ordre quelconque, on a

$$\varphi'(z) = \sum_1^{+\infty} n a_n z^{n-1}, \quad \varphi''(z) = \sum_2^{+\infty} n(n-1) a_n z^{n-2}, \ldots, \\ \varphi^{(k)}(z) = \sum_k^{+\infty} n(n-1)\ldots(n-k+1) a_n z^{n-k}$$

donc on obtient

$$f(z) = b_0\, \varphi(z) + b_1 z\, \varphi'(z) + b_2 z^2\, \varphi''(z) + \ldots + b_k z^k\, \varphi^{(k)}(z)$$

En particulier

$$\sum_0^{+\infty} \frac{P(n)}{n\,!} z^n = (b_0 + b_1 z + b_2 z^2 + \ldots + b_k z^k)\, e^z \qquad (\forall z \in \mathbb{C})$$

$$P(n) = (n + 1)^3 = 1 + 3\,n + 3\,n^2 + n^3 = b_0 + b_1 n + b_2 n(n-1) + b_3 n(n-1)(n-2)$$

on obtient par identification $b_0 = 1$, $b_1 = 7$, $b_2 = 6$, $b_3 = 1$,

$$\sum_0^{+\infty} \frac{(n+1)^3}{n\,!} z^n = (z^3 + 6\,z^2 + 7\,z + 1)\,e^z$$

c) A des coefficients non nuls près, $P^{(k)}(x)$, $P^{(k-1)}(x)$, ..., $P(x)$ sont de la forme $Q_0(x)$, $Q_1(x)$, ..., $Q_k(x)$, suivant les notations du (*a*), donc constituent une base de E. La décomposition unique de $P(x + a)$, polynôme de degré k, dans cette base, est fournie par la formule de Taylor (sans reste, puisque $P^{(k+1)}(x) = 0$) :

$$P(x + a) = P(x) + aP'(x) + \frac{a^2}{2\,!} P''(x) + \dots + \frac{a^k}{k\,!} P^{(k)}(x)$$

■ 5. *a*) Dans l'ensemble des suites réelles, on pose $(u_n) + (v_n) = (u_n + v_n)$, $\lambda(u_n) = (\lambda u_n)$. Montrer que cet ensemble E est ainsi muni d'une structure d'espace vectoriel sur $\mathbb{R}$.

b) a et b étant deux réels donnés, montrer que les suites vérifiant la relation de récurrence $u_n = au_{n-1} + bu_{n-2}$ $(n \geqslant 2)$ constituent un sous-espace F de E.

c) A quelle condition existe-t-il deux suites géométriques indépendantes, éléments de F ? Application : soit $u_0 = 1$, $u_1 = 1$, $u_n = u_{n-1} + u_{n-2}$ (suite de Fibonacci); calculer u_n en fonction de n, ainsi que le rayon de convergence et la somme de la série entière $\sum_0^{+\infty} u_n z^n$.

---

a) Immédiat : il suffit de vérifier les axiomes de la définition d'un espace vectoriel.

b) On vérifie que si $u_n = au_{n-1} + bu_{n-2}$ et $v_n = av_{n-1} + bv_{n-1}$, et si on pose $w_n = \lambda u_n + \mu v_n$, alors $w_n = aw_{n-1} + bw_{n-2}$, donc toute combinaison linéaire de deux suites, éléments de F, est elle-même un élément de F, qui, ainsi, est un sous-espace de E.

c) Une suite géométrique $(u_n = \alpha r^n)$ est un élément de F si

$$\alpha r^n = a\alpha r^{n-1} + b\alpha r^{n-2}$$

α et r n'étant pas nuls puisqu'on cherche des suites indépendantes, il faut donc $r^2 - ar - b = 0$. Il nous faut deux valeurs réelles distinctes de r, d'où la condition $a^2 + 4\,b > 0$. On obtient ainsi deux suites (r_1^n) et (r_2^n), indépendantes car $\lambda r_1^n + \mu r_2^n = 0$ implique $\lambda + \mu = 0$ (pour $n = 0$) et $\lambda r_1 + \mu r_2 = 0$ (pour $n = 1$), d'où

$$\lambda = \mu = 0$$

Une suite (u_n) de F étant complètement définie par la donnée de u_0 et u_1, donc dépendant de deux paramètres, il semble que la dimension de F

soit 2. Il suffit, pour s'en assurer, de vérifier que (r_1^n) et (r_2^n) constituent une base de F, c'est-à-dire qu'il existe un seul couple (α, β) tel que $u_n = \alpha r_1^n + \beta r_2^n$: $n = 0$ et $n = 1$ donnent les relations $u_0 = \alpha + \beta$ et $u_1 = \alpha r_1 + \beta r_2$ qui définissent α, β puisque $\begin{vmatrix} 1 & 1 \\ r_1 & r_2 \end{vmatrix} = r_2 - r_1 \neq 0$.

La suite ainsi définie $(\alpha r_1^n + \beta r_2^n)$ est bien un élément de F, et tout élément de F est de cette forme : F est de dimension 2 et a pour base $[(r_1^n), (r_2^n)]$.

Application numérique : $u_n = u_{n-1} + u_{n-2}$ donne l'équation en r (appelée équation caractéristique de la suite récurrente linéaire envisagée) $r^2 - r - 1 = 0$, de racines $r_1 = \dfrac{1 + \sqrt{5}}{2}$, $r_2 = \dfrac{1 - \sqrt{5}}{2}$. u_n est donc de la forme $\alpha\left(\dfrac{1 + \sqrt{5}}{2}\right)^n + \beta\left(\dfrac{1 - \sqrt{5}}{2}\right)^n$, et $u_0 = u_1 = 1$ fournissent le système :

$$\begin{cases} \alpha + \beta = 1 \\ (\alpha + \beta) + (\alpha - \beta)\sqrt{5} = 2 \end{cases} \qquad \text{d'où } \alpha - \beta = \frac{1}{\sqrt{5}}$$

$$\alpha = \frac{1 + \sqrt{5}}{2\sqrt{5}}, \qquad \beta = -\frac{1 - \sqrt{5}}{2\sqrt{5}}$$

d'où
$$u_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1 + \sqrt{5}}{2}\right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2}\right)^{n+1}\right]$$

Remarquons que malgré les apparences u_n est un nombre entier.

La série entière $\sum_0^{+\infty} u_n z^n$ a pour rayon de convergence

$$R = \lim_{n \to \infty} \frac{u_n}{u_{n+1}}$$

si cette limite existe. Comme

$$\left|\frac{1 + \sqrt{5}}{2}\right| > 1 \quad \text{et} \quad \left|\frac{1 - \sqrt{5}}{2}\right| < 1,$$

on a, pour $n \to +\infty$,

$$u_n \sim \frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^{n+1} \quad \text{et} \quad \frac{u_n}{u_{n+1}} \to \frac{2}{1 + \sqrt{5}} = \frac{\sqrt{5} - 1}{2}$$

donc
$$R = \frac{\sqrt{5} - 1}{2}$$

Enfin

$$f(z) = \frac{1}{\sqrt{5}}\left(\sum_0^{+\infty} r_1^{n+1} z^n - \sum_0^{+\infty} r_2^{n+1} z^n\right) = \frac{1}{\sqrt{5}}\left(\frac{r_1}{1 - r_1 z} - \frac{r_2}{1 - r_2 z}\right)$$

$$= \frac{1}{\sqrt{5}} \cdot \frac{r_1 - r_2}{1 - (r_1 + r_2)z + r_1 r_2 z^2}$$

Or $r_1 - r_2 = \sqrt{5}$, $r_1 + r_2 = 1$ et $r_1 r_2 = -1$,

donc
$$f(z) = \frac{1}{1 - z - z^2}$$

Le lecteur pourra, prenant le problème « à l'envers », vérifier que la division suivant les puissances croissantes de 1 par $1 - z - z^2$ qui donne, sous réserve de convergence, le développement en série entière de cette fraction, fait bien apparaître des coefficients obéissant à la relation de récurrence $u_n = u_{n-1} + u_{n-2}$ avec $u_0 = u_1 = 1$:

$$f(z) = 1 + z + 2z^2 + 3z^3 + 5z^4 + \dots$$

■ **6.** f et g étant deux applications linéaires d'un espace vectoriel E de dimension finie dans un espace vectoriel F, établir les relations

$$|\operatorname{rang} f - \operatorname{rang} g| \leqslant \operatorname{rang}(f + g) \leqslant \operatorname{rang} f + \operatorname{rang} g$$

$$\forall \vec{u} \in E, \quad (f + g)(\vec{u}) = f(\vec{u}) + g(\vec{u}) \in f(E) + g(E) = \operatorname{Im} f + \operatorname{Im} g$$

Donc
$$\operatorname{Im}(f + g) \subset \operatorname{Im} f + \operatorname{Im} g$$

Or ces images sont de dimensions finies, puisque inférieures ou égales à n, dimension de E. Donc

$$\operatorname{rang}(f + g) \leqslant \operatorname{rang} f + \operatorname{rang} g \tag{2}$$

La seconde inégalité est donc établie. Pour la première, on peut, pour des raisons de symétrie, supposer $\operatorname{rang} f \geqslant \operatorname{rang} g$ et établir

$$\operatorname{rang} f \leqslant \operatorname{rang} g + \operatorname{rang}(f + g) \tag{1}$$

Or ceci résulte de (2) en posant $f + g = \varphi$ donc $f = \varphi - g$

$$(2) \Rightarrow \operatorname{rang}(\varphi - g) \leqslant \operatorname{rang} \varphi + \operatorname{rang}(-g)$$

et $\operatorname{rang}(-g) = \operatorname{rang} g$ car d'une façon générale

$$\operatorname{Im}(\lambda g) = \operatorname{Im} g, \quad \ker(\lambda g) = \ker g \quad \forall \lambda \neq 0$$

(1) est donc établie, et on a bien

$$|\operatorname{rang} f - \operatorname{rang} g| \leqslant \operatorname{rang}(f + g) \leqslant \operatorname{rang} f + \operatorname{rang} g$$

■ **7.** *a*) f et g étant deux endomorphismes de l'espace E de dimension n, établir la relation entre les rangs de f et g tels que $g \circ f = 0$.

b) On suppose $f \circ f = 0$. Quel est le rang maximum de f ?

c) Montrer que

$$\operatorname{Im} f = \ker f \Leftrightarrow \begin{cases} n \text{ pair} \\ \operatorname{rang} f = \dfrac{n}{2} \\ f \circ f = 0 \end{cases}$$

Indiquer une forme simple de la matrice A d'un tel endomorphisme f, en choisissant une base convenable.

---

a) $\forall\, \vec{u} \in E, \quad g\,[f(\vec{u})] = \vec{0}$ donc $\operatorname{Im} f \subset \ker g$ d'où

$$\operatorname{rang} f + \operatorname{rang} g \leqslant n \tag{1}$$

Un cas remarquable est celui où f et g sont deux projecteurs de E, l'un sur F parallèlement à un supplémentaire de G, l'autre G parallèlement à F : alors dans (1) il y a égalité. Mais (1) n'est pas suffisante pour que

$$g \circ f = 0 \quad \text{et} \quad g \circ f = 0 \not\Rightarrow f \circ g = 0$$

b) $\operatorname{Im} f = \ker f \Rightarrow 2 \dim \operatorname{Im} f = n$ donc n pair et rang $f = \dfrac{n}{2}$.

De plus $\forall\, \vec{u} \in E$, $f(\vec{u}) \in \operatorname{Im} f = \ker g$, donc $(f \circ f)\,(\vec{u}) = \vec{0}$, et $f \circ f = 0$. La réciproque est immédiate d'après (*a*).

c) Soit $n = 2\,p$. Soit $F = \operatorname{Im} f = \ker f$ de dimension p. Une base de E peut être obtenue par l'union d'une base de F et d'une base d'un supplémentaire F' de F dans E, qui a aussi pour dimension p. Choisissons cette base de F' : $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_p})$. Nous pouvons nous donner $f(\vec{e_1}) = \overrightarrow{e_{p+1}}$, $f(\vec{e_2}) = \overrightarrow{e_{p+2}}, \ldots, f(\vec{e_p}) = \overrightarrow{e_{2p}}$, vecteurs indépendants dans F, puisque l'image de F' par f est une partie de F. On a ainsi construit une base de E, $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_p}, \overrightarrow{e_{p+1}}, \ldots, \overrightarrow{e_{2p}})$ telle que $f(\vec{e_i}) = \overrightarrow{e_{p+i}}$ si $0 \leqslant i \leqslant p$; il suffit de se donner $f(\vec{e_i}) = \vec{0}$ si $p + 1 \leqslant i \leqslant 2\,p$, pour exprimer que $F = \ker f$.

Dans une telle base, la matrice de f est

$$\begin{array}{cl}
A = \begin{pmatrix}
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & & & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & \vdots & & \vdots \\
\vdots & & \ddots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & \cdots & 0
\end{pmatrix} &
\begin{array}{l}
\\ \\ \leftarrow \text{ligne } p \\ \leftarrow \text{ligne } p+1 \\ \\ \\ \\
\end{array} \\
\begin{array}{cc} \uparrow & \uparrow \\ \text{colonne } p & \text{colonne } p+1 \end{array} &
\end{array}$$

Les seuls termes non nuls de A — et alors égaux à 1 — sont ceux de la forme $a_{i+p,i}$, $i = 1, 2, \ldots, p$. On peut aussi écrire une telle matrice sous la forme « décomposée en blocs » :

$$A = \left(\begin{array}{c|c} 0 & 0 \\ \hline I_p & 0 \end{array}\right)$$

On peut encore remarquer que, par construction même, $F' = f(F)$. Nous avons construit un endomorphisme d'image et de noyau donnés dans E. L'exercice suivant généralise ce problème.

L'exemple le plus simple de matrice A est celui obtenu en dimension 2,

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

■ **8.** Soit E un espace vectoriel de dimension finie n.

a) Soit F et G deux sous-espaces de E, de même dimension p; déterminer un isomorphisme φ de E (c'est-à-dire un endomorphisme bijectif) dans lequel l'image de F par φ est G.

b) φ étant un isomorphisme de E, et f un endomorphisme quelconque de E, montrer que $\operatorname{rang}(\varphi \circ f) = \operatorname{rang}(f \circ \varphi) = \operatorname{rang} f$.

c) Soit F et K deux sous-espaces de E, de dimensions p et q; à quelle condition existe-t-il un endomorphisme g de E tel que

$$\operatorname{Im} g = F, \qquad \ker g = K\ ?$$

---

a) F' et G' étant deux supplémentaires de F et G respectivement, une base de E est l'union d'une base de F et d'une base de F'. Si nous donnons, à une base de F, pour image, une base de G (ce qui est possible puisque $\dim F \geqslant \dim G$), et, à une base de F', pour image, une base de G', l'endomorphisme φ ainsi défini est bijectif car l'image d'une base de E est une base de E. De plus $\varphi(\vec{u}) \in G$ si et seulement si $\vec{u} \in F$, donc $G = \varphi(F)$. φ répond donc à la question.

b) Soit $\operatorname{Im} f = F$. $\forall\, \varphi$, endomorphismes de E,

$$\dim \varphi(F) \leqslant \dim F = p$$

Si $\dim \varphi(F) < p$, une base de E, union d'une base de F et d'une base d'un de ses supplémentaires F', aura pour image l'union de deux systèmes de vecteurs de $\varphi(F)$ et de $\varphi(F')$, dont le rang sera $\leqslant \dim \varphi(F) + \dim \varphi(F') < p + \dim \varphi(F') \leqslant p + n - p = n$, ce ne sera pas une base de E. On a donc

$$\dim \varphi(F) = p = \dim [(\varphi \circ f)(E)] = \dim \operatorname{Im}(\varphi \circ f) = \dim \operatorname{Im} f$$

Ainsi $\operatorname{rang}(\varphi \circ f) = \operatorname{rang} f$.

Démonstration analogue pour $\operatorname{rang}(f \circ \varphi) = \operatorname{rang} f$.

Il en résulte que la multiplication à droite ou à gauche par une matrice régulière respecte le rang d'une matrice.

c) Une condition nécessaire pour que $F = \operatorname{Im} g$ et $K = \ker g$ est évidemment $\dim F + \dim K = n$, soit $p + q = n$.

On peut alors considérer deux cas :

– un cas particulier où F et K sont supplémentaires, une solution du problème est alors le projecteur de E sur F parallèlement à K;

– le cas général : on peut alors désigner par G un supplémentaire de K dans E. On a $\dim G = \dim F = n - q = p$. Il existe un projecteur f de E sur G parallèlement à K et un isomorphisme φ de E tel que $\varphi(G) = F$.

Si nous composons f par φ, l'endomorphisme $g = \varphi \circ f$ a même rang p que f ; on a $\forall\ \vec{u} \in E,\ f(\vec{u}) \in G$, donc $\varphi\ [f(\vec{u})] \in \varphi(G) = F$ c'est-à-dire $\operatorname{Im} g \subset F$ et comme $\dim \operatorname{Im} g = p$, $\operatorname{Im} g = F$.

De plus $\forall\ \vec{v} \in K,\ f(\vec{v}) = \vec{0}$ donc $\varphi\ [f(\vec{v})] = \vec{0}$, $\vec{v} \in \ker g$; donc $K \subset \ker g$ et comme $\dim \ker g = n - p = q$, $\ker g = K$.

Nous avons donc déterminé un endomorphisme g de E tel que

$$\operatorname{Im} g = F, \qquad \ker g = K$$

sous-espaces donnés de dimensions p et $q = n - p$.

■ **9.** Déterminer le noyau et l'image de l'application linéaire définie par

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 2 \\ 4 & 5 & -7 \\ 0 & 3 & -5 \end{pmatrix}$$

Le système

$$\begin{cases} 2x + y - z = 0 \\ x - y + 2z = 0 \\ 4x + 5y - 7z = 0 \\ \phantom{4x + {}} 3y - 5z = 0 \end{cases}$$

se réduit à deux équations $x - y + 2z = 0$, $3y - 5z = 0$.

Le noyau, dans $\mathbb{R}^3$, est la droite vectorielle de base $\vec{u}\ (1, -5, -3)$.
L'image est un sous-espace de dimension 2 dans $\mathbb{R}^4$. On obtient ses équations en X, Y, Z, T en éliminant x, y, z entre

$$\left\{\begin{array}{lll} 2x + y - z = X & (1) & \\ x - y + 2z = Y & (2) & (\times(-2))^{+} \qquad (\times(-4))^{+} \qquad (*) \\ 4x + 5y - 7z = Z & (3) & \\ 3y - 5z = T & (4) & \\ (1) \text{ et } (2) \Rightarrow 3y - 5z = X - 2Y, & \text{d'où } X - 2Y = T & \\ (2) \text{ et } (3) \Rightarrow 9y - 15z = Z - 4Y, & \text{d'où } Z - 4Y = 3T & \end{array}\right.$$

(les flèches ci-dessus indiquent les combinaisons effectuées).

L'image est donc définie par $3X - 6Y = -4Y + Z = 3T$ c'est bien un sous-espace de dimension 2 dans $\mathbb{R}^4$, dont une base est constituée par les vecteurs $\vec{v_1}\ (3, 0, 3, 1)$, $\vec{v_2}\ (2, 1, 4, 0)$.

(*) La notation $(\times(-2))^{+}$ signifie que l'on multiplie la ligne (2) par -2 et que l'on ajoute le résultat à la ligne (1).

■ **10.** Soit f un projecteur de l'espace vectoriel E. A quelle condition l'endomorphisme $\varphi(t) = e + tf$, où t est un scalaire et e l'endomorphisme-unité, est-il inversible ?

Déterminer alors $\varphi^{-1}(t)$. On supposera $f \neq 0$ et $f \neq e$.

On sait que $f \circ f = f$. D'autre part si l'inverse de φ existe, il est unique.

Cherchons s'il est de la forme $\varphi(\lambda) = e + \lambda f$

$$\varphi(t) \circ \varphi(\lambda) = (e + tf) \circ (e + \lambda f) = e + tf + \lambda f + t\lambda f$$
$$= e + (t + \lambda + t\lambda) f$$

On veut $\varphi(t) \circ \varphi(\lambda) = e$ donc $t + \lambda + t\lambda = 0$ $\lambda(1 + t) = -t$.

Donc si $t \neq -1$ $\lambda = -\dfrac{t}{1+t}$ donne l'inverse de $\varphi(t)$:

$$[\varphi(t)]^{-1} = e - \frac{t}{1+t} f$$

Si $t = -1$, $\varphi(t) = e - f$ n'est pas inversible car c'est un diviseur de zéro : $f \circ (e - f) = f - f \circ f = 0$.

■ **11.** Soit a, b, c trois réels tels que $a^2 + b^2 + c^2 = 1$.

On pose $P = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix}$ $Q = I - P$ (I = matrice-unité).

a) Déterminer $\ker P$ et $\operatorname{Im} P$ dans $\mathbb{R}^3$.

b) Calculer P^2, PQ, QP et Q^2.

c) Montrer que l'ensemble des matrices $A(\alpha, \beta) = \alpha P + \beta Q$ (α, β réels) est un sous-anneau de $\mathcal{M}_3$, trouver A^{-1} lorsqu'elle existe, et indiquer les matrices A singulières.

d) Caractériser géométriquement P et Q dans $\mathbb{R}^3$ supposé orthonormé, (a, b, c) étant les coordonnées d'un vecteur $\vec{u}$.

Réponses :

a) $\ker P$ est le plan $ax + by + cz = 0$,

$\operatorname{Im} P$ la droite $$\frac{x}{a} = \frac{y}{b} = \frac{z}{c} \qquad (1)$$

b) $P^2 = P$, $PQ = QP = 0$, $Q^2 = Q$.

c) L'addition et la multiplication sont internes; $A^{-1} = \dfrac{P}{\alpha} + \dfrac{Q}{\beta}$, si $\alpha\beta \neq 0$. Les matrices singulières sont celles de la forme αP ou βQ.

d) P est le projecteur orthogonal sur Δ, de base $\vec{u}$; Q le projecteur orthogonal sur le plan π, perpendiculaire à Δ.

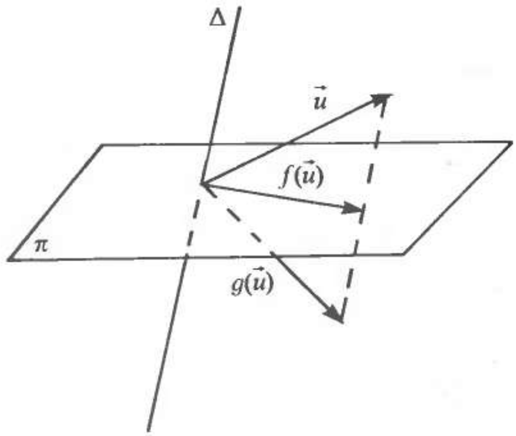
■ **12.** f étant un projecteur de E, soit $g = 2f - e$, e étant l'endomorphisme-unité. Montrer que g est involutif. Réciproque. Interprétation géométrique dans $\mathbb{R}^3$. Généraliser cette interprétation.

Réponses : $f \circ f = f \Rightarrow g \circ g = e$.

Réciproquement soit g involutif, on peut poser $f = \dfrac{e+g}{2}$ et vérifier que $g \circ g = e \Rightarrow f \circ f = f$.

Si f est un projecteur sur π (droite ou plan) parallèlement à Δ (plan ou droite), g est la symétrie par rapport à π, parallèlement à Δ (symétrie orthogonale si $\Delta \perp \pi$).

En dimension n quelconque, à tout projecteur sur un sous-espace F de E parallèlement à un supplémentaire G de F, on peut associer la symétrie par rapport à F parallèlement à E, définie par $g(\vec{u}) = 2f(\vec{u}) - \vec{u}$.



■ **13.** $\mathbb{R}^3$ étant rapporté à un repère orthonormé direct, soit $\vec{u}$ le vecteur (a, b, c), non nul. Montrer que l'application φ définie par $\varphi(\vec{v}) = \vec{u} \wedge \vec{v}$ est un endomorphisme de $\mathbb{R}^3$. Déterminer $\ker \varphi$ et $\operatorname{Im} \varphi$. Déterminer $\varphi^2(\vec{v})$, $\varphi^3(\vec{v})$, etc. Généraliser. Vérifier ces résultats en utilisant la matrice de φ.

Réponses : $\varphi(\alpha\vec{v} + \beta\vec{w}) = \alpha\varphi(\vec{v}) + \beta\varphi(\vec{w})$ φ est un endomorphisme. $\ker \varphi$ est la droite vectorielle de base $\vec{u}$, $\operatorname{Im} \varphi$ le plan vectoriel perpendiculaire.

$$\varphi^2(\vec{v}) = \vec{u} \wedge (\vec{u} \wedge \vec{v}) = (\vec{u} \cdot \vec{v})\,\vec{u} - u^2 \cdot \vec{v}$$

(formule du double produit vectoriel).

$\varphi^3(\vec{v}) = -u^2 . \varphi(\vec{v})$ donc $\varphi^3 = -u^2 . \varphi = -(a^2 + b^2 + c^2)\varphi$

D'où $\varphi^{2k+1} = -(a^2 + b^2 + c^2)^k\varphi$, $\varphi^{2k} = -(a^2 + b^2 + c^2)^{k-1}\varphi^2$.

La matrice de φ est

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \qquad \text{(elle est antisymétrique)}$$

$$A^2 = \begin{pmatrix} -c^2 - b^2 & ab & ac \\ ab & -c^2 - a^2 & bc \\ ac & bc & -a^2 - b^2 \end{pmatrix}$$

$$A^3 = \begin{pmatrix} 0 & c(a^2 + b^2 + c^2) & -b(a^2 + b^2 + c^2) \\ -c(a^2 + b^2 + c^2) & 0 & a(a^2 + b^2 + c^2) \\ b(a^2 + b^2 + c^2) & -a(a^2 + b^2 + c^2) & 0 \end{pmatrix}$$

$= -u^2A$, etc.

■ **14.** Montrer que toute matrice carrée est la somme d'une matrice symétrique et d'une matrice antisymétrique.

Réponse : $A = \dfrac{1}{2}(A + {}^tA) + \dfrac{1}{2}(A - {}^tA)$.

■ **15.** A et B étant deux matrices carrées d'ordre n, I la matrice-unité, on suppose que $I - AB$ est inversible. Vérifier que

$$(I - BA)^{-1} = I + B(I - AB)^{-1}A$$

Réponse : il suffit de vérifier $(I - BA)[I + B(I - AB)^{-1}A] = I$.

■ **16.** M étant un point (x, y, z) de $\mathbb{R}^3$ rapporté à un repère orthonormé direct, A une matrice d'ordre n, caractériser les matrices A telles que le champ vectoriel $\vec{V}(M) = A\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ soit à circulation et à flux conservatifs.

Réponse : En posant $A = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}$, on obtient les conditions

$$b'' = c', \quad a'' = c, \quad a' = b$$

(provenant de $\overrightarrow{\text{rot}\,\vec{V}} = \vec{0}$) et $a + b' + c'' = 0$ (provenant de $\text{div}\,\vec{V} = 0$).

A doit donc être une matrice symétrique de trace nulle (voir exercice suivant).

■ **17.** On appelle trace d'une matrice carrée A, et on note tr A, la somme des termes diagonaux de A : $\operatorname{tr} A = \sum_{1}^{n} a_{ii}$.

Démontrer que $\operatorname{tr}(BA) = \operatorname{tr}(AB)$. En déduire que

$$A' \sim A \Rightarrow \operatorname{tr} A' = \operatorname{tr} A$$

Réponse :
$$\operatorname{tr}(AB) = \sum_{i=1}^{i=n}\left(\sum_{k=1}^{k=n} a_{ik}b_{ki}\right) = S = \sum_{\substack{1 \leq i \leq n \\ 1 \leq k \leq n}} a_{ik}b_{ki}$$

$$\operatorname{tr}(BA) = \sum_{i=1}^{i=n}\left(\sum_{k=1}^{k=n} b_{ik}a_{ki}\right) = S$$

Soit $A' = P^{-1}AP$, $\operatorname{tr} A' = \operatorname{tr}\left[(P^{-1}A)\,P\right] = \operatorname{tr}\left[P\,(P^{-1}A)\right] = \operatorname{tr} A$.

■ **18.** Déterminer l'inverse de la matrice

$$M = \begin{pmatrix} 1 & 2 & 3\dots & & n \\ 0 & 1 & 2\dots & & n-1 \\ 0 & 0 & 1\dots & & n-2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0\dots & 0 & 1 \end{pmatrix}$$

On cherche à résoudre le système

$$(S) \begin{cases} x_1 + 2x_2 + 3x_3 + \dots + nx_n = y_1 & (1) \\ \quad x_2 + 2x_3 + \dots + (n-1)\,x_n = y_2 & (2) \\ \qquad x_3 + \dots + (n-2)\,x_n = y_3 & (3) \\ \qquad \dots\dots\dots\dots & \\ \qquad\quad x_{n-1} + 2x_n = y_{n-1} & (n-1) \\ \qquad\qquad x_n = y_n & (n) \end{cases}$$

C'est un système « triangulaire » ou « en cascade ».

En retranchant (2) de (1), puis (3) de (2), ..., (n) de $(n-1)$, et conservant (n), on forme le nouveau système équivalent (S') :

$$(S') \begin{cases} x_1 + x_2 + x_3 + \dots + x_n = y_1 - y_2 & (1)' \\ \quad x_2 + x_3 + \dots + x_n = y_2 - y_3 & (2)' \\ \qquad x_3 + \dots + x_n = y_3 - y_4 & (3)' \\ \qquad \dots\dots\dots\dots\dots\dots\dots\dots & \\ \qquad x_{n-1} + x_n = y_{n-1} - y_n & (n-1)' \\ \qquad\quad x_n = y_n & (n) \end{cases}$$

En retranchant encore (2)′ de (1)′, (3)′ de (2)′ et ainsi de suite, on obtient :

$$\begin{cases} x_1 = y_1 - 2\,y_2 + y_3 \\ x_2 = y_2 - 2\,y_3 + y_4 \\ \dots\dots\dots\dots\dots\dots\dots \\ x_{n-2} = y_{n-2} - 2\,y_{n-2} + y_n \\ x_{n-1} = y_{n-1} - 2\,y_n \\ x_n = y_n \end{cases}$$

d'où la matrice cherchée :

$$M^{-1} = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & -2 & 1 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \dots & \dots & 0 & 1 & -2 \\ 0 & 0 & 0 & \dots & \dots & 0 & 0 & 1 \end{pmatrix}$$

Cette matrice est triangulaire, comme on pouvait le prévoir, puisque l'ensemble des matrices triangulaires supérieures est un sous-anneau de l'anneau $\mathcal{M}_n$.

■ **19.** Soit f un endomorphisme de l'espace vectoriel E de dimension n.

Posant $f \circ f = f^2$, $f \circ f \circ f = f^3$, etc., $\operatorname{Im} f^p = J_p$, $\ker f^p = K_p$, démontrer les inclusions $J_{p+1} \subset J_p$, $K_p \subset K_{p+1}$ (au sens large).

En posant $u_p = \dim J_p$ montrer que la suite (u_p) est convergente et en déduire l'existence d'un entier s tel que

$$E = J_s \oplus K_s$$

$f^{p+1} = f \circ f^p = f^p \circ f$ donc $\vec{u} \in J_{p+1} \Rightarrow \exists\ \vec{v}$ tel que $\vec{u} = f^p\ [f(\vec{v})]$ donc $\exists\ \vec{v_1} = f(\vec{v}) \in E$ tel que $\vec{u} = f^p(\vec{v_1})$, d'où

$$\vec{u} \in J_p \quad \text{et} \quad J_{p+1} \subset J_p$$

De même

$$\vec{z} \in K_p \Rightarrow f^p(\vec{z}) = \vec{0} \Rightarrow (f \circ f^p)(\vec{z}) = \vec{0} \Rightarrow \vec{z} \in K_{p+1}$$

donc
$$K_p \subset K_{p+1}$$

En posant $u_p = \dim J_p$, $J_{p+1} \subset J_p \Rightarrow u_{p+1} \leqslant u_p$.

La suite (u_p) est donc décroissante (au sens large) et minorée par 0, donc convergente ; soit l sa limite : $\forall\ \varepsilon > 0$, en particulier $\varepsilon \in]0, 1[$, $\exists\ s$ entier tel que $p \geqslant s \Rightarrow |u_p - l| < \varepsilon$; u_p étant entier, l l'est forcément aussi et $u_p = l$ pour $p \geqslant s$.

Donc si $p \geqslant s$, $u_p = u_s = l$, et comme $J_p \subset J_{p-1} \subset \dots \subset J_s$

$$\dim J_p = \dim J_s \Rightarrow J_p = J_s$$

De même $K_s \subset K_p$ et $\dim K_s = n - \dim J_s = n - \dim J_p = \dim K_p$ entraînent $K_s = K_p$.

Ainsi il y a stabilisation des images et noyaux des f^p à partir de $p = s$.

Soit alors $\vec{v} \in K_s \cap J_s$. On a $\vec{v} \in K_s$ donc $f^s(\vec{v}) = \vec{0}$, $\vec{v} \in J_s$, donc $\exists \vec{w}$ tel que $\vec{v} = f^s(\vec{w})$.

Il s'ensuit $f^{2s}(\vec{w}) = \vec{0}$ donc $\vec{w} \in K_{2s}$.

Mais $K_{2s} = K_s$ donc $\vec{w} \in K_s$ et $f^s(\vec{w}) = \vec{v} = \vec{0}$. Ainsi

$$K_s \cap J_s = \{ \vec{0} \}$$

et comme $\dim K_s + \dim J_s = n$, K_s et J_s sont supplémentaires, autrement dit $E = J_s \oplus K_s$.

Remarques. Deux cas étaient *a priori* immédiats, celui où $f = 0$, et celui où f est un isomorphisme : dans le premier $J_p = \{ \vec{0} \}$ et $K_p = E$, dans le second $J_p = E$ et $K_p = \{ \vec{0} \}$; dans les deux cas $s = 0$.

Un autre cas intéressant est celui où $l = 0$, c'est-à-dire où il existe s tel que $f^s = 0$, $f^{s-1} \neq 0$. Alors on dit que f est un endomorphisme nilpotent (« de puissance nulle ») d'ordre s, $J_s = \{ \vec{0} \}$ et $K_s = E$. Enfin si f est un projecteur, $f^p = f$, $\forall\, p \geqslant 1$ et $s = 1$, effectivement on sait que $E = \operatorname{Im} f \oplus \ker f$.

■ **20.** *a)* Soit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Démontrer qu'il existe deux scalaires α et β tels que $A^2 = \alpha A + \beta I$.

b) Soit $A = \begin{pmatrix} 2 & 1 \\ -5 & -4 \end{pmatrix}$. Démontrer que A^n est de la forme $u_n A + v_n I$ où u_n et v_n sont des scalaires que l'on déterminera en fonction de n.

c) A gardant la valeur précédente, démontrer que la série $(\sum U_n)$, où U_n est la matrice $\dfrac{A^n}{n\,!}$, est convergente, et calculer sa somme, que l'on notera $e^A = \sum_0^{+\infty} \dfrac{A^n}{n\,!}$.

Réponses :

a) On est conduit au système

$$\begin{aligned} a^2 + bc &= \alpha a + \beta \\ c\,(a + d - \alpha) &= 0 \\ b\,(a + d - \alpha) &= 0 \\ d^2 + bc &= \alpha d + \beta \end{aligned}$$

Si $(b, c, a - d) \neq (0, 0, 0)$ on obtient $\alpha = a + d$, $\beta = bc - ad$.

Si $b = c = 0$ et $a = d$, $A = aI$, $A^2 = a^2 I$, α est arbitraire et $\beta = a^2 - \alpha a$.

Dans tous les cas $A^2 = (a + d)\,A + (bc - ad)\,I$.

b) $A^0 = I$, $u_0 = 0$, $v_0 = I$; $A^1 = A$, $u_1 = 1$, $v_1 = 0$; $A^2 = -2A + 3I$, $u_2 = -2$, $v_2 = 3$.

Par récurrence on obtient, en posant $A^{n-1} = u_{n-1}A + v_{n-1}I$

$$A^n = u_{n-1}A^2 + v_{n-1}A$$

qui donne

$$u_n = -2u_{n-1} + v_{n-1}, \quad v_n = 3u_{n-1}$$

d'où $$u_n + v_n = u_{n-1} + v_{n-1} = \ldots = u_0 + v_0 = 1,$$

donc $$v_n = -3v_{n-1} + 3$$

que l'on écrit $v_n - \lambda = -3(v_{n-1} - \lambda)$, d'où $\lambda = \frac{3}{4}$.

Il en résulte $v_n = \frac{1}{4}[3 - (-3)^n]$, $u_n = \frac{1}{4}[1 + (-3)^n]$ et

$$A^n = \frac{1}{4}([1 + (-3)^n]A + [3 - (-3)^n]I)$$

c) $U_n = \frac{1}{4}\frac{A + 3I}{n!} + \frac{1}{4}\frac{(A - I)(-3)^n}{n!}$

Comme $\sum_0^{+\infty}\frac{1}{n!} = e$ et $\sum_0^{+\infty}\frac{(-3)^n}{n!} = e^{-3}$, $\sum_0^{+\infty} U_n$ existe et vaut $\frac{1}{4}(A + 3I)e + \frac{1}{4}(A - I)e^{-3}$, par définition on aura donc

$$e^A = \sum_0^{+\infty}\frac{A^n}{n!} = \frac{1}{4}[(A + 3I)e + (A - I)e^{-3}]$$

$$e^A = \frac{1}{4}\begin{pmatrix} 5e + e^{-3} & e + e^{-3} \\ -5(e + e^{-3}) & -e - 5e^{-3} \end{pmatrix}$$

On peut évidemment concevoir une généralisation sous forme de définition d'une fonction d'une variable matricielle par une série entière.

■ **21.** Soit $\alpha = e^{\frac{2i\pi}{n}}$ (racine $n^{\text{ième}}$ de l'unité).

On pose

$$M(\alpha) = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \alpha & \alpha^2 & \ldots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \ldots & \alpha^{2(n-1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha^{n-1} & \alpha^{2(n-1)} & \ldots & \alpha^{(n-1)^2} \end{pmatrix}$$

Calculer $M(\alpha) \,.\, M\left(\frac{1}{\alpha}\right)$ et en déduire $[M(\alpha)]^{-1}$.

Réponse : en observant que tous les termes du produit $M(\alpha) \,.\, M\left(\frac{1}{\alpha}\right)$ sont des sommes de progressions géométriques de la forme

$$1 + \alpha^p + (\alpha^p)^2 + \dots + (\alpha^p)^{n-1}$$

on obtient

$$M(\alpha)\, M\left(\frac{1}{\alpha}\right) = nI$$

et

$$[M(\alpha)]^{-1} = \frac{1}{n} M\left(\frac{1}{\alpha}\right)$$

Comme $\frac{1}{\alpha} = \bar{\alpha}$, $M\left(\frac{1}{\alpha}\right)$ est la conjuguée de $M(\alpha)$, ou $\overline{M(\alpha)}$.

■ **22.** On appelle endomorphisme nilpotent d'ordre k un endomorphisme φ tel que $\varphi^k = 0$, $\varphi^{k-1} \neq 0$. Même définition pour une matrice nilpotente d'ordre k.

a) Soit A une matrice triangulaire supérieure dont les éléments diagonaux sont nuls, c'est-à-dire de la forme

$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & & \vdots & \ddots & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \qquad \text{(ordre } n\text{)}$$

Soit φ l'endomorphisme associé. En utilisant $(\varphi \circ \varphi)(\vec{e_i})$ pour $i = 1, 2, \dots, n$, trouver la forme de A^2, puis celle de A^3 et ainsi de suite. En déduire que A est nilpotente d'ordre au plus égal à n.

b) Application : soit la matrice d'ordre 4

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Calculer B^p pour p entier quelconque.

---

a) $\varphi(\vec{e_1}) = \vec{0}$, $\varphi(\vec{e_2}) \in \text{Vect}\,(\vec{e_1})$, $\varphi(\vec{e_3}) \in \text{Vect}\,(\vec{e_1}, \vec{e_2})$, ...,
$\varphi(\vec{e_n}) \in \text{Vect}\,(\vec{e_1}, \vec{e_2}, \dots, \overrightarrow{e_{n-1}})$

Posons $\text{Vect}\,(\vec{e_1}, \vec{e_2}, \dots, \vec{e_i}) = E_i$ on a donc $\varphi(\overrightarrow{e_{i+1}}) \in E_i$, $i \geqslant 1$.

Donc $\varphi^2(\vec{e_1}) = \vec{0}$, $\varphi^2(\vec{e_2}) = \vec{0}$, $\varphi^2(\vec{e_3}) \in E_1$, ..., $\varphi^2(\vec{e_n}) \in E_{n-2}$.

Dans la matrice A^2, la diagonale de zéros est elle-même bordée supérieurement par des zéros. De même

$$\varphi^3(\vec{e_1}) = \varphi^3(\vec{e_2}) = \varphi^3(\vec{e_3}) = \vec{0}, \quad \varphi^3(\vec{e_4}) \in E_1, \ldots, \varphi^3(\vec{e_n}) \in E_{n-3}$$

Dans A^3, les zéros bordant supérieurement la diagonale seront eux-mêmes bordés de zéros, et ainsi de suite. Raisonnons par récurrence, en supposant

$$\varphi^p(\vec{e_1}) = \varphi^p(\vec{e_2}) = \ldots = \varphi^p(\vec{e_p}) = \vec{0}$$

$$\varphi^p(\overrightarrow{e_{p+1}}) \in E_1, \quad \varphi^p(\overrightarrow{e_{p+2}}) \in E_2, \ldots, \varphi^p(\vec{e_n}) \in E_{n-p}$$

alors

$$\varphi^{p+1}(\vec{e_1}) = \varphi^{p+1}(\vec{e_2}) = \ldots = \varphi^{p+1}(\overrightarrow{e_{p+1}}) = \vec{0}$$

$$\varphi^{p+1}(\overrightarrow{e_{p+2}}) \in E_1, \ldots, \varphi^{p+1}(\vec{e_n}) \in E_{n-p-1}$$

Il y a donc « remontée » progressive des zéros, on aura
$\varphi^{n-1}(\vec{e_1}) = \varphi^{n-1}(\vec{e_2}) = \ldots = \varphi^{n-1}(\overrightarrow{e_{n-1}}) = \vec{0}, \quad \varphi^{n-1}(\vec{e_n}) \in E_1$ et $\varphi^{(n)}(\vec{e_i}) = \vec{0}, \forall i$, donc $\varphi^n = 0$ et $A^n = 0$; $A^p = 0, \forall p \geqslant n$.

A est bien nilpotente d'ordre au plus égal à n.

b) Écrivons $B = I + A$ avec

$$A = \begin{pmatrix} 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A^2 = \begin{pmatrix} 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A^3 = \begin{pmatrix} 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad A^4 = 0$$

$$(I + A)^p = I + C_p^1 A + C_p^2 A^2 + C_p^3 A^3$$

valable $\forall p \in \mathbb{N}$ car, si $p < q$, $C_p^q = 0$.

Donc $\forall p \in \mathbb{N}$

$$B^p = \begin{pmatrix} 1 & 2p & p(2p+1) & \frac{2p}{3}(p+1)(2p+1) \\ 0 & 1 & 2p & p(2p+1) \\ 0 & 0 & 1 & 2p \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

En fait le lecteur pourra vérifier que cette formule est encore valable lorsque p est un entier négatif.

■ **23.** Soit la matrice d'ordre n

$$A = \begin{pmatrix} b & a & a \dots a \\ a & b & a \dots a \\ a & a & b \dots a \\ \vdots & \vdots & \vdots \ddots \vdots \\ a & a & a \dots b \end{pmatrix}$$

a) Calculer A^{-1} lorsqu'elle existe.

b) Calculer A^k, k entier.

---

a) En écrivant le système

$$\begin{cases} bx_1 + ax_2 + ax_3 + \dots + ax_n = y_1 & (1) \\ ax_1 + bx_2 + ax_3 + \dots + ax_n = y_2 & (2) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ ax_1 + ax_2 + ax_3 + \dots + bx_n = y_n & (n) \end{cases}$$

on obtient par addition de toutes ces équations

$$[b + (n-1)a] \sum_1^n x_i = \sum_1^n y_i \qquad \text{(I)}$$

Or $$(1) \Leftrightarrow (b-a)x_1 + a\sum_1^n x_i = y_i$$

Donc, si $b \neq (1-n)a$, $\displaystyle\sum_1^n x_i = \frac{\sum_1^n y_i}{b + (n-1)a}$

Si, en outre, $b \neq a$ $\displaystyle x_1 = \frac{1}{b-a}\left[y_1 - \frac{a}{b+(n-1)a}\sum_1^n y_i\right]$

$$x_1 = \frac{1}{(b-a)[b+(n-1)a]} \left[(b + (n-2)a)y_1 - a\sum_2^n y_i\right]$$

De même on obtient

$$x_j = \frac{1}{(b-a)[b+(n-1)a]}\left[(b+(n-2)a)y_j - a\sum_{i \neq j} y_i\right]$$

La matrice inverse est donc, si $b \neq a$ et $b \neq (1-n)a$,

$$A^{-1} = \frac{1}{(b-a)[b+(n-1)a]} \begin{pmatrix} b+(n-2)a & -a & -a & \dots & -a \\ -a & b+(n-2)a & -a & \dots & -a \\ -a & -a & b+(n-2)a & \dots & -a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a & -a & -a & \dots & b+(n-2)a \end{pmatrix}$$

Si $b = (1-n)a$ ou $b = a$ (I) ou (1) est impossible pour des y_i quelconques, la matrice A n'est pas inversible : d'ailleurs on constate facilement, alors, que ses vecteurs-colonnes sont liés.

b) Si $b = a$, $A = aJ$, avec

$$J = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \qquad (a_{ij} = 1, \quad \forall\, i, j)$$

Or $J^2 = nJ$, $J^3 = n^2 J, \dots, J^k = n^{k-1} J$.

Donc, si $b = a$, $A^k = a^k n^{k-1} J$.

Pour b quelconque, on peut écrire $A = (b-a)\,I + aJ$ et la formule du binôme s'applique puisque I commute avec toute matrice. Donc

$$A^k = (b-a)^k I + C_k^1\, a(b-a)^{k-1} J + C_k^2\, a^2(b-a)^{k-2} J^2 + \dots$$

$$A^k = (b-a)^k I + C_k^1\, a(b-a)^{k-1} J + C_k^2\, a^2(b-a)^{k-2} nJ + \dots$$
$$+ C_k^p\, a^p(b-a)^{k-p} n^{p-1} J + \dots + C_k^k a^k n^{k-1} J$$

$$A_k = (b-a)^k I + \frac{1}{n}\left(\sum_{p=1}^{k} C_k^p\,(na)^p\,(b-a)^{k-p}\right) J$$

soit

$$A^k = (b-a)^k I + \frac{1}{n}\left[(b + (n-1)a)^k - (b-a)^k\right] J$$

A^k présente donc la même forme que A :

$$A^k = \begin{pmatrix} \beta_k & \alpha_k & \dots & \alpha_k \\ \alpha_k & \beta_k & \dots & \alpha_k \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_k & \alpha_k & \dots & \beta_k \end{pmatrix}$$

avec

$$\alpha_k = \frac{1}{n}\left[(b + (n-1)\,a)^k - (b-a)^k\right]$$

$$\beta_k = \frac{1}{n}\left[(b + (n-1)\,a)^k + (n-1)\,(b-a)^k\right]$$

On vérifie que ces formules sont encore valables pour $k = -1$ (question (*a*)), et on pourrait démontrer qu'elles le sont encore pour k entier négatif.

CHAPITRE 7

DÉTERMINANTS. SYSTÈMES D'ÉQUATIONS LINÉAIRES

A. PERMUTATIONS D'UN ENSEMBLE FINI

1. Rappel

Soit E un ensemble fini, que l'on peut noter $\{ 1, 2, \ldots, n \}$. On appelle permutation de E (ou aussi « substitution ») une **application bijective de E dans lui-même**. Elle est définie quand on se donne

$$\varphi(i) = \alpha_i, \qquad i = 1, 2, \ldots, n$$

les α_i étant tous distincts et égaux, à l'ordre près, à $1, 2, \ldots, n$.

On peut écrire conventionnellement la permutation φ sous la forme

$$\varphi = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}$$

Rappelons qu'il y a n ! permutations de E, parmi lesquelles la permutation « identique », ou « neutre », ou « unité », telle que $\varphi(i) = i$, $\forall\ i$: on la désigne par e :

$$e = \begin{pmatrix} 1 & 2 \ldots i \ldots n \\ 1 & 2 \ldots i \ldots n \end{pmatrix}$$

Il est manifeste que, pour la composition (ou multiplication), les permutations de E forment un groupe d'elément neutre e, où l'inverse de φ est

$$\varphi^{-1} = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ 1 & 2 & \ldots & n \end{pmatrix}$$

2. Transpositions

Une transposition de E est une permutation qui échange deux éléments i et j distincts de E, en laissant les autres invariants. On peut le noter T_{ij}, avec la convention $T_{ii} = e$.

On a donc

$$T_{ij}(i) = j, \quad T_{ij}(j) = i, \quad T_{ij}(p) = p \quad \text{si} \quad p \neq i \quad \text{et} \quad p \neq j$$

Exemple.

Si $E = \{ 1, 2, 3, 4, 5 \}$

$$T_{3,5} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 4 & 3 \end{pmatrix}$$

Il va de soi que $T_{ij} = T_{ji} = (T_{ij})^{-1}$.

Nous allons établir le **théorème fondamental :**

Toute permutation de l'ensemble fini E est un produit de transpositions.

Il suffit d'établir qu'une permutation laissant invariants au moins p éléments de E est un produit de transpositions : ceci est vrai pour $p = n$, alors $\varphi = e = T_{ij} \circ T_{ij}$. On peut raisonner par récurrence « descendante », c'est-à-dire supposer que toute permutation laissant invariants p éléments au moins est un tel produit, et montrer qu'alors il en est de même pour celles qui en laissent $p - 1$ invariants.

Soit φ_{p-1} une telle permutation (distincte de e), il existe j tel que $\varphi_{p-1}(j) = k \neq j$. Composons φ_{p-1} par T_{jk}, qui échange j et k et laisse les autres éléments invariants : si $\psi = T_{jk} \circ \varphi_{p-1}$, $\psi(i) = i$ pour les $p - 1$ éléments invariants dans φ_{p-1}

$$\psi(j) = T_{jk} \left[\varphi_{p-1}(j)\right] = T_{jk}(k) = j$$

ψ laisse donc invariants au moins p éléments de E, et est un produit de transpositions, il en est de même de $\varphi_{p-1} = T_{jk} \circ \psi$. Le théorème est démontré.

3. Inversions d'une permutation. Parité. Signature

Nous supposons ici E totalement ordonné, l'ordre croissant ou «naturel » étant précisément $(1, 2, 3, \ldots, n)$. Alors, étant donné la permutation

$$\varphi = \begin{pmatrix} 1 & 2 & 3 \ldots n \\ \alpha_1 & \alpha_2 & \alpha_3 \ldots \alpha_n \end{pmatrix}$$

on dit que α_i et α_j présentent une inversion dans φ si $i < j$ et $\alpha_i > \alpha_j$.

Par définition, une permutation φ est dite paire si le nombre total des inversions qu'elle présente est pair, elle est dite impaire si ce nombre est impair.

Si $I(\varphi)$ est ce nombre d'inversions, le nombre $\sigma(\varphi) = (-1)^{I(\varphi)}$ est appelé signature de φ. Il vaut 1 ou -1 suivant que φ est paire ou impaire.

Exemple.

$$\text{Soit} \quad \varphi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 4 & 3 & 6 & 1 \end{pmatrix}$$

$$I(\varphi) = 8, \qquad \varphi \text{ est paire,} \qquad \sigma(\varphi) = 1$$

On peut démontrer, et nous admettrons dans un souci d'allégement, le théorème suivant :

a) **Quand on compose une permutation quelconque par une transposition, on obtient une nouvelle permutation, de parité différente.**

Il en résulte les propriétés suivantes :

b) Pour qu'une permutation soit paire (respectivement, impaire), il faut et il suffit qu'elle soit le produit d'un nombre pair (respectivement, impair) de transpositions.

c) Si φ_1 et φ_2 sont deux permutations de même parité, $\varphi_1 \circ \varphi_2$ est paire; si elles sont de parités différentes, $\varphi_1 \circ \varphi_2$ est impaire. Ainsi

$$\sigma(\varphi_1 \circ \varphi_2) = \sigma(\varphi_1)\, \sigma(\varphi_2)$$

d) Deux permutations inverses l'une de l'autre ont même parité (ceci résultant de (*c*), puisque $\sigma(\varphi)\, \sigma(\varphi^{-1}) = \sigma(\varphi \circ \varphi^{-1}) = \sigma(e) = 1$).

B. DÉTERMINANTS

1. Forme multilinéaire alternée

Soit E un espace vectoriel de dimension n, sur un corps de scalaires K.

a) On appelle **forme *p*-linéaire** sur E une application de

$$E^p = E \times E \times \ldots \times E$$

ensemble des systèmes de p vecteurs $(\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_p)$ de E, dans K

$$(\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_p) \in E^p \underset{f}{\rightarrow} f(\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_p) \in K$$

cette application étant linéaire par rapport à chacun des vecteurs $\vec{V}_i$.

Ainsi

$$\vec{V}_i = \alpha \vec{U}_i + \beta \vec{W}_i \Rightarrow f(\vec{V}_1, \vec{V}_2, \ldots, \vec{V}_i, \ldots, \vec{V}_p) = \alpha f(\vec{V}_1, \vec{V}_2, \ldots, \vec{U}_i, \ldots, \vec{V}_p) + \beta f(\vec{V}_1, \vec{V}_2, \ldots, \vec{W}_i, \ldots, \vec{V}_p)$$

Pour $p = 1$, on retrouve évidemment les formes linéaires.

Pour $p = 2$, on a une forme bilinéaire, cas sur lequel nous reviendrons plus tard.

Ces deux cas écartés, le cas le plus fréquent est celui où $p = n$, f est une **forme** n-linéaire, ou simplement **multilinéaire**. C'est ce cas que nous envisageons désormais.

b) Une **forme n-linéaire** f sur E est dite alternée si l'échange de deux vecteurs dans la suite $(\vec{V_1}, \ldots, \vec{V_n})$ change le signe de $f(\vec{V_1}, \ldots, \vec{V_n})$, c'est-à-dire si

$$f(\vec{V_1}, \ldots, \vec{V_i}, \ldots, \vec{V_j}, \ldots, \vec{V_n}) = -f(\vec{V_1}, \ldots, \vec{V_j}, \ldots, \vec{V_i}, \ldots, \vec{V_n})$$

En particulier $\vec{V_i} = \vec{V_j} \Rightarrow f(\vec{V_1}, \ldots, \vec{V_n}) = 0$.

2. Expression d'une forme multilinéaire alternée dans une base de E

Soit $B = (\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ une base de E. Dans cette base, $\vec{V_i}$ a pour coordonnées $a_{i1}, a_{i2}, \ldots, a_{in}$, c'est-à-dire

$$\vec{V_i} = \sum_{j=1}^{j=n} a_{ij}\,\vec{e_j}$$

Du fait de la multilinéairité de la forme f, nous aurons

$$f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n}) = \sum_{j_1=1}^{j_1=n} a_{1j_1}\, f(\vec{e_{j_1}}, \vec{V_2}, \ldots, \vec{V_n})$$

$$= \sum_{j_1=1}^{j_1=n} a_{1j_1}\left[\sum_{j_2=1}^{j_2=n} a_{2j_2}\, f(\vec{e_{j_1}}, \vec{e_{j_2}}, \vec{V_3}, \ldots, \vec{V_n})\right]$$

donc, de proche en proche

$$f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n}) = \sum a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}\, f(\vec{e_{j_1}}, \vec{e_{j_2}}, \ldots, \vec{e_{j_n}})$$

cette somme portant sur tous les systèmes d'indices $j_1, j_2, \ldots, j_n$.

Mais du fait de l'alternance de la forme f, deux indices égaux conduisent à un terme nul, donc les j_i résultent d'une permutation φ de

$$\{1, 2, \ldots, n\}$$

et la somme ci-dessus porte sur ces permutations φ, on peut la noter $\sum\limits_{\varphi}$.

La permutation φ telle que $j_i = \varphi(i)$ est le produit de k transpositions qui, chacune, changent le signe de f, donc

$$f(\vec{e_{j_1}}, \vec{e_{j_2}}, \ldots, \vec{e_{j_n}}) = (-1)^k\, f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$$

Or d'après la propriété (A.3.*b*) de ce chapitre $(-1)^k = \sigma(\varphi)$ on obtient donc la formule

$$\boxed{f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n}) = f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) \sum_{\varphi} \sigma(\varphi)\, a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}} \qquad \text{(I)}$$

On constate que la valeur de f en un « point » $(\vec{V_1}, \ldots, \vec{V_n})$ de E^n est complètement connue dès que l'on se donne la valeur de $f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$, soit λ.

3. Définition d'un déterminant

Cette valeur λ peut *a priori* être quelconque, en particulier $\lambda = 0$ donnerait $f(\vec{V_1}, \ldots, \vec{V_n}) = 0$ quels que soient les $\vec{V_i}$, c'est-à-dire $f = 0$. Ce cas est sans intérêt.

Pour $\lambda \neq 0$, le rapport $\Delta = \frac{1}{\lambda} f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n})$ qui ne dépend que des a_{ij}, est lui-même une forme multilinéaire alternée des vecteurs $\vec{V_i}$. Ce scalaire

$$\boxed{\Delta = \sum_{\varphi} \sigma(\varphi)\, a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}} \qquad (1) \qquad (n\ !\ \text{produits})$$

où φ est la permutation $\varphi = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix}$ est le déterminant des vecteurs $\vec{V_1}, \ldots, \vec{V_n}$ relativement à la base $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$.

Pratiquement il est commode de choisir $\lambda = f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) = 1$ ce qui conduit à $\Delta = f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n})$.

On dit aussi (et essentiellement) que Δ est le déterminant de la matrice carrée

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

et on écrit

$$\boxed{\Delta = \text{dét}\,(A) = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}}$$

Ce tableau carré de n^2 éléments, encadré de deux barres verticales, représente donc un scalaire dépendant de la matrice A; deux matrices égales ont des déterminants égaux, la réciproque est évidemment fausse puisqu'une matrice dépend de n^2 scalaires.

On emploiera cependant en partie le même langage pour les déterminants que pour les matrices, à savoir :

vecteurs-lignes $(a_{i1}, a_{i2}, \ldots, a_{in})$ (ici, coordonnées des $\vec{V}_i$);
vecteurs-colonnes $(a_{1j}, a_{2j}, \ldots, a_{nj})$;
diagonale dite « principale » $(a_{11}, a_{22}, \ldots, a_{nn})$.

Notons encore qu'il est facile de vérifier d'après (1) que

$$\boxed{\text{dét}\, I = 1}$$

On peut aussi démontrer qu'il existe une forme multilinéaire alternée et une seule qui prend la valeur 1 pour les vecteurs $(\vec{e_1}, \ldots, \vec{e_n})$ donnés.

4. Exemples. Déterminants d'ordres 2 et 3. Règle de Sarrus

L'application directe de la définition (1) de B.3, donnant naissance à $n\,!$ produits du type $a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}$, est évidemment fastidieuse dès que n dépasse 3, et nous verrons d'autres méthodes de calcul. Cependant, avant de considérer $n = 2$ et $n = 3$, il est intéressant de remarquer que (1) exprime la décomposition de Δ en somme, avec signes alternés (il y a évidemment autant de permutations σ paires que de permutations impaires), de tous les produits obtenus en prenant dans le tableau Δ (ou la matrice A) un élément et un seul par rangée, en procédant par lignes « descendantes », le signe étant fourni par la signature de la permutation des rangs de colonnes correspondants. Ainsi, dans

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

le terme $a_{12}\, a_{23}\, a_{31}\, a_{44}$ (correspondant aux flèches ci-dessus) est à précéder du coefficient $\sigma(2, 3, 1, 4) = (-1)^2 = 1$ (deux inversions). Il y a en tout 24 termes de cette forme dans le développement de Δ.

Pour $n = 1$, cas pratiquement sans intérêt, dét $(a) = a$ (la notation $|a|$ pourrait prêter à confusion avec une valeur absolue).

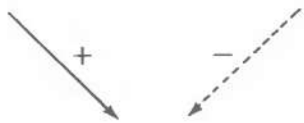
Pour $n = 2$ $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{12}\, a_{21}$ résultat déjà connu.

Pour $n = 3$ la définition donne

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} - a_{11}\, a_{23}\, a_{32} - a_{12}\, a_{21}\, a_{33} - a_{13}\, a_{22}\, a_{31}$$

On peut obtenir rapidement ce développement par la **règle de Sarrus** qui consiste à répéter, au-dessous du tableau précédent, les deux premières lignes, et à affecter du signe + les produits obtenus parallèlement à la diagonale principale, du signe − ceux obtenus parallèlement à la diagonale non principale :

$$\begin{array}{|ccc|} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ \hline a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array}$$



5. Propriétés des déterminants

Elles sont liées en majeure partie à celles des formes multilinéaires alternées.

a) **Théorème fondamental. Pour que *n* vecteurs d'un espace vectoriel *E* de dimension *n* forment un système lié, c'est-à-dire soient linéairement dépendants, il faut et il suffit que leur déterminant dans une base de *E* soit nul.**

Démonstration.

– Si les $\vec{V_i}$ sont dépendants, l'un d'eux, par exemple $\vec{V_1}$, est combinaison linéaire des $n - 1$ autres, et

$$\text{dét}\,(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n}) = \sum_2^n \lambda_i \,\text{dét}\,(\vec{V_i}, \vec{V_2}, \ldots, \vec{V_n}) = 0$$

puisque chacun de ces déterminants a deux vecteurs-lignes égaux.

– Réciproquement, soit $\Delta = \text{dét}\,(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n}) = 0$. Si les $\vec{V_i}$ sont indépendants, ils constituent une base de *E*, sur laquelle on peut décomposer les vecteurs $\vec{e_i}$ de la base donnée, et d'après la formule (I) de B.2

$$\text{dét}\,(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) = \left(\sum_{\varphi} \sigma(\varphi)\, b_{1j_1}\, b_{2j_2} \ldots b_{nj_n}\right) \text{dét}\,(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n})$$

Or $\text{dét}\,(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) = 1$ c'est incompatible avec $\Delta = 0$. Il est donc impossible que les $\vec{V_i}$ soient indépendants. D'où le théorème énoncé plus haut.

Sous une autre forme, et plus simplement, on peut aussi conclure pour une matrice *A* :

$$A \text{ singulière} \Leftrightarrow \text{dét}\, A = 0$$
$$A \text{ régulière} \Leftrightarrow \text{dét}\, A \neq 0$$

b) **On ne modifie pas un déterminant quand on ajoute à une ligne, une combinaison linéaire des autres lignes.**

En effet
$$\text{dét}\left(\vec{V_1}+\sum_2^n \lambda_i \vec{V_i}, \vec{V_2}, \ldots, \vec{V_n}\right) = \text{dét}\,(\vec{V_1}, \ldots, \vec{V_n}) + \sum_2^n \Delta_i$$

où les déterminants Δ_i sont nuls.

c) **Si on échange deux lignes d'un déterminant, celui-ci est changé en son opposé.**

Ceci résulte de la définition d'une forme n-linéaire alternée.

d) **Si on multiplie tous les éléments d'une ligne par λ, le déterminant est multiplié par λ.**

Ceci résulte de la multilinéarité.

Conséquence : en dimension n, $\text{dét}\,(\lambda A) = \lambda^n\,\text{dét}\,(A)$.

e) **Le déterminant d'une matrice carrée est égal au déterminant de la matrice transposée.**

En effet soit,
$$A = \begin{pmatrix} & \vdots & \\ \ldots & a_{ij} & \ldots \\ & \vdots & \end{pmatrix} \quad \text{et} \quad {}^tA = \begin{pmatrix} & \vdots & \\ \ldots & b_{ij} & \ldots \\ & \vdots & \end{pmatrix} \quad \text{avec} \quad b_{ij} = a_{ji}$$

$$\Delta' = \text{dét}\,({}^tA) = \sum_{\varphi} \sigma(\varphi)\, b_{1j_1}\, b_{2j_2} \ldots b_{nj_n}$$

où $\varphi = \begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix}$

$$\Delta' = \sum_{\varphi} \sigma(\varphi)\, a_{j_1 1}\, a_{j_2 2} \ldots a_{j_n n}$$

Pour comparer Δ' à $\Delta = \text{dét}\,(A)$, il faut replacer les indices $j_1, j_2, \ldots, j_n$ dans l'ordre naturel, donc effectuer sur les deux systèmes d'indices la permutation φ^{-1}, et comme $\sigma(\varphi^{-1}) = \sigma(\varphi)$ on aura

$$\Delta' = \sum \sigma(\varphi^{-1})\, a_{1k_1}\, a_{2k_2} \ldots a_{nk_n} \qquad \text{où } k_i = \varphi^{-1}(i)$$

En posant $\varphi^{-1} = \psi$, l'ensemble des ψ est identique à l'ensemble des φ, donc

$$\Delta' = \sum_{\psi} \sigma(\psi)\, a_{1k_1}\, a_{2k_2} \ldots a_{nk_n} = \text{dét}\, A$$

On a donc bien $\boxed{\text{dét}\,({}^tA) = \text{dét}\, A}$

Il en résulte que tout ce qui a été établi concernant les lignes d'un déterminant est vrai concernant les colonnes. Comme, en dimension 3, le déterminant de 3 vecteurs est l'expression dans une base de leur produit mixte $(\vec{V_1}, \vec{V_2}, \vec{V_3})$, on peut étendre cette notation et désigner par $(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n})$ le déterminant de n vecteurs-colonnes (ou aussi bien vecteurs-lignes).

f) **Le déterminant d'un produit de matrices carrées A et B est égal au produit des déterminants de ces matrices.**

$$\boxed{\text{dét}\,(AB) = \text{dét}\,A \,.\, \text{dét}\,B} \qquad (2)$$

Démonstration.

– Si B est singulière, AB est singulière car, f et g étant les endomorphismes correspondants, $\dim \operatorname{Im}\,(f \circ g) \leqslant \dim \operatorname{Im}\, g \leqslant n-1$ donc $\text{dét}\,(AB) = 0$, $\text{dét}\,B = 0$, (2) est vérifiée.

– Soit B régulière. Soit $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ une base de E, f la forme multilinéaire alternée telle que $f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) = 1$. Soit $(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n})$ une autre base de E, définie, par rapport à la base $(\vec{e_i})$, par la matrice B. Soit $(\vec{U_1}, \vec{U_2}, \ldots, \vec{U_n})$ un système de vecteurs de E, défini, par rapport à la base $(\vec{V_i})$, par la matrice A. La matrice des $(\vec{U_i})$ dans la base initiale est donc AB.

D'après la formule (I) du paragraphe (B.2), on a

$$f(\vec{U_1}, \vec{U_2}, \ldots, \vec{U_n}) = (\text{dét}\,A)\,.\,f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n})$$
$$f(\vec{U_1}, \vec{U_2}, \ldots, \vec{U_n}) = (\text{dét}\,(AB))\,.\,f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) = \text{dét}\,(AB)$$
$$f(\vec{V_1}, \vec{V_2}, \ldots, \vec{V_n}) = \text{dét}\,B\,.\,f(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n}) = \text{dét}\,B$$

Il s'ensuit que $f(\vec{U_1}, \vec{U_2}, \ldots, \vec{U_n}) = \text{dét}\,A\,.\,\text{dét}\,B = \text{dét}\,(AB)$.

La relation (2) est donc établie dans tous les cas. On a ainsi

$$\text{dét}\,(BA) = \text{dét}\,B\,.\,\text{dét}\,A = \text{dét}\,A\,.\,\text{dét}\,B \quad \text{donc} \quad \text{dét}\,(BA) = \text{dét}\,(AB)$$

Compte tenu de $\text{dét}\,({}^tA) = \text{dét}\,A$, on voit que le produit de deux déterminants peut être effectué ligne par colonne, ligne par ligne, colonne par ligne ou colonne par colonne.

g) Si A est une matrice régulière

$$\boxed{\text{dét}\,(A^{-1}) = \frac{1}{\text{dét}\,A}}$$

Ceci est immédiat puisque $A\,.\,A^{-1} = I$ donc

$$\text{dét}\,A\,.\,\text{dét}\,(A^{-1}) = \text{dét}\,I = 1$$

h) **Deux matrices semblables ont même déterminant.**

En effet $A' = P^{-1}AP \Rightarrow \text{dét } A' = \text{dét } P^{-1} . \text{dét } A . \text{dét } P$
$= \text{dét } A . \text{dét } P^{-1} . \text{dét } P = \text{dét } A.$

On peut donc définir le **déterminant d'un endomorphisme** de E, qui est indépendant de la base utilisée dans E.

6. Développement d'un déterminant suivant une rangée

Le déterminant $\Delta = \begin{vmatrix} & \vdots & \\ \dots & a_{ij} & \dots \\ & \vdots & \end{vmatrix}$, d'ordre n, est une fonction linéaire des éléments de chacune de ses rangées. Il peut donc être développé de $2\,n$ façons suivant une de ses rangées, sous la forme

$$\Delta = \sum_{j=1}^{j=n} a_{ij}\, X_{ij} \qquad \text{développement suivant la ligne } i;$$

$$\Delta = \sum_{i=1}^{i=n} a_{ij}\, Y_{ij} \qquad \text{développement suivant la colonne } j.$$

Pour un couple (i, j) fixé, on a $Y_{ij} = X_{ij}$, c'est le coefficient avec lequel figure a_{ij} dans Δ : il ne contient que des éléments des lignes de rang autre que i, et des colonnes de rang autre que j, X_{ij} s'appelle le cofacteur de l'élément a_{ij}. Tout revient donc à déterminer ce cofacteur.

a) Calcul de X_{11}.

Dans $\Delta = \sum_{\varphi} \sigma(\varphi)\, a_{1j_1}\, a_{2j_2} \dots a_{nj_n}$ les termes contenant a_{11} sont ceux contenus dans

$$\sum_{\varphi} \sigma(\varphi)\, a_{11}\, a_{2j_2} \dots a_{nj_n} \qquad (j_1 = 2)$$

où $\varphi = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & j_2 & \dots & j_n \end{pmatrix}$.

Donc $X_{11} = \sum_{\varphi} \sigma(\varphi)\, a_{2j_2}\, a_{3j_3} \dots a_{nj_n}$ les permutations φ envisagées ne différant deux à deux que par les images de 2, 3, ..., n, puisque $\varphi(1) = 1$. Donc si $\varphi' = \begin{pmatrix} 2 & 3 & \dots & n \\ j_2 & j_3 & \dots & j_n \end{pmatrix}$, $\sigma(\varphi) = \sigma(\varphi')$ (mêmes inversions, et

$$X_{11} = \sum_{\varphi'} \sigma(\varphi')\, a_{2j_2}\, a_{3j_3} \dots a_{nj_n}$$

On voit que X_{11} n'est autre que le déterminant d'ordre $n - 1$ obtenu en supprimant dans Δ sa première ligne et sa première colonne. On l'appelle déterminant mineur, ou simplement **mineur** de l'élément a_{11}, et on le note Δ_{11}. Ainsi $X_{11} = \Delta_{11}$.

b) Calcul de X_{ij}.

On se ramène au cas précédent en développant le déterminant Δ' déduit de Δ par les échanges successifs de sa $i^{\text{ième}}$ ligne avec les $i - 1$ précédentes, et de sa $j^{\text{ième}}$ colonne avec les $j - 1$ précédentes. Ces échanges amènent a_{ij} à la place initiale de a_{11} en conservant l'ordre des autres rangées. Chaque échange de rangées parallèles changeant le signe du déterminant, le signe a changé en tout $(i - 1) + (j - 1)$ fois, donc

$$\Delta = (-1)^{i+j-2}\,\Delta' = (-1)^{i+j}\,\Delta'$$

Or dans Δ' le cofacteur de a_{ij} est, d'après le cas (*a*), le déterminant mineur (d'ordre $n - 1$) obtenu en supprimant dans Δ' (donc aussi dans Δ où l'ordre des lignes et des colonnes ne contenant pas a_{ij} est le même que dans Δ') la $i^{\text{ième}}$ ligne et la $j^{\text{ième}}$ colonne. Soit Δ_{ij} ce mineur, nous avons donc $X_{ij} = (-1)^{i+j}\,\Delta_{ij}$. D'où le théorème :

Un déterminant d'ordre *n* est la somme des produits de chacun des éléments d'une même rangée par son cofacteur.

$$\Delta = \sum_{j=1}^{j=n} a_{ij}\,X_{ij}, \qquad \text{développement suivant la ligne } i;$$

$$\Delta = \sum_{i=1}^{i=n} a_{ij}\,X_{ij}, \qquad \text{développement suivant la colonne } j;$$

$$\boxed{X_{ij} = (-1)^{i+j}\,\Delta_{ij}}$$

Δ_{ij} étant le mineur de a_{ij}, déterminant d'ordre $n - 1$ obtenu en supprimant dans Δ la ligne et la colonne contenant a_{ij}.

Ce résultat est fondamental car il permet de ramener un déterminant d'ordre n à n déterminants d'ordre $n - 1$.

La répartition des signes à prendre devant les mineurs est alternée à partir du signe + pour l'élément a_{11}, par exemple pour un déterminant d'ordre 5 :

$$\begin{vmatrix} + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \end{vmatrix}$$

7. Calcul pratique d'un déterminant

Le principe général consiste à abaisser l'ordre grâce au développement suivant une rangée. Mais le plus souvent on cherchera préalablement, grâce à l'addition à une rangée d'une combinaison linéaire des rangées parallèles, à faire apparaître soit des zéros sur une même rangée, soit un facteur commun. Même un déterminant d'ordre 3, pour lequel la règle de Sarrus donne un développement simple, pourra être traité ainsi. Notons aussi que, sans abaissement d'ordre, la réduction à un déterminant triangulaire donne un calcul immédiat :

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ 0 & 0 & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots 0 & a_{nn} \end{vmatrix} = a_{11}\, a_{22} \dots a_{nn}$$

(obtenu, par exemple, en développant chaque fois suivant la première colonne).

Exemples.

a) Calculer

$$\Delta = \begin{vmatrix} 2 & 1 & 4 & -4 \\ -1 & -1 & 5 & 6 \\ 3 & 2 & -3 & 1 \\ 4 & 0 & 2 & 3 \end{vmatrix} \; +$$

La seconde colonne contient un zéro, il est facile d'en faire apparaître un autre en ajoutant la seconde ligne à la première (schéma ci-dessus), d'où

$$\Delta = \begin{vmatrix} 1 & 0 & 9 & 2 \\ -1 & -1 & 5 & 6 \\ 3 & 2 & -3 & 1 \\ 4 & 0 & 2 & 3 \end{vmatrix} \times 2 \; +$$

On peut encore multiplier par 2 la seconde ligne et l'ajouter à la troisième

$$\Delta = \begin{vmatrix} 1 & 0 & 9 & 2 \\ -1 & -1 & 5 & 6 \\ 1 & 0 & 7 & 13 \\ 4 & 0 & 2 & 3 \end{vmatrix}$$

On développe suivant la seconde colonne : le seul élément non nul est $a_{22} = -1$, son cofacteur est $(-1)^{2+2}\,\Delta_{22} = \Delta_{22}$, mineur de a_{22}, d'où

$$\Delta = -\begin{vmatrix} 1 & 9 & 2 \\ 1 & 7 & 13 \\ 4 & 2 & 3 \end{vmatrix} \times (-1) \; + \; \times (-4) \; +$$

En effectuant les deux opérations indiquées, qui conservent la seconde ligne,

$$\Delta = -\begin{vmatrix} 0 & 2 & -11 \\ 1 & 7 & 13 \\ 0 & -26 & -49 \end{vmatrix}$$

On développe enfin suivant la première colonne, où le terme a_{21} a pour cofacteur $(-1)^{2+1}\Delta_{21} = -\Delta_{21}$, de sorte que

$$\Delta = \begin{vmatrix} 2 & -11 \\ -26 & -49 \end{vmatrix} = -384$$

Le calcul est évidemment moins simple lorsque les éléments de départ ne sont pas entiers, il relève alors de l'analyse numérique et de l'utilisation des machines, et n'entre pas dans le cadre de cet ouvrage.

b) Calculer le déterminant d'ordre n

$$\Delta = \begin{vmatrix} 1 & x & x^2 \ldots x^{n-2} & x^{n-1} \\ x^{n-1} & 1 & x \ldots x^{n-3} & x^{n-2} \\ x^{n-2} & x^{n-1} & 1 \ldots x^{n-4} & x^{n-3} \\ x & x^2 & x^3 \ldots x^{n-1} & 1 \end{vmatrix}$$

Cette forme de déterminant, dit « circulant » (où on passe d'une ligne à la suivante par permutation circulaire, le dernier élément de la première ligne passant en première position dans la seconde, et ainsi de suite) fera l'objet d'un exercice plus complet dans le chapitre suivant. Ici, remarquons qu'en multipliant chaque ligne à partir de la seconde par x, et en la retranchant à la précédente (ces opérations étant successives et respectant toutes les lignes sauf une, mais pouvant être réunies en une seule transformation), on obtient

$$\Delta = \begin{vmatrix} 1-x^n & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1-x^n & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1-x^n & \ldots & 0 & 0 \\ \hdashline 0 & 0 & 0 & \ldots & 1-x^n & 0 \\ x & x^2 & x^3 & \ldots & x^{n-1} & 1 \end{vmatrix}$$

Ce déterminant est triangulaire, et égal au produit des termes diagonaux, donc

$$\Delta = (1 - x^n)^{n-1}$$

8. Déterminant de Van der Monde

Ce déterminant, très classique, est de la forme générale

$$\Delta = \begin{vmatrix} 1 & x_1 & x_1^2 \ldots x_1^{n-1} \\ 1 & x_2 & x_2^2 \ldots x_2^{n-1} \\ \ldots & \ldots & \ldots \\ 1 & x_n & x_n^2 \ldots x_n^{n-1} \end{vmatrix}$$

On peut le calculer facilement en remarquant que, par rapport à chacune des variables x_i, c'est un polynôme de degré $n-1$; par rapport à l'ensemble des variables $x_1, x_2, \ldots, x_n$, c'est un polynôme homogène de degré $p = 1 + 2 + 3 + \ldots + (n-1) = \dfrac{n(n-1)}{2}$ (puisque chaque terme du développement $\sum_{\varphi} \sigma(\varphi)\, a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}$ est un produit empruntant un facteur et un seul à chaque rangée).

En tant que polynôme $P(x_n)$, Δ s'annule pour

$$x_n = x_1, x_n = x_2, \ldots, x_n = x_{n-1},$$

il est donc divisible par $\prod_{i=1}^{i=n-1} (x_n - x_i)$. Le quotient ne contient plus x_n car le produit précédent est de même degré que P, en x_n. Ce quotient s'annule pour $x_{n-1} = x_1,\ x_{n-1} = x_2,\ \ldots,\ x_{n-1} = x_{n-2}$, il est donc divisible par $\prod_{j=1}^{j=n-2} (x_{n-1} - x_j)$ et ainsi de suite. On a donc

$$\Delta = \lambda \,.\, \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)$$

λ étant un polynôme ne contenant plus aucune des variables, donc une constante; ce que confirme le fait qu'il y a $C_n^2 = \dfrac{n(n-1)}{2}$ facteurs de la forme $x_i - x_j$, alors que Δ est précisément de degré $\dfrac{n(n-1)}{2}$.

La constante λ peut être obtenue en remarquant que dans Δ le produit des éléments diagonaux est $x_2\, x_3^2\, x_4^3 \ldots x_n^{n-1}$, c'est le seul terme de cette forme, et c'est, avec le même coefficient, le seul dans

$$(x_n - x_1)(x_n - x_2) \ldots (x_n - x_{n-1})(x_{n-1} - x_1) \ldots (x_{n-1} - x_{n-2}) \ldots (x_2 - x_1)$$

Donc $\lambda = 1$ et nous obtenons pour le **déterminant de Van der Monde d'ordre n :**

$$\boxed{\begin{vmatrix} 1 & x_1 & x_1^2 \ldots x_1^{n-1} \\ 1 & x_2 & x_2^2 \ldots x_2^{n-1} \\ \hdashline 1 & x_n & x_n^2 \ldots x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)}$$

C. INVERSION D'UNE MATRICE. SYSTÈMES D'ÉQUATIONS LINÉAIRES

1. Inversion d'une matrice

D'après ce qui précède, une matrice carrée est inversible si et seulement si son déterminant est non nul. Soit alors $\Delta = \text{dét } A$, soit a_{ij} le terme général de A (ou de Δ), X_{ij} son cofacteur. Le développement $\Delta = \sum_{j=1}^{j=n} a_{ij} X_{ij}$ peut être considéré comme la somme des produits des termes de la $i^{\text{ième}}$ ligne de A par ceux de la $i^{\text{ième}}$ ligne de la matrice des cofacteurs X_{ij}, matrice que nous appellerons comatrice de A, notée com A. On peut donc introduire $B = {}^t(\text{com } A)$ et former le produit AB. Dans ce produit, le terme diagonal c_{ii} est le produit de la $i^{\text{ième}}$ ligne de A par la $i^{\text{ième}}$ colonne de B, donc Δ.

Un terme non diagonal de C, c_{ik}, où $i \neq k$, est de la forme

$$c_{ik} = \sum_{j=1}^{j=n} a_{ij} b_{jk} \qquad \text{avec} \qquad b_{jk} = X_{kj}$$

puisque $B = {}^t(\text{com } A)$.

Or

$$c_{kk} = \Delta = \sum_{j=1}^{j=n} a_{kj} X_{kj}$$

c_{ik} est donc le déterminant obtenu en remplaçant dans Δ la ligne k par la ligne i, elle-même n'étant pas modifiée : ce déterminant possède deux lignes identiques, donc est nul. On a donc $c_{ik} = 0$ si $i \neq k$, $c_{ik} = \Delta$ si $i = k$, c'est-à-dire $AB = \Delta I$ (relation qui reste d'ailleurs vraie si $\Delta = 0$).

Lorsque $\Delta \neq 0$, on a $A \cdot \dfrac{B}{\Delta} = I$ donc $A^{-1} = \dfrac{B}{\Delta}$.

• En conclusion, si la matrice A est inversible, son inverse est donnée par la formule

$$\boxed{A^{-1} = \frac{1}{\text{dét } A} \times {}^t(\text{com } A)} \qquad (3)$$

formule dans laquelle com A (comatrice de A) est la matrice obtenue en remplaçant chaque élément de A par son cofacteur.

La matrice $B = {}^t(\text{com } A)$ est parfois appelée matrice adjointe de A, et notée A^*.

Remarquons que l'application systématique de (3), dans le cas général, exige, pour une matrice d'ordre n, le calcul de n^2 déterminants d'ordre

$n - 1$, donc est peu praticable. Elle est par contre facile lorsque $n = 2$ ou $n = 3$, et lorsque A présente un certain nombre de zéros. Pour $n = 2$, on retrouve

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

2. Systèmes d'équations linéaires. Généralités

On appelle système de n équations linéaires à p inconnues sur le corps K (généralement $\mathbb{R}$ ou $\mathbb{C}$), un système de la forme

$$(S)\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2p}x_p = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{np}x_p = b_p \end{array}\right.$$

où les coefficients a_{ij} et b_j sont donnés dans K, et où les x_i sont des inconnues à valeurs dans K.

(S) peut s'écrire sous la forme abrégée $\sum\limits_{j=1}^{j=p} a_{ij}\,x_j = b_i,\ i = 1, 2, \ldots, n,$ ou sous la forme matricielle

$$AX = B \tag{4}$$

où A est la matrice des coefficients a_{ij}, matrice à n lignes et p colonnes, B la matrice-colonne des b_j, X celle des x_i.

L'écriture (4) est encore l'expression relativement à des bases de E, espace vectoriel de dimension p, et de F, espace vectoriel de dimension n, de la relation

$$f(\vec{u}) = \vec{v} \tag{5}$$

où $\vec{v}$ est un vecteur donné dans F, et $\vec{u}$ un vecteur inconnu dans E, f étant l'application linéaire de E dans F définie par la matrice A.

(5) est la forme vectorielle du système (S).

Deux cas, de ce point de vue, peuvent se présenter : ou bien $\vec{v} \in \operatorname{Im} f$, il admet alors au moins un antécédent dans E, le système (S) a des solutions; ou bien $\vec{v} \notin \operatorname{Im} f$, le système (S) est alors impossible. On voit se poser à nouveau le problème de la bijectivité de f. On sait que f ne peut être bijective, éventuellement, que si $p = n$, il suffit alors que A soit inversible pour que (S) ait une solution unique. Nous traiterons d'abord ce cas.

3. Système de Cramer

On appelle ainsi un système de n équations à n inconnues dont la matrice est inversible, c'est-à-dire telle que $\Delta = \text{dét } A \neq 0$. Δ est d'ailleurs appelé déterminant du système (S).

Alors (S) admet la solution unique

$$X = A^{-1}B = \frac{1}{\Delta}\,{}^t(\mathrm{com}\ A)\,.\,B$$

La première ligne, dans cette relation matricielle, s'écrit

$$x_1 = \frac{1}{\Delta}\sum_{j=1}^{j=n} X_{j1}\,b_j$$

X_{ij} étant, dans Δ, le cofacteur de a_{ij}.

Si on rapproche cette relation de $\Delta = \sum_{j=1}^{j=n} X_{j1}\,a_{j1}$ développement de Δ suivant sa première colonne, on voit que $\sum_{j=1}^{j=n} X_{j1}\,b_j$ n'est autre que le déterminant d'ordre n obtenu en remplaçant, dans Δ, la première colonne par la matrice-colonne B; nous le désignerons par Δ_1. Ainsi $x_1 = \dfrac{\Delta_1}{\Delta}$, avec

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{2n} & \dots & a_{nn} \end{vmatrix}$$

De même $x_i = \dfrac{1}{\Delta}\sum_{j=1}^{j=n} X_{ji}\,b_j$, $\forall\, i = 1, 2, 3, \dots, n$. Donc

$$\boxed{x_i = \frac{\Delta_i}{\Delta}} \quad i = 1, 2, 3, \dots, n \qquad \text{(II)}$$

en désignant par Δ_i le déterminant déduit de Δ par remplacement de la colonne de rang i (celle des coefficients de x_i dans le système S) par la colonne des « seconds membres » b_j.

Les formules (II) s'appellent formules de Cramer. Elles résolvent dans le cas général un système de Cramer.

Comme la formule d'inversion d'une matrice, les formules de Cramer ont, dans le cas général, un intérêt plus théorique que pratique : elles exigent le calcul de $n + 1$ déterminants d'ordre n. Elles sont cependant fondamentales et permettent notamment, lorsque Δ n'est pas nul, de calculer une des inconnues du système sans faire intervenir les autres.

Exemple.

Calculons y dans le système

$$\begin{cases} 2x + 3y - z + t = 1 \\ x - y + 3z - 2t = 0 \\ 3x - 4y + z - 3t = 2 \\ x + 5y - 2z + 2t = -3 \end{cases}$$

On a

$$\Delta = \begin{vmatrix} 2 & 3 & -1 & 1 \\ 1 & -1 & 3 & -2 \\ 3 & -4 & 1 & -3 \\ 1 & 5 & -2 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 3 & -1 & 0 \\ 1 & -1 & 3 & 1 \\ 3 & -4 & 1 & -2 \\ 1 & 5 & -2 & 0 \end{vmatrix} (\times 2) +$$

$$= \begin{vmatrix} 2 & 3 & -1 & 0 \\ 1 & -1 & 3 & 1 \\ 5 & -6 & 7 & 0 \\ 1 & 5 & -2 & 0 \end{vmatrix}$$

$$= (-1)^{4+2} \times 1 \times \begin{vmatrix} 2 & 3 & -1 \\ 5 & -6 & 7 \\ 1 & 5 & -2 \end{vmatrix} = -26$$

Δ_2, ou encore

$$\Delta_y = \begin{vmatrix} 2 & 1 & -1 & 1 \\ 1 & 0 & 3 & -2 \\ 3 & 2 & 1 & -3 \\ 1 & -3 & -2 & 2 \end{vmatrix} = \begin{vmatrix} 2 & 1 & -7 & 5 \\ 1 & 0 & 0 & 0 \\ 3 & 2 & -8 & 3 \\ 1 & -3 & -5 & 4 \end{vmatrix}$$

$(\times -3)$ + ; $(\times 2)$ +

$$= (-1)^{2+1} \times 1 \times \begin{vmatrix} 1 & -7 & 5 \\ 2 & -8 & 3 \\ -3 & -5 & 4 \end{vmatrix} = 68$$

Il en résulte $y = -\dfrac{26}{68} = -\dfrac{13}{34} \simeq -0{,}382.$

4. Rang d'un système linéaire

a) On appelle déterminant d'ordre q extrait d'une matrice (n, p), soit A, le déterminant d'une matrice carrée d'ordre q formée en supprimant dans A $(n-q)$ lignes et $(p-q)$ colonnes.

On appelle rang d'une matrice A l'ordre du déterminant non nul d'ordre le plus élevé, extrait de A. Il résulte de cette définition que A et ${}^tA$ ont le même rang.

Exemple.

Soit

$$A = \begin{pmatrix} 1 & 1 & 3 & 5 \\ 1 & 2 & 5 & 9 \\ 2 & 3 & 8 & 14 \end{pmatrix}$$

Les quatre déterminants d'ordre 3 extraits de A sont nuls, par contre le déterminant d'ordre 2 extrait $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ n'est pas nul, A est de rang 2.

b) Par définition, on appelle rang r d'un système d'équations linéaires le rang de la matrice A de ce système.

Nous allons comparer ce rang à celui, q, du système des vecteurs-colonnes de A. Comme des échanges de lignes ou de colonnes dans A ne modifient pas les valeurs absolues des déterminants extraits, ils ne modifient pas non plus le rang de A, et nous pouvons supposer que le déterminant formé des r premières lignes et des r premières colonnes de A est non nul :

$$A = \begin{pmatrix} a_{11} & a_{12} \dots a_{1r} \dots a_{1p} \\ a_{21} & a_{22} \dots a_{2r} \dots a_{2p} \\ \dots & \dots\dots\dots\dots\dots \\ a_{r1} & a_{r2} \dots a_{rr} \dots a_{rp} \\ \dots & \dots\dots\dots\dots\dots \\ a_{n1} & a_{n2} \dots a_{nr} \dots a_{np} \end{pmatrix} \qquad \Delta = \begin{vmatrix} a_{11} & a_{12} \dots a_{1r} \\ a_{21} & a_{22} \dots a_{2r} \\ \hdashline a_{r1} & a_{r2} \dots a_{rr} \end{vmatrix} \neq 0$$

Appelons $(\vec{e_1}, \vec{e_2}, \dots, \vec{e_n})$ une base de l'espace E de dimension n; A définit une application linéaire de F (dimension p) dans E; appelons $\vec{V_j}$ le vecteur-colonne $\sum_{i=1}^{i=n} a_{ij}\vec{e_i}$ et soit $\vec{U_j} = \sum_{i=1}^{i=r} a_{ij}\vec{e_j}$ $(j = 1, 2, \dots, r)$

Les r vecteurs $\vec{U_j}$ sont indépendants puisque leur déterminant Δ est non nul, ils forment une base du sous-espace G de E engendré par $\vec{e_1}, \vec{e_2}, \dots, \vec{e_r}$, base que l'on peut compléter par $\overrightarrow{e_{r+1}}, \dots, \vec{e_n}$ pour obtenir une base de E.

Voyons si les $\vec{V_j}$ $(j = 1, 2, \dots, r)$ sont indépendants :

$$\vec{V_j} = \vec{U_j} + \sum_{i=r+1}^{i=n} a_{ij}\,\vec{e_i}$$

donc une relation

$$\sum_{j=1}^{j=r} \alpha_j \vec{V_j} = \vec{0} \quad \text{ou} \quad \sum_{j=1}^{j=r} \alpha_j \vec{U_j} + \sum_{i=r+1}^{i=n} \left(\sum_{j=1}^{j=r} \alpha_j a_{ij}\right) \vec{e_i} = \vec{0}$$

entraîne, puisque $(\vec{U_1}, \dots, \vec{U_r}, \overrightarrow{e_{r+1}}, \dots, \vec{e_n})$ forment une base de E,

$$\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$$

$\vec{V_1}, \vec{V_2}, \dots, \vec{V_r}$ sont donc indépendants, et on a $q \geqslant r$.

Un raisonnement analogue permettrait de démontrer que $q \leqslant r$.

On a donc $q = r$, et on peut énoncer le théorème :

Le rang d'une matrice A, à n lignes et p colonnes, qui est aussi le rang d'un système linéaire de matrice A, est égal au rang du système des vecteurs-colonnes de A.

Remarquons que ce théorème permet de ramener la recherche du rang d'un système de vecteurs, et aussi du rang d'une application linéaire, à l'examen des déterminants extraits de la matrice du système ou de l'application linéaire envisagés.

5. Cas général du système de n équations à p inconnues. Théorème de Fontené-Rouché

Soit le système (S), que nous supposerons de rang r et écrit de telle façon que le déterminant Δ des coefficients des r premières inconnues dans les r premières équations soit non nul

$$(S)\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r + \ldots + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2r}x_r + \ldots + a_{2p}x_p = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r + \ldots + a_{rp}x_p = b_r \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nr}x_r + \ldots + a_{np}x_p = b_n \end{cases}$$

Les r premières équations s'appellent équations principales, les r premières inconnues $x_1, x_2, \ldots, x_r$ sont les inconnues principales.

On peut envisager deux cas :

a) $r = n$; alors $p \geqslant n$, et si on donne aux $p - r$ inconnues non principales des valeurs arbitraires, les r équations principales se confondent avec le système (S) et s'écrivent

$$(S)\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1p}x_p = c_1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - \ldots - a_{rp}x_p = c_r \end{cases}$$

Par rapport aux inconnues $x_1, \ldots, x_r$, (S) est un système de Cramer, il a une solution unique pour $p - r$ valeurs données des paramètres $x_{r+1}, x_{r+2}, \ldots, x_p$.

On dit que (S) est un système indéterminé à $p - r$ paramètres.

b) $r < n$.

Si nous revenons à la forme vectorielle de (S), $f(\vec{u}) = \vec{v}$, où $\vec{v}$ est donné dans F, espace de dimension n, nous avons

$$r = \dim \operatorname{Im} f < \dim F$$

L'application f n'est plus surjective et (S) a des solutions si et seulement si $\vec{v} \in \operatorname{Im} f$. Or $\operatorname{Im} f$ a pour base $\{ f(\vec{e_1}), f(\vec{e_2}), \ldots, f(\vec{e_r}) \}$ donc $\vec{v} \in \operatorname{Im} f$ si et seulement si la famille de $r+1$ vecteurs $\{ f(\vec{e_1}), \ldots, f(\vec{e_r}), \vec{v} \}$ est liée dans F, c'est-à-dire si la matrice B de ces vecteurs est de rang au plus égal à r.

Comme $$B = \begin{pmatrix} a_{11} & \ldots & a_{1r} & b_1 \\ a_{r1} & \ldots & a_{rr} & b_r \\ a_{n1} & \ldots & a_{nr} & b_n \end{pmatrix}$$

le déterminant Δ n'étant pas nul, B est de rang au moins égal à r. Il en résulte que (S) a des solutions si et seulement si B est de rang r, c'est-à-dire si tous les déterminants d'ordre $r+1$ extraits de B sont nuls. Un tel déterminant est de la forme :

$$D_k = \begin{vmatrix} a_{11} & \ldots & a_{1r} & b_1 \\ \ldots & \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} & b_r \\ a_{k1} & \ldots & a_{kr} & b_k \end{vmatrix}, \; k = r+1, r+2, \ldots, n$$

Ces $n - r$ déterminants sont appelés déterminants caractéristiques de (S).

On peut résumer ce qui précède par le théorème suivant.

Théorème de Fontené-Rouché.

Soit (S) un système de n équations à p inconnues, de rang r.

a) Si $r = n$, le système (S) est indéterminé à $p - r$ paramètres : on attribue des valeurs arbitraires aux $p - r$ inconnues non principales, les r inconnues principales sont alors données par un système de Cramer.

b) Si $r < n$, et si l'un au moins des déterminants caractéristiques de (S) est non nul, (S) n'a pas de solution.

c) Si $r < n$, et si les $n - r$ déterminants caractéristiques de (S) sont nuls, (S) se réduit aux r équations principales et se résout comme dans le cas (*a*).

Exemple.

Soit à résoudre le système (S) de quatre équations à trois inconnues.

$$(S) \begin{cases} x + y + 2z = -2 & (1) \\ x + 2y + 3z = a & (2) \\ 3x + 5y + 8z = 2 & (3) \\ 5x + 9y + 14z = b & (4) \end{cases}$$

(a, b supposés donnés).

La matrice de ce système est la transposée de celle de l'exemple (4, *a*). Elle est de rang 2. On peut prendre pour inconnues principales x et y, et pour équations principales (1) et (2). Il y a deux déterminants caractéristiques :

$$D_1 = \begin{vmatrix} 1 & 1 & -2 \\ 1 & 2 & a \\ 3 & 5 & 2 \end{vmatrix} \qquad D_2 = \begin{vmatrix} 1 & 1 & -2 \\ 1 & 2 & a \\ 5 & 9 & b \end{vmatrix}$$

$$D_1 = -2a + 4 \qquad D_2 = -4a + b + 2$$

Pour que (S) ait des solutions, il faut et il suffit que $D_1 = D_2 = 0$, donc $a = 2$, $b = 6$.

Alors le système principal s'écrit

$$\begin{cases} x + y = -2z - 2 \\ x + 2y = -3z + 2 \end{cases}$$

Par soustraction $y = -z + 4$ d'où $x = -z - 6$.

En résumé, si $(a, b) \neq (2, 6)$ (S) est impossible.

Si $a = 2, b = 6$ (S) est indéterminé à un paramètre :

$$x = -z - 6 \qquad y = -z + 4 \qquad z \text{ arbitraire}$$

6. Système homogène

Un système linéaire est dit homogène lorsque les « seconds membres » b_i sont tous nuls. Sa forme vectorielle est $f(\vec{u}) = \vec{0}$.

Ses solutions sont constituées par les vecteurs de $\ker f$, donc (S) possède toujours au moins la solution nulle $\vec{u} = \vec{0}$, soit

$$x_1 = x_2 = \ldots = x_p = 0$$

Si le rang de (S) est p, il n'y a pas d'autre solution.

Si ce rang est inférieur à p, il y a une infinité de solutions et on résout (S) par la méthode générale.

Notons en particulier que pour qu'un système homogène de n équations à n inconnues admette des solutions autres que $(0, 0, \ldots, 0)$, il faut et il suffit que son déterminant soit nul.

D. NOTIONS SUR L'ÉLIMINATION

1. Résultant de deux polynômes à une variable

Soit $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \qquad a_0 \neq 0$

$g(x) = b_0x^p + b_1x^{p-1} + \ldots + b_p \qquad b_0 \neq 0$

deux polynômes sur $\mathbb{C}$.

Éliminer x entre f et g, c'est trouver une relation nécessaire et suffisante entre les coefficients de f et g exprimant que les équations $f(x) = 0$

et $g(x) = 0$ ont au moins une racine commune. Ceci équivaut à dire que f et g ont au moins un facteur commun du premier degré $x - \alpha$, $\alpha \in \mathbb{C}$.

a) Nous allons établir tout d'abord que cette condition équivaut encore à la suivante : il existe deux polynômes non nuls $P(x)$ de degré $\leqslant p - 1$, $Q(x)$ de degré $\leqslant n - 1$, tels que $\quad Pf + Qg = 0 \qquad (6)$

En effet si f et g sont divisibles par $x - \alpha$, (6) est réalisée en prenant $P = \dfrac{g(x)}{x - \alpha}$, $\quad Q = -\dfrac{f(x)}{x - \alpha}$.

Réciproquement, $(6) \Leftrightarrow Pf = -Qg$. En décomposant sur $\mathbb{C}$ ces deux produits en facteurs du premier degré, les deux décompositions devant être identiques, et P étant de degré au plus égal à $p - 1$, on voit que g contient au plus $p - 1$ facteurs provenant de P, comme il est de degré p, il contient au moins un facteur provenant de f, autrement dit f et g ont au moins un facteur commun du premier degré. La réciproque est donc établie.

b) Explicitons P et Q dans (6) :

$$P(x) = \alpha_0 x^{p-1} + \alpha_1 x^{p-2} + \ldots + \alpha_{p-1}$$
$$Q(x) = \beta_0 x^{n-1} + \beta_1 x^{n-2} + \ldots + \beta_{n-1}$$

$$(6) \Leftrightarrow \alpha_0 x^{p-1} f + \alpha_1 x^{p-2} f + \ldots + \alpha_{p-1} f + \beta_0 x^{n-1} g + \beta_1 x^{n-2} g + \ldots + \beta_{n-1} = 0$$

Cette relation, où les α_i et β_j ne sont pas tous nuls, exprime que les $n + p$ polynômes $x^{p-1}f, x^{p-2}f, \ldots, xf, f; x^{n-1}g, x^{n-2}g, \ldots, xg, g$, « vecteurs » de l'espace vectoriel de dimension $n + p$, de base

$$(x^{n+p-1}, x^{n+p-2}, \ldots, x, 1)$$

sont liés. Leur matrice dans cette base peut s'écrire

$$S = \begin{pmatrix}
a_0 & 0 & \cdots & 0 & b_0 & 0 & \cdots & 0 \\
a_1 & a_0 & \cdots & 0 & b_1 & b_0 & \cdots & 0 \\
a_2 & a_1 & \ddots & \vdots & b_2 & b_1 & \ddots & \vdots \\
\vdots & & \ddots & 0 & \vdots & & \ddots & \\
a_n & a_{n-1} & \cdots a_1 & a_0 & b_p & \cdots & \cdots & b_0 \\
0 & a_n & \cdots & a_1 & 0 & \ddots & & b_1 \\
\vdots & \ddots & \ddots & a_{n-1} & \vdots & & \ddots & \vdots \\
0 & \cdots & 0 & a_n & 0 & \cdots & 0 & b_p
\end{pmatrix}$$

Cette matrice carrée d'ordre $n + p$ s'appelle **matrice de Sylvester** des deux polynômes f et g. On voit que la condition nécessaire et suffisante pour que les équations $f = 0$ et $g = 0$ aient une racine commune est $\text{dét } S = 0$. Ce déterminant s'appelle **déterminant de Sylvester**, ou plus couramment **résultant** des deux polynômes f et g. On le note $\; R(f, g)$.

Exemple.

Soit $f(x) = ax^2 + bx + c$, $g(x) = a'x^2 + b'x + c'$,

$$R(f, g) = \begin{vmatrix} a & 0 & a' & 0 \\ b & a & b' & a' \\ c & b & c' & b' \\ 0 & c & 0 & c' \end{vmatrix}$$

La méthode précédente, comme toutes celles faisant intervenir des déterminants d'ordre élevé, conduit à des calculs souvent pénibles. On peut souvent faire apparaître le résultant à l'aide de combinaisons de $f(x) = 0$ et $g(x) = 0$ conduisant à un abaissement de degré, combinaisons qui introduisent parfois des solutions étrangères, faciles à éliminer dans la plupart des cas.

Revenons sur le système

$$\begin{cases} a x^2 + b x + c = 0 & (7) \qquad (a \neq 0) \\ a'x^2 + b'x + c' = 0 & (8) \qquad (a' \neq 0) \end{cases}$$

En multipliant (7) par a' et (8) par $-a$, et ajoutant les équations obtenues, on a :

$$(ba' - ab')\,x + ca' - ac' = 0 \tag{9}$$

Si $ab' - ba' \neq 0$, une racine commune est $x = \dfrac{ca' - ac'}{ab' - ba'}$ ce qui, reporté dans (7), donne

$$a\,(ca' - ac')^2 + b\,(ca' - ac')\,(ab' - ba') + c\,(ab' - ba')^2 = 0$$

$$a\,(ca' - ac')^2 + (ab' - ba')\,a\,(cb' - bc') = 0$$

a étant non nul, il reste

$$R = (ca' - ac')^2 - (ab' - ba')\,(bc' - cb') = 0 \tag{10}$$

Cette relation (10) reste valable si $ab' - ba' = 0$ car il faut alors $ca' - ac' = 0$ et les équations (7) et (8), identiques à un coefficient près, ont alors deux racines communes. On vérifie que R, polynôme homogène du quatrième degré en a, b, c, a', b', c', est bien égal au déterminant $R(f, g)$ écrit plus haut.

2. Discriminant d'un polynôme à une variable

Envisageant de généraliser la notion de discriminant d'un polynôme du second degré, cherchons à exprimer la **condition pour qu'un polynôme $f(x)$ ait une racine multiple :** il faut et il suffit pour cela que les équations $f(x) = 0$ et $f'(x) = 0$ aient au moins une racine commune, donc que le résultant $R(f, f')$ soit nul. Ce résultant, obtenu soit par l'élimination de x entre $f(x)$ et sa dérivée $f'(x)$ soit comme déterminant de Sylvester, est appelé discriminant de f, et souvent noté $\Delta(f)$.

Exemples.

a) $f(x) = ax^2 + bx + c \quad \Delta(f) = \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = a(4ac - b^2)$

a étant supposé $\neq 0$, la condition pour que $f(x) = 0$ ait une racine double est effectivement $4ac - b^2 = 0$.

b) $f(x) = x^3 + px + q \quad \Delta(f) = \begin{vmatrix} 1 & 0 & 3 & 0 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ p & 0 & p & 0 & 3 \\ q & p & 0 & p & 0 \\ 0 & q & 0 & 0 & p \end{vmatrix} = 4p^3 + 27q^2$

On retrouve un résultat classique : si $4p^3 + 27q^2 = 0$, la racine double est $-\dfrac{3q}{2p}$, la racine simple $\dfrac{3q}{p}$.

Résumé du chapitre 7

• **Permutation. Transposition.**

Une permutation de l'ensemble $E = \{1, 2, \ldots, n\}$ est une application bijective de E dans lui-même. Une transposition est une permutation qui échange deux éléments i et j de E, laissant les autres invariants.

• Toute permutation de E est un **produit de transpositions**.

• On dit que dans une permutation φ, où $\varphi(i) = \alpha_i$, $\varphi(j) = \alpha_j$, α_i et α_j présentent une inversion si $i < j$ et $\alpha_i > \alpha_j$.

$I(\varphi)$ étant le nombre total d'inversions d'une permutation,

$\sigma(\varphi) = (-1)^{I(\varphi)}$ est la **signature de** φ

• **Déterminant.**

On appelle déterminant de la matrice $A = \begin{pmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{pmatrix}$, carrée d'ordre n, le nombre $\Delta = \sum_{\varphi} \sigma(\varphi)\, a_{1j_1}\, a_{2j_2} \ldots a_{nj_n}$ où φ est la permutation $\begin{pmatrix} 1 & 2 & \ldots & n \\ j_1 & j_2 & \ldots & j_n \end{pmatrix}$.

On écrit

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \qquad \text{déterminant d'ordre } n.$$

Pour $n = 2$, $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{21}\, a_{12}$.

Pour $n = 3$, règle de Sarrus.

• **Propriété fondamentale.**

Pour que n vecteurs d'un espace vectoriel de dimension n forment un système lié, c'est-à-dire pour que leur matrice A soit singulière, il faut et il suffit que dét $A = 0$.

• **Calcul des déterminants.**

Pour effectuer ce calcul, on utilise les propriétés suivantes :

– dét $({}^tA) =$ dét A.

– On ne modifie pas un déterminant quand on ajoute à une rangée une combinaison linéaire des rangées parallèles.

– Si on échange deux rangées parallèles d'un déterminant, celui-ci est changé en son opposé.

– Si on multiplie tous les éléments d'une rangée par λ, le déterminant est multiplié par λ.

– On peut développer un déterminant $\Delta = \begin{vmatrix} & \vdots & \\ \dots & a_{ij} & \dots \\ & \vdots & \end{vmatrix}$ suivant la ligne de rang i ou la colonne de rang j, par les formules :

$$\Delta = \sum_{j=1}^{j=n} a_{ij}\, X_{ij}, \qquad \Delta = \sum_{i=1}^{i=n} a_{ij}\, X_{ij}$$

X_{ij} est le cofacteur de l'élément a_{ij}, donné par

$$X_{ij} = (-1)^{i+j}\, \Delta_{ij}$$

où Δ_{ij}, mineur de a_{ij}, est le déterminant d'ordre $n - 1$ obtenu en supprimant dans Δ les deux rangées contenant a_{ij}.

• **dét $(AB) =$ dét A . dét B.**

• Un **déterminant triangulaire** (déterminant d'une matrice triangulaire) est égal au produit des termes diagonaux.

• **Déterminant de Van der Monde.**

$$\begin{vmatrix} 1 & x_1 & x_1^2 \dots x_1^{n-1} \\ 1 & x_2 & x_2^2 \dots x_2^{n-1} \\ \dots & \dots & \dots \\ 1 & x_n & x_n^2 \dots x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)$$

• **Inverse d'une matrice.**

A inversible $\Leftrightarrow$ dét $A \neq 0$, alors $A^{-1} = \dfrac{1}{\text{dét } A} \cdot {}^t(\text{com } A)$,

com A étant la comatrice ou matrice des cofacteurs de A.

• **Système de Cramer.**

Système de n équations linéaires à n inconnues, de déterminant non nul :

$$(S) \quad \begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases}$$

Ce système a une solution unique donnée par les formules de Cramer :

$$x_i = \frac{\Delta_i}{\Delta}, \qquad i = 1, 2, \dots, n$$

Δ_i étant le déterminant déduit de celui de (S) en remplaçant la colonne des coefficients de x_i par celle des seconds membres $b_1, \dots, b_n$.

• **Rang d'une matrice.**

Le rang d'une matrice A est l'ordre du déterminant non nul, d'ordre le plus élevé, extrait de A. C'est aussi le rang d'un système linéaire de matrice A, et le rang du système des vecteurs-colonnes de A.

• Le cas le plus général du **système de n équations à p inconnues** conduit à la notion d'inconnues et d'équations principales, et de déterminants caractéristiques (voir cours détaillé).

• La condition pour que deux polynômes $f(x)$, de degré n, et $g(x)$, de degré p, aient un zéro commun, s'exprime en annulant un polynôme de degré $n + p$ par rapport aux coefficients de f et g, appelé **résultant** de f et g et obtenu sous forme d'un **déterminant dit de Sylvester**.

Le résultant Δ du polynôme f et de sa dérivée f' s'appelle discriminant de f : $\Delta = 0$ exprime que l'équation $f(x) = 0$ admet une racine multiple.

EXERCICES

■ **1.** Discuter suivant les entiers n et p la signature de la permutation circulaire

$$\varphi = \begin{pmatrix} 1 & 2 \dots p & & p+1 \dots n \\ p+1 & p+2 \dots\dots n & & 1 \dots p \end{pmatrix}$$

$p + 1, p + 2, \dots, n$ présentent des inversions avec chacun des éléments $1, 2, \dots, p$, le nombre d'inversions de φ est $p(n-p)$ et

$$\sigma(\varphi) = (-1)^{p(n-p)}$$

Donc $\sigma = 1$ si p est pair, ou si n et p sont tous deux impairs;
$\sigma = -1$ si p est impair et n pair.

■ **2.** E étant l'ensemble $\{1, 2, 3, \dots, n\}$, déterminer :

a) le nombre maximum d'inversions que peut présenter une permutation de E;

b) le nombre total des inversions présentées par toutes les permutations de E.

---

a) Ce maximum est atteint lorsque tous les couples (i, j) sont en inversion, donc pour

$$\varphi = \begin{pmatrix} 1 & 2 & 3 & \dots n-1 & n \\ n & n-1 & n-2 \dots & 2 & 1 \end{pmatrix}$$

alors $I(\varphi) = C_n^2 = \dfrac{n(n-1)}{2}$.

b) On associe à une permutation φ de E la permutation φ' obtenue en renversant l'ordre des termes : $\varphi \cup \varphi'$ présente $\dfrac{n(n-1)}{2}$ inversions, car un couple (i, j) en présente une soit dans φ, soit dans φ'.

Il y a donc au total $\dfrac{n(n-1)}{2} \cdot \dfrac{n!}{2} = \dfrac{1}{4} n!\, n(n-1)$ inversions.

■ **3.** Calculer le déterminant

$$\Delta = \begin{vmatrix} 2 & 5 & 6 & 1 \\ 3 & 1 & 5 & 8 \\ 7 & 9 & 2 & 4 \\ 2{,}37 & 5{,}19 & 6{,}52 & 1{,}84 \end{vmatrix}$$

Réponse : $\Delta = 0$ (les lignes sont liées linéairement).

■ **4.** Factoriser sur $\mathbb{R}$, puis sur $\mathbb{C}$, le déterminant

$$\Delta = \begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$$

on introduira j et j^2, racines cubiques complexes de 1).

Caractériser les triangles ABC du plan complexe tels que les affixes a, b, c des points A, B, C vérifient $\Delta = 0$.

$$\Delta = (a+b+c)\begin{vmatrix} 1 & b & c \\ 1 & a & b \\ 1 & c & a \end{vmatrix} = (a+b+c)\begin{vmatrix} 1 & b & c \\ 0 & a-b & b-c \\ 0 & c-b & a-c \end{vmatrix}$$

$$= (a+b+c)(a^2+b^2+c^2-bc-ca-ab)$$

En factorisant $A = a^2 - (b+c)a + b^2 + c^2 - bc$, un calcul simple donne $A = (a + jb + j^2c)(a + j^2b + jc)$, donc

$$\Delta = (a+b+c)(a+jb+j^2c)(a+j^2b+jc)$$

Δ est nul dans trois cas :

- $a + b + c = 0$, l'origine est centre de gravité du triangle ABC
- $a + jb + j^2c = 0$, ou $a + jb - (1+j)c = 0$ (car $1 + j + j^2 = 0$)

$$a - c = -j(b-c)$$

Ceci exprime que l'on passe du vecteur $\overrightarrow{CB}$ au vecteur $\overrightarrow{CA}$ par une rotation d'angle $\arg(-j) = \arg j + \pi = \dfrac{5\pi}{3}$, ou $-\dfrac{\pi}{3}$, le triangle ABC est donc équilatéral « direct ».

- $a + j^2b + jc = 0$, ou $a - c = -j^2(b-c)$, le triangle ABC est alors équilatéral « indirect ».

En résumé, $\Delta = 0$ caractérise l'ensemble des triangles équilatéraux et des triangles admettant l'origine pour centre de gravité.

■ **5.** Calculer le déterminant d'ordre n

$$\Delta_n = \begin{vmatrix} n & n-1 & n-2 & \dots 3 & 2 & 1 \\ n-1 & n-2 & n-3 & \dots 2 & 1 & 0 \\ n-2 & n-3 & n-4 & \dots 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 3 & 2 & 1 & 0 \dots 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \dots 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \dots 0 & 0 & 0 \end{vmatrix}$$

En développant Δ_n suivant la dernière colonne, on obtient

$$\Delta_n = (-1)^{n+1}\,\Delta_{n-1}$$

De même

$$\begin{aligned} \Delta_{n-1} &= (-1)^n\,\Delta_{n-2} \\ &\dots\dots\dots\dots \\ \Delta_3 &= (-1)^4\,\Delta_2 \\ \Delta_2 &= (-1)^3\,\Delta_1 \\ \Delta_1 &= 1 \end{aligned}$$

Par multiplication,

$$\Delta_n = (-1)^{3+4+5+\dots+(n+1)} = (-1)^{\frac{(n+1)(n+2)}{2}-3} = (-1)^{\frac{(n+1)(n-4)}{2}}$$

ou, en retranchant de l'exposant un nombre pair,

$$\Delta_n = (-1)^{\frac{n(n-1)}{2}}$$

■ **6.** Calculer le déterminant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{vmatrix}$$

En ajoutant à la première ligne les trois autres, on obtient

$$\Delta = \begin{vmatrix} 4 & 0 & 0 & 0 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{vmatrix} = 4 \begin{vmatrix} i & -1 & -i \\ -1 & 1 & -1 \\ -i & -1 & i \end{vmatrix}$$

$$= 4 \begin{vmatrix} 0 & -2 & 0 \\ -1 & 1 & -1 \\ -i & -1 & i \end{vmatrix}$$

soit

$$\Delta = -16\,i$$

On peut aussi remarquer que Δ est un déterminant de Van der Monde de la forme

$$D(x_1, x_2, x_3, x_4) = \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \\ 1 & x_4 & x_4^2 & x_4^3 \end{vmatrix} = (x_4 - x_1)(x_4 - x_2)\dots(x_2 - x_1)$$

Ici

$$\begin{aligned} \Delta = D(1, i, -1, -i) &= (-i-1)(-2i)(-i+1)(-2)(-1-i)(i-1) \\ &= -16\,i \end{aligned}$$

■ **7.** Calculer les déterminants

$$A_n = \begin{vmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ a_1 & a_1 & a_2 & \dots & a_{n-1} \\ a_1 & a_1 & a_1 & \dots & a_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ a_1 & a_1 & a_1 & \dots & a_1 \end{vmatrix} \qquad \text{(ordre } n\text{)}$$

$$B_{n+1} = \begin{vmatrix} a_n & a_{n-1} & a_{n-2} & \dots\dots & a_0 \\ -1 & x & 0 & \dots\dots & 0 \\ 0 & -1 & x & \dots\dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots -1 & x \end{vmatrix} \qquad \text{(ordre } n+1\text{)}$$

$$D_n = \begin{vmatrix} 1 & C_k^1 & C_k^2 & \dots & C_k^{n-1} \\ 1 & C_{k+1}^1 & C_{k+2}^2 & \dots & C_{k+1}^{n-1} \\ 1 & C_{k+2}^1 & C_{k+2}^2 & \dots & C_{k+2}^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & C_{k+n-1}^1 & C_{k+n-1}^2 & \dots & C_{k+n-1}^{n-1} \end{vmatrix} \qquad \text{(ordre } n\text{)}$$

Réponses :

$A_n = a_1 (a_1 - a_2)^{n-1}$ (on soustrait la première colonne des autres)

$B_{n+1} = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$

$D_n = 1$ (en retranchant chaque ligne de la suivante, on obtient $D_n = D_{n-1}$)

■ **8.** Étant donné n réels $\alpha_1, \alpha_2, \dots, \alpha_n$, calculer le déterminant Δ_n, d'ordre n, dont le terme général est $a_{ij} = \sin(\alpha_i + \alpha_j)$

Réponse : les vecteurs-colonnes de Δ_n sont tous de la forme $\lambda X + \mu Y$, où X et Y sont les matrices-colonnes :

$$X = \begin{pmatrix} \cos \alpha_1 \\ \cos \alpha_2 \\ \vdots \\ \cos \alpha_n \end{pmatrix} \qquad Y = \begin{pmatrix} \sin \alpha_1 \\ \sin \alpha_2 \\ \vdots \\ \sin \alpha_n \end{pmatrix}$$

Pour $n \geqslant 3$, ils sont liés et $\Delta_n = 0$.

Par contre $\Delta_1 = \sin 2\alpha_1$, $\Delta_2 = -\sin^2(\alpha_1 - \alpha_2)$, en général non nuls.

■ **9.** Démontrer que toute matrice antisymétrique d'ordre impair est singulière.

$$A \text{ antisymétrique} \Leftrightarrow {}^tA = -A$$

Comme $\quad$ dét $(\lambda A) = \lambda^n$ dét A, $\quad$ dimension n, $\quad$ on a

$$\text{dét}\,({}^tA) = \text{dét}\,A = (-1)^n\,\text{dét}\,A$$

donc, si n est impair, $\quad$ dét $A = 0$ $\quad$ et $\quad A$ est singulière.

■ **10.** Calculer le déterminant

$$\Delta = \begin{vmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{vmatrix}$$

en calculant d'abord Δ^2.

Le produit de deux déterminants de même ordre peut être effectué ligne par ligne, ou ligne par colonne, ou colonne par ligne, ou colonne par colonne. En opérant ici colonne par colonne, on obtient :

$$\Delta^2 = \begin{vmatrix} a^2+b^2+c^2+d^2 & 0 & 0 & 0 \\ 0 & a^2+b^2+c^2+d^2 & 0 & 0 \\ 0 & 0 & a^2+b^2+c^2+d^2 & 0 \\ 0 & 0 & 0 & a^2+b^2+c^2+d^2 \end{vmatrix}$$
$$= (a^2+b^2+c^2+d^2)^4$$

Il en résulte $\quad \Delta = \varepsilon\,(a^2+b^2+c^2+d^2)^2 \quad$ où $\varepsilon = \pm 1$.

La considération du terme en a^4 montre que $\varepsilon = 1$, ainsi

$$\Delta = (a^2+b^2+c^2+d^2)^2$$

■ **11.** Calculer le déterminant circulant

$$D_n = \begin{vmatrix} 1 & 2 & 3 \ldots & n-1 & n \\ n & 1 & 2 \ldots & n-2 & n-1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 2 & 3 & 4 \ldots & n & 1 \end{vmatrix}$$

En ajoutant à la première ligne toutes les autres, on obtient

$$D_n = \frac{n\,(n+1)}{2} \begin{vmatrix} 1 & 1 & 1 \ldots & 1 & 1 \\ n & 1 & 2 \ldots & n-2 & n-1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 2 & 3 & 4 \ldots & n & 1 \end{vmatrix}.$$

En retranchant à la $n^{\text{ième}}$ colonne la $(n-1)^{\text{ième}}$, à celle-ci la $(n-2)^{\text{ième}}$, ..., à la 2^e la 1^{re}

$$D_n = \frac{n(n+1)}{2}\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ n & 1-n & 1 & & 1 \\ n-1 & 1 & 1-n & & 1 \\ & & & & \\ 3 & 1 & 1 & \dots & 1 \\ 2 & 1 & 1 & \dots & 1-n \end{vmatrix} = \frac{n(n+1)}{2}\Delta_{n-1}$$

où

$$\Delta_{n-1} = \begin{vmatrix} 1-n & 1 & \dots & 1 \\ 1 & 1-n & & 1 \\ 1 & 1 & \dots & 1 \\ & & & \\ 1 & 1 & \dots & 1-n \end{vmatrix} \qquad (\text{ordre } n-1)$$

En effectuant sur Δ_{n-1} les mêmes opérations que sur D_n, il vient

$$\Delta_{n-1} = -\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1-n & 1 & \dots & 1 \\ 1 & 1 & 1-n & \dots & 1 \\ \vdots & \vdots & \vdots & & \\ 1 & 1 & 1 & \dots & 1-n \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & -n & 0 & \dots & 0 \\ 1 & 0 & -n & \dots & 0 \\ \vdots & \vdots & \vdots & & \\ 1 & 0 & 0 & \dots & -n \end{vmatrix}$$

$$= -(-n)^{n-2} = (-1)^{n-1} n^{n-2}$$

d'où

$$D_n = (-1)^{n-1}\frac{n^{n-1}(n+1)}{2}$$

■ **12.** On donne n nombres $a_1, a_2, \dots, a_n$, et on considère les polynômes

$$P_i(x) = \prod_{k \neq i}(x - a_k) \qquad i = 1, 2, \dots, n$$

On forme le déterminant

$$\Delta(x) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_n \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \\ a_1^{n-2} & a_2^{n-2} & \dots & a_n^{n-2} \\ P_1(x) & P_2(x) & \dots & P_n(x) \end{vmatrix}$$

a) Montrer que $\Delta(x)$ est indépendant de x.

b) En supposant les a_i distincts, déterminer n nombres b_i tels que

$$\sum_1^n b_i P_i(x) = 1$$

a) Supposons d'abord les a_i distincts, $P_i(x)$ est un polynôme de degré $n-1$, dont les zéros sont les a_k tels que $k \neq i$. Donc

$$P_i(a_k) = 0 \qquad \text{si } k \neq i,$$
$$P_i(a_i) = \prod_{k \neq i} (a_i - a_k)$$

Pour montrer que $\Delta(x)$ est indépendant de x, comparons $\Delta(a_1)$, $\Delta(a_2)$, $\ldots$, $\Delta(a_n)$. On a $P_2(a_1) = P_3(a_1) = \ldots = P_n(a_1) = 0$, soit en développant $\Delta(a_1)$ suivant sa dernière ligne :

$$\Delta(a_1) = (-1)^{n+1} \prod_{k \neq 1} (a_1 - a_k) \begin{vmatrix} 1 & 1 & \ldots & 1 \\ a_2 & a_3 & \ldots & a_n \\ a_2^2 & a_3^2 & \ldots & a_n^2 \\ a_2^{n-2} & a_3^{n-2} & \ldots & a_n^{n-2} \end{vmatrix}$$
$$= \prod_{k \neq 1} (a_k - a_1) . \prod_{2 \leqslant j < i \leqslant n} (a_i - a_j)$$
$$= \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j) = \lambda$$

λ n'est autre que le déterminant de Van der Monde des a_i.

Calculons de même $\Delta(a_i)$: le seul terme non nul dans la dernière ligne sera $P_i(a_i)$, sur la $i^{\text{ième}}$ colonne, et nous aurons par développement suivant la dernière ligne :

$$\Delta(a_i) = (-1)^{n+i} \prod_{k \neq i} (a_i - a_k) \begin{vmatrix} 1 & \ldots & 1 & 1 & \ldots & 1 \\ a_1 & & a_{i-1} & a_{i+1} & & a_n \\ a_1^2 & \ldots & a_{i-1}^2 & a_{i+1}^2 & \ldots & a_n^2 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^{n-2} & & a_{i-1}^{n-2} & a_{i+1}^{n-2} & & a_n^{n-2} \end{vmatrix}$$
$$= (-1)^{n+i} \prod_{k=1}^{k=i-1} (a_i - a_k) . (-1)^{n-i} \prod_{k=i+1}^{k=n} (a_k - a_i) = \prod_{\substack{1 \leqslant j < k \leqslant n \\ j \neq i \\ k \neq i}} (a_k - a_j)$$
$$= \lambda$$

Ceci vaut $\forall\, i = 1, 2, \ldots, n$. Ainsi le polynôme $\Delta(x) - \lambda$, qui, d'après le développement de $\Delta(x)$ suivant sa dernière ligne, est de degré au plus égal à $n-1$, s'annule pour n valeurs distinctes $a_1, \ldots, a_n$ de x. Il est donc identique à 0, et $\Delta(x) = \lambda \quad \forall\, x$.

Ce raisonnement n'est plus valable si les a_i ne sont pas tous distincts. Mais alors si nous supposons par exemple $a_1 = a_2$, on a aussi $P_1(x) = P_2(x)$, les deux premières colonnes du déterminant $\Delta(x)$ sont identiques, et celui-ci est nul. Comme alors $\lambda = 0$, on peut conclure :

$$\forall\, x, \qquad \Delta(x) = \lambda = \prod_{1 \leqslant j < i \leqslant n} (a_i - a_j)$$

b) Les a_i étant distincts, donc $\lambda \neq 0$, le développement suivant la dernière ligne donne $X_{n1}P_1(x) + X_{n2}P_2(x) + \ldots + X_{nn}P_n(x) = \lambda$, soit, en posant $b_i = \dfrac{X_{ni}}{\lambda}$

$$\sum_1^n b_i P_i(x) = 1$$

X_{ni} est le cofacteur de l'élément $P_i(x)$ dans $\Delta(x)$, donc

$$X_{ni} = (-1)^{n+i} \begin{vmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ a_1 & & a_{i-1} & a_{i+1} & & a_n \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^{n-2} & \dots & a_{i-1}^{n-2} & a_{i+1}^{n-2} & \dots & a_n^{n-2} \end{vmatrix}$$

$$= (-1)^{n+i} \prod_{\substack{1 \leqslant j < k \leqslant n \\ j \neq i \\ k \neq i}} (a_k - a_j)$$

Donc on trouve pour valeur de b_i :

$$(-1)^{n+i} \frac{1}{(a_i - a_1)(a_i - a_2)\dots(a_i - a_{i-1})(a_n - a_i)(a_{n-1} - a_i)\dots(a_{i+1} - a_i)}$$

ou $$b_i = \frac{(-1)^{n+1}}{\prod\limits_{j \neq i} (a_j - a_i)} = \frac{1}{\prod\limits_{j \neq i} (a_i - a_j)}$$

On a donc la relation $\sum\limits_{i=1}^{i=n} \left(\prod\limits_{j \neq i} \dfrac{x - a_j}{a_i - a_j} \right) = 1$ facile à vérifier directement puisque le polynôme du premier membre, de degré au plus égal à $n - 1$, prend la valeur 1 pour n valeurs $x = a_1$, $x = a_2$, $\dots$, $x = a_n$.

■ **13.** Calculer en fonction de $\cos \theta_1$, $\cos \theta_2$, $\dots$, $\cos \theta_n$ le déterminant d'ordre n

$$\Delta = \begin{vmatrix} 1 & \cos \theta_1 & \cos 2\,\theta_1 & \dots & \cos (n-1)\,\theta_1 \\ 1 & \cos \theta_2 & \cos 2\,\theta_2 & \dots & \cos (n-1)\,\theta_2 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \cos \theta_n & \cos 2\,\theta_n & \dots & \cos (n-1)\,\theta_n \end{vmatrix}$$

Posons $\cos \theta = x$, on a $\cos 2\,\theta = 2\,x^2 - 1$

$$\cos 3\,\theta = 4\,x^3 - 3\,x.$$

D'une façon générale $\cos k\theta$ est un polynôme de degré k en x, dont le terme de plus haut degré peut être calculé par récurrence en écrivant

$$\cos k\theta + \cos (k-2)\,\theta = 2 \cos (k-1)\,\theta \cos \theta \tag{1}$$

Si a_k est le coefficient de x^k dans $\cos k\theta$, l'égalité des termes en x^k dans (1) donne $a_k = 2\,a_{k-1}$ d'où aussitôt $a_k = 2^{k-1} a_1 = 2^{k-1}$ puisque $\cos \theta = x$.

En posant $\cos \theta_i = x_i$ la troisième colonne de Δ s'obtient en retranchant de $2 \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix}$ la colonne $\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, que l'on peut donc supprimer comme identique à la première colonne de Δ. Plus généralement la $(k+1)^{\text{ième}}$ colonne de Δ s'obtient en ajoutant à la colonne $2^{k-1}\begin{pmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{pmatrix}$ une combinaison linéaire des k premières colonnes de Δ, combinaison que l'on peut supprimer.

Il reste ainsi

$$\Delta = \begin{vmatrix} 1 & x_1 & 2\,x_1^2 & \dots & 2^{n-1}\,x_1^n \\ 1 & x_2 & 2\,x_2^2 & & 2^{n-1}\,x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & 2\,x_n^2 & & 2^{n-1}\,x_n^n \end{vmatrix}$$

$$= 2^{1+2+3+\dots+(n-1)} \prod_{1 \leqslant j < i \leqslant n} (x_i - x_j)$$

soit
$$\Delta = 2^{\frac{n(n-1)}{2}} \prod_{1 \leqslant j < i \leqslant n} (\cos\theta_i - \cos\theta_j)$$

■ **14.** On considère le déterminant, d'ordre n, $\Delta(x)$, dont le terme général est une fonction $a_{ij}(x)$ dérivable sur un domaine $\mathbb{D}$. Établir une formule donnant sa dérivée.

Application : calculer
$$\Delta(x) = \begin{vmatrix} x+a_1 & x & x & \dots & x \\ x & x+a_2 & x & & x \\ x & x & x+a_3 & & x \\ \vdots & \vdots & \vdots & & \vdots \\ x & x & x & & x+a_n \end{vmatrix}$$

Réponses :

– En appliquant la définition du déterminant

$$\Delta = \sum_{\varphi} \sigma(\varphi)\, a_{1j_1}\, a_{2j_2} \dots a_{nj_n}$$

on obtient, en désignant par $\overrightarrow{V_i}(x)$ le $i^{\text{ième}}$ vecteur-colonne de Δ

$$\Delta'(x) = \sum_{i=1}^{i=n} |\,\overrightarrow{V_1}, \overrightarrow{V_2}, \dots, \overrightarrow{V_{i-1}}, \overrightarrow{V_i'}, \overrightarrow{V_{i+1}}, \dots, \overrightarrow{V_n}\,|$$

formule qui généralise celle d'un produit.

– On obtient facilement $\Delta''(x) = 0$

donc $\Delta'(x) = \alpha \qquad \Delta(x) = \alpha x + \beta$

α et β étant des constantes que l'on calcule en donnant à x la valeur 0.

On obtient ainsi :

$$\Delta(x) = x \sum_{i=1}^{i=n} \left(\prod_{j \neq i} a_j \right) + \prod_{j=1}^{j=n} a_j$$

ou, si les a_j sont non nuls,

$$\Delta(x) = \left(\prod_{j=1}^{j=n} a_j\right)\left(1 + x\sum_{j=1}^{j=n} \frac{1}{a_j}\right)$$

■ **15.** Soit A une matrice carrée d'ordre n, $\Delta = \text{dét}\, A$.

a) Exprimer en fonction de Δ le déterminant D de la comatrice de A.

b) En dimension 3, le repère étant supposé orthonormé, traduire vectoriellement la relation ainsi obtenue, et retrouver cette relation par un calcul purement vectoriel.

---

a) On sait que $A\,.\,{}^t(\text{com}\, A) = \Delta\,.\,I$, I étant la matrice unité d'ordre n. En prenant les déterminants des deux membres, on obtient :

$$\Delta D = \text{dét}\,(\Delta I) = \Delta^n\, \text{dét}\, I = \Delta^n$$

Deux cas se présentent.

Si $\Delta \neq 0$, $D = \Delta^{n-1}$.

Si $\Delta = 0$, la relation précédente ne donne rien, mais A est singulière. Alors si $A = 0$, les cofacteurs de A, éléments de com A, sont nuls, et $D = 0$. Si $A \neq 0$, com $A \neq 0$; on a aussi $B = {}^t(\text{com}\, A) \neq 0$, mais $AB = \Delta I = 0$, A et B sont des diviseurs de zéro et sont des matrices singulières (ce que l'on savait déjà concernant A), et $D = \text{dét}\, B = 0$. On a donc toujours $D = \Delta^{n-1}$.

Enfin, si $A \neq 0$ et com $A = 0$, $B = 0$ et dét $B = 0 = \Delta^{n-1}$.

En conclusion, pour une matrice carrée d'ordre n,

$$\text{dét}\,(\text{com}\, A) = (\text{dét}\, A)^{n-1} \tag{1}$$

b) En dimension 3, on a donc $D = \Delta^2$. Si on désigne par $\vec{U}$, $\vec{V}$, $\vec{W}$ les vecteurs de coordonnées (a, b, c), (a', b', c'), (a'', b'', c''), colonnes de A, dans un repère orthonormé, Δ est le produit mixte $(\vec{U}, \vec{V}, \vec{W})$. Les cofacteurs de a, b, c dans $\Delta = \begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}$ sont

$$\alpha = b'c'' - c'b'', \quad \beta = c'a'' - a'c'', \quad \gamma = a'b'' - b'a'',$$

le vecteur (α, β, γ) est le produit vectoriel $\vec{V} \wedge \vec{W}$, et, en raisonnant de même sur les autres cofacteurs, on voit que la relation (1) du (*a*) s'écrit

$$(\vec{V} \wedge \vec{W}, \vec{W} \wedge \vec{U}, \vec{U} \wedge \vec{V}) = (\vec{U}, \vec{V}, \vec{W})^2$$

On peut aussi l'établir ainsi :

$$P = (\vec{V} \wedge \vec{W}, \vec{W} \wedge \vec{U}, \vec{U} \wedge \vec{V}) = [(\vec{V} \wedge \vec{W}) \wedge (\vec{W} \wedge \vec{U})]\,.\,(\vec{U} \wedge \vec{V})$$
$$= \vec{S}\,.\,(\vec{U} \wedge \vec{V})$$

D'après la formule du double produit vectoriel

$$(\vec{a} \wedge \vec{b}) \wedge \vec{c} = (\vec{a}\,.\,\vec{c})\,\vec{b} - (\vec{b}\,.\,\vec{c})\,\vec{a}$$

$$\vec{S} = [\vec{V} \cdot (\vec{W} \wedge \vec{U})]\,\vec{W} - [\vec{W} \cdot (\vec{W} \wedge \vec{U})]\,\vec{V} = (\vec{U}, \vec{V}, \vec{W})\,\vec{W}$$

car le second produit scalaire est nul, et le produit mixte est invariant par permutation circulaire.

Donc

$$P = (\vec{U}, \vec{V}, \vec{W})\,[\vec{W} \cdot (\vec{U} \wedge \vec{V})] = (\vec{U}, \vec{V}, \vec{W})^2$$

■ **16.** Démontrer qu'un polynôme $P(x)$ de degré n est complètement déterminé quand on donne ses valeurs numériques pour $n + 1$ valeurs distinctes de la variable x : $P(x_i) = \alpha_i, \quad i = 1, 2, \ldots, n, n + 1$.

Réponse : le déterminant du système donnant les coefficients de $P(x)$ est un déterminant de Van der Monde $\prod_{1 \leq j < i \leq n+1} (x_i - x_j)$, non nul, on a donc affaire à un système de Cramer, et les coefficients de $P(x)$ sont déterminés sans ambiguïté.

■ **17.** Écrire à l'aide d'un déterminant l'équation, dans $\mathbb{R}^2$, d'un cercle (ou d'une droite) passant par trois points donnés $A_1(x_1, y_1)$, $A_2(x_2, y_2)$, $A_3(x_3, y_3)$. Même question dans $\mathbb{R}^3$ pour une sphère passant par quatre points donnés.

En exprimant que le système homogène en (a, b, c, d)

$$\begin{cases} a(x^2 + y^2) + bx + cy + d = 0 \\ a(x_1^2 + y_1^2) + bx_1 + cy_1 + d = 0 \\ a(x_2^2 + y_2^2) + bx_2 + cy_2 + d = 0 \\ a(x_3^2 + y_3^2) + bx_3 + cy_3 + d = 0 \end{cases}$$

admet d'autres solutions que $(0, 0, 0, 0)$, on obtient l'équation cherchée :

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0$$

Les points A_1, A_2, A_3 sont alignés si et seulement si

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

On obtient de même dans $\mathbb{R}^3$ l'équation de la sphère passant par $A_1(x_1, y_1, z_1), \ldots, A_4(x_4, y_4, z_4)$

$$\begin{vmatrix} x^2 + y^2 + z^2 & x & y & z & 1 \\ x_1^2 + y_1^2 + z_1^2 & x_1 & y_1 & z_1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_4^2 + y_4^2 + z_4^2 & x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0$$

Les quatre points A_1, A_2, A_3, A_4 sont coplanaires si et seulement si

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0$$

On pourrait obtenir de même, dans $\mathbb{R}^2$, l'équation d'une courbe du second ordre passant par 5 points donnés; dans $\mathbb{R}^3$, l'équation d'une surface du second ordre passant par 9 points donnés.

18. Résoudre et discuter, sur $\mathbb{C}$, le système d'équations

$$(S)\quad \left|\begin{array}{l} ax_1 + x_2 + x_3 + \ldots + x_n = 1 \\ x_1 + ax_2 + x_3 + \ldots + x_n = b \\ x_1 + x_2 + ax_3 + \ldots + x_n = b^2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1 + x_2 + x_3 + \ldots + ax_n = b^{n-1} \end{array}\right|$$

Par addition on obtient

$$(a + n - 1)\sum_1^n x_j = 1 + b + b^2 + \ldots + b^{n-1} \qquad \text{(I)}$$

D'où la discussion suivante : si $a \neq 1 - n$

$$\text{(I)} \Rightarrow \sum_1^n x_j = \frac{\sum_0^{n-1} b^j}{a + n - 1} \qquad \text{(II)}$$

En retranchant (II) successivement des n équations de (S), on obtient

$$(a - 1)\, x_j = b^{j-1} - \frac{\sum_0^{n-1} b_j}{a + n - 1} \qquad (i = 1, 2, \ldots, n)$$

Donc si $a \neq 1$ et $a \neq 1 - n$, le système (S) admet la solution unique

$$x_j = \frac{b^{j-1}}{a - 1} - \frac{\sum_0^{n-1} b^j}{(a - 1)(a + n - 1)} \qquad (i = 1, 2, \ldots, n)$$

Remarquons que si $b \neq 1$, $x_j = \dfrac{b^{j-1}}{a - 1} - \dfrac{1 - b^n}{(1 - b)(a - 1)(a + n - 1)}$

$$\text{et si } b = 1,\ x_j = \frac{1}{a - 1} - \frac{n}{(a - 1)(a + n - 1)}$$

$$= \frac{1}{a + n - 1} \quad (x_1 = x_2 = \ldots = x_n)$$

Cas particuliers :

- $a = 1 - n$; (I) $\Rightarrow 0 = 1 + b + b^2 + \ldots + b^{n-1}$

Cette équation est impossible sauf si b est racine $n^{\text{ième}}$ de l'unité. autre que 1, c'est-à-dire si $b \neq e^{2i\frac{k\pi}{n}}$, $k = 1, 2, \ldots, n-1$. (S) est alors également impossible.

Si $b = e^{2i\frac{k\pi}{n}}$, les équations de (S) sont liées et on est ramené à $n - 1$ équations à n inconnues, par exemple

$$(S') \begin{cases} (1-n)\,x_1 + x_2 + \ldots + x_n = 1 \\ x_1 + (1-n)\,x_2 + \ldots + x_n = b \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_1 + x_2 + \ldots + (1-n)\,x_{n-1} + x_n = b^{n-2} \end{cases}$$

On peut prendre x_n arbitraire, et, par addition, obtenir

$$-\sum_1^{n-1} x_j + (n-1)\,x_n = 1 + b + b^2 + \ldots + b^{n-2} = -b^{n-1} = -\frac{1}{b}$$

donc
$$\sum_1^{n-1} x_j = (n-1)\,x_n + \frac{1}{b}$$

En retranchant de cette équation, successivement, celles de (S'), on a

$$nx_1 = nx_n + \frac{1}{b} - 1, \quad x_1 = x_n + \frac{1}{n}\left(e^{-2i\frac{k\pi}{n}} - 1\right)$$

et de même
$$x_j = x_n + \frac{1}{n}\left(e^{-2i\frac{k\pi}{n}} - e^{2i(j-1)\frac{k\pi}{n}}\right)$$

- $a = 1$; par report direct dans (S), on voit que (S) est impossible si $b \neq 1$; indéterminé à $n - 1$ paramètres (système de rang 1) si $b = 1$: on peut alors fixer arbitrairement $x_1, x_2, \ldots, x_{n-1}$, et obtenir

$$x_n = 1 - (x_1 + x_2 + \ldots + x_{n-1})$$

Remarquons que le calcul du déterminant de (S) n'a pas été utilisé ici. Le lecteur pourra établir que ce déterminant, du type circulant, est

$$\Delta = (a + n - 1)\,(a - 1)^{n-1}$$

ainsi (S) est un système de Cramer lorsque $a \neq 1$ et $a \neq 1 - n$. Cependant l'application des formules de Cramer serait malaisée ici.

■ **19.** Résoudre et discuter sur $\mathbb{C}$ le système

$$\begin{cases} x - ay + a^2z = a \\ ax - a^2y + az = 1 \\ ax + y - a^3z = 1 \end{cases}$$

Réponse : on trouve le déterminant $\Delta = a\,(a^2 + 1)\,(a^2 - 1)$.

Si $a \notin \{0, 1, -1, i, -i\}$, système de Cramer : on trouve $x = a$, $y = 1$, $z = \dfrac{1}{a}$.

Si $a = 0$, système impossible.

Si $a = \pm 1$ et $a = \pm i$, système indéterminé de rang 2 (un paramètre).

20. λ étant un réel, on considère le système de n équations à n inconnues.

$$(S)\left\{\begin{array}{r} 2\lambda x_1 + x_2 = 1 \\ x_1 + 2\lambda x_2 + x_3 = 0 \\ x_2 + 2\lambda x_3 + x_4 = 0 \\ \dots\dots\dots\dots\dots\dots \\ x_{n-2} + 2\lambda x_{n-1} + x_n = 0 \\ x_{n-1} + 2\lambda x_n = 0 \end{array}\right.$$

Soit Δ_n le déterminant de (S).

a) Exprimer Δ_n à l'aide de Δ_{n-1} et Δ_{n-2}.

b) Établir que si $\Delta_n \neq 0$, $x_k = (-1)^{k+1} \dfrac{\Delta_{n-k}}{\Delta_n}$, $k = 1, 2, \dots, n$.

c) On suppose $|\lambda| < 1$ et on pose $\lambda = \cos\theta$. Calculer Δ_n, et calculer x_k lorsque c'est possible.

Réponses :

a) Par développement suivant la première ligne de

$$\Delta_n = \begin{vmatrix} 2\lambda & 1 & 0 & \dots\dots & 0 \\ 1 & 2\lambda & 1 & \dots\dots & 0 \\ 0 & 1 & 2\lambda & 1 \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots 1 & 2\lambda \end{vmatrix}$$

on obtient $\Delta_n = 2\lambda\Delta_{n-1} - \Delta_{n-2}$.

b) Cette expression de x_k s'obtient par la formule de Cramer.

c) On vérifie $\Delta_1 = \dfrac{\sin 2\theta}{\sin\theta}$ et $\Delta_2 = \dfrac{\sin 3\theta}{\sin\theta}$. La relation de récurrence du (*a*) permet alors d'établir que $\Delta_n = \dfrac{\sin(n+1)\theta}{\sin\theta}$ d'où

$$x_k = (-1)^{k+1}\frac{\sin(n-k+1)\theta}{\sin(n+1)\theta} \qquad \text{si } \theta \neq \frac{h\pi}{n+1} \quad (h \text{ entier})$$

■ **21.** Étudier, en appliquant le théorème de Fontené-Rouché, le système de trois équations à quatre inconnues, où a et b sont des paramètres :

$$(S)\begin{cases} x + y - z + 2t = 1 \\ ax - 3y + z + t = 2 \\ 5x - 5y + z + 4t = b \end{cases}$$

Réponse : si $a \neq 2$, le système est de rang 3; en prenant pour inconnues principales x, y, z (S) est un système de Cramer et a une solution unique pour toute valeur de t qui peut être fixée arbitrairement.

Si $a = 2$, (S) est de rang 2; son unique déterminant caractéristique (avec x, y inconnues principales) s'annule pour $b = 5$; (S) est impossible pour $a = 2, b \neq 5$.

Si $a = 2$, $b = 5$, c'est un système de Cramer en x, y, où z et t peuvent être choisis arbitraires.

CHAPITRE 8

VALEURS PROPRES. RÉDUCTION D'UN ENDOMORPHISME

1. VALEURS PROPRES, VECTEURS PROPRES, SOUS-ESPACES PROPRES

Définition. f étant un endomorphisme d'un espace vectoriel E sur un corps K, on dit qu'un nombre λ, élément de K, est une **valeur propre** de f, s'il existe un vecteur $\vec{u}$ non nul tel que

$$\boxed{f(\vec{u}) = \lambda\vec{u}} \qquad (1)$$

soit $f(\vec{u})$ colinéaire à $\vec{u}$.

On dit alors que $\vec{u}$ est un **vecteur propre** de f, associé à la valeur propre λ. Comme $(1) \Leftrightarrow (f - \lambda I)(\vec{u}) = \vec{0}$, on voit que l'ensemble des vecteurs propres associés à une valeur propre λ n'est autre que le noyau de l'endomorphisme $f - \lambda I$, c'est un sous-espace de E, appelé **sous-espace propre** associé à la valeur propre λ, que nous noterons E_λ : il contient par définition, un vecteur non nul, il a donc une dimension $\geqslant 1$. Nous retiendrons

$$\boxed{E_\lambda = \ker(f - \lambda I)} \qquad (I = \text{endormorphisme-unité})$$

Exemples.

a) Soit, dans l'espace euclidien $\mathbb{R}^3$, la symétrie f par rapport à un plan P : $f(\vec{u})$ est colinéaire à $\vec{u}$ si $\vec{u} \in P$ ou $\vec{u} \perp P$, et dans ces deux cas seulement.

Dans le premier cas, $f(\vec{u}) = \vec{u}$, $\lambda = 1$, $E_1 = P$.

Dans le second, $f(\vec{u}) = -\vec{u}$, $\lambda = -1$, $E_{-1} = \Delta$, droite vectorielle orthogonale à P. f admet les deux valeurs propres 1 et -1.

b) Soit, dans l'espace E des fonctions $u(x)$ définies et dérivables sur $\mathbb{R}$, l'application $u(x) \underset{f}{\to} u'(x)$. Pour tout λ réel (ou même complexe, si l'on envisage u fonction complexe d'une variable réelle),

$$u'(x) = \lambda u(x) \Leftrightarrow u(x) = ke^{\lambda x}$$

on voit que f admet tout nombre λ pour valeur propre et que les sous-espaces propres correspondants sont les « droites vectorielles » de bases respectives $e^{\lambda x}$.

2. RECHERCHE DES VALEURS PROPRES EN DIMENSION FINIE. POLYNOME CARACTÉRISTIQUE

a) Si $\dim E = n$, et si nous choisissons dans E une base $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$, f est représenté par une matrice carrée d'ordre n, et les valeurs propres sont les scalaires λ tels que $ker\,(f - \lambda I) \neq \{\vec{0}\}$ donc

$$\text{dét}\,(f - \lambda I) = \boxed{\text{dét}\,(A - \lambda I) = 0}$$

Si $A = \begin{pmatrix} & \vdots & \\ \ldots & a_{ij} & \ldots \\ & \vdots & \end{pmatrix}$ ce déterminant, d'ailleurs indépendant de la base choisie dans E, s'écrit

$$\varphi(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{vmatrix}$$

Il s'obtient à partir de celui de A en retranchant λ aux éléments diagonaux a_{ii}. Un développement suivant, par exemple, la première ligne montre que $\varphi(\lambda)$ est un polynôme de degré n. On le nomme **polynôme caractéristique** de l'endomorphisme f, ou de la matrice A.

L'équation $\varphi(\lambda) = 0$, dite équation caractéristique, a pour racines les valeurs propres de f. Si le corps de référence de E est $\mathbb{R}$, f admet donc au maximum n valeurs propres, distinctes ou confondues. Si ce corps de référence est $\mathbb{C}$, ce qui est le cas usuel dans les applications, **tout endomorphisme de E admet n valeurs propres, distinctes ou confondues.**

Exemple.

Déterminons les valeurs propres et les vecteurs propres de f défini dans $\mathbb{R}^2$ par $A = \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}$.

$$\varphi(\lambda) = \begin{vmatrix} 2-\lambda & 3 \\ 4 & 1-\lambda \end{vmatrix} = \lambda^2 - 3\lambda - 10 = 0 \quad \text{pour} \quad \lambda_1 = 5, \quad \lambda_2 = -2$$

λ étant un de ces deux nombres, un vecteur propre associé à λ a pour matrice

$$\begin{pmatrix} x \\ y \end{pmatrix} \quad \text{telle que} \quad \begin{cases} (2-\lambda)\,x + 3\,y = 0 \\ 4\,x + (1-\lambda)\,y = 0 \end{cases}$$

système qui se réduit à une seule équation puisqu'il est homogène et n'est pas « de Cramer »; pour $\lambda = 5$, nous obtenons

$$\begin{cases} -3\,x + 3\,y = 0 \\ 4\,x - 4\,y = 0 \end{cases}$$

d'où la droite vectorielle $x = y$, sous-espace propre associé à λ_1.

On obtient de même pour $\lambda = -2$ la droite vectorielle $4\,x + 3\,y = 0$.

b) Revenons au cas général et cherchons à expliciter en partie le polynôme caractéristique.

Son terme de plus haut degré provient, dans le développement du déterminant comme forme multilinéaire alternée, du produit des éléments diagonaux $(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda)$ c'est donc $(-1)^n \lambda^n$.

Cherchons le coefficient de λ^{n-1} : dans le produit qui précède, c'est

$$(-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn}) = (-1)^{n-1} \sum_1^n a_{ii} = \alpha_1$$

Or un des produits, apparaissant dans la forme multilinéaire alternée $\varphi(\lambda)$, contenant un élément non diagonal, par exemple a_{21}, devra contenir un autre élément non diagonal au moins (car $a_{22} - \lambda$ est à exclure), donc ne pourra faire apparaître λ que $n - 2$ fois au plus, et sera de degré $\leqslant n - 2$, en λ.

Donc α_1 constitue tout le coefficient de λ^{n-1} dans $\varphi(\lambda)$. La somme des termes diagonaux de A s'appelle la trace de A, notée tr A, on a donc $\alpha_1 = \operatorname{tr} A$.

Les coefficients suivants $\alpha_2, \alpha_3, \dots, \alpha_n$ de $\varphi(\lambda)$ sont moins simples, sauf le terme constant α_n qui n'est autre que $\varphi(0) = \text{dét } A = \Delta$. Donc

$$\varphi(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\left(\sum_1^n a_{ii}\right)\lambda^{n-1} + \alpha_2 \lambda^{n-2} + \dots + \alpha_{n-1}\lambda + \Delta$$

Or, d'après les relations entre les coefficients et les racines d'un polynôme de degré n, la somme de ces racines est égale à $-\dfrac{\alpha_1}{\alpha_0}$ et leur produit à $(-1)^n \dfrac{\alpha_n}{\alpha_0}$ avec $\alpha_0 = (-1)^n$ coefficient de λ^n. Comme ces racines sont les valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$ de f, on en déduit

$$\boxed{\sum_1^n \lambda_i = \sum_1^n a_{ii} = \operatorname{tr} A} \qquad \boxed{\prod_1^n \lambda_i = \Delta = \text{dét } A}$$

Remarques.

$\Delta = 0 \Leftrightarrow 0$ est une valeur propre, donc A est singulière, résultat bien connu.

– La trace d'une matrice est invariante par changement de base.

3. RÉDUCTION D'UN ENDOMORPHISME POSSÉDANT n VALEURS PROPRES DISTINCTES

a) Revenons sur l'exemple du paragraphe précédent. Deux vecteurs propres, par exemple $\vec{u_1}$ (1, 1) et $\vec{u_2}$ (3, − 4), associés respectivement à $\lambda_1 = 5$, $\lambda_2 = -2$, forment une base de E, et comme par définition

$$f(\vec{u_1}) = 5\,\vec{u_1} \qquad \text{et} \qquad f(\vec{u_2}) = -2\,\vec{u_2}$$

la matrice de f dans cette base est $D = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$, elle est diagonale.

On dit que la **base propre** $(\vec{u_1}, \vec{u_2})$ permet de diagonaliser f, et aussi A.

On a $D = P^{-1}AP$ avec $P = \begin{pmatrix} 1 & 3 \\ 1 & -4 \end{pmatrix}$.

b) Plus généralement, soit f un endomorphisme de l'espace E de dimension n, possédant n valeurs propres distinctes $\lambda_1, \lambda_2, \ldots, \lambda_n$.

D'une façon générale considérons p valeurs propres distinctes de f, et des vecteurs propres associés $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p}$: ils engendrent un sous-espace F de E, et $\dim F = q \leqslant p$. Si $\vec{v} \in F$, il existe des scalaires $\alpha_1, \alpha_2, \ldots, \alpha_p$ tels que

$$\vec{v} = \sum_1^p \alpha_i\,\vec{u_i} \qquad \text{et} \qquad f(\vec{v}) = \sum_1^p \alpha_i\, f(\vec{u_i}) = \sum_1^p \alpha_i \lambda_i\,\vec{u_i}$$

donc $f(\vec{v}) \in F$ et $f(F) \subset F$. La restriction g de f à F est un endomorphisme de F dont $\lambda_1, \lambda_2, \ldots, \lambda_p$ sont des valeurs propres, elles sont donc racines du polynôme caractéristique de g qui est de degré q, on a donc $p \leqslant q$. Comme on a aussi $q \leqslant p$, il en résulte $p = q$, $f(F) = F$, par conséquant $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_p})$ est une base de F, d'où le théorème :

Si $\lambda_1, \lambda_2, \ldots, \lambda_p$ sont p valeurs propres distinctes de l'endomorphisme f, p vecteurs propres associés respectivement à ces valeurs propres forment un système libre.

Si nous appliquons ce résultat à $p = n$, nous en déduisons que n vecteurs propres de f, associés aux n valeurs propres distinctes $\lambda_1, \lambda_2, \ldots, \lambda_n$, forment une base de E. Comme, pour chacun de ces vecteurs $\vec{u_i}$, on a $f(\vec{u_i}) = \lambda_i \vec{u_i}$, la matrice de f dans cette base s'écrit

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

Les sous-espaces E_{λ_i} sont les droites vectorielles de bases $(\vec{u_i})$, chacune est définie par un système de rang $n-1$, $(A - \lambda_i I)\, X_i = 0$. P étant la matrice des vecteurs propres $\vec{u_i}$, on a $D = P^{-1}AP$ ou $A = PDP^{-1}$. On peut donc énoncer le théorème :

Si un endomorphisme f d'un espace E de dimension finie n, a n valeurs propres distinctes, c'est-à-dire si son polynôme caractéristique a toutes ses racines simples, il existe une base de E constituée de vecteurs propres de f et, dans cette base propre, f s'exprime par une matrice diagonale

$$D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

les λ_i étant les valeurs propres.

On dit que f et sa matrice dans une base de E sont diagonalisables.

Remarque.

L'ordre où l'on dispose dans D les n valeurs propres n'a aucune importance, mais, une fois choisi, impose le même ordre pour les vecteurs propres dans P.

4. ÉTUDE GÉNÉRALE DE LA DIAGONALISATION

a) Supposons que l'endomorphisme f soit diagonalisable, c'est-à-dire qu'il existe une base $(\vec{u_1}, \ldots, \vec{u_n})$ de l'espace E où f s'exprime par une matrice diagonale (semblable à A)

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

On a, par construction même, $f(\vec{u_1}) = \lambda_1 \vec{u_1}, \ldots, f(\vec{u_n}) = \lambda_n \vec{u_n}$. Les $\vec{u_i}$ sont donc des vecteurs propres de E. La réciproque étant évidente, on voit que pour que f soit diagonalisable, il faut et il suffit qu'il existe une base de E formée de vecteurs propres de f. Une condition suffisante pour qu'il en soit ainsi est celle du paragraphe 3 : l'existence de valeurs propres distinctes. Mais cette condition n'est pas nécessaire puisque, par exemple, toute matrice A semblable à $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$ est diagonalisable, bien qu'admettant la valeur propre double $\lambda = 1$.

b) Nous allons établir le résultat suivant : pour que f soit diagonalisable, il faut et il suffit que, E_i étant le sous-espace propre, de dimension q_i, associé à la valeur propre λ_i, d'ordre k_i

$$\left(i = 1, 2, \ldots, p;\quad 1 \leqslant k_i \leqslant n;\quad 1 \leqslant p \leqslant n;\quad \sum_1^p k_i = n\right)$$

on ait $\dim E_i = k_i$.

— Observons d'abord que si λ_i et λ_j sont des valeurs propres distinctes, $\vec{u} \in E_i \cap E_j \Rightarrow f(\vec{u}) = \lambda_i \vec{u} = \lambda_j \vec{u}$, donc $(\lambda_i - \lambda_j)\vec{u} = \vec{0}$ et $\vec{u} = \vec{0}$. Donc $E_i \cap E_j = \{\vec{0}\}$ et $\dim(E_i + E_j) = q_i + q_j$, puisque $E_i + E_j = E_i \oplus E_j$.

— Comparons q_i à k_i : si dans E nous prenons une base $\mathscr{B}_i$ dont les q_i premiers vecteurs sont ceux d'une base de E_i, la matrice de f dans $\mathscr{B}_i$ est de la forme

$$C = \begin{pmatrix} \lambda_i & 0 & \ldots & 0 & C_{1,q_i+1} & \ldots & C_{1n} \\ 0 & \lambda_i & & 0 & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_i & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & \vdots & & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \ldots & 0 & C_{n,q_i+1} & \ldots & C_{nn} \end{pmatrix}$$

Le polynôme caractéristique $\varphi(\lambda) = \text{dét}\,(C - \lambda I)$ contient en facteur $(\lambda_i - \lambda)^{q_i}$, λ_i est donc valeur propre d'ordre au moins égal à q_i et $1 \leqslant q_i \leqslant k_i$.

— Une base éventuelle de E, constituée de vecteurs propres de f, renfermera r_1 vecteurs de E_1, r_2 vecteurs de E_2, $\ldots$, r_p vecteurs de E_p, avec les conditions

$$\sum_1^p r_i = n, \quad r_i \leqslant \dim E_i = q_i \leqslant k_i \quad \text{Or} \quad \sum_1^p k_i = n$$

il faut donc que $\forall\, i = 1, 2, \ldots, p$, $r_i = q_i = k_i$.

— Voyons si cette condition est suffisante : si $q_i = k_i$, $\forall\, i$, une base $\mathscr{B}_i$ de vecteurs de E_i comporte k_i vecteurs. L'union de ces p bases $\mathscr{B}_i$, union comportant n vecteurs, sera une base de E — et alors une base propre — si et seulement si ces n vecteurs sont indépendants. Or une relation de dépendance entre eux est une relation de la forme

$$\sum_{i=1}^{i=p} \alpha_i \vec{u_i} = \vec{0} \tag{1}$$

où $\vec{u_i}$ est un vecteur de E_i, non nul (combinaison de vecteurs de $\mathscr{B}_i$).

Alors $(1) \Rightarrow \sum_1^p \alpha_i f(\vec{u_1}) = \vec{0}$ donc $\sum_1^p \alpha_i \lambda_i \vec{u_i} = \vec{0}$

De même $\sum_1^p \alpha_i \lambda_i f(\vec{u_i}) = \vec{0}$ donc $\sum_1^p \alpha_i \lambda_i^2 \vec{u_i} = \vec{0}$

Ainsi, de proche en proche $\sum_1^p \alpha_i \lambda_i^s \vec{u_i} = 0, \quad \forall s$ entier.

Dans une base quelconque de E, $(\vec{e_1}, \ldots, \vec{e_n})$, les coordonnées des p vecteurs $\alpha_i \vec{u_i}$ suivant $\vec{e_j}$ sont donc solutions d'un système de la forme

$$(S) \begin{cases} x_1 + x_2 + \ldots + x_p = 0 \\ \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_p x_p = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1^{p-1} x_1 + \lambda_2^{p-1} x_2 + \ldots + \lambda_p^{p-1} x_p = 0 \end{cases}$$

Le déterminant de (S) est un déterminant de Van der Monde non nul, puisque les λ_i sont tous distincts, et la solution unique de (S) est

$$x_1 = x_2 = \ldots = x_p = 0$$

Donc $\alpha_i \vec{u_i}$ est nul, ayant toutes ses coordonnées nulles. Comme $\vec{u_i} \neq \vec{0}$, on a $\alpha_i = 0$, les $\vec{u_i}$ sont indépendants, et il existe une base de E formée de vecteurs propres de f.

Nous en déduisons le **théorème fondamental sur la diagonalisation :**

Pour qu'un endomorphisme f d'un espace vectoriel E de dimension finie n (ou la matrice A de f dans une base de E) soit diagonalisable, il faut et il suffit que chaque sous-espace propre de f ait pour dimension l'ordre de multiplicité de la valeur propre correspondante.

Remarques :

— Cet énoncé équivaut encore à dire que la somme des dimensions des sous-espaces propres E_i est égale à n, ou que ceux-ci ont une somme directe égale à E.

— Un endomorphisme, autre qu'une homothétie, ayant une seule valeur propre, λ_1, d'ordre n, n'est pas diagonalisable, car la condition précédente donnerait

$$\dim E_1 = n, \quad \text{donc} \quad E_1 = E \quad \text{et} \quad f(\vec{u}) = \lambda_1 \vec{u} \qquad \forall \vec{u} \in E$$

— Tout projecteur de E est diagonalisable : en effet si f est le projecteur sur l'espace E_1 de dimension p, parallèlement à un supplémentaire E_2

de E_1 dans E. E_1 est un sous-espace propre pour la valeur propre $\lambda_1 = 1$, E_2 un sous-espace propre pour la valeur propre $\lambda_2 = 0$, et comme

$$E_1 \oplus E_2 = E$$

il n'y a pas d'autre sous-espace propre; 1 est valeur propre d'ordre p, 0 valeur propre d'ordre $n - p$, et le polynôme caractéristique est

$$(-1)^n \lambda^{n-p} (\lambda - 1)^p$$

5. TRIGONALISATION

Trigonaliser (on dit aussi « trianguler ») un endomorphisme f (ou une matrice) de E, c'est trouver une base de E dans laquelle f s'exprime par une matrice triangulaire. On dit qu'un endomorphisme est trigonalisable quand il existe une telle base.

La diagonalisation est évidemment un cas particulier de la trigonalisation. Si f est trigonalisable, par exemple sous forme de matrice triangulaire supérieure

$$T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots 0 & a_{nn} \end{pmatrix}$$

dét $(T - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \ldots (a_{nn} - \lambda)$ montre que les a_{ii} sont les valeurs propres de T, donc de f. S'ils sont tous distincts, et même peut-être si plusieurs sont confondus, f est diagonalisable, mais nous allons établir que, d'une façon générale, tout endomorphisme est trigonalisable.

Raisonnons par récurrence sur n : ce résultat est évidemment exact pour $n = 1$. Supposons-le exact dans un espace de dimension $n - 1$.

Soit f un endomorphisme de l'espace vectoriel E de dimension n, sur $\mathbb{C}$. f possède au moins une valeur propre λ_1 et un vecteur propre associé $\vec{u_1}$ non nul : $f(\vec{u_1}) = \lambda_1 \vec{u_1}$.

Soit F la droite vectorielle de base $\vec{u_1}$, et soit G un supplémentaire de F dans E (c'est un hyperplan, sous-espace de dimension $n - 1$); soit g le projecteur de E sur G parallèlement à F. Pour tout $\vec{v}$ de E, il existe un scalaire α tel que $\vec{v} = g(\vec{v}) + \alpha\vec{u_1}$, $g(\vec{v}) \in G$.

En particulier, $\vec{u}$ étant un vecteur quelconque de E, $\vec{v} = f(\vec{u})$ donne

$$f(\vec{u}) = (g \circ f)(\vec{u}) + \alpha\vec{u_1}$$

Si nous posons $g \circ f = h$, la restriction de h à G est un endomorphisme de G, trigonalisable puisque G est de dimension $n - 1$. On peut déterminer une base de G, $\mathscr{B} = (\vec{u_2}, \vec{u_3}, \ldots, \vec{u_n})$ où h s'exprime par une matrice triangulaire; comme F est un supplémentaire de G dans E, $\mathscr{B}' = (\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n})$ est une base de E, et on a :

$$\begin{aligned}
f(\vec{u_1}) &= \lambda_1 \vec{u_1} \\
f(\vec{u_2}) &= h(\vec{u_2}) + \alpha_2 \vec{u_1} \\
f(\vec{u_3}) &= h(\vec{u_3}) + \alpha_3 \vec{u_1} \\
&\ldots\ldots\ldots\ldots\ldots \\
f(\vec{u_n}) &= h(\vec{u_n}) + \alpha_n \vec{u_1}
\end{aligned}$$

Comme la matrice de h dans $\mathscr{B}'$, qui n'est autre que celle obtenue en supprimant la première ligne et la première colonne de celle de f dans $\mathscr{B}$, est triangulaire, cette dernière est de la forme

$$M(f) = \begin{pmatrix} \lambda_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_n \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \ldots & a_{2n} \end{pmatrix}$$

Elle est triangulaire. Nous pouvons donc énoncer le **théorème** :

Tout endomorphisme d'un espace vectoriel de dimension finie est trigonalisable sur $\mathbb{C}$.

— Pratiquement, la trigonalisation, comme la diagonalisation éventuelle, commence par la recherche des valeurs propres et des sous-espaces propres : dans ceux qui sont de dimension égale à l'ordre de multiplicité de la valeur propre associée, on choisit une base de vecteurs propres; dans les autres, une base incomplète de vecteurs propres : on complète ce système par des vecteurs non propres, souvent des vecteurs de la base canonique. Nous allons étudier un exemple.

Soit la matrice

$$A = \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$$

$$\varphi(\lambda) = \text{dét}\,(A - \lambda I) = \begin{vmatrix} 1-\lambda & -3 & 4 \\ 4 & -7-\lambda & 8 \\ 6 & -7 & 7-\lambda \end{vmatrix} \times 2 \, \big) -$$

$$= \begin{vmatrix} 1-\lambda & -3 & 4 \\ 2+2\lambda & -1-\lambda & 0 \\ 6 & -7 & 7-\lambda \end{vmatrix}$$

d'où factorisation de $\lambda + 1$ et, après calcul :

$$\varphi(\lambda) = -(\lambda + 1)^2(\lambda - 3)$$

Cherchons le sous-espace associé à $\lambda_1 = 3$ qui est *a priori* une droite vectorielle représentée par

$$(A - 3I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \quad \text{donc} \quad \begin{cases} -2x - 3y + 4z = 0 \\ 4x - 10y + 8z = 0 \\ 6x - 7y + 4z = 0 \end{cases}$$

soit $2x = y = z$.

D'où le vecteur propre $\vec{u_1}\,(1, 2, 2)$.

Le sous-espace associé à $\lambda_2 = -1$, valeur propre double, peut *a priori* être de dimension 1 ou 2. Le système qui le définit est

$$(A + I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \quad \text{donc} \quad \begin{cases} 2x - 3y + 4z = 0 \\ 4x - 6y + 8z = 0 \\ 6x - 7y + 8z = 0 \end{cases}$$

Il est de rang 2 et définit la droite vectorielle d'équation $z = x$, $y = 2x$, donc de base $\vec{u_2}\,(1, 2, 1)$. Ce sous-espace étant de dimension inférieure à 2, ordre de la valeur propre -1, A n'est pas diagonalisable.

En prenant pour base $(\vec{u_1}, \vec{u_2}, \vec{u_3})$ où $\vec{u_3}$ est un vecteur quelconque indépendant de $\vec{u_1}$ et $\vec{u_2}$, on aura la matrice triangulaire semblable à A :

$$T = \begin{pmatrix} 3 & 0 & a \\ 0 & -1 & b \\ 0 & 0 & -1 \end{pmatrix}$$

Pour préciser la matrice de passage ainsi que a et b, prenons pour $\vec{u_3}$ par exemple le vecteur $\vec{e_1}$ de la base canonique (base où A définit un endomorphisme f) : ce choix est acceptable car dans cette base canonique

$$\text{dét}\,(\vec{u_1}, \vec{u_2}, \vec{e_1}) = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 0 \\ 2 & 1 & 0 \end{vmatrix} = -2 \neq 0$$

Remarquons que $\vec{u_3} = \vec{e_2}$ convient aussi, mais pas $\vec{u_3} = \vec{e_3}$.

La matrice de passage est donc $P = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 0 \\ 2 & 1 & 0 \end{pmatrix}$.

On peut trouver a et b en écrivant

$$T = P^{-1}AP \tag{2}$$

mais il est plus simple de remarquer que

$$\begin{aligned} f(\vec{e_1}) &= a\vec{u_1} + b\vec{u_2} - \vec{e_1} \\ &= \vec{e_1} + 4\,\vec{e_2} + 6\,\vec{e_3} \end{aligned}$$

avec

$$\vec{u_1} = \vec{e_1} + 2\,\vec{e_2} + 2\,\vec{e_3}, \qquad \vec{u_2} = \vec{e_1} + 2\,\vec{e_2} + \vec{e_3}, \qquad \vec{u_3} = \vec{e_1}$$

donc $\vec{e_3} = \vec{u_1} - \vec{u_2}$, $\vec{e_1} = \vec{u_3}$ et

$$\vec{u_2} = \vec{u_3} + 2\,\vec{e_2} + \vec{u_1} - \vec{u_2} \Rightarrow \vec{e_2} = -\frac{1}{2}\vec{u_1} + \vec{u_2} - \frac{1}{2}\vec{u_3}$$

donc

$$\begin{aligned} f(\vec{e_1}) = f(\vec{u_3}) &= \vec{u_3} + 4\left(-\frac{1}{2}\vec{u_1} + \vec{u_2} - \frac{1}{2}\vec{u_3}\right) + 6\,(\vec{u_1} - \vec{u_2}) \\ &= 4\,\vec{u_1} - 2\,\vec{u_2} - \vec{u_3} \end{aligned}$$

Ainsi $a = 4$, $b = -2$ et $T = \begin{pmatrix} 3 & 0 & 4 \\ 0 & -1 & -2 \\ 0 & 0 & -1 \end{pmatrix}$.

Il est alors intéressant de vérifier (2) ou sa forme équivalente

$$PT = AP$$

qui évite, dans certains cas, le calcul de P^{-1} : en fait, celui-ci a été effectué plus haut puisqu'on a exprimé $\vec{e_1}, \vec{e_2}, \vec{e_3}$ en fonction de $\vec{u_1}, \vec{u_2}, \vec{u_3}$.

Notons qu'il y a une infinité de changements de base permettant de trigonaliser une matrice donnée.

6. MATRICES DE JORDAN. RÉDUCTION DE JORDAN

a) Revenons à l'exemple qui précède. Le choix de $\vec{u_1}$, $\vec{u_2}$, vecteurs propres constituant une base incomplète, étant fait, on peut se demander si un choix judicieux de $\vec{u_3}$ pourrait conduire à une forme réduite de A, plus simple que T. La forme $\begin{pmatrix} 3 & 0 & a \\ 0 & -1 & b \\ 0 & 0 & -1 \end{pmatrix}$ étant imposée, et $a = b = 0$ étant *a priori* irréalisable, on peut chercher $\vec{u_3}$ tel que $a = 0$, $b = 1$, c'est-à-dire $f(\vec{u_3}) = \vec{u_2} - \vec{u_3}$, donc, si

$$\vec{u_3} = \alpha\vec{e_1} + \beta\vec{e_2} + \gamma\vec{e_3}, \quad \begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

soit
$$\begin{cases} 2\,\alpha - 3\,\beta + 4\,\gamma = 1 \\ 4\,\alpha - 6\,\beta + 8\,\gamma = 2 \\ 6\,\alpha - 7\,\beta + 8\,\gamma = 1 \end{cases}$$

Ce système se réduit à deux équations et est indéterminé à un paramètre, on peut prendre par exemple $\gamma = 0$, $2\,\alpha - 3\,\beta = 6\,\alpha - 7\,\beta = 1$ donc $\alpha = -1$, $\beta = -1$. La matrice de passage est alors

$$Q = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix}$$

et la forme réduite de A est

$$J = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} = Q^{-1}AQ$$

Une telle matrice est appelée matrice de Jordan, et le calcul qui précède s'appelle réduction de Jordan de la matrice A.

b) D'une façon générale, on appelle **matrice de Jordan** d'ordre n une matrice triangulaire supérieure dans laquelle la diagonale est bordée supérieurement par des éléments égaux à 0 ou 1, ces derniers étant eux-mêmes surmontés exclusivement par des zéros. La forme générale d'une telle matrice est donc

$$J = \begin{pmatrix} \lambda_1 & \alpha_2 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \alpha_3 & \ddots & \vdots \\ 0 & 0 & \lambda_3 & \ddots & 0 \\ 0 & & & \ddots & \alpha_n \\ \vdots & & & & \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

où $\alpha_i = 0$ ou 1, $(i = 2, 3, \ldots, n)$; les λ_i sont les valeurs propres de J et de l'endomorphisme que représente J.

Une matrice diagonale est évidemment une matrice de Jordan particulière.

On démontre, et nous admettrons, que tout endomorphisme d'un espace vectoriel de dimension finie n peut être représenté, dans une base convenable (dite base de Jordan) par une matrice de Jordan. Autrement dit, toute matrice carrée A admet une matrice de Jordan qui lui est semblable, appelée « réduite de Jordan » de A.

Il va de soi que, la diagonalisation étant *a priori* préférée à toute autre réduction, le nombre des α_i égaux à 1 pourra être pris égal à la différence entre n et la somme des dimensions des sous-espaces propres de E dont, ainsi que cela a été fait dans l'exemple (*a*), l'union des bases est prise pour base incomplète de E. Cependant, en général, le choix des vecteurs propres faisant partie d'une base de Jordan n'est pas arbitraire, car, par exemple dans le cas ci-dessus, il faut $(f+I)(\overrightarrow{u_3}) = \overrightarrow{u_2}$ donc $\overrightarrow{u_2} \in \operatorname{Im}(f+I) \cap \ker(f+I)$, condition qui s'est avérée remplie ici puisque le calcul a été possible, mais qui aurait pu imposer des restrictions sur $\overrightarrow{u_1}$. L'étude complète du cas $n = 3$ sera proposée en exercice.

7. APPLICATIONS DE LA RÉDUCTION AU CALCUL DES PUISSANCES D'UNE MATRICE ET AUX SUITES RÉCURRENTES

On a vu que $A' = P^{-1}AP \Rightarrow A'^k = P^{-1}A^kP$, $\forall\ k$ entier positif. On a aussi $A^k = PA'^kP^{-1}$ de sorte que calculer A^k revient à calculer A'^k, A' étant une matrice réduite de A, de préférence diagonale, à défaut triangulaire ou de Jordan.

a) Si A est diagonalisable, de valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$, pas nécessairement toutes distinctes,

$$A' = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} \Rightarrow A'^k = \begin{pmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n^k \end{pmatrix}$$

d'où le calcul de A^k si l'on connaît une matrice de passage P.

b) Si A est seulement trigonalisable, le calcul est moins simple : on peut écrire $A' = D + T$, D étant une matrice diagonale et T une matrice triangulaire d'éléments diagonaux nuls, donc nilpotente, mais en général $DT \neq TD$ et la formule du binôme ne s'applique pas, donc la nilpotence de T est peu utilisable. Par contre si D et T commutent et si l'ordre de A n'est pas trop élevé, le calcul est la plupart du temps facile à partir d'une réduite triangulaire, ou mieux de Jordan.

c) A ce calcul des puissances d'une matrice se ramène celui du terme général d'une suite récurrente vectorielle du type

$$\overrightarrow{u_n} = f(\overrightarrow{u_{n-1}}) \quad \text{ou} \quad \overrightarrow{u_n} = f(\overrightarrow{u_{n-1}}) + \vec{v},$$

f étant un endomorphisme, et $\overrightarrow{u_0}$ et $\vec{v}$ étant des vecteurs donnés.

On peut ramener à cette forme une suite récurrente numérique du type

$$u_n = au_{n-1} + bu_{n-2}, \qquad n \geqslant 2; \qquad u_0, u_1 \text{ donnés}$$

avec $X = PY$, $\Phi = P\Psi$, $A = PBP^{-1}$, P étant la matrice de changement de base.

a) Si B est diagonale, chaque équation du système (S) représenté par (5) est de la forme

$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = \lambda_i y_i + \psi_i(t) \tag{6}$$

elle se résout facilement lorsque les $\psi_i(t)$ et par suite les $\varphi_i(t)$ sont des fonctions « simples » au sens des équations différentielles linéaires à coefficients constants; le calcul de Ψ exige cependant le calcul de P^{-1}, sauf évidemment si Φ est nul, auquel cas on a un système homogène diagonalisable, de solution générale $X = PY$ avec

$$Y = \begin{pmatrix} \alpha_1 \; e^{\lambda_1 t} \\ \alpha_2 \; e^{\lambda_2 t} \\ \cdots\cdots \\ \alpha_n \; e^{\lambda_n t} \end{pmatrix}$$

les α_i étant les constantes d'intégration.

La seule difficulté est la résolution de l'équation caractéristique de f, que l'on appelle aussi équation caractéristique de (S).

b) Si B est une matrice de Jordan, ou même une matrice triangulaire quelconque, la dernière équation du système (S') est encore de la forme (6), la précédente est encore, compte tenu de la connaissance de $y_n(t)$ de la même forme, et le système se résout de proche en proche en remontant de la dernière à la première équation.

Exemple.

En utilisant à nouveau la matrice A des paragraphes 5 et 6, cherchons à résoudre le système différentiel

$$(S) \quad \begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = x - 3\,y + 4\,z \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = 4\,x - 7\,y + 8\,z \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = 6\,x - 7\,y + 7\,z \end{cases} \quad \text{ou} \quad \frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Dans la base de Jordan définie par

$$Q = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix}$$

(notations du 6. *a*) (*S*) s'écrit

$$(S') \quad \begin{cases} \dfrac{dx_1}{dt} = 3\,x_1 \\ \dfrac{dy_1}{dt} = -\,y_1 + z_1 \\ \dfrac{dz_1}{dt} = -\,z_1 \end{cases} \qquad \text{où} \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = Q \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

La première et la dernière équations nous donnent

$$x_1 = \alpha e^{3t}, \qquad z_1 = \gamma e^{-t},$$

α et γ étant des constantes.

L'équation restante s'écrit alors $\dfrac{dy_1}{dt} + y_1 = \gamma e^{-t}$.

La solution générale de l'équation sans second membre est βe^{-t} et on peut chercher une solution particulière de l'équation complète sous la forme $u(t)\,e^{-t}$ d'où

$$(u' - u)\,e^{-t} + ue^{-t} = \gamma e^{-t}, \quad u' = \gamma, \quad u = \gamma t$$

Ainsi $y_1 = (\beta + \gamma t)\,e^{-t}$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha e^{3t} \\ (\beta + \gamma t)\,e^{-t} \\ \gamma e^{-t} \end{pmatrix}$$

soit

$$\begin{cases} x = \alpha e^{3t} + (\gamma t + \beta - \gamma)\,e^{-t} \\ y = 2\,\alpha e^{3t} + (2\,\gamma t + 2\,\beta - \gamma)\,e^{-t} \\ z = 2\,\alpha e^{3t} + (\gamma t + \beta)\,e^{-t} \end{cases}$$

Remarque.

Une équation différentielle unique, linéaire d'ordre n, à coefficients constants, à une seule fonction inconnue, se ramène par le choix d'inconnues auxiliaires à un système de la forme précédente.

Dans

$$\frac{d^n x}{dt^n} + a_1 \frac{d^{n-1}x}{dt^{n-1}} + \dots + a_{n-1}\frac{dx}{dt} + a_n x = f(t) \qquad \text{(E)}$$

nous posons

$$\begin{cases} x = x_1 \\ \dfrac{dx}{dt} = x_2 \\ \dots\dots\dots\dots \\ \dfrac{d^{n-1}x}{dt^{n-1}} = x_n \end{cases}$$

et (E) s'écrit

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ f(t) \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{pmatrix}$$

On vérifie que le polynôme caractéristique de A est, au signe près,

$$\lambda^n + a_1\lambda^{n-1} + \ldots + a_n\lambda + a_n$$

que l'on retrouve en cherchant pour (E) des solutions $e^{\lambda t}$ et aussi en utilisant la transformation de Laplace.

9. POLYNOMES D'ENDOMORPHISMES. THÉORÈME DE CAYLEY-HAMILTON

a) En désignant par I à la fois la matrice-unité et l'endomorphisme-unité, par f un endomorphisme quelconque, dans un espace vectoriel E, de dimension n, sur $\mathbb{C}$, on peut, nous l'avons vu, associer à tout polynôme sur $\mathbb{C}$

$$Q(x) = a_0x^m + a_1x^{m-1} + \ldots + a_{m-1}x + a_m$$

le polynôme d'endomorphisme (ou endomorphisme-polynôme)

$$Q(f) = a_0f^m + a_1f^{m-1} + \ldots + a_{m-1}f + a_mI$$

dont la matrice dans une base de E, où f a pour matrice A, est

$$Q(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mI$$

Une réduction de f, diagonalisation ou trigonalisation, a pour effet la même réduction sur $Q(f)$, car les matrices diagonales et les matrices triangulaires constituent des sous-anneaux de l'anneau des matrices (n, n).

Par exemple $A = PDP^{-1}$ où D est diagonale

$$\Rightarrow Q(A) = PQ(D)P^{-1}$$

et $Q(D)$ est évidemment diagonale. Remarquons que ceci n'est pas vrai pour une réduction de Jordan, car en général le produit de deux matrices de Jordan n'est pas une matrice de Jordan. Notons aussi que deux endomorphismes-polynômes de f commutent.

b) Parmi les polynômes Q, on peut définir les polynômes annulateurs de f tels que $Q(f) = 0$. De tels polynômes existent forcément : puisque $\mathscr{L}(E)$ est un espace vectoriel de dimension n^2, les endomorphismes $I, f, f^2, \ldots, f^{n^2}$ sont nécessairement liés et il existe au moins un polynôme annulateur, de degré n^2.

c) Cherchons à caractériser un polynôme annulateur de f, de degré minimum m : il a m racines $\alpha_1, \alpha_2, \ldots, \alpha_m$ et

$$Q(x) = a_0 (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_m) \qquad a_0 \neq 0;$$

donc $$Q(f) = a_0 (f - \alpha_1 I)(f - \alpha_2 I) \ldots (f - \alpha_m I) = 0$$

chacun des endomorphismes $f - \alpha_j I$ est un diviseur de zéro, car $\prod_{i \neq j} (f - \alpha_i I) \neq 0$ (degré $\leqslant m - 1$) et les α_j sont valeurs propres de f.

Nous pouvons donc conclure :

Un polynôme annulateur de f, de degré minimum m, a pour zéros des valeurs propres de f.

On démontre, et nous admettrons, qu'il existe un seul polynôme (à un coefficient constant près) annulateur de f et de degré minimum : on l'appelle le polynôme minimum de f. Appelons encore Q ce polynôme, où nous pouvons prendre $a_0 = 1$. Voyons si toute valeur propre de f est un zéro de Q. Soit λ une telle valeur propre et $\vec{u}$ un vecteur propre de f associé à λ :

$$f(\vec{u}) = \lambda \vec{u} \Rightarrow f^k(\vec{u}) = \lambda^k \vec{u} \qquad \forall\, k \in \mathbb{N}$$

et $$\begin{aligned}(Q(f))(\vec{u}) &= (f^m + a_1 f^{m-1} + \ldots + a_{m-1} f + a_m I)(\vec{u}) \\ &= (\lambda^m + a_1 \lambda^{m-1} + \ldots + a_{m-1}\lambda + a_m)\,\vec{u} = Q(\lambda)\,.\,\vec{u}\end{aligned}$$

Comme $Q(f) = 0$, $Q(\lambda)\,.\,\vec{u} = \vec{0}$, or $\vec{u} \neq \vec{0}$, donc $Q(\lambda) = 0$.

Ainsi l'ensemble des zéros du polynôme minimum de l'endomorphisme f est l'ensemble des valeurs propres de f.

Il en résulte que son degré est au moins égal au nombre des valeurs propres distinctes de f. Notons qu'il est *a priori* au plus égal à n^2.

d) Envisageons alors le cas simple où f admet n valeurs propres distinctes, $\lambda_1, \ldots, \lambda_n$; son polynôme minimum a un degré $m \geqslant n$. Considérons le polynôme caractéristique

$$\varphi(f) = (-1)^n (f - \lambda_1 I)(f - \lambda_2 I) \ldots (f - \lambda_n I)$$

f est diagonalisable et, D étant sa matrice dans une base propre $(\vec{u_1}, \ldots, \vec{u_n})$

$$\varphi(D) = (-1)^n (D - \lambda_1 I)(D - \lambda_2 I) \ldots (D - \lambda_n I)$$

Comme

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \qquad D - \lambda_1 I = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

$(D - \lambda_1 I)(D - \lambda_2 I) =$

$$\begin{pmatrix} 0 & 0 & 0 \dots 0 \\ 0 & \lambda_2 - \lambda_1 & 0 \dots 0 \\ \vdots & \vdots & \ddots \; 0 \\ 0 & 0 & \dots \; \lambda_n - \lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1 - \lambda_2 & 0 & 0 \dots 0 \\ 0 & 0 & 0 \dots 0 \\ \vdots & & \ddots \; \vdots \\ 0 & \dots & \dots \; \lambda_n - \lambda_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \alpha_3 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \dots & \dots & \alpha_n \end{pmatrix}$$

De proche en proche, on montre que

$$(D - \lambda_1 I)(D - \lambda_2 I) \dots (D - \lambda_p I)$$

où $1 \leqslant p \leqslant n$, est une matrice diagonale où les p premiers termes de la diagonale sont nuls. Donc $\varphi(D) = 0$ et $\varphi(f) = 0$.

Le polynôme caractéristique de f est donc, dans ce cas, polynôme annulateur de f et, d'après (*c*), n'est autre, au coefficient $(-1)^n$ près, que le polynôme minimum.

e) Dans le cas général où f n'est pas forcément diagonalisable, mais seulement trigonalisable, le raisonnement effectué ci-dessus peut être en partie repris.

Si

$$T = \begin{pmatrix} \lambda_1 & \dots & \dots & \vdots \\ 0 & \lambda_2 & a_{ij} & \vdots \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \qquad T - \lambda_1 I \quad \text{et} \quad T - \lambda_2 I$$

sont deux matrices triangulaires où respectivement la première et la deuxième colonnes n'ont de termes non nuls qu'au-dessus de la diagonale, leur produit présente ce même caractère dans les deux premières colonnes, et le produit

$$\varphi(T) = (-1)^n (T - \lambda_1 I) \dots (T - \lambda_n I)$$

ne contient plus que des zéros. Ainsi, là encore, $\varphi(f) = 0$

Nous pouvons donc énoncer le **théorème de Cayley-Hamilton** :

Tout endomorphisme (ou toute matrice) est un zéro de son polynôme caractéristique.

Énoncé équivalent : le polynôme caractéristique de f est un polynôme annulateur de f. Ainsi, dans le cas général, k_i étant l'ordre de la valeur propre λ_i de f, avec $i = 1, 2, \dots, p$ $(1 \leqslant p \leqslant n)$ on a :

$$\varphi(f) = (-1)^n (f - \lambda_1 I)^{k_1} (f - \lambda_2 I)^{k_2} \dots (f - \lambda_p I)^{k_p}$$

polynôme caractéristique;

$$Q(f) = (f - \lambda_1 I)^{\alpha_1} (f - \lambda_2 I)^{\alpha_2} \dots (f - \lambda_p I)^{\alpha_p}$$

polynôme minimum,

avec $1 \leqslant \alpha_i \leqslant k_i$.

En d'autres termes, le polynôme minimum est un diviseur du polynôme caractéristique, possédant les mêmes zéros que ce polynôme. Les polynômes annulateurs de f sont les multiples du polynôme minimum, donc constituent un idéal de polynômes. Nous n'insisterons pas davantage sur le cas général.

Un exemple simple montre qu'en cas de valeurs propres multiples le polynôme minimum peut différer du polynôme caractéristique : nous avons vu (fin du paragraphe 4) qu'un projecteur f de E, de dimension n, sur un sous-espace de dimension p, a pour polynôme caractéristique $\varphi(\lambda) = (-1)^n \lambda^{n-p}(\lambda - 1)^p$ donc $\varphi(f) = 0$; mais on sait que $f^2 = f$, soit $f(f - I) = 0$.

Comme le polynôme minimum admet nécessairement 0 et 1 pour racines, ce polynôme minimum est $f(f - I)$.

— Enfin nous admettrons le résultat général suivant :

Pour qu'un endomorphisme soit diagonalisable, il faut et il suffit que son polynôme minimum n'ait que des racines simples.

Le cas du projecteur en est une illustration.

Résumé du chapitre 8

• On appelle **valeur propre** et **vecteur propre associé** d'un endomorphisme f d'un espace E, un scalaire λ et un vecteur $\vec{u}$ non nuls tels que

$$f(\vec{u}) = \lambda \vec{u}$$

L'ensemble des vecteurs propres associés à une valeur propre λ de f est un sous-espace de E, égal à $\ker(f - \lambda I)$ et appelé **sous-espace propre** de f associé à λ.

• A étant la matrice de f dans une base quelconque de E, les valeurs propres de f sont les racines de l'équation $\varphi(\lambda) = \text{dét}\,(A - \lambda I) = 0$ appelée équation caractéristique de f : $\varphi(\lambda)$ est le **polynôme caractéristique** de f.

• **Tout endomorphisme d'un espace vectoriel** E de dimension n **admet** n **valeurs propres**, réelles ou imaginaires, **distinctes ou confondues.** Leur somme est égale à la trace de la matrice A de f, leur produit est égal au déterminant de A.

• **Tout endomorphisme possédant** n **valeurs propres distinctes est diago-trigonalisable,** c'est-à-dire peut être représenté dans une base conve-la base correspondante est une base de vecteurs propres ou **base propre.**

• D'une façon générale, **pour qu'un endomorphisme d'un espace vectoriel de dimension finie soit diagonalisable, il faut et il suffit que chaque sous-espace propre de** f **ait pour dimension l'ordre de multiplicité de la valeur propre correspondante.**

• **Tout endomorphisme d'un espace vectoriel de dimension finie est trigonalisable,** c'est-à-dire peut être représenté dans une base convenable par une matrice triangulaire. Il peut aussi être représenté par une **matrice de Jordan**, matrice triangulaire supérieure où tous les termes non diagonaux sont nuls, sauf ceux qui bordent supérieurement la diagonale, lesquels peuvent être égaux à 1 ou 0.

Les éléments diagonaux d'une matrice triangulaire représentant f sont les valeurs propres de f.

• **La réduction d'une matrice,** en particulier la diagonalisation quand elle est possible, **permet le calcul des puissances successives d'un endomorphisme,** par la formule

$$A = PA'P^{-1} \Rightarrow A^k = PA'^kP^{-1}$$

• Un système de n équations différentielles linéaires du premier ordre à coefficients constants peut s'écrire sous forme matricielle

$$\frac{\mathrm{d}X}{\mathrm{d}t} = AX + \Phi \qquad (1)$$

ou, par un changement de base convenable,

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = BY + \Psi \qquad (2)$$

où B est une **matrice** diagonale, ou triangulaire, ou **de Jordan**.

(2) se résout sous forme d'équations successives comportant chacune une seule fonction inconnue, puis on a $X = PY$ P étant la **matrice de changement de base.**

• **Un polynôme annulateur d'un endomorphisme f est un polynôme Q tel que $Q(f) = 0$.** Toute valeur propre de f est un zéro de Q.

Le polynôme Q, annulateur de f, de plus bas degré, s'appelle **polynôme minimum de Q.**

Le polynôme caractéristique est un polynôme annulateur de f, en d'autres termes, **tout endomorphisme est un zéro de son polynôme caractéristique (théorème de Cayley-Hamilton).** Le polynôme minimum est un diviseur du polynôme caractéristique.

• **Pour qu'un endomorphisme soit diagonalisable, il faut et il suffit que son polynôme minimum n'admette que des racines simples.**

EXERCICES

■ **1.** Diagonaliser les matrices

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix} \qquad B = \begin{pmatrix} i & -1 & -i & 1 \\ -1 & 1 & -1 & 1 \\ -i & -1 & i & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

---

a) On calcule aisément $\text{dét}\,(A - \lambda I) = -\lambda(\lambda - 1)(\lambda - 4)$.

Les valeurs propres sont $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 4$: comme elles sont distinctes, A est *a priori* diagonalisable. Déterminons les sous-espaces propres :

$$E_1 = \ker A \quad \text{est défini par } x + y + 2z = 0 \qquad (1)$$
$$2x + y + z = 0 \qquad (2)$$
$$y + 3z = 0 \qquad (3)$$

(3) résulte de l'élimination de x entre (1) et (2), on peut ne conserver que (2) et (3) qui donnent $x = z$, $y = -3z$, d'où la droite vectorielle de base $\vec{u_1}\,(1, -3, 1)$.

De même $E_2 = \ker(A - I)$ est défini par $y + 2z = 0$

$$2x + z = 0$$
$$y + 2z = 0$$

d'où la droite vectorielle de base $\vec{u_2}\,(1, 4, -2)$.

Enfin $E_3 = \ker(A - 4I)$ est défini par
$$\begin{aligned} x + y - 2z &= 0 \\ 2x + y - 3z &= 0 \\ y - z &= 0 \end{aligned}$$
d'où la droite vectorielle de base $\vec{u_3}\,(1, 1, 1)$.

Une matrice diagonale semblable à A est
$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$
avec matrice de passage
$$P = \begin{pmatrix} 1 & 1 & 1 \\ -3 & 4 & 1 \\ 1 & -2 & 1 \end{pmatrix}$$
On vérifie $D = P^{-1}AP$.

b) $\varphi(\lambda) = \text{dét}\,(B - \lambda I)$
$$= \begin{vmatrix} i-\lambda & -1 & -i & 1 \\ -1 & 1-\lambda & -1 & 1 \\ -i & -1 & i-\lambda & 1 \\ 1 & 1 & 1 & 1-\lambda \end{vmatrix}$$
$$= \begin{vmatrix} i-\lambda & -1 & -i & 1 \\ 0 & 2-\lambda & 0 & 2-\lambda \\ \lambda - 2i & 0 & 2i-\lambda & 0 \\ 1 & 1 & 1 & 1-\lambda \end{vmatrix}$$
$$= (2-\lambda)(\lambda - 2i)\begin{vmatrix} i-\lambda & -1 & -i & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 1-\lambda \end{vmatrix}$$
$$= (2-\lambda)(\lambda - 2i)\begin{vmatrix} i-\lambda & -1 & -i & 2 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & -\lambda \end{vmatrix}$$
d'où en développant suivant la seconde ligne :
$$\varphi(\lambda) = (2-\lambda)(\lambda - 2i)(4 - \lambda^2) = (\lambda - 2)^2(\lambda + 2)(\lambda - 2i)$$
Il y a une valeur propre double $\lambda_1 = 2$, la diagonalisation n'est donc pas assurée : il faut voir si le sous-espace propre associé à λ_1 est de dimension 1 ou 2. Ce sous-espace est défini par le système :
$$\begin{cases} (i-2)x - y - iz + t = 0 & (1) \\ -x - y - z + t = 0 & (2) \\ -ix - y + (i-2)z + t = 0 & (3) \\ x + y + z - t = 0 & (4) \end{cases}$$

(2) et (4) sont identiques, (3) peut s'obtenir en multipliant (2) par 2 et en lui retranchant (1), ce système est donc de rang 2 et définit un sous-espace de dimension 2, ordre de la valeur propre λ_1, il en résulte que B est diagonalisable : en écrivant la matrice diagonale semblable sous la forme

$$B' = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 2i \end{pmatrix}$$

on prendra une matrice de passage $(\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4})$ dans laquelle $\vec{u_1}$, $\vec{u_2}$ sont des vecteurs du sous-espace E_1 défini par les équations (1) et (2) ci-dessus, par exemple $\vec{u_1}\,(0, 1, 0, 1)$ et $\vec{u_2}\,(1, 0, 1, 2)$; $\vec{u_3}$ et $\vec{u_4}$ sont des vecteurs propres associés aux valeurs propres simples $\lambda_3 = -2$, $\lambda_4 = 2i$.

On obtient facilement $\vec{u_3}\,(1, 1, 1, -1)$ et $\vec{u_4}\,(1, 0, -1, 0)$.

On a donc $B' = P^{-1}BP$ avec

$$P = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 \\ 1 & 2 & -1 & 0 \end{pmatrix}$$

■ **2.** Réduire à leurs expressions les plus simples les matrices

$$A = \begin{pmatrix} 0 & a & a^2 \\ \dfrac{1}{a} & 0 & a \\ \dfrac{1}{a^2} & \dfrac{1}{a} & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 0 & 0 & -1 \\ 1 & -1 & -1 \\ 0 & 1 & -2 \end{pmatrix}$$

Déterminer le polynôme minimum de B.

Pour A, on trouve le polynôme caractéristique

$$\varphi(\lambda) = -(\lambda + 1)^2 (\lambda - 2)$$

Le sous-espace propre associé à $\lambda_1 = -1$ est le plan d'équation

$$x + ay + a^2 z = 0$$

celui associé à $\lambda_2 = 2$ est la droite $x = ay = a^2 z$; A est diagonalisable, sous la forme

$$A' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

avec matrice de passage

$$P = \begin{pmatrix} a & 0 & a^2 \\ -1 & a & a \\ 0 & -1 & 1 \end{pmatrix}$$

Pour B, le polynôme caractéristique est $-(\lambda + 1)^3$. La seule valeur propre étant -1 et B étant différente de $-I$, on peut affirmer *a priori* que B n'est pas diagonalisable. Son unique sous-espace est défini par le système

$$\begin{cases} x - z = 0 \\ x - z = 0 \\ y - z = 0 \end{cases} \qquad \text{d'où} \quad x = y = z$$

Ce sous-espace a pour dimension 1, on dispose pour une base de E d'un seul vecteur propre $\vec{u_1}$ (1, 1, 1), et la réduction la plus achevée pour B sera la matrice de Jordan

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

Cherchons $\vec{u_2}$ et $\vec{u_3}$, complétant $\vec{u_1}$ pour constituer une base de E, dans laquelle l'endomorphisme f considéré ait pour matrice J. Il faut

$$f(\vec{u_2}) = \vec{u_1} - \vec{u_2}, \quad f(\vec{u_3}) = \vec{u_2} - \vec{u_3}$$

donc $$(f + I)(\vec{u_2}) = \vec{u_1}, \quad (f + I)(\vec{u_3}) = \vec{u_2}$$

et $\vec{u_1}$, $\vec{u_2}$ doivent appartenir à $\operatorname{Im}(f + I)$ qui est l'ensemble des vecteurs

$$(B + I)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - z \\ x - z \\ y - z \end{pmatrix}$$

c'est-à-dire le plan d'équation $X - Y = 0$. $\vec{u_1}$ appartient à ce plan. Prenons $\vec{u_2}$ dans ce plan, soit $\vec{u_2}$ (a, a, b), ces coordonnées étant prises dans l'ancienne base.

La condition $f(\vec{u_2}) = \vec{u_1} - \vec{u_2}$ conduit à la relation $a - b = 1$. Donc $\vec{u_2} = \begin{pmatrix} a \\ a \\ a - 1 \end{pmatrix}$. Si $\vec{u_3}$ a pour coordonnées α, β, γ, la condition $f(\vec{u_3}) = \vec{u_2} - \vec{u_3}$ conduit à $\alpha - \gamma = a$, $\beta - \gamma = a - 1$.

On voit que a et γ peuvent être choisis arbitrairement, par exemple $a = 1$ et $\gamma = 0$, alors $\alpha = 1$, $\beta = 0$. Une base de Jordan est donc définie par la matrice de passage

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

Le lecteur pourra vérifier que l'on a bien $J = P^{-1}AP$.

D'après le cours, le polynôme minimum de B (ou de f) est un diviseur du polynôme caractéristique $(\lambda + 1)^3$. Il nous faut déterminer la plus petite valeur de l'entier k tel que $(f + I)^k = 0$ ou $(J + I)^k = 0$.

$$J + I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

(matrice « nilpotente » du type de Jordan).

On a évidemment $J + I \neq 0$ mais aussi

$$(J + I)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq 0$$

On a donc $k = 3$.

Dans cet exemple le polynôme minimum n'est autre que le polynôme caractéristique.

■ **3.** Une matrice A d'ordre n ayant pour valeurs propres $\lambda_1, \lambda_2, \ldots, \lambda_n$ (distinctes ou non), quelles sont les valeurs propres de A^2 ? Calculer en fonction des éléments de A la somme $S = \sum_1^n \lambda_i^2$ lorsque A est une matrice symétrique.

L'endomorphisme f, de matrice A, est trigonalisable sur $\mathbb{C}$, et sa forme triangulaire est

$$T = \begin{pmatrix} \lambda_1 & c_{12} & \cdots & \cdots & c_{1n} \\ 0 & \lambda_2 & & & c_{2n} \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

dans une base convenable.

f^2, dans cette base, a pour matrice T^2 qui est également triangulaire supérieure. L'élément diagonal de rang i dans T^2 est

$$\alpha_{ii} = (0\,0\ \ldots\ 0\ \lambda_i\ c_{i,i+1}\ \ldots\ c_{in}) \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ \lambda_i \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_i^2$$

Donc

$$T^2 = \begin{pmatrix} \lambda_1^2 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \lambda_2^2 & \cdots & \alpha_{2n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n^2 \end{pmatrix}$$

ses valeurs propres sont $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$. Ce sont aussi celles de A^2.

La somme S est la trace de A^2, c'est-à-dire la somme de ses termes diagonaux

Si $A = \begin{pmatrix} & \vdots & \\ \ldots & a_{ij} & \ldots \\ & \vdots & \end{pmatrix}$ un terme diagonal de A^2 s'écrit

$$b_{ii} = (a_{i1}\ a_{i2}\ \ldots\ a_{in}) \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} = \sum_{j=1}^{j=n} a_{ij}\, a_{ji}$$

Lorsque A est symétrique, $a_{ji} = a_{ij}$, $b_{ii} = \sum_{j=1}^{j=n} a_{ij}^2$ et

$$S = \operatorname{tr}(A^2) = \sum_{i,j} a_{ij}^2$$

somme des carrés de tous les éléments de A.

Le lecteur pourra vérifier numériquement cette relation sur la matrice B de l'exercice 1.

■ **4.** Soit la matrice

$$M(x, y, z, t) = \begin{pmatrix} x & z & y & t \\ z & x & t & y \\ t & y & x & z \\ y & t & z & x \end{pmatrix}$$

où x, y, z et t sont des réels.

Soit $A = M(0, 1, 0, 0)$ et I la matrice-unité d'ordre 4.

a) Montrer que l'ensemble des matrices M est un espace vectoriel dont on précisera une base à l'aide de I et A.

b) Vérifier que A est diagonalisable; en déduire les valeurs propres et le déterminant de M.

---

a) $$M = xI + yA + zB + tC$$

avec

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Comme

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

on vérifie aisément que $A^2 = B$ et que $A^3 = C$.

On a donc $M = xI + yA + zA^2 + tA^3$

L'ensemble des matrices M est un espace vectoriel engendré par I, A, A^2, A^3 qui sont indépendants puisque $xI + yA + zA^2 + tA^3 = 0$ exige $x = y = z = t = 0$ c'est donc un espace vectoriel de dimension 4 et de base (I, A, A^2, A^3). Notons que d'après ce qui précède le polynôme minimum de A est de degré au moins égal à 4, donc en fait égal à 4 et confondu avec le polynôme caractéristique.

b) Le calcul des valeurs propres de A revient à celui du polynôme caractéristique, c'est-à-dire ici du polynôme minimum; on vérifie immédiatement que $A^3 . A = A^4 = I$ ce polynôme est donc $\lambda^4 - 1$ et les valeurs propres sont 1, -1, i et $-i$. On aurait pu aussi calculer

$$\text{dét}\,(A - \lambda I) = \lambda^4 - 1$$

A possédant quatre valeurs propres distinctes, elle est diagonalisable et il existe une matrice régulière P telle que $A = PDP^{-1}$ avec

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

Il n'est pas indispensable ici de calculer P. Nous aurons

$$M = P\,(xI + yD + zD^2 + tD^3)\,P^{-1} = P\Delta P^{-1}$$

Δ étant une matrice diagonale puisque combinaisons linéaires de telles matrices.

Les éléments diagonaux de Δ sont $\alpha_k = x + y\lambda_k + z\lambda_k^2 + t\lambda_k^3$, $k = 1, 2, 3, 4$; les λ_k étant les valeurs propres de A. Les α_k sont les valeurs propres de M, soit

$$\begin{aligned} \alpha_1 &= x + y + z + t && (\text{car } \lambda_1 = 1) \\ \alpha_2 &= x - y + z - t && (\text{car } \lambda_2 = -1) \\ \alpha_3 &= x + iy - z - it && (\text{car } \lambda_3 = i) \\ \alpha_4 &= x - iy - z + it && (\text{car } \lambda_4 = -i) \end{aligned}$$

Leur produit est le déterminant de M, soit sous forme réelle

$$\text{dét}\,M = (x + y + z + t)\,(x - y + z - t)\,[(x - z)^2 + (y - t)^2]$$

■ 5. Soit la matrice d'ordre n, où a et b sont des réels :

$$M = \begin{pmatrix} 0 & a & a & \dots & \dots & a \\ b & 0 & a & \dots & \dots & a \\ b & b & 0 & \ddots & \dots & a \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & a \\ b & b & b & \dots & b & 0 \end{pmatrix}$$

Montrer que son polynôme caractéristique peut s'écrire, si $a \neq b$,

$$\varphi_n(\lambda) = (-1)^n \frac{a(\lambda + b)^n - b(\lambda + a)^n}{a - b} \tag{1}$$

Quelle est sa valeur lorsque $a = b$?

Démontrer que M est diagonalisable sur $\mathbb{C}$.
On pourra, pour établir (1), utiliser la dérivation d'un déterminant (exercice 14, chap. 7).

$a)$ La relation (1) se vérifie immédiatement pour $n = 2$. Raisonnons par récurrence et supposons que

$$\varphi_{n-1}(\lambda) = (-1)^{n-1} \frac{a(\lambda + b)^{n-1} - b(\lambda + a)^{n-1}}{a - b}$$

Calculons $\varphi_n'(\lambda)$, somme des déterminants obtenus en remplaçant successivement chaque ligne par sa dérivée. On obtient

$$\varphi_n'(\lambda) = -n\,\varphi_{n-1}(\lambda) = (-1)^n . \frac{an(\lambda + b)^{n-1} - bn(\lambda + a)^{n-1}}{a - b}$$

Par intégration, on obtient

$$\varphi_n(\lambda) = (-1)^n \frac{a(\lambda + b)^n - b(\lambda + a)^n}{a - b} + C \tag{2}$$

C étant une constante indépendante de λ.

Il reste à s'assurer que $C = 0$. Pour cela donnons à λ la valeur $-a$. On a, en retranchant, dans ce déterminant, de chaque ligne la précédente en partant du bas,

$$\varphi_n(-a) = (-1)^{n+1} a(b - a)^{n-1}$$

ce qui, portée dans (2), donne bien $C = 0$.

La relation (1) est donc démontrée.

$b)$ Cette forme n'est plus valable lorsque $b = a$; cependant il est manifeste que $\varphi_n(\lambda)$ qui est un polynôme en a, b, λ, est continue par rapport à chacune de ces variables, en particulier par rapport à b lorsque l'on fixe λ et a.

L'application de la règle de l'Hospital quand b tend vers a donne lorsque $b = a$

$$\varphi_n(\lambda) = (-1)^n \cdot \lim_{b \to a} \frac{an(\lambda+b)^{n-1} - (\lambda+a)^n}{-1}$$

soit $$\varphi_n(\lambda) = (-1)^n (\lambda+a)^{n-1}[\lambda - (n-1)a]$$

Ce résultat s'obtient aisément à partir du déterminant

$$\begin{vmatrix} -\lambda & a & \dots & a \\ a & -\lambda & \dots & a \\ a & a & \dots & a \\ \vdots & \vdots & & \\ a & a & \dots & -\lambda \end{vmatrix}$$

c) Pour $b \neq a$, les valeurs propres de M s'obtiennent en écrivant

$$a(\lambda+b)^n - b(\lambda+a)^n = 0$$

$\lambda = -b$ n'étant pas solution, ceci équivaut à $\left(\dfrac{\lambda+a}{\lambda+b}\right)^n = \dfrac{a}{b}$.

Comme $\dfrac{a}{b}$ admet n racines $n^{\text{ième}}$ distinctes, et comme l'applicatio $\lambda \to \dfrac{\lambda+a}{\lambda+b}$ est bijective, M admet n valeurs propres distinctes et est diagc nalisable.

d) Pour $b = a$, on pourrait établir que M est encore diagonalisabl en raisonnant par un passage à la limite analogue à celui effectué plus hau Mais il est immédiat de constater que le sous-espace de M relatif à la valeı propre d'ordre $n-1$, $\lambda = -a$, est l'hyperplan d'équation

$$x_1 + x_2 + \dots + x_n = 0$$

donc est de dimension $n-1$, ce qui justifie la possibilité de diagonaliser M

■ 6. On considère, dans un espace vectoriel E de dimension n, un endomo phisme f diagonalisable possédant seulement deux valeurs propres di tinctes λ_1 et λ_2, de sous-espaces associés E_1 et E_2. On appelle φ_1 le proje teur de E sur E_1, parallèlement à E_2; φ_2 le projecteur de E sur E_2, parall lement à E_1.

Évaluer $\varphi_1 + \varphi_2$, $\varphi_1 \circ \varphi_2$, $\varphi_2 \circ \varphi_1$, $\lambda_1\varphi_1 + \lambda_2\varphi_2$.

Quel est le polynôme minimum de f ?

Réponse : f étant diagonalisable, $E = E_1 \oplus E_2$ ce qui justil l'existence des projecteurs φ_1 et φ_2. On obtient facilement

$$\varphi_1 + \varphi_2 = I,$$
$$\varphi_1 \circ \varphi_2 = \varphi_2 \circ \varphi_1 = 0$$
$$\lambda_1\varphi_1 + \lambda_2\varphi_2 = f$$

On établit

$$\varphi_1 = \frac{f - \lambda_2 I}{\lambda_1 - \lambda_2}, \qquad \varphi_2 = \frac{f - \lambda_1 I}{\lambda_2 - \lambda_1}$$

d'où $(f - \lambda_1 I)(f - \lambda_2 I) = 0$, le polynôme minimum est

$$Q(x) = (x - \lambda_1)(x - \lambda_2)$$

■ **7.** Soit M et A deux matrices carrées d'ordre n, à termes complexes, telles que $MA = AM$; on suppose que M a toutes ses valeurs propres distinctes.

a) Montrer que tout vecteur propre de M est vecteur propre de A.

b) Montrer qu'il existe des complexes α_k tels que $A = \sum_{k=0}^{n-1} \alpha_k M^k$.

---

a) Soit, dans la base canonique utilisée, X la matrice-colonne d'un vecteur propre de M : on a $MX = \lambda X$, λ étant la valeur propre associée. Donc $AMX = A \, . \, \lambda X$ et $M \, . \, AX = \lambda \, . \, AX$.

AX est donc vecteur propre de M associé à λ. Or, les valeurs propres de M étant toutes distinctes, chaque sous-espace propre de M est une droite vectorielle, AX et X font partie d'une telle droite et il existe un scalaire μ tel que $AX = \mu X$. X est donc un vecteur propre de A, la valeur propre associée étant μ : tout vecteur propre de M est vecteur propre de A.

b) M admet n vecteurs propres indépendants, qui sont aussi vecteurs propres de A; M et A (qui peut avoir des valeurs propres multiples) sont donc diagonalisables, et il suffit d'établir la relation demandée sur les matrices diagonales M' et A' semblables à M et A, avec même matrice de passage P. Comme $M' = \text{diag}\,(\lambda_1, \ldots, \lambda_n)$ et $A' = \text{diag}\,(\mu_1, \ldots, \mu_n)$, il suffit de montrer que l'on peut résoudre le système, d'inconnues α_0, $\alpha_1, \ldots, \alpha_{n-1}$:

$$(S) \begin{cases} \mu_1 = \alpha_0 + \alpha_1\lambda_1 + \alpha_2\lambda_1^2 + \ldots + \alpha_{n-1}\lambda_1^{n-1} \\ \mu_2 = \alpha_0 + \alpha_1\lambda_2 + \alpha_2\lambda_2^2 + \ldots + \alpha_{n-1}\lambda_2^{n-1} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \mu_n = \alpha_0 + \alpha_1\lambda_n + \alpha_2\lambda_n^2 + \ldots + \alpha_{n-1}\lambda_n^{n-1} \end{cases}$$

Or les λ_j étant toutes distinctes, le déterminant de ce système est le déterminant de Van der Monde, non nul, $\prod_{1 \leq i < j \leq n} (\lambda_j - \lambda_i)$ et (S) est un système de Cramer. Il existe bien n complexes α_k vérifiant la relation

$$A = \sum_{k=0}^{n-1} \alpha_k M^k$$

Ainsi tout endomorphisme commutant avec un endomorphisme φ à valeurs propres distinctes est un polynôme $P(\varphi)$, de degré inférieur ou égal à $n - 1$.

■ **8.** On considère les matrices $A = \begin{pmatrix} & \vdots & \\ \dots & a_{ij} & \dots \\ & \vdots & \end{pmatrix}$ d'ordre n telles que, dans chaque rangée, la somme des termes soit égale à 1 :

$$\sum_{i=1}^{i=n} a_{ij} = \sum_{j=1}^{j=n} a_{ij} = 1$$

a) Montrer que l'on peut obtenir une telle matrice en complétant une matrice donnée quelconque, d'ordre $n - 1$.

b) Montrer que, sur l'ensemble E de ces matrices, la multiplication est une opération interne.

c) Montrer que 1 est une valeur propre de A, et que le sous-espace propre associé contient une droite vectorielle indépendante des a_{ij}.

d) Établir que si, en outre, les a_{ij} sont positifs, toutes les valeurs propres de A ont des modules au plus égaux à 1.

---

a) Partant de

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1,n-1} \\ b_{21} & b_{22} & \dots & b_{2,n-1} \\ \dots & \dots & \dots & \dots \\ b_{n-1,1} & b_{n-1,2} & \dots & b_{n-1,n-1} \end{pmatrix}$$

on peut border la dernière ligne et la dernière colonne par des termes

$$a_{nj} = 1 - \sum_{i=1}^{i=n-1} b_{ij}, \qquad j = 1, 2, \dots, n-1$$

et

$$a_{in} = 1 - \sum_{j=1}^{j=n-1} b_{ij}, \qquad i = 1, 2, \dots, n-1$$

Le dernier terme a_{nn} devra être égal à la fois à

$$1 - \sum_{i=1}^{i=n-1} a_{in} \quad \text{et} \quad 1 - \sum_{j=1}^{j=n-1} a_{nj}$$

or

$$\sum_{i=1}^{i=n-1} a_{in} = n - 1 - \sum_{i=1}^{i=n-1} \left(\sum_{j=1}^{j=n-1} b_{ij} \right) = n - 1 - S$$

S étant la somme de tous les termes de B, et on a de même

$$\sum_{j=1}^{j=n-1} a_{nj} = n - 1 - S$$

Les deux conditions imposées à d_{nn} sont donc compatibles, et

$$a_{nn} = 2 - n + S$$

On obtient ainsi une matrice A du type demandé.

b) A et B étant deux matrices éléments de E, formons dans la matrice

$C = AB = \left(\dots \begin{matrix} \vdots \\ c_{ij} \\ \vdots \end{matrix} \dots \right)$ la somme $s_i = \sum_{j=1}^{j=n} c_{ij}$.

Comme $c_{ij} = \sum_{k=1}^{k=n} a_{ik}\, b_{kj}$ on a

$$\begin{aligned} s_i = c_{i1} + c_{i2} + \dots + c_{in} = {} & a_{i1}b_{11} + a_{i2}b_{21} + \dots + a_{in}b_{n1} \\ & + a_{i1}b_{12} + a_{i2}b_{22} + \dots + a_{in}b_{n2} \\ & + \dots\dots\dots\dots\dots\dots \\ & + a_{i1}b_{1n} + a_{i2}b_{2n} + \dots + a_{in}b_{nn} \end{aligned}$$

$$= a_{i1} \cdot \sum_1^n b_{1j} + a_{i2} \sum_1^n b_{2j} + \dots + a_{in} \sum_1^n b_{nj} = a_{i1} + a_{i2} + \dots + a_{in} = 1$$

On obtiendrait de même $\sum_{i=1}^{i=n} c_{ij} = 1$ donc C est élément de E.

c) A étant élément de E,

$$\text{dét}\,(A - I) = \begin{vmatrix} a_{11} - 1 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - 1 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - 1 \end{vmatrix} = 0$$

car on obtient une ligne de zéros en ajoutant à la première ligne toutes les autres, 1 est donc valeur propre.

Le système

$$\begin{cases} (a_{11} - 1)\, x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + (a_{nn} - 1)\, x_n = 0 \end{cases}$$

admet la solution évidente $x_1 = x_2 = \dots = x_n$ qui représente une droite vectorielle faisant partie du sous-espace propre associé à $\lambda = 1$.

d) Soit λ une valeur propre, et $(x_1, x_2, \dots, x_n)$ un vecteur propre associé à λ, de la matrice A où $a_{ij} \geqslant 0$.

On a

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = \lambda x_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = \lambda x_n \end{cases}$$

En écrivant que dans chacune de ces relations le module d'une somme est inférieur ou égal à la somme des modules, on a :

$$|\lambda|\,|x_i| \leqslant \sum_{j=1}^{n} |a_{ij}x_j| = \sum_{j=1}^{n} a_{ij}\,|x_j| \qquad i = 1, 2, \dots, n$$

En ajoutant ces relations, on obtient, en posant $x = \sum_1^n |x_i|$

$$|\lambda| \sum_1^n |x_i| \leqslant \sum_{j=1}^{j=n} \left(\sum_{i=1}^{i=n} a_{ij} \right) |x_j| = \sum_{j=1}^{j=n} |x_j|$$

Donc $|\lambda| x \leqslant x$ or x est strictement positif puisqu'un vecteur propre possède au moins une coordonnée non nulle. On a donc $|\lambda| \leqslant 1$.

■ **9.** Soit la matrice

$$K = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

où les seuls termes non nuls sont

$$a_{1n} = 1 \qquad \text{et} \qquad a_{12} = a_{23} = \dots = a_{n-1,n} = 1$$

a) Calculer $K^2, K^3, \dots, K^{n-1}, K^n$.

b) Quelles sont les valeurs propres de K ?

c) En déduire les valeurs propres de la « matrice circulante »

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \dots\dots & a_n \\ a_n & a_1 & a_2 & \dots\dots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \dots\dots & a_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ a_2 & a_3 & a_4 & \dots\ a_n & a_1 \end{pmatrix}$$

et le déterminant de cette matrice.

d) Retrouver, par la formule obtenue, la valeur du déterminant

$$\Delta = \begin{vmatrix} 1 & 2 & 3 & \dots & n \\ n & 1 & 2 & \dots & n-1 \\ n-1 & n & 1 & \dots & n-2 \\ \dots & \dots & \dots & \dots & \dots \\ 2 & 3 & 4 & \dots & 1 \end{vmatrix}$$

(voir exercice 11, chap. 6)

Réponse :

a) On obtient

$$K^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & & & 0 \\ \vdots & \vdots & \vdots & & & 1 \\ 1 & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \end{pmatrix}$$

et ainsi de suite, les rangées obliques de 1 « remontant » de proche en proche. On obtient aussi $K^n = I$.

b) $K^n - I = 0$ montre que le polynôme $\lambda^n - 1$ est à la fois le polynôme minimum et le polynôme caractéristique de K. En posant $\alpha = e^{\frac{2i\pi}{n}}$, les valeurs propres de K sont $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ (racines $n^{\text{ièmes}}$ de l'unité).

c) K est diagonalisable, A aussi (voir exercice 4), et ses valeurs propres sont

$$\begin{aligned}
\lambda_1 &= a_1 + a_2 + \ldots + a_n \\
\lambda_2 &= a_1 + a_2\alpha + \ldots + a_n\alpha^{n-1} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\lambda_j &= a_1 + a_2\alpha^{j-1} + a_3\alpha^{2(j-1)} + \ldots + a_n\alpha^{(n-1)(j-1)} \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\lambda_n &= a_1 + a_2\alpha^{n-1} + a_3\alpha^{2(n-1)} + \ldots + a_n\alpha^{(n-1)^2}
\end{aligned}$$

Le déterminant de A est le produit de ces valeurs propres.

d) On en déduit

$$\begin{aligned}
\Delta &= (1 + 2 + 3 + \ldots + n)(1 + 2\alpha + 3\alpha^2 + \ldots + n\alpha^{n-1}) \ldots \\
&\qquad (1 + 2\alpha^{n-1} + 3\alpha^{2(n-1)} + \ldots + n\alpha^{(n-1)^2}) \\
&= \frac{n(n+1)}{2} \prod_{j=1}^{j=n-1} (1 + 2x_j + 3x_j^2 + \ldots + nx_1^{n-j})
\end{aligned}$$

où les x_j sont les racines $n^{\text{ièmes}}$ de l'unité, autres que 1, c'est-à-dire les racines de l'équation

$$x^{n-1} + x^{n-2} + \ldots + x + 1 = 0 \tag{1}$$

Or

$$\begin{aligned}
1 + 2x + 3x^2 + \ldots + nx^{n-1} &= \frac{\mathrm{d}}{\mathrm{d}x}(1 + x + x^2 + \ldots + x^n) \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1 - x^{n+1}}{1 - x}\right) = \frac{nx^{n+1} - (n+1)x^n + 1}{(x-1)^2}
\end{aligned}$$

Comme $x_j^n = 1$, $1 + 2x_j + \ldots + nx_j^{n-1} = \dfrac{n}{x_j - 1}$

et
$$\Delta = \frac{n^n(n+1)}{2} \cdot \frac{1}{\prod_1^{n-1}(x_j - 1)}$$

En posant $x_j - 1 = y_j$, les y_j sont les racines de l'équation déduite de (1) par la translation $x = 1 + y$ soit

$$(y+1)^{n-1} + (y+1)^{n-2} + \ldots + (y+1) + 1 = 0$$

soit
$$y^{n-1} + c_1y^{n-2} + \ldots + c_{n-2}y + n = 0$$

donc $\prod_1^{n-1} y_j = (-1)^{n-1}n$ et $\Delta = (-1)^{n-1}\dfrac{n^{n-1}(n+1)}{2}$

■ **10.** Résoudre le système différentiel

$$(S)\quad \begin{cases} \dfrac{dx}{dt} = 4x - 3y + 9z + 1 \\ \dfrac{dy}{dt} = -3x + 4y - 9z + t \\ \dfrac{dz}{dt} = -3x + 3y - 8z + t \end{cases}$$

(S) s'écrit $\dfrac{dX}{dt} = AX + \Phi$ avec $A = \begin{pmatrix} 4 & -3 & 9 \\ -3 & 4 & -9 \\ -3 & 3 & -8 \end{pmatrix}$ et $\Phi(t) = \begin{pmatrix} 1 \\ t \\ t \end{pmatrix}$.

On trouve

$$\text{dét}\,(A - \lambda I) = -(\lambda - 1)^2(\lambda + 2)$$

Les vecteurs propres associés à $\lambda = 1$ satisfont à

$$\begin{cases} 3x - 3y + 9z = 0 \\ -3x + 3y - 9z = 0 \\ -3x + 3y - 9z = 0 \end{cases}$$

soit l'unique équation $x - y + 3z = 0$, d'où deux vecteurs propres $\vec{u_1}\,(1, 1, 0)$, $\vec{u_2}\,(0, 3, 1)$.

Pour $\lambda = -2$ on trouve $\vec{u_3}\,(1, -1, -1)$.

A est diagonalisable par la matrice de passage

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 3 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

et le système « diagonalisé » s'écrit : avec $X = PX_1$

$$(S_1)\quad \frac{dX_1}{dt} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} X_1 + P^{-1}\begin{pmatrix} 1 \\ t \\ t \end{pmatrix} \qquad X_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

On calcule facilement

$$P^{-1} = \frac{1}{\text{dét}\,P}\cdot {}^t(\text{com}\,P) = -\begin{pmatrix} -2 & 1 & -3 \\ 1 & -1 & 2 \\ 1 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 3 \\ -1 & 1 & -2 \\ -1 & 1 & -3 \end{pmatrix}$$

$$P^{-1}\begin{pmatrix} 1 \\ t \\ t \end{pmatrix} = \begin{pmatrix} 2(1+t) \\ -(1+t) \\ -(1+2t) \end{pmatrix}$$

(S_1) se décompose donc en $\begin{cases} \dfrac{dx_1}{dt} = x_1 + 2(1+t) \\ \dfrac{dy_1}{dt} = y_1 - (1+t) \\ \dfrac{dz_1}{dt} = -2z_1 - (1+2t) \end{cases}$

La méthode générale donne $\begin{cases} x_1 = \alpha e^t + x_0(t) \\ y_1 = \beta e^t + y_0(t) \\ z_1 = \gamma e^{-2t} + z_0(t) \end{cases}$

où $x_0(t)$, $y_0(t)$, $z_0(t)$ sont des polynômes du premier degré calculables par identification. On obtient

$$\begin{cases} x_1 = \alpha e^t - 2t - 4 \\ y_1 = \beta e^t + t + 2 \\ z_1 = \gamma e^{-2t} - t \end{cases}$$

Enfin, on revient à x, y, z, donc aux solutions de (S), par la relation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$

qui nous donne $\begin{cases} x = \alpha e^t + \gamma e^{2t} - 3t - 4 \\ y = (\alpha + 3\beta) e^t - \gamma e^{-2t} + 2t + 2 \\ z = \beta e^t - \gamma e^{-2t} + 2t + 2 \end{cases}$

où α, β, γ sont les trois constantes d'intégration.

Remarque : on aurait pu éviter le calcul de P^{-1} en résolvant d'abord le système « sans second membre » et en ajoutant aux valeurs trouvées trois polynômes du premier degré $x = at + b$, $y = ct + d$, $z = et + f$, calculés par identification dans (S).

■ **11.** Résoudre le système différentiel

$$(S) \begin{cases} \dfrac{dx}{dt} = -z \\ \dfrac{dy}{dt} = x - y - z \\ \dfrac{dz}{dt} = y - 2z \end{cases}$$

(S) est homogène ou « sans second membre ». Sa matrice associée est la matrice B de l'exercice 2, elle est réductible à la matrice de Jordan.

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

par la matrice de passage

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

Le système réduit s'écrit

$$\left\{\begin{aligned} \frac{dx_1}{dt} &= -x_1 + y_1 && (1) \\ \frac{dy_1}{dt} &= -y_1 + z_1 && (2) \\ \frac{dz_1}{dt} &= -z_1 && (3) \end{aligned}\right.$$

$$(3) \Rightarrow z_1 = \gamma e^{-t}; \qquad (2) \text{ s'écrit } \frac{dy_1}{dt} + y_1 = \gamma e^{-t}$$

$y_0 = ue^{-t}$ conduit à $u' = \gamma$, $u = \gamma t$, donc $y_1 = (\gamma t + \beta)\, e^{-t}$.

(1) s'écrit $$\frac{dx_1}{dt} + x_1 = (\gamma t + \beta)\, e^{-t}$$

$x_0 = ve^{-t}$ conduit à

$$v' = \gamma t + \beta, \quad v = \frac{\gamma}{2} t^2 + \beta t, \quad x_1 = \left(\frac{\gamma}{2} t^2 + \beta t + \alpha\right) e^{-t}$$

Enfin $$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$$ donne

$$\left\{\begin{aligned} x &= \left[\frac{\gamma}{2} t^2 + (\beta + \gamma)\, t + (\alpha + \beta + \gamma)\right] e^{-t} \\ y &= \left[\frac{\gamma}{2} t^2 + (\beta + \gamma)\, t + \alpha + \beta\right] e^{-t} \\ z &= \left(\frac{\gamma}{2} t^2 + \beta t + \alpha\right) e^{-t} \end{aligned}\right.$$

α, β, γ étant trois constantes.

■ **12.** Résoudre le système différentiel

$$(S) \left\{\begin{aligned} \frac{dx}{dt} &= ax + by + cz \\ \frac{dy}{dt} &= cx + ay + bz \\ \frac{dz}{dt} &= bx + cy + az \end{aligned}\right.$$

où a, b, c sont trois réels donnés. On pourra utiliser l'exercice 9.

Déterminer, sous forme réelle, la solution particulière telle que

$$x(0) = y(0) = z(0) = 1$$

La matrice de (S), $A = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ est une matrice circulante, ses valeurs propres, d'après l'exercice 9, sont

$$\lambda_1 = a + b + c, \qquad \text{réelle;}$$
$$\lambda_2 = a + bj + cj^2, \qquad \text{complexe;}$$
$$\lambda_3 = a + bj^2 + cj = \overline{\lambda_2}$$

avec $j = e^{2i\frac{\pi}{3}}$, racine cubique de l'unité. Un calcul direct serait également possible. A est donc diagonalisable et ses sous-espaces propres, d'après l'exercice 9, sont ceux de $K = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

Pour $\lambda_1 = a + b + c$, la valeur propre de K associée est 1, et les vecteurs propres sont donnés par

$$\begin{cases} -x + y = 0 \\ -y + z = 0 \\ x - z = 0 \end{cases}$$

d'où $\overrightarrow{u_1}\,(1, 1, 1)$.

Pour $\lambda_2 = a + bj + cj^2$, la valeur propre de K est j, on a donc

$$\begin{pmatrix} -jx + y = 0 \\ -jy + z = 0 \\ x - jz = 0 \end{pmatrix}$$

d'où $\overrightarrow{u_2}\,(1, j, j^2)$.

Un vecteur propre associé à $\lambda_3 = \overline{\lambda_2}$ est le conjugué de $\overrightarrow{u_2}$, soit $\overrightarrow{u_3}\,(1, j^2, j)$. D'où une matrice de passage

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{pmatrix}$$

et le système réduit

$$\begin{cases} \dfrac{dx_1}{dt} = (a + b + c)\, x_1 \\[2mm] \dfrac{dy_1}{dt} = (a + bj + cj^2)\, y_1 \\[2mm] \dfrac{dz_1}{dt} = (a + bj^2 + cj)\, z_1 \end{cases}$$

qui a pour solution générale

$$\begin{cases} x_1 = \alpha e^{(a+b+c)t} \\ y_1 = \beta e^{(a+bj+cj^2)t} \\ z_1 = \gamma e^{(a+bj^2+cj)t} \end{cases}$$

d'où

$$\begin{cases} x = x_1 + y_1 + z_1 = e^{at}\left[\alpha e^{(b+c)t} + \beta e^{(bj+cj^2)t} + \gamma e^{(bj^2+cj)t}\right] \\ y = x_1 + jy_1 + j^2 z_1 = e^{at}\left[\alpha e^{(b+c)t} + j\beta e^{(bj+cj^2)t} + j^2\gamma e^{(bj^2+cj)t}\right] \\ z = x_1 + j^2 y_1 + jz_1 = e^{at}\left[\alpha e^{(b+c)t} + j^2\beta e^{(bj+cj^2)t} + j\gamma e^{(bj^2+cj)t}\right] \end{cases}$$

Dans le cas général nous n'expliciterons pas cette solution, qui sera réelle si l'on prend α réel et β, γ complexes conjugués.

Les conditions initiales $x(0) = y(0) = z(0) = 1$ conduisent au système

$$\begin{cases} \alpha + \beta + \gamma = 1 \\ \alpha + j\beta + j^2\gamma = 1 \\ \alpha + j^2\beta + j\gamma = 1 \end{cases}$$

Par addition on obtient $\alpha = \dfrac{1}{3}$, il reste $\beta + \gamma = \dfrac{2}{3}$, $j\beta + j^2\gamma = \dfrac{2}{3}$

d'où $\beta = -\dfrac{2}{3}j$, $\gamma = -\dfrac{2}{3}j^2$

$$x = \frac{1}{3}e^{at}\left[e^{(b+c)t} - 2\,je^{(bj+cj^2)t} - 2\,j^2 e^{(bj^2+cj)t}\right]$$

$$y = \frac{1}{3}e^{at}\left[e^{(b+c)t} - 2\,j^2 e^{(bj+cj^2)t} - 2\,je^{(bj^2+cj)t}\right]$$

$$z = \frac{1}{3}e^{at}\left[e^{(b+c)t} - 2\,e^{(bj+cj^2)t} - 2\,e^{(bj^2+cj)t}\right]$$

Pour revenir à $\mathbb{R}$, il faut remplacer j par

$$\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = \frac{-1 + i\sqrt{3}}{2}, \quad j^2 \text{ par } \frac{-1 - i\sqrt{3}}{2}$$

d'où

$$bj + cj^2 = -\frac{b+c}{2} + \frac{i\sqrt{3}}{2}(b-c)$$

$$bj^2 + cj = -\frac{b+c}{2} - \frac{i\sqrt{3}}{2}(b-c)$$

Il s'ensuit un calcul assez long dont nous donnerons seulement le résultat :

$$\begin{cases} x = \dfrac{1}{3}e^{at}\left[e^{(b+c)t} + 2\,e^{-\frac{b+c}{2}t}\left(\cos\dfrac{\sqrt{3}}{2}(b-c)\,t + \sqrt{3}\sin\dfrac{\sqrt{3}}{2}(b-c)\,t\right)\right] \\ y = \dfrac{1}{3}e^{at}\left[e^{(b+c)t} + 2\,e^{-\frac{b+c}{2}t}\left(\cos\dfrac{\sqrt{3}}{2}(b-c)\,t - \sqrt{3}\sin\dfrac{\sqrt{3}}{2}(b-c)\,t\right)\right] \\ z = \dfrac{1}{3}e^{at}\left[e^{(b+c)t} - 4\,e^{-\frac{b+c}{2}t}\cos\dfrac{\sqrt{3}}{2}(b-c)\,t\right] \end{cases}$$

■ **13.** Dans ce problème, on envisage les différentes réductions possibles d'un endomorphisme f d'un espace vectoriel de dimension 3 : matrice diagonale si possible; à défaut, matrice de Jordan. La diagonalisation étant assurée lorsque f admet trois valeurs propres distinctes, nous envisagerons seulement les deux autres cas et leurs subdivisions suivant le degré du polynôme minimum de f.

1) f admet une valeur propre triple α : le polynôme minimum de f est $(f - \alpha I)^h$, où $h = 1$, 2 ou 3. Étudier dans ces divers cas les noyaux et images de $f - \alpha I$ et de ses puissances, en déduire les formes réduites de f.

2) Par des raisonnements analogues, étudier le cas où f admet une valeur propre double α et une valeur propre simple β.

---

1) *a*) $h = 1$: alors $f = \alpha I$, de matrice diagonale dans toute base.

b) $h = 2$: $(f - \alpha I)^2 = 0$, $f \neq \alpha I$.

$\forall \vec{u} \in E$, $(f - \alpha I)[(f - \alpha I)\vec{u}] = \vec{0}$, $(f - \alpha I)\vec{u} \in \ker(f - \alpha I)$.

Posons, pour simplifier, dans tout ce qui suit $\varphi = f - \alpha I$. On a donc

$$\operatorname{Im}\varphi \subset \ker\varphi \tag{1}$$

Compte tenu de $\dim \operatorname{Im}\varphi + \dim\ker\varphi = 3$, on a

$$\dim \operatorname{Im}\varphi = 1, \quad \dim\ker\varphi = 2$$

Il est donc possible de former une base de E avec deux vecteurs propres $\vec{u_1}$, $\vec{u_2}$ relatifs à la valeur propre α, et un troisième vecteur $\vec{u_3}$ non propre : cherchons à déterminer u_1, $\vec{u_2}$, $\vec{u_3}$ tels que dans cette base la matrice de f soit une matrice de Jordan.

Il faut pour cela $f(\vec{u_3}) = \vec{u_2} + \alpha\vec{u_3}$, soit $\varphi(\vec{u_3}) = \vec{u_2}$.

C'est possible si $\vec{u_2} \in \operatorname{Im}\varphi$. On choisira donc $\vec{u_2} \in \operatorname{Im}\varphi$, $\vec{u_1} \in \ker\varphi$, avec $\vec{u_1}$, $\vec{u_2}$ non colinéaires. Il existe alors $\vec{u_3} \neq 0$ tel que $\varphi(\vec{u_3}) = \vec{u_2}$ et comme $\vec{u_3} \notin \ker\varphi$, $\vec{u_1}, \vec{u_2}, \vec{u_3}$ forment une base de E.

Dans cette base, f s'exprime par $J_1 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$.

c) $k = 3$: $(f - \alpha I)^3 = 0$, $(f - \alpha I)^2 \neq 0$.

On ne peut plus avoir $\dim\ker\varphi = 2$ car f pourrait s'exprimer sous la forme triangulaire $T = \begin{pmatrix} \alpha & 0 & a \\ 0 & \alpha & b \\ 0 & 0 & \alpha \end{pmatrix}$ qui conduit à $(T - \alpha I)^2 = 0$ donc $\varphi^2 = 0$, ce qui est exclu.

Donc, nécessairement $\dim\ker\varphi = 1$ et on dispose, à un coefficient près, d'un seul vecteur propre u_1. Cherchons à déterminer $\vec{u_2}$, $\vec{u_3}$ formant avec $\vec{u_1}$ une base où f s'exprime par la matrice de Jordan

$$J_2 = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$$

Ceci équivaut à

$$f(\vec{u_2}) = \vec{u_1} + \alpha\vec{u_2}, \quad f(\vec{u_3}) = \vec{u_2} + \alpha\vec{u_3}, \tag{2}$$

donc $\quad \varphi(\vec{u_2}) = \vec{u_1}, \quad \varphi(\vec{u_3}) = \vec{u_2}$ et par suite $\varphi^2(\vec{u_3}) = \vec{u_1}$

Cette dernière relation signifie $\vec{u_1} \in \operatorname{Im} \varphi^2$ donc $\ker \varphi \subset \operatorname{Im} \varphi^2$. Or $\varphi^3 = 0$ montre que, $\forall \vec{u} \in E$, $\varphi[\varphi^2(\vec{u})] = \vec{0}$, donc $\operatorname{Im} \varphi^2 \subset \ker \varphi$. Comme $\dim \ker \varphi = 1$ et comme $\operatorname{Im} \varphi^2 = \{\vec{0}\}$ est impossible puisque $\varphi^2 \neq 0$, on a $\operatorname{Im} \varphi^2 = \ker \varphi$ et $\vec{u_1} \in \operatorname{Im} \varphi^2$.

Ainsi $\exists\ \vec{u_3}$, indépendant de $\vec{u_1}$, tel que $\varphi^2(\vec{u_3}) = \vec{u_1}$. Soit alors $\vec{u_2} = \varphi(\vec{u_3})$, $\vec{u_2}$ n'est pas colinéaire à $\vec{u_3}$ qui n'est pas un vecteur propre de f. Vérifions que $\vec{u_1}, \vec{u_2}, \vec{u_3}$ ne sont pas liés : en effet

$$\lambda_1\vec{u_1} + \lambda_2\vec{u_2} + \lambda_3\vec{u_3} = \vec{0} \Rightarrow \lambda_1\,\varphi(\vec{u_1}) + \lambda_2\,\varphi(\vec{u_2}) + \lambda_3\,\varphi(\vec{u_3}) = \vec{0}$$

soit $\lambda_1 . \vec{0} + \lambda_2\vec{u_1} + \lambda_3 . \vec{u_2} = \vec{0}$ et

$$\lambda_2\,\varphi(\vec{u_1}) + \lambda_3\,\varphi(\vec{u_2}) = \vec{0}, \qquad \lambda_2 . \vec{0} + \lambda_3\vec{u_1} = \vec{0}$$

Il en résulte $\lambda_3 = 0$, donc $\lambda_2 = 0$, donc $\lambda_1 = 0$.

$\vec{u_1}, \vec{u_2}, \vec{u_3}$ forment une base de E, satisfaisant aux relations (2). Dans cette base, f s'exprime par la matrice $J_2 = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$.

2) Lorsque f admet une valeur propre double α et une valeur propre simple β, son polynôme caractéristique est $(x - \alpha)^2\,(x - \beta)$, son polynôme minimum $Q(x)$ peut être $(x - \alpha)^2\,(x - \beta)$ ou $(x - \alpha)\,(x - \beta)$. Étudions ces deux cas.

a) $Q(x) = (x - \alpha)(x - \beta)$.

Alors $(f - \alpha I)(f - \beta I) = 0$, et $\forall \vec{u} \in E$

$$(f - \alpha I)[(f - \beta I)(\vec{u})] = \vec{0}$$
$$(f - \beta I)[(f - \alpha I)(\vec{u})] = \vec{0}$$

Donc

$$\operatorname{Im}(f - \beta I) \subset \ker(f - \alpha \mathrm{I})$$
$$\operatorname{Im}(f - \alpha I) \subset \ker(f - \beta I)$$

Mais, β étant valeur propre simple, $\dim \ker(f - \beta I) = 1$.

Comme $\dim \operatorname{Im}(f - \alpha I) = 0$ est exclu, car on aurait $f = \alpha I$, on a forcément $\dim \operatorname{Im}(f - \alpha I) = 1$, $\dim \ker(f - \alpha I) = 2$. On sait qu'alors f est diagonalisable.

b) $Q(x) = (x - \alpha)^2\,(x - \beta)$.

Alors, en raisonnant comme ci-dessus, on obtient

$$\operatorname{Im}(f - \alpha I)^2 \subset \ker(f - \beta I) \text{ qui est de dimension 1}$$

donc $\quad \operatorname{Im}(f-\alpha I)^2 = \ker(f-\beta I)$ puisque $(f-\alpha I)^2 \neq 0$

De même $\operatorname{Im}(f-\beta I) \subset \ker(f-\alpha I)^2$. Comme ces deux sous-espaces ont même dimension 2,

$$\operatorname{Im}(f-\beta I) = \ker(f-\alpha I)^2$$

Mais $(f-\alpha I)(f-\beta I) \neq 0$ donc $\operatorname{Im}(f-\alpha I) \not\subset \ker(f-\beta I)$

$$\operatorname{Im}(f-\alpha I) \not\subset \operatorname{Im}(f-\alpha I)^2$$

Comme $\quad \operatorname{Im}(f-\alpha I)^2 \subset \operatorname{Im}(f-\alpha I)$

on a $\quad \operatorname{Im}(f-\alpha I)^2 \neq \operatorname{Im}(f-\alpha I)$

$$\dim \operatorname{Im}(f-\alpha I) > \dim \operatorname{Im}(f-\alpha I)^2 = 1$$

Il s'ensuit $\dim \operatorname{Im}(f-\alpha I) \geqslant 2$, $\dim \ker(f-\alpha I) \leqslant 1$ et comme α est une valeur propre, $\dim \ker(f-\alpha I) = 1$.

Il en résulte que f n'est pas diagonalisable. On peut choisir, d'une seule façon à des coefficients près, deux vecteurs propres indépendants, $\vec{u_1}$, $\vec{u_2}$ associés respectivement aux valeurs propres β et α, et chercher un troisième vecteur $\vec{u_3}$ formant avec $\vec{u_1}$, $\vec{u_2}$ une base où f s'exprime par la matrice de Jordan $J_3 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$

Il faut pour cela $f(\vec{u_3}) = \vec{u_2} + \alpha\vec{u_3}$, $\vec{u_3}$ existe si $\vec{u_2} \in \operatorname{Im}(f-\alpha I)$ c'est-à-dire si $\ker(f-\alpha I) \subset \operatorname{Im}(f-\alpha I)$. Vu les dimensions, il suffit d'établir que ces deux sous-espaces ne sont pas supplémentaires, donc qu'il existe un vecteur $\vec{u}$ non nul qui leur est commun $(f-\alpha I)(\vec{u}) = \vec{0}$ avec $\vec{u} = (f-\alpha I)(\vec{v})$, $\vec{v} \in E$, donc $(f-\alpha I)^2(\vec{v}) = \vec{0}$, $\vec{v} \in \ker(f-\alpha I)^2$; un tel vecteur peut être choisi dans $\ker(f-\alpha I)^2$ qui est de dimension 2, et n'appartenant pas à $\ker(f-\alpha I)$ qui est de dimension 1, de sorte que $\vec{u} \neq \vec{0}$.

Donc $\vec{u_3}$ existe. Il reste à voir si $\vec{u_1}, \vec{u_2}, \vec{u_3}$ sont indépendants.

$$\lambda_1\vec{u_1} + \lambda_2\vec{u_2} + \lambda_3\vec{u_3} = \vec{0} \Rightarrow \lambda_1(f-\alpha I)(\vec{u_1}) + \lambda_2(f-\alpha I)(\vec{u_2}) + \lambda_3(f-\alpha I)(\vec{u_3}) = \vec{0}$$

$$\lambda_1(\beta-\alpha)\vec{u_1} + \lambda_2 . \vec{0} + \lambda_3\vec{u_2} = 0$$

Or $\vec{u}_1$ et $\vec{u_2}$ sont indépendants, donc $\lambda_3 = 0$ et $\lambda_1 = 0$ (car $\beta - \alpha \neq 0$). Il en résulte $\lambda_2 = 0$.

Dans la base $(\vec{u_1}, \vec{u_2}, \vec{u_3})$ ainsi constituée, f s'exprime par la matrice de Jordan $J_3 = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$.

En conclusion, les trois réductions les plus « achevées » que l'on peut obtenir pour un endomorphisme de dimension 3 sont les suivantes :

$D = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$ matrice diagonale, où α, β, γ ne sont pas nécessairement distincts.

$$J = \begin{pmatrix} \beta & 0 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$$ où on peut avoir $\alpha = \beta$ (forme commune de J_1 et J_3 dans la terminologie précédente).

$$J_2 = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}.$$

Remarquons que cet exercice constitue aussi, en dimension 3, une démonstration du théorème que nous avons admis en fin de chapitre, suivant lequel un endomorphisme est diagonalisable si et seulement si son polynôme minimum n'a que des racines simples.

CHAPITRE 9

FORMES QUADRATIQUES. ESPACES EUCLIDIENS. ISOMÉTRIES

On limite souvent l'algèbre linéaire proprement dit au contenu des trois chapitres qui précèdent. Dans ce chapitre et dans le suivant, nous aborderons sommairement les structures plus riches que peuvent conférer à un espace vectoriel sur $\mathbb{R}$ les notions de formes quadratiques et de produit scalaire, et à un espace vectoriel sur $\mathbb{C}$ celles de formes hermitiennes et de produit scalaire hermitien. Plusieurs résultats seront admis sans démonstration.

A. FORMES BILINÉAIRES. FORMES QUADRATIQUES

1. Forme bilinéaire. Définition et représentation matricielle

a) Par définition, une forme bilinéaire d'un espace vectoriel E sur un corps K ($\mathbb{R}$, habituellement), est une application φ de $E \times E$ dans K,

$$(\vec{u}, \vec{v}) \to \varphi(\vec{u}, \vec{v})$$

telle que, pour tout vecteur et pour tout scalaire,

$$\varphi(\vec{u_1} + \vec{u_2}, \vec{v}) = \varphi(\vec{u_1}, \vec{v}) + \varphi(\vec{u_2}, \vec{v})$$
$$\varphi(\vec{u}, \vec{v_1} + \vec{v_2}) = \varphi(u, \vec{v_1}) + \varphi(\vec{u}, \vec{v_2})$$
$$\varphi(\lambda\vec{u}, \vec{v}) = \varphi(\vec{u}, \lambda\vec{v}) = \lambda\varphi(\vec{u}, \vec{v})$$

φ est donc linéaire à gauche et linéaire à droite.

b) Si E est de dimension finie n, rapporté à une base $\mathscr{B} = (\vec{e_1}, \ldots, \vec{e_n})$, et si $\vec{u}$, $\vec{v}$ sont, dans cette base, définis par les matrices-colonnes

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

la bilinéarité permet d'écrire

$$\varphi(\vec{u}, \vec{v}) = \varphi\left(\sum_1^n x_i\vec{e_i}, \sum_1^n y_j\vec{e_j}\right) = \sum_{i,j} x_i y_j\, \varphi(\vec{e_i}, \vec{e_j})$$

φ est donc parfaitement définie si l'on se donne

$$\varphi(\vec{e_i}, \vec{e_j}) = a_{ij} \qquad (\forall\, i, j,\ 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n)$$

c'est-à-dire la matrice carrée $A = \begin{pmatrix} & \vdots & \\ \dots & a_{ij} & \dots \\ & \vdots & \end{pmatrix}$.

On a donc

$$\varphi(\vec{u}, \vec{v}) = \sum_{i,j} a_{ij}x_iy_j = \sum_{j=1}^{j=n}\left(\sum_{i=1}^{i=n} a_{ij}x_i\right) y_j$$

c'est-à-dire

$$\boxed{\varphi(\vec{u}, \vec{v}) = {}^tXAY}$$

A est la matrice de φ dans la base $\mathscr{B}$.

Notons que l'on a également $\varphi(\vec{u}, \vec{v}) = {}^tY\,{}^tAX$.

Par exemple pour $n = 2$, $A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$ on obtient

$$\varphi(\vec{u}, \vec{v}) = (x_1 \quad x_2)\begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
$$= x_1y_1 - x_1y_2 + 2\,x_2y_1 + 3\,x_2y_2$$

c) Soit P la matrice d'un changement de base de E : X', Y' et A' étant les matrices de φ dans la nouvelle base $\mathscr{B}'$, on a

$$X = PX', \qquad Y = PY'$$
$$\varphi(\vec{u}, \vec{v}) = {}^tXAY = {}^tX'\,{}^tPAPY' = {}^tX'A'Y'$$

il en résulte que

$$\boxed{A' = {}^tPAP}$$

(formule de changement de base pour une application bilinéaire).

Comme P et ${}^tP$ sont régulières, A et A' ont même rang, et le rang commun à toutes les matrices représentant φ s'appelle le rang de φ.

Remarquons qu'en général A' et A ne sont pas semblables.

2. Forme bilinéaire symétrique, forme quadratique associée

On dit que la forme bilinéaire φ est symétrique si

$$\forall\, \vec{u}, \vec{v} \qquad \boxed{\varphi(\vec{v}, \vec{u}) = \varphi(\vec{u}, \vec{v})}$$

On appelle **forme quadratique associée à** φ l'application $\vec{u} \to Q(\vec{u})$ telle que

$$\boxed{Q(\vec{u}) = \varphi(\vec{u}, \vec{u})}$$

En dimension finie, et dans une base $\mathscr{B}$, φ est symétrique si et seulement si, $\forall\ i, j,\ \ \varphi(\vec{e_i}, \vec{e_j}) = \varphi(\vec{e_j}, \vec{e_i})$ donc si A est symétrique.

La forme matricielle de Q dans une base est donc

$$Q(\vec{u}) = {}^tXAX$$

A étant une matrice symétrique.

Un changement de base dans E s'effectue encore par la formule $A' = {}^tPAP$.

En explicitant $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ on a donc

$$Q(\vec{u}) = \sum_1^n a_{ii}x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_{ij}x_ix_j$$

On peut donc encore définir une forme quadratique comme une application qui, à un vecteur de E, associe un polynôme homogène de degré 2 des coordonnées de ce vecteur. On vérifie aisément que la matrice A de Q est celle des n formes linéaires $\dfrac{1}{2}\dfrac{\partial Q}{\partial x_i}$. La bilinéarité permet d'établir facilement les propriétés suivantes :

$\forall\ \vec{u} \in E,\ \ \forall\ \lambda$ scalaire $\boxed{Q(\lambda\vec{u}) = \lambda^2 Q(\vec{u})}$ (1)

$\forall\ \vec{u}, \vec{v}$ de E

$$\boxed{\varphi(\vec{u}, \vec{v}) = \frac{1}{2}[Q(\vec{u} + \vec{v}) - Q(\vec{u}) - Q(\vec{v})]} \qquad (2)$$

La propriété (1) et le fait que l'application φ définie par (2) est bilinéaire caractérisent une forme quadratique.

Ceci est valable en dimension finie ou infinie.

En dimension finie n, le passage de φ à Q s'effectue évidemment en écrivant $y_i = x_i,\ \ \forall\ i$; le passage de Q à φ s'effectue, pratiquement, par la règle (facile à vérifier en partant de (2)) dite du « dédoublement des termes » : on remplace x_i^2 par x_iy_i

$$x_ix_j \quad \text{par} \quad \frac{1}{2}(x_iy_j + x_jy_i).$$

La forme bilinéaire symétrique φ associée à Q est souvent appelée **forme polaire** de Q.

3. Orthogonalité par rapport à une forme quadratique

a) Q étant une forme quadratique de E, φ sa forme polaire, deux vecteurs $\vec{u}$, $\vec{v}$ de E sont dits orthogonaux (ou conjugués) par rapport à Q si $\varphi(\vec{u}, \vec{v}) = 0$; on note $\vec{u} \perp \vec{v}$ (ou $\vec{v} \perp \vec{u}$).

F étant un sous-espace de E, on appelle orthogonal de F, et on note $F^{\perp}$, l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F. Ainsi

$$\vec{v} \in F^{\perp} \Leftrightarrow \varphi(\vec{u}, \vec{v}) = 0, \qquad \forall \vec{u} \in F$$

Il faut et il suffit pour cela que $\vec{v}$ soit orthogonal à chacun des vecteurs d'une base de F.

La bilinéarité de φ montre que

$$\varphi(\vec{u}, \vec{v}) = 0 \Rightarrow \varphi(\vec{u}, \lambda\vec{v}) = 0 \qquad \text{donc} \quad \vec{v} \in F^{\perp} \Rightarrow \lambda\vec{v} \in F^{\perp}$$

$$\varphi(\vec{u}, \vec{v_1}) = 0 \quad \text{et} \quad \varphi(\vec{u}, \vec{v_2}) = 0 \Rightarrow \varphi(\vec{u}, \vec{v_1} + \vec{v_2}) = 0$$

donc $$\vec{v_1} \in F^{\perp}, \vec{v_2} \in F^{\perp} \Rightarrow \vec{v_1} + \vec{v_2} \in F^{\perp}$$

$F^{\perp}$ est donc un sous-espace vectoriel de E.

b) On appelle noyau de la forme quadratique Q, de forme polaire φ, l'ensemble des vecteurs $\vec{v}$ de E tels que $\varphi(\vec{u}, \vec{v}) = 0 \quad \forall \vec{u} \in E$. Le noyau n'est donc autre que $E^{\perp}$.

Une forme quadratique est dite non dégénérée si et seulement si son noyau est réduit au vecteur nul.

On démontre, et nous admettrons, que pour une forme quadratique non dégénérée d'un espace E de dimension finie n, on a

$$\dim F + \dim F^{\perp} = n \qquad (F^{\perp})^{\perp} = F$$

Cependant, en général, F et $F^{\perp}$ ne sont pas supplémentaires.

4. Base orthogonale. Diagonalisation. Décomposition en carrés

Dans ce qui suit, le corps K associé à E est $\mathbb{R}$.

a) E étant un espace vectoriel de dimension n, Q une forme quadratique de E, de forme polaire φ, une base $(\vec{e_1}, \ldots, \vec{e_n})$ de E est orthogonale par rapport à Q si, pour

$$i \neq j \qquad \varphi(\vec{e_i}, \vec{e_j}) = 0$$

Nous admettrons l'existence d'une telle base.

Dans une telle base, la matrice de Q est diagonale, et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ étant la matrice d'un vecteur $\vec{u}$, on a $Q(\vec{u}) = \sum_{1}^{n} \alpha_i y_i^2$.

Le rang de Q est égal à celui de la matrice

$$\begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_n \end{pmatrix}$$

donc au nombre r d'éléments non nuls parmi les α_i.

On dit que la connaissance d'une base orthogonale permet de diagonaliser la forme quadratique Q. Elle permet aussi de la décomposer « en carrés », plus précisément de l'exprimer sous forme d'une somme algébrique de carrés de formes linéaires, indépendantes*, des coordonnées de $\vec{u}$ dans une base quelconque. Si A est la matrice de Q dans cette base et A' la matrice diag $(\alpha_1, \alpha_2, \dots, \alpha_n)$, on sait que A' n'est pas semblable à A, les α_i ne sont donc pas les valeurs propres de φ (sauf exception), et diagonaliser Q ne signifie pas diagonaliser A. Cependant nous verrons ultérieurement que la diagonalisation de A conduit à un cas particulier de la diagonalisation de Q.

b) Réciproquement, si on peut trouver une base de E dans laquelle $Q(\vec{u})$ peut s'exprimer sous la forme $\sum_{1}^{r} \alpha_i y_i^2$ où les y_i sont des formes linéaires, indépendantes, des coordonnées x_j de $\vec{u}$ dans la base canonique, la base où est ainsi exprimée Q est une base orthogonale $(\vec{f_1}, \vec{f_2}, \dots, \vec{f_n})$ puisque $\varphi(\vec{f_i}, \vec{f_j}) = 0$ pour $i \neq j$.

r étant le rang de φ, les α_i sont tous non nuls et on peut les supposer ordonnés de telle sorte que

$$\alpha_1 > 0, \alpha_2 > 0, \dots, \alpha_p > 0, \quad \alpha_{p+1} < 0, \alpha_{p+2} < 0, \dots, \alpha_r < 0$$

avec $1 \leqslant p \leqslant r \leqslant n$.

En posant $\sqrt{\alpha_i}\, y_i = z_i$ si $\alpha_i > 0$, $\sqrt{-\alpha_i}\, y_i = z_i$ si $\alpha_i < 0$, on a donc

$$\boxed{Q(\vec{u}) = z_1^2 + z_2^2 + \dots + z_p^2 - z_{p+1}^2 - z_{p+2}^2 - \dots - z_r^2} \qquad (3)$$

* Indépendantes, car les y_i s'expriment au moyen des « anciennes » coordonnées x_i par la matrice régulière P^{-1}, inverse de la matrice de changement de base.

L'expression (3), dont l'existence est assurée par celle d'une base orthogonale dans E, où les z_i sont r formes linéaires des x_j, indépendantes, est une décomposition en carrés de Q. On démontre, et nous l'admettrons, que les diverses décompositions en carrés de Q font apparaître le même nombre de signes + et le même nombre de signes −, autrement dit p est un entier attaché à la forme quadratique Q : on dit que Q est une forme quadratique de **signature** $(p, r - p)$, ou encore de type $(p, r - p)$.

Ce résultat est connu sous le nom de « loi d'inertie de Sylvester ».

Lorsque $p = r$, Q est une forme quadratique positive.

Lorsque $p = r = n$, Q est une **forme quadratique définie positive** :

$$Q(\vec{u}) \geqslant 0, \quad \forall \vec{u}, \qquad Q(\vec{u}) = 0 \Leftrightarrow \vec{u} = \vec{0}$$

On définit de même une forme négative, une forme définie négative.

c) Nous allons établir un procédé algébrique de décomposition en carrés d'une forme quadratique exprimée comme polynôme homogène du second degré $Q(x_1, x_2, \ldots, x_n)$. Pour $n = 1$, l'existence d'une telle décomposition est évidente.

Raisonnons par récurrence et supposons que toute forme quadratique de $n - 1$ variables est décomposable en carrés de formes linéaires indépendantes.

— Si Q contient un terme en x_1^2, on peut écrire (avec $\alpha_1 \neq 0$)

$$Q(x_1, x_2, \ldots, x_n) = \alpha_1 x_1^2 + 2\, x_1 f(x_2, x_3, \ldots, x_n) + R(x_2, x_3, \ldots, x_n)$$

où f est une forme linéaire et R une forme quadratique. Donc

$$Q = \alpha_1 \left(x_1 + \frac{f}{\alpha_1}\right)^2 - \frac{f^2}{\alpha_1} + R = \alpha_1 \left(x_1 + \frac{f}{\alpha_1}\right)^2 + S$$

où S est une forme quadratique des variables $x_2, \ldots, x_n$.

S est décomposable en carrés de formes linéaires indépendantes, et $x_1 + \dfrac{f}{\alpha_1}$ est indépendante de celles-ci puisqu'elles ne contiennent pas x_1. La propriété envisagée est donc vraie pour n variables.

— Si Q ne contient pas x_1^2, on peut écrire, avec $a \neq 0$,

$$Q = a x_1 x_2 + x_1 f(x_3, x_4, \ldots, x_n) + x_2 g(x_3, \ldots, x_n) + T(x_3, \ldots, x_n)$$

où f et g sont des formes linéaires et T une forme quadratique.

Donc

$$\begin{aligned} Q &= a\left[\left(x_1 + \frac{g}{a}\right)\left(x_2 + \frac{f}{a}\right) - \frac{fg}{a^2}\right] + T \\ &= \frac{a}{4}\left[\left(x_1 + x_2 + \frac{f+g}{a}\right)^2 - \left(x_1 - x_2 + \frac{g-f}{a}\right)^2\right] + T - \frac{fg}{a} \end{aligned}$$

$T - \dfrac{fg}{a}$ étant une forme quadratique à $n-2$ variables, le lecteur achèvera facilement la démonstration.

Cette méthode, connue sous le nom de **méthode de Gauss**, est en général d'application simple.

Exemple.

$$\begin{aligned}
Q(x, y, z, t) &= x^2 + y^2 + 4t^2 + 2xy - 4xt + 2yz - 4zt \\
&= x^2 + 2(y - 2t)x + y^2 + 4t^2 + 2yz - 4zt \\
&= (x + y - 2t)^2 - (y - 2t)^2 + y^2 + 4t^2 + 2yz - 4zt \\
&= (x + y - 2t)^2 + 2yz + 4yt - 4zt \\
&= (x + y - 2t)^2 + 2[(y - 2t)(z + 2t) + 4t^2] \\
&= (x + y - 2t)^2 + \frac{1}{2}(y + z)^2 - \frac{1}{2}(y - z - 4t)^2 + 8t^2
\end{aligned}$$

On a donc affaire à une forme quadratique de rang 4, de signature (3, 1).

La matrice P telle que

$$\begin{pmatrix} x + y - 2t \\ y + z \\ y - z - 4t \\ t \end{pmatrix} = P^{-1} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \quad \text{soit} \quad P^{-1} = \begin{pmatrix} 1 & 1 & 0 & -2 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & -4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

est une matrice de changement de base permettant la diagonalisation de Q.

La matrice de Q dans la base canonique étant celle de

$$\frac{1}{2}\frac{\partial Q}{\partial x}, \ldots, \frac{1}{2}\frac{\partial Q}{\partial t}$$

soit

$$A = \begin{pmatrix} 1 & 1 & 0 & -2 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -2 \\ -2 & 0 & -2 & 4 \end{pmatrix}$$

on pourra vérifier que

$${}^tPAP = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

B. ESPACE EUCLIDIEN

1. Définition d'un espace euclidien. Produit scalaire

Soit E un espace vectoriel sur lequel on peut choisir une forme quadratique définie positive Q, de forme bilinéaire associée φ. Le nombre réel $\varphi(\vec{u}, \vec{v})$ s'appelle produit scalaire des deux vecteurs $\vec{u}$, $\vec{v}$, associé à Q, et se note

$$\boxed{(\vec{u} \mid \vec{v})} \qquad \text{ou} \qquad < \vec{u}, \vec{v} >$$

ou plus simplement dans les cas usuels $\vec{u} \cdot \vec{v}$ (« $\vec{u}$ scalaire $\vec{v}$ »).

L'espace E muni d'un produit scalaire s'appelle espace vectoriel euclidien.

2. Norme associée à un produit scalaire. Espace normé

a) Q et φ étant choisies sur E, le « carré scalaire » $(\vec{u} \mid \vec{u})$ est positif ou nul. Sa racine carrée s'appelle norme de $\vec{u}$, et se note $\| \vec{u} \|$. Ainsi, par définition

$$\boxed{\| \vec{u} \| = \sqrt{\vec{u} \mid \vec{u}}}$$

D'après les propriétés des formes quadratiques définies positives,

$$\| \lambda\vec{u} \| = | \lambda | \; \| \vec{u} \| \qquad \forall \lambda \text{ scalaire}$$

et

$$\| \vec{u} \| = 0 \Rightarrow \vec{u} = \vec{0}$$

Nous allons établir une autre propriété : par définition $\forall \lambda, \vec{u}, \vec{v}$,

$$\begin{aligned} Q(\vec{u} + \lambda\vec{v}) &= \varphi(\vec{u} + \lambda\vec{v}, \vec{u} + \lambda\vec{v}) \\ &= \varphi(\vec{u}, \vec{u}) + \lambda\varphi(\vec{v}, \vec{u}) + \lambda\varphi(\vec{u}, \vec{v}) + \lambda^2\varphi(\vec{v}, \vec{v}) \\ &= \| \vec{u} \|^2 + 2\lambda(\vec{u} \mid \vec{v}) + \lambda^2 \| \vec{v} \|^2 \geqslant 0 \end{aligned}$$

le discriminant de ce trinôme est donc négatif ou nul, et nous avons

$$(\vec{u} \mid \vec{v})^2 - \| \vec{u} \|^2 \| \vec{v} \|^2 \leqslant 0$$

Donc

$$\boxed{|(\vec{u} \mid \vec{v})| \leqslant \| \vec{u} \| \; \| \vec{v} \|} \tag{4}$$

C'est **l'inégalité de Schwarz** (ou de Cauchy-Schwarz). On peut en déduire

$$\begin{aligned} \| \vec{u} + \vec{v} \|^2 &= (\vec{u} + \vec{v} \mid \vec{u} + \vec{v}) \\ &= (\vec{u} \mid \vec{u}) + (\vec{v} \mid \vec{v}) + 2(\vec{u} \mid \vec{v}) \\ &\leqslant (\vec{u} \mid \vec{u}) + (\vec{v} \mid \vec{v}) + 2 \| \vec{u} \| \; \| \vec{v} \| = (\| \vec{u} \| + \| \vec{v} \|)^2 \end{aligned}$$

Donc
$$\boxed{\| \vec{u} + \vec{v} \| \leqslant \| \vec{u} \| + \| \vec{v} \|} \tag{5}$$

C'est **l'inégalité de Minkowski** ou **inégalité triangulaire.**

Cette dernière dénomination évoque le fait que le produit scalaire et la norme dans l'espace $\mathbb{R}^3$ de la « géométrie vectorielle » ordinaire sont des cas particuliers des notions générales précédentes et qu'alors (5) exprime qu'un côté d'un triangle est inférieur ou égal à la somme des deux autres.

b) D'une façon générale, une **norme** sur un espace vectoriel E est une application de E dans $\mathbb{R}^+$ $\vec{u} \to \| \vec{u} \|$ satisfaisant aux trois conditions suivantes

$$\boxed{\begin{array}{c} \| \lambda\vec{u} \| = | \lambda | \, \|\vec{u}\| \\ \| \vec{u} \| = 0 \Rightarrow \vec{u} = \vec{0} \\ \| \vec{u} + \vec{v} \| \leqslant \| \vec{u} \| + \| \vec{v} \| \end{array}}$$

Un espace vectoriel muni d'une norme est un **espace vectoriel normé.**

Notons encore qu'en analyse on appelle espace de Banach un espace vectoriel normé « complet », c'est-à-dire sur lequel toute suite de Cauchy (suite $\vec{u_n}$ telle que $\forall \varepsilon > 0$, $\exists n_0 \in \mathbb{N}$ tel que

$$p \geqslant n \geqslant n_0 \Rightarrow \| \vec{u_p} - \vec{u_n} \| < \varepsilon)$$

est convergente. $\mathbb{R}$ et $\mathbb{C}$ sont des espaces de Banach et l'étude classique de la convergence d'une suite ou d'une série s'étend aux espaces de Banach.

3. Base orthonormée d'un espace euclidien. Procédé d'orthonormalisation de Schmidt

a) Par définition, **$\vec{u}$ et $\vec{v}$ sont orthogonaux si $(\vec{u} \mid \vec{v}) = 0$.**

b) Rappelons qu'une forme quadratique définie positive, sur un espace vectoriel E de dimension n, a pour rang n et pour signature $(n, 0)$, elle peut donc s'écrire dans une base $\mathscr{B}$ sous la forme

$$Q(\vec{u}) = \sum_1^n x_i^2$$

où $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est la matrice de $\vec{u}$ dans $\mathscr{B}$.

La forme bilinéaire associée, d'après la règle du dédoublement des termes, s'écrit

$$\varphi(\vec{u}, \vec{v}) = \sum_1^n x_i y_i \qquad \text{où} \qquad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

est la matrice de $\vec{v}$. La base $\mathscr{B}$ étant constituée de $(\vec{e_1}, \ldots, \vec{e_n})$, de matrices respectives $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ est telle que

$$\forall\, i \neq j \qquad \varphi(\vec{e_i}, \vec{e_j}) = 0$$
$$\forall\, i \qquad \varphi(\vec{e_i}, \vec{e_i}) = 1$$

Q définissant sur E le produit scalaire noté $(\vec{u} \mid \vec{v})$, on a

$$(\vec{e_i} \mid \vec{e_j}) = 0 \quad \text{si} \quad i \neq j \qquad \| \vec{e_i} \| = 1$$

Les vecteurs de base sont deux à deux orthogonaux et ont pour norme 1, la base $\mathscr{B}$ est dite orthonormée (ou orthonormale). Les $\vec{e_i}$ sont des vecteurs unitaires. Relativement à une **base orthonormée**, la norme et le produit scalaire s'écrivent

$$\boxed{\begin{gathered} \| \vec{u} \| = \sqrt{\sum_1^n x_i^2} \\ (\vec{u} \mid \vec{v}) = \sum_1^n x_i y_i \end{gathered}}$$

Ceci peut également s'écrire $\vec{u} \,.\, \vec{v} = {}^tXY$.

D'une façon générale une famille de p vecteurs de E, non nuls et deux à deux orthogonaux, est nécessairement une famille libre; en effet

$$\sum_1^p \lambda_i \vec{u_i} = \vec{0} \Rightarrow \sum_1^p \lambda_i (\vec{u_i} \mid \vec{u_j}) = 0, \qquad \forall\, j = 1, 2, \ldots, p$$

Or $(\vec{u_i} \mid \vec{u_j}) = 0$ si $i \neq j$ et $(\vec{u_j} \mid \vec{u_j}) = \| \vec{u_j} \|^2 \neq 0$ donc $\lambda_j = 0$.

c) Nous allons montrer qu'on peut construire, à partir d'une suite finie ou non $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}, \ldots$ de vecteurs indépendants de E (la dimension ici n'étant pas nécessairement finie), une suite orthonormée $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}, \ldots$ Il suffit pour cela de construire une suite de vecteurs $\vec{w_1}, \vec{w_2}, \ldots, \vec{w_n} \ldots$ deux à deux orthogonaux, la normalisation étant ensuite évidente en prenant $\vec{v_n} = \dfrac{\vec{w_n}}{\| \vec{w_n} \|}$. Nous noterons ici, pour simplifier, $\vec{u} \,.\, \vec{v}$ le produit scalaire.

Prenons $\vec{w_1} = \vec{u_1}$ et cherchons $\vec{w_2} = \vec{u_2} + \alpha \vec{w_1}$.

$\vec{w_1} \cdot \vec{w_2} = 0$ donne la condition $\vec{u_2} \cdot \vec{w_1} + \alpha \vec{w_1}^2 = 0$ qui détermine α puisque $\vec{w_1}^2 \neq 0$.

$\vec{w_1}, \vec{w_2}$ sont orthogonaux et indépendants.

Cherchons ensuite $\overrightarrow{w_3}$ sous la forme $\overrightarrow{w_3} = \overrightarrow{u_3} + \beta\overrightarrow{w_1} + \gamma\overrightarrow{w_2}$.

Les relations $\overrightarrow{w_1} \cdot \overrightarrow{w_3} = 0$ et $\overrightarrow{w_2} \cdot \overrightarrow{w_3} = 0$ s'écrivent

$$\overrightarrow{u_3} \cdot \vec{w} + \beta\overrightarrow{w_1}^2 = 0 \quad \text{qui fournit } \beta$$
$$\overrightarrow{u_3} \cdot \overrightarrow{w_2} + \gamma\overrightarrow{w_2}^2 = 0 \quad \text{qui fournit } \gamma$$

$\overrightarrow{w_1}, \overrightarrow{w_2}, \overrightarrow{w_3}$ sont orthogonaux et non nuls, donc indépendants.

Une récurrence est alors immédiate : si $\overrightarrow{w_1}, \overrightarrow{w_2}, \ldots, \overrightarrow{w_{n-1}}$ ainsi construits sont deux à deux orthogonaux et non nuls.

$$\overrightarrow{w_n} = \overrightarrow{u_n} + \lambda_1\overrightarrow{w_1} + \lambda_2\overrightarrow{w_2} + \ldots + \lambda_{n-1}\overrightarrow{w_{n-1}}$$

peut être déterminé comme orthogonal à $\overrightarrow{w_1}, \overrightarrow{w_2}, \ldots, \overrightarrow{w_{n-1}}$, par les relations $\overrightarrow{w_n} \cdot \overrightarrow{w_i} = \overrightarrow{u_n} \cdot \overrightarrow{w_i} + \lambda_i\overrightarrow{w_i}^2 = 0$ qui fournissent les λ_i ($i = 1, 2, \ldots, n-1$).

La méthode précédente, qui justifie en dimension finie l'existence d'une base orthonormée et permet de la construire de proche en proche, est connue sous le nom de **procédé d'orthonormalisation de Schmidt**.

Remarque importante.

La donnée d'une base orthonormée de l'espace E de dimension finie, définit complètement le produit scalaire sur E, et confère ainsi à E une structure d'espace euclidien.

4. Changement de base orthonormée. Matrice orthogonale, endomorphisme orthogonal

a) Soit $\mathscr{B} = (\overrightarrow{e_1}, \overrightarrow{e_2}, \ldots, \overrightarrow{e_n})$ et $\mathscr{B}' = (\overrightarrow{f_1}, \overrightarrow{f_2}, \ldots, \overrightarrow{f_n})$ deux bases orthonormées de E. La matrice A de $\mathscr{B}'$ par rapport à $\mathscr{B}$ définit un endomorphisme φ tel que $\varphi(\overrightarrow{e_i}) = f_i$, $i = 1, 2, \ldots, n$. Soit

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

Les vecteurs-colonnes de A sont unitaires et deux à deux orthogonaux, donc la somme des carrés des éléments d'une colonne est égale à 1, et la somme des produits deux à deux des éléments de deux colonnes distinctes, somme de la forme $a_{1i}a_{1j} + a_{2i}a_{2j} + \ldots + a_{ni}a_{nj}$ est nulle.

Il en résulte ${}^tA \cdot A = I$.

Ceci confirme que les $\overrightarrow{f_i}$ constituent une base de E, puisque A est inversible, et montre que $A^{-1} = {}^tA$.

Une telle matrice est dite orthogonale. Sa transposée est aussi orthogonale. Par définition, **on appelle matrice orthogonale toute matrice carrée qui a pour inverse sa transposée**. C'est la matrice d'un changement de base orthonormée, ou de l'endomorphisme φ appliquant une base orthonormée sur une autre base orthonormée (éventuellement, la même, alors $A = I$ qui est bien une matrice orthogonale).

b) Étudions les propriétés de φ et de A.

Comme $\varphi(\vec{e_i}) = \vec{f_i}$

$$\varphi\left(\sum_1^n x_i\vec{e_i}\right) = \sum_1^n x_i\vec{f_i} \qquad \vec{u} = \sum_1^n x_i\vec{e_i}$$

$$\varphi\left(\sum_1^n y_i\vec{e_i}\right) = \sum_1^n y_i\vec{f_i} \qquad \vec{v} = \sum_1^n y_i\ \vec{e_i}$$

les deux bases étant orthonormées

$$\varphi(\vec{u}) \,.\, \varphi(\vec{v}) = \sum_1^n x_i y_i = \vec{u} \,.\, \vec{v}$$

L'application φ conserve donc le produit scalaire de deux vecteurs. Elle conserve donc aussi la norme d'un vecteur. On dit que φ est une isométrie. Ainsi, par définition, **on appelle isométrie un endomorphisme qui conserve le produit scalaire.**

Un tel endormorphisme transformant une base orthonormée en base orthonormée, sa matrice est orthogonale.

Si λ est une valeur propre de φ, $\vec{u}$ un vecteur propre associé à λ, $\varphi(\vec{u}) = \lambda\vec{u} \Rightarrow \|\varphi(\vec{u})\| = |\lambda|\,\|\vec{u}\|$ donc $\|\vec{u}\| = |\lambda|\,\|\vec{u}\|$ et $|\lambda| = 1$. Ceci est établi dans $\mathbb{R}$, mais vaut également dans $\mathbb{C}$ en considérant pour $\vec{u}$ et $\varphi(\vec{u})$ les « normes hermitiennes » (chap. suivant).

Ainsi **toute valeur propre d'une isométrie a pour module 1.**

On démontre, et nous admettrons qu'**une isométrie** (donc aussi une matrice orthogonale) **est diagonalisable sur** $\mathbb{C}$.

Enfin

$${}^tA \,.\, A = I \Rightarrow (\text{dét}\ A)^2 = 1 \qquad \boxed{\text{dét}\ A = \pm 1}$$

si A est orthogonale.

5. Notions sur les isométries d'un espace vectoriel réel de dimension finie

a) En géométrie, une isométrie est une transformation qui conserve la distance de deux points quelconques, donc la norme d'un vecteur; elle conserve donc aussi l'angle de deux vecteurs et leur produit scalaire.

Cette définition, on le voit, reste la même dans un espace euclidien de dimension quelconque, elle conduit à généraliser la notion d'angle et à définir l'angle θ de deux vecteurs $\vec{u}$ et $\vec{v}$ par la formule

$$\theta = \text{Arc}\cos\frac{\vec{u} \,.\, \vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}$$

b) A tout sous-espace de E, soit F, on peut associer un orthogonal et un seul de F, qui est l'un des sous-espaces supplémentaires de F, noté $F^{\perp}$. On peut compléter une base orthonormée de F par une base orthonormée de $F^{\perp}$ et former ainsi une base orthonormée de E. Le projecteur de E sur F parallèlement à $F^{\perp}$ est appelé **projecteur orthogonal** de E sur F.

La décomposition unique $\vec{u} = \vec{u_1} + \vec{u_2}$, $\vec{u_1} \in F$, $\vec{u_2} \in F^{\perp}$ fait apparaître les projections orthogonales $\vec{u_1}$ et $\vec{u_2}$ de $\vec{u}$ sur F et $F^{\perp}$. On a

$$\vec{u} \cdot \vec{u} = \| \vec{u} \|^2 = \| \vec{u_1} \|^2 + 2\, \vec{u_1} \cdot \vec{u_2} + \| \vec{u_2} \|^2$$

or $\vec{u_1} \cdot \vec{u_2} = 0$ d'où le **théorème de Pythagore :**

$$\boxed{\| \vec{u} \|^2 = \| \vec{u_1} \|^2 + \| \vec{u_2} \|^2}$$

où $\vec{u_1}$, $\vec{u_2}$ sont les projections orthogonales de $\vec{u}$ sur deux sous-espaces orthogonaux de E.

c) En généralisant encore des propriétés de géométrie élémentaire, et en utilisant les notations qui précèdent, considérons l'endomorphisme φ tel que $\varphi(\vec{u}) = \vec{u_1} - \vec{u_2}$ c'est-à-dire $\varphi = p - q$, p étant le projecteur orthogonal sur F et q le projecteur orthogonal sur $F^{\perp}$.

Comme $p + q = I$ on a aussi

$$\varphi = 2\,p - I, \qquad \varphi^2 = 4\,p^2 - 4\,p + I = I$$

φ est un endomorphisme involutif. Comme

$$\varphi(\vec{u_1}) = \vec{u_1}, \qquad \forall\, \vec{u_1} \in F,$$

et

$$\varphi(\vec{u_2}) = -\,\vec{u_2}, \qquad \forall\, \vec{u_2} \in F$$

φ a pour valeurs propres 1 et -1, pour sous-espaces propres F et $F^{\perp}$, supplémentaires, et si $\dim F = r$ (rang de p), sa matrice dans une base orthonormée de E, union de bases orthonormées de F et $F^{\perp}$, est

$$A = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & & & & & \vdots \\ \vdots & & \ddots & & & & \vdots \\ 0 & \cdots & \cdots & 1 & & & \vdots \\ \vdots & & & 0 & -1 & & \vdots \\ \vdots & & & \vdots & & -1 & 0 \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & -1 \end{pmatrix}$$

$$\underbrace{\qquad\qquad}_{\substack{r \\ \text{colonnes}}} \qquad \underbrace{\qquad\qquad}_{\substack{n-r \\ \text{colonnes}}}$$

φ est, par définition, la **symétrie orthogonale par rapport à F.**

C'est, pour des raisons évidentes et aussi par définition (A étant orthogonale et étant la matrice de φ dans une base orthonormée), une isométrie. Son déterminant est $(-1)^{n-r}$, donc 1 si $F^{\perp}$ est de dimension paire, -1 si $F^{\perp}$ est de dimension impaire.

Réciproquement une isométrie qui a pour seules valeurs propres 1 et -1, étant diagonalisable, a pour polynôme minimum $\lambda^2 - 1$, est involutive et il est facile de l'identifier à une symétrie orthogonale par rapport au sous-espace de ses vecteurs invariants.

d) Une isométrie φ admettant une valeur propre complexe $e^{i\theta}$, $\theta \neq k\pi$, admet aussi la valeur propre conjuguée $e^{-i\theta}$, et est réductible à la forme diagonale

$$A = \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\
0 & 1 & & & & & & & & & \vdots \\
\vdots & & \ddots & 0 & & & & & & & \vdots \\
\vdots & & & 1 & 0 & & & & & & \vdots \\
\vdots & & & 0 & -1 & & 0 & & & & \vdots \\
\vdots & & & & 0 & \ddots & -1 & 0 & & & \vdots \\
\vdots & & & & & & 0 & e^{i\theta_1} & & & \vdots \\
\vdots & & & & & & & & e^{-i\theta_1} & & \vdots \\
\vdots & & & & & & & & & e^{i\theta_s} & 0 \\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & \cdots & 0 & e^{-i\theta_s}
\end{pmatrix}$$

$$\underbrace{\qquad\qquad}_{\substack{r \\ \text{colonnes}}} \quad \underbrace{\qquad\qquad}_{\substack{n-r-2s \\ \text{colonnes}}} \quad \underbrace{\qquad\qquad\qquad}_{\substack{2s \\ \text{colonnes}}}$$

Le déterminant de φ est encore $(-1)^{n-r}$.

Dans les $2s$ dernières colonnes, on peut, moyennant un changement de base dans chacun des plans tels que $(\vec{u_1}, \vec{\bar{u}_1})$ où $\vec{u_1}$ et $\vec{\bar{u}_1}$ sont des vecteurs propres complexes unitaires relatifs aux valeurs propres $e^{i\theta_1}$ et $e^{-i\theta_1}$, remplacer la sous-matrice

$$\begin{pmatrix} e^{i\theta_1} & 0 \\ 0 & e^{-i\theta_1} \end{pmatrix} \quad \text{par} \quad R_1 = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}$$

qui est la matrice d'une rotation de l'angle θ_1 dans ce plan.

On arrive ainsi à la forme canonique la plus générale, sous forme réelle, d'une isométrie

$$M = \begin{pmatrix}
1 & & & & & & \\
& \ddots & & & & 0 & \\
& & 1 & & & & \\
& & & -1 & & & \\
& & & & \ddots & & \\
& & & & & -1 & \\
& 0 & & & & & R_1 \\
& & & & & & & \ddots \\
& & & & & & & & R_s
\end{pmatrix}$$

où $R_2, R_3, \ldots, R_s$ ont même forme que R_1.

Nous appellerons isométrie directe une isométrie de déterminant égal à 1, isométrie indirecte une isométrie de déterminant égal à -1. On démontre, et nous admettrons, que toute isométrie de dimension n est un produit de symétries en nombre au plus égal à n.

D'autre part, il est immédiat de vérifier que l'ensemble des isométries de E constitue un groupe, dont l'ensemble des isométries directes est un sous-groupe.

6. Isométries de $\mathbb{R}^2$ et $\mathbb{R}^3$

a) Dans $\mathbb{R}^2$, rapporté à une base orthonormée $(\vec{i}, \vec{j})$, les formes canoniques d'une isométrie φ sont les suivantes :

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$ identité; $\quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$; ici $\lambda_1 = \lambda_2 = \pm 1$

$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ symétrie par rapport à la droite D de base $\vec{i}$; ici $\lambda_1 = 1$, $\lambda_2 = -1$

$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ rotation d'angle θ; ici $\lambda_1 = e^{i\theta}$, $\lambda_2 = e^{-i\theta}$

Cette rotation est le produit de deux symétries, par rapport à une droite quelconque D et par rapport à la droite D' qui s'en déduit par rotation de $\dfrac{\theta}{2}$.

b) Dans $\mathbb{R}^3$, rapporté à une base orthonormée $\vec{i}, \vec{j}, \vec{k}$, et en admettant pour θ les valeurs possibles 0 et π, les formes réduites sont les suivantes :

$$M_1 = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{isométrie directe}$$

c'est une rotation d'angle θ autour de l'axe de vecteur unitaire $\vec{k}$;

$$M_2 = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{isométrie indirecte}$$

c'est le produit, commutatif, de la rotation $(\vec{k}, \theta)$ et de la symétrie par rapport au plan $(\vec{i}, \vec{j})$. Tout ceci concerne en principe les espaces vectoriels mais s'adapte facilement aux espaces affines : dans ce dernier cas, l'isométrie directe a une droite invariante point par point, son axe, tandis que l'isométrie indirecte a seulement un point invariant, commun à l'axe de rotation et au plan de symétrie.

Il est important, pratiquement, de déterminer l'axe et l'angle d'une rotation définie dans $\mathbb{R}^3$, par une matrice A, de déterminant égal à 1, dans une base orthonormée quelconque. Le support de l'axe est l'ensemble des vecteurs (ou des points) invariants, défini par $A\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}$.

L'angle θ de rotation est connu, au signe près, par l'invariance de la trace dans un changement de base : $\mathbf{1 + 2\cos\theta = tr\ A = t}$, qui fournit $\cos\theta$. Le sens de l'axe peut être obtenu en cherchant le transformé d'un vecteur $\vec{u}$ perpendiculaire à l'axe, ce sens est celui de $\vec{u} \wedge \varphi(\vec{u})$, et la rotation a pour angle $\theta = \text{Arc}\cos\dfrac{t-1}{2}$.

7. Réduction d'une matrice symétrique réelle. Application aux formes quadratiques

a) Soit E un espace vectoriel euclidien, rapporté à une base orthonormée $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$. Soit Q une forme quadratique de E, de matrice A (symétrique). Soit φ l'endomorphisme de matrice A, u un vecteur propre de φ, relatif à la valeur propre λ, X la matrice-colonne de $\vec{u}$ dans la base $(\vec{e_1}, \ldots, \vec{e_n})$.

$$\varphi(\vec{u}) = \lambda\vec{u} \Leftrightarrow AX = \lambda X$$

On démontre à l'aide du produit scalaire hermitien (chap. suivant) que λ est réelle et que X l'est également.

Ainsi **une matrice symétrique réelle n'a que des valeurs propres réelles.**

En outre, on démontre aussi, et nous admettrons, que **pour toute matrice symétrique réelle il existe une base orthonormée de vecteurs propres :** une telle matrice est donc diagonalisable dans $\mathbb{R}$.

b) Soit D une matrice diagonale exprimant φ dans une base propre orthonormée, on a $A = PDP^{-1}$, P étant une matrice orthogonale, et $X = PY$ donc $Q(\vec{u}) = {}^tXAX = {}^tY{}^tPPDP^{-1}X$.

Or ${}^tP = P^{-1}$ et $P^{-1}X = Y$ donc $Q(\vec{u}) = {}^tYDY$.

Diagonalisant A, on a donc aussi diagonalisé Q; alors que, rappelons-le, diagonaliser une forme quadratique ne signifie pas forcément diagonaliser sa matrice. Si $y_1, y_2, \ldots, y_n$ sont les coordonnées de $\vec{u}$ dans la base propre utilisée, et $\lambda_1, \lambda_2, \ldots, \lambda_n$ les valeurs propres correspondantes de A, on a

$$Q(\vec{u}) = \sum_1^n \lambda_i y_i^2$$

Le rang de Q est égal au nombre de valeurs propres non nulles, et la signature de Q est $(p, r-p)$, p étant le nombre de valeurs propres strictement positives. On voit que la diagonalisation de A fournit pour Q une décomposition en carrés particulière, et qu'une décomposition en carrés quelconque de Q fournit pour A la répartition des signes des valeurs propres.

8. Réduction des coniques et quadriques à centres

L'étude complète de la réduction d'une conique ou d'une quadrique peut être entreprise en partant de l'équation vectorielle

$$Q(\vec{u}) + f(\vec{u}) + k = 0$$

où $\vec{u} = \overrightarrow{OM}$, M étant le point générique de l'ensemble envisagé, O l'origine, Q une forme quadratique et f une forme linéaire. Dans un souci de brièveté, nous nous bornerons ici aux coniques ou quadriques à centres en partant d'une équation de la forme $Q(\vec{u}) = 1$ (le cas où il n'y a pas de terme constant conduirait à deux droites dans $\mathbb{R}^2$, à un cône ou à deux plans dans $\mathbb{R}^3$, et ne pose pas de problème). O est alors centre de symétrie, puisque $Q(-\vec{u}) = Q(\vec{u})$.

a) Soit, dans $\mathbb{R}^2$, la conique d'équation $Q(\vec{u}) = 1$. On sait diagonaliser Q sur une base propre orthonormée et écrire l'équation sous la forme

$$\lambda_1 x^2 + \lambda_2 y^2 = 1$$

λ_1 et λ_2 étant les valeurs propres de la matrice Q, et la discussion est immédiate :

$\lambda_1 > 0, \lambda_2 > 0$: ellipse

λ_1 et λ_2 de signes contraires : hyperbole

$\lambda_1 \leqslant 0, \lambda_2 \leqslant 0$: aucun point réel

λ_1 ou λ_2 nul, l'autre étant positif : deux droites parallèles

b) Soit, dans $\mathbb{R}^3$, la quadrique d'équation $Q(\vec{u}) = 1$. Si on sait calculer effectivement les trois valeurs propres de la matrice A de Q, on peut diagonaliser Q sur une base propre orthonormée, l'équation s'écrit alors

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 = 1 \tag{6}$$

Même si on ne sait pas calculer λ_1, λ_2, λ_3, on connaît leurs signes par la signature de Q, et une décomposition en carrés de Q donne une équation de la forme (6) où le repère n'est pas orthonormé et où les λ_i ne sont pas les valeurs propres.

Dans tous les cas on peut discuter la nature de la quadrique (en supposant $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$)

λ_1, λ_2, λ_3 positifs	signature (3, 0) : ellipsoïde
λ_1, λ_2, positifs, λ_3 négatif	signature (2, 1) : hyperboloïde à une nappe
λ_1 positif, λ_2 et λ_3 négatifs	signature (1, 2) : hyperboloïde à deux nappes
$\lambda_1 > 0, \lambda_2 > 0, \lambda_3 = 0$	signature (2, 0) : cylindre elliptique

$\lambda_1 > 0, \lambda_2 = 0, \lambda_3 < 0$ signature (1, 1) : cylindre hyperbolique

$\lambda_1 > 0, \lambda_2 = \lambda_3 = 0$ signature (1, 0) : deux plans parallèles

Les autres cas donnent des surfaces sans point réel.

9. Produit mixte et produit vectoriel dans un espace euclidien de dimension *n*

En généralisant les notions de $\mathbb{R}^3$, nous appellerons, en dimension n :

a) produit mixte de n vecteurs $\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}$ leur déterminant dans un repère orthonormé : on le note $(\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n})$.

b) produit vectoriel de $n - 1$ vecteurs $\vec{u_1}, \vec{u_2}, \ldots, \overrightarrow{u_{n-1}}$ le vecteur unique $\vec{v}$ tel que, pour tout vecteur $\vec{w}$

$$(\vec{w}, \vec{u_1}, \vec{u_2}, \ldots, \ldots, \overrightarrow{u_{n-1}}) = \vec{w} \cdot \vec{v}$$

On le note : $\vec{v} = \vec{u_1} \wedge \vec{u_2} \wedge \ldots \wedge \overrightarrow{u_{n-1}}$.

Ses coordonnées sont les cofacteurs des éléments de la première colonne d'un déterminant où les $n - 1$ colonnes suivantes sont constituées par les coordonnées de $\vec{u_1}, \vec{u_2}, \ldots, \overrightarrow{u_{n-1}}$, dans un repère orthonormé direct (c'est-à-dire de produit mixte égal à 1).

Résumé du chapitre 9

- **Forme bilinéaire de $E \times E$ dans le corps K de référence de E.**

$(\vec{u}, \vec{v}) \to \varphi(\vec{u}, \vec{v})$ telle que
$$\varphi(\vec{u_1} + \vec{u_2}, \vec{v}) = \varphi(\vec{u_1}, \vec{v}) + \varphi(\vec{u_2}, \vec{v})$$
$$\varphi(\vec{u}, \vec{v_1} + \vec{v_2}) = \varphi(\vec{u}, \vec{v_1}) + \varphi(\vec{u}, \vec{v_2})$$
$$\varphi(\lambda\vec{u}, \vec{v}) = \varphi(\vec{u}, \lambda\vec{v}) = \lambda\varphi(\vec{u}, \vec{v})$$

Il s'ensuit que si, dans une base de E, $\vec{u}$ et $\vec{v}$ ont pour matrices X et Y

$$\varphi(\vec{u}, \vec{v}) = {}^tXAY$$

A est la matrice de φ dans la base utilisée.

- Un **changement de base** se traduit par $A' = {}^tPAP$.

- La **forme bilinéaire** φ est **symétrique** si $\varphi(\vec{v}, \vec{u}) = \varphi(\vec{u}, \vec{v})$.

  Sa matrice est alors symétrique ${}^tA = A$.

- La **forme quadratique associée** à φ est

$$Q(\vec{u}) = \varphi(\vec{u}, \vec{u}) = {}^tXAX$$

où A est une matrice symétrique.

On a $$Q(\lambda\vec{u}) = \lambda^2 Q(\vec{u})$$
$$\varphi(\vec{u}, \vec{v}) = \frac{1}{2}[Q(\vec{u} + \vec{v}) - Q(\vec{u}) - Q(\vec{v})]$$

φ est la **forme polaire** de Q.

La matrice A d'une forme quadratique $Q(x_1, x_2, \ldots, x_n)$ est celle des n formes linéaires $\frac{1}{2}\frac{\partial Q}{\partial x_i}$.

- Deux **vecteurs** $\vec{u}$ et $\vec{v}$ sont **orthogonaux par rapport à la forme quadratique** Q si $\varphi(\vec{u}, \vec{v}) = 0$.

- **Base orthogonale.**

Pour toute forme quadratique Q, il existe au moins une base orthogonale, et dans cette base $Q(\vec{u})$ s'écrit

$$Q(\vec{u}) = \sum_1^r \alpha_i y_i^2 \qquad 1 \leqslant r \leqslant n$$

les y_i étant des formes linéaires, indépendantes, des anciennes coordonnées x_j.

On a aussi
$$Q(\vec{u}) = z_1^2 + z_2^2 + \ldots + z_p^2 - z_{p+1}^2 - z_{p+2}^2 - \ldots - z_r^2$$
(les z_i étant r formes linéaires indépendantes).

Q est ainsi **décomposée en carrés**. r est le rang de Q et sa signature est $(p, r - p)$; p et r sont indépendants de la base orthogonale utilisée. r est aussi le rang de A.

- **Décomposition en carrés par la méthode de Gauss.**

On utilise les relations
$$x_1^2 + 2x_1 f(x_2, \ldots, x_n) = (x_1 + f)^2 - f^2$$
ou
$$x_1x_2 + x_1 f(x_3, \ldots, x_n) + x_2 g(x_3, \ldots, x_n) = (x_1 + g)(x_2 + f) - fg$$

- **Espace euclidien.**

E est un espace euclidien s'il est muni d'une forme quadratique définie positive (telle que $p = r = n$), la forme polaire s'appelant alors produit scalaire
$$\varphi(\vec{u}, \vec{v}) = (\vec{u} \mid \vec{v}) = \vec{u} \cdot \vec{v}$$

La norme associée à ce produit scalaire est $\|\vec{u}\| = \sqrt{\vec{u} \mid \vec{u}}$.

- **Inégalité de Schwarz.**

$$|(\vec{u} \mid \vec{v})| \leqslant \|\vec{u}\| \, \|\vec{v}\|$$

- **Inégalité triangulaire.**

$$\|\vec{u} + \vec{v}\| \leqslant \|\vec{u}\| + \|\vec{v}\|$$

- **$\vec{u}$ et $\vec{v}$ sont orthogonaux si $\vec{u} \cdot \vec{v} = 0$.**

- **Base orthonormée.**

Il existe dans l'espace euclidien E au moins une base $(\vec{e_1}, \ldots, \vec{e_n})$ telle que $\vec{e_i} \cdot \vec{e_j} = \delta_{ij}$ (symbole de Kronecker), une telle base est dite orthonormée.

Dans une telle base

$$\vec{u} \cdot \vec{v} = {}^tXY = \sum_1^n x_i y_i$$

$$\|\vec{u}\|^2 = \sum_1^n x_i^2$$

- **Procédé d'orthonormalisation de Schmidt.**

Voir cours détaillé.

- **Matrice orthogonale.**

A est orthogonale si $A^{-1} = {}^tA$.
C'est la matrice d'un changement de base orthonormée dans E.

- **Isométrie.**

On appelle isométrie ou matrice orthogonale un endomorphisme qui conserve le produit scalaire (et par suite aussi la norme).

La matrice d'une isométrie est orthogonale. Ses valeurs propres ont pour module 1, son déterminant est égal à ± 1.

Une isométrie est réductible à la forme canonique

$$M = \begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & R_1 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & R_s \end{pmatrix}$$

les matrices R_k étant de la forme $\begin{pmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{pmatrix}$.

• **Dans $\mathbb{R}^2$, une isométrie est une symétrie par rapport à une droite ou une rotation.**

• **Dans $\mathbb{R}^3$, une isométrie est soit une rotation, soit le produit d'une rotation et d'une symétrie par rapport à un plan.** Dans le cas d'une rotation, l'angle de rotation θ est tel que $1 + 2\cos\theta$ est égal à la trace de la matrice de la rotation.

• **Réduction d'une matrice symétrique réelle et d'une forme quadratique.**

Une matrice symétrique réelle a toutes ses valeurs propres réelles et on peut déterminer pour cette matrice une base propre orthonormée, qui permet de diagonaliser la forme quadratique associée.

Ceci s'applique en particulier à la réduction des coniques et quadriques à centre.

EXERCICES

■ **1.** Décomposer en carrés la forme quadratique

$$Q(\vec{u}) = Q(x, y, z, t) = x^2 + y^2 + z^2 - 2t^2 - 2xy + 2xz - 2xt + 2yz - 4yt$$

La méthode de Gauss donne

$$Q(\vec{u}) = (x - y + z - t)^2 - 3\left(t + y - \frac{z}{3}\right)^2 + 3\left(y + \frac{z}{3}\right)^2$$

La forme est de rang 3, sa signature est (2, 1).

■ **2.** Montrer que l'application $(P, Q) \to \displaystyle\int_0^{+\infty} P(x)\,Q(x)\,e^{-x}\,\mathrm{d}x = (P\,|\,Q)$ définit un produit scalaire sur l'espace vectoriel des polynômes à une variable.

Trouver une base orthonormée du sous-espace constitué par les polynômes de degré inférieur ou égal à 2.

---

a) L'intégrale ci-dessus est convergente quels que soient les polynômes P et Q, car pour x infini $|P(x)\,Q(x)\,e^{-x}| \sim ax^n e^{-x} < \frac{1}{x^2}$. L'application est donc définie.

C'est bien une forme bilinéaire symétrique car, l'intégration étant linéaire,

$$(P_1 + P_2 \mid Q) = (P_1 \mid Q) + (P_2 \mid Q)$$
$$(P \mid Q_1 + Q_2) = (P \mid Q_1) + (P \mid Q_2)$$
$$(\lambda P \mid Q) = (P \mid \lambda Q) = \lambda\,(P \mid Q)$$

et de plus $$(P \mid Q) = (Q \mid P)$$

En outre $(P \mid P) = \int_0^{+\infty} P^2(x)\,e^{-x}\,dx > 0 \qquad \forall\, P \neq 0.$

On a donc bien affaire à un produit scalaire.

b) Appliquons la méthode d'orthonormalisation de Schmidt en partant de la base canonique $P_0 = 1$, $P_1 = x$, $P_2 = x^2$, et cherchant une base Q_0, Q_1, Q_2 constituée de polynômes deux à deux orthogonaux, que l'on pourra facilement normer.

Prenons $Q_0 = P_0$ et cherchons α tel que $Q_1 = P_1 + \alpha P_0 = x + \alpha$ on veut

$$(Q_0 \mid Q_1) = \int_0^{+\infty} (x + \alpha)\,e^{-x}\,dx = 0$$

soit

$$[-(x + \alpha)\,e^{-x}]_0^{+\infty} + \int_0^{+\infty} e^{-x}\,dx = 0$$

d'où $\alpha + 1 = 0$, $\alpha = -1$, $Q_1 = x - 1$.

Cherchons $Q_2 = P_2 + \beta Q_1 + \gamma Q_0$ on veut

$$(Q_2 \mid Q_0) = (P_2 \mid Q_0) + \beta\,(Q_1 \mid Q_0) + \gamma\,(Q_0 \mid Q_0) = 0$$

Or $(Q_1 \mid Q_0) = 0$ donc

$$\int_0^{+\infty} x^2\,e^{-x}\,dx + \gamma \int_0^{+\infty} e^{-x}\,dx = 0$$

$$[-x^2 e^{-x}]_0^{+\infty} + \int_0^{+\infty} 2\,xe^{-x}\,dx + \gamma = 0$$

$$2\,[-xe^{-x}]_0^{+\infty} + 2\int_0^{+\infty} e^{-x}\,dx + \gamma = 0 \qquad \gamma = -2$$

On veut aussi

$$(Q_2 \mid Q_1) = (P_2 \mid Q_1) + \beta\,(Q_1 \mid Q_1) + \gamma\,(Q_0 \mid Q_1) = 0$$

donc

$$\int_0^{+\infty} x^2\,(x - 1)\,e^{-x}\,dx + \beta \int_0^{+\infty} (x - 1)^2\,e^{-x}\,dx = 0$$

De nouvelles intégrations par parties nous donnent $\beta = -4$, d'où

$$Q_2 = x^2 - 4x - 2$$

Nous avons une base orthogonale, il reste à la normer en prenant

$$U_k = \frac{Q_k}{\| Q_k \|}, \qquad k = 0, 1, 2$$

Calculons

$$\| Q_0 \|^2 = \int_0^{+\infty} e^{-x}\,\mathrm{d}x = 1 \qquad \text{il vient} \quad U_0 = Q_0 = P_0 = 1$$

$$\| Q_1 \|^2 = \int_0^{+\infty} (x-1)^2 e^{-x}\,\mathrm{d}x = 1 \qquad \text{il vient} \quad U_1 = Q_1 = x - 1$$

$$\| Q^2 \|^2 = \int_0^{+\infty} (x^2 - 4x - 2)^2 e^{-x}\,\mathrm{d}x = 68 \quad U_2 = \frac{Q_2}{\sqrt{68}} = \frac{x^2 - 4x - 2}{2\sqrt{17}}$$

La base orthonormée obtenue est donc

$$1, \quad x - 1, \quad \frac{x^2 - 4x - 2}{2\sqrt{17}}$$

■ **3.** Dans $\mathbb{R}^4$, supposé rapporté à une base orthonormée, déterminer la projection orthogonale du vecteur $\vec{u}\,(2, 3, -1, -4)$ sur le plan défini par les vecteurs $\vec{a}\,(2, 1, -1, 1)$, $\vec{b}\,(3, 1, 1, 0)$.

Il suffit d'écrire $\vec{u} = \alpha\vec{a} + \beta\vec{b} + \vec{v}$, $\vec{v}$ étant le vecteur (x, y, z, t), et d'exprimer $\vec{a}\,.\,\vec{v} = \vec{b}\,.\,\vec{v} = 0$ on obtient un système de 6 équations à 6 inconnues dans lequel on élimine x, y, z, t, d'où un système de Cramer en α, β. On trouve $\alpha = -\dfrac{4}{41}$, $\beta = \dfrac{32}{41}$ et la projection demandée est

$$\vec{w} = \alpha\vec{a} + \beta\vec{b} = \frac{1}{41}(88, 28, 36, -4)$$

■ **4.** Comparer une matrice orthogonale à sa comatrice.

A, orthogonale, est inversible et $A^{-1} = \dfrac{{}^t\mathrm{com}\,A}{\det A} = {}^tA$ or $\det A = \pm 1$

donc $\mathrm{com}\,A = A$ si A est orthogonale directe,

$\mathrm{com}\,A = -A$ si A est orthogonale indirecte.

■ 5. Caractériser, dans un espace euclidien de dimension n, l'endomorphisme φ défini dans une base orthonormée par la matrice

$$M = \begin{pmatrix} 0 & \dots & 0 & 1 \\ \vdots & & 1 & 0 \\ 0 & & & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}$$

Réponse : φ est une isométrie car M est orthogonale.

$\Delta = \text{dét}\, M = (-1)^{\frac{(n+1)(n+2)}{2}} = 1$ si $n \equiv 2$ ou $n \equiv 3$ (mod. 4) (isométrie directe).

$\Delta = -1$ si $n \equiv 0$ ou $n \equiv 1$ (mod. 4) (isométrie indirecte).

De plus $M^2 = I$, les valeurs propres sont 1 et -1. Leurs ordres de multiplicité sont les dimensions des sous-espaces propres associés, puisque φ est diagonalisable. Le lecteur achèvera facilement cette discussion.

■ 6. E étant un espace vectoriel de dimension finie n, de base $\mathscr{B}$, démontrer que $N(\vec{u}) = \sum_{1}^{n} |x_i|$ où $(x_1, x_2, \dots, x_n)$ sont les coordonnées de $\vec{u}$, définit une norme sur E.

Il suffit de vérifier

$$N(\lambda\vec{u}) = |\lambda|\, N(\vec{u})$$
$$N(\vec{u}) = 0 \Rightarrow \vec{u} = \vec{0}$$
$$N(\vec{u} + \vec{v}) \leqslant N(\vec{u}) + N(\vec{v})$$

ce qui est immédiat ici.

■ 7. Soit $\mathscr{M}$ l'ensemble des matrices carrées réelles d'ordre n, $A = \begin{pmatrix} & \vdots & \\ \dots & a_{ij} & \dots \\ & \vdots & \end{pmatrix}$ une de ces matrices.

a) Montrer que $N(A) = \sup |a_{ij}|$ définit une norme sur $\mathscr{M}$.

b) Établir une inégalité liant $N(AB)$, $N(A)$, $N(B)$.

c) Montrer que la série $\sum_{0}^{+\infty} \frac{A^p}{p\,!}$ (ou $A^0 = I$) est convergente. Elle définit $\exp A$.

d) En admettant pour le produit de deux séries de matrices $\sum_{0}^{+\infty} U_p$ et $\sum_{0}^{+\infty} V_p$ supposées convergentes en normes (c'est-à-dire $\sum_{0}^{+\infty} N(U_p)$

et $\sum_{0}^{+\infty} N(V_p)$ convergentes), la même loi de formation que pour le produit de deux séries à termes réels absolument convergentes, démontrer que ${}^tA = -A \Rightarrow \exp A$ est orthogonale.

Réponse :

a) Vérification immédiate des trois conditions rappelées dans l'exercice 6.

b) $N(AB) \leqslant n\, N(A)\, N(B)$.

c) $N(A^p) \leqslant n^{p-1}\, [N(A)]^p$ chaque terme de la matrice $\dfrac{A^p}{p\,!}$ est majoré en module par le terme général de la série convergente

$$\sum_{p=0}^{+\infty} \frac{n^{p-1}\,[N(A)]^p}{p\,!}$$

donc $\exp A$ existe.

d) $(\exp A)(\exp {}^tA) = (\exp A)[\exp(-A)] = \exp 0 = I$
(se démontre par la même méthode que $e^{z_1}\, e^{z_2} = e^{z_1 + z_2}$).

Or $\exp {}^tA = {}^t(\exp A)$ donc ${}^t(\exp A) = (\exp A)^{-1}$ et $\exp A$ est orthogonale.

■ **8.** *a*) Étudier le rang d'une forme quadratique produit de deux formes linéaires sur un espace vectoriel E de dimension n.

b) A et B étant deux matrices-colonnes réelles à n éléments, étudier le rang de la matrice $C = A\,.\,{}^tB + B\,.\,{}^tA$.

---

a) $Q(\vec{u}) = \varphi(\vec{u})\,\psi(\vec{u}) = \dfrac{1}{4}[\varphi(\vec{u}) + \psi(\vec{u})]^2 - \dfrac{1}{4}[\varphi(\vec{u}) - \psi(\vec{u})]^2$

est de rang 2 si φ et ψ sont indépendantes,
de rang 1 si elles sont dépendantes non nulles,
de rang 0 si φ ou ψ est nulle.

b) C est symétrique et on vérifie qu'elle est la matrice d'une forme quadratique du type précédent, qui a même rang que C. Donc C est de rang 2 si A et B sont indépendantes, de rang 1 si elles sont dépendantes non nulles, de rang 0 si l'une est nulle.

■ **9.** Soit, dans un espace vectoriel euclidien rapporté à une base orthonormée $(\vec{e_1}, \ldots, \vec{e_n})$, où un vecteur $\vec{u}$ a pour coordonnées $x_1, x_2, \ldots, x_n$, la forme quadratique

$$Q(\vec{u}) = \sum_{1 \leqslant i < j \leqslant n} (x_i - x_j)^2$$

Déterminer pour Q une base orthonormée de vecteurs propres.

Q a pour matrice

$$A = \begin{pmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ -1 & \cdots & -1 & n-1 \end{pmatrix} \quad (a_{ii} = n-1 \,;\, a_{ij} = -1 \text{ si } i \neq j)$$

dont le polynôme caractéristique est $(-1)^n \lambda (\lambda - n)^{n-1}$.

Les valeurs propres sont
$\lambda_1 = 0$, de sous-espace associé E_1 $(x_1 = x_2 = \ldots = x_n)$, de dimension 1,
$\lambda_2 = n$, de sous-espace associé $E_2 \left(\sum_1^n x_i = 0\right)$, de dimension $n-1$.

On a $E_2 \perp E_1$, une base propre orthonormée est constituée de $\overrightarrow{u_1}\left(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right)$ et d'une base orthonormée de E_2; elle-même formée des $n-1$ vecteurs

$$\overrightarrow{u_{k+1}}\left(x_1 = x_2 = \ldots = x_k = \frac{1}{\sqrt{k(k+1)}}, \quad x_{k+1} = -\frac{k}{\sqrt{k(k+1)}}, \quad x_{k+2} = \ldots = x_n = 0\right)$$

avec $k = 1, 2, \ldots, n-1$.

Si $y_1, y_2, \ldots, y_n$ sont les coordonnées de $\vec{u}$ dans cette base, Q s'exprime par

$$Q(\vec{u}) = n \sum_2^n y_k^2$$

le rang de Q est $n-1$, sa signature est $(n-1, 0)$.

10. Déterminer les axes et l'équation réduite de la quadrique définie dans $\mathbb{R}^3$, par rapport à un repère orthonormé, par l'équation

$$x^2 + y^2 + z^2 + 4xy + 4xz - 4yz - 3 = 0$$

Si on écrit cette équation sous la forme $Q(x, y, z) = 3$, la matrice de Q est celle des formes linéaires $\frac{1}{2}\frac{\partial Q}{\partial x}, \frac{1}{2}\frac{\partial Q}{\partial y}, \frac{1}{2}\frac{\partial Q}{\partial z}$, soit

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

ses valeurs propres sont $\lambda_1 = \lambda_2 = 3$, $\lambda_3 = -3$.

On peut prendre pour base propre orthonormée directe

$$\vec{u_1}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \quad \vec{u_2}\left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right), \quad \vec{u_3}\left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$$

Dans cette base, l'origine étant conservée, l'équation réduite est

$$X^2 + Y^2 - Z^2 = 1$$

la quadrique est un hyperboloïde de révolution à une nappe, ses axes sont OX, OY, OZ de vecteurs unitaires $\vec{u_1}$, $\vec{u_2}$, $\vec{u_3}$.

■ **11.** Montrer que l'aire d'une ellipse et le volume d'un ellipsoïde peuvent être calculés sans réduction de l'équation, en utilisant dans un repère orthonormé l'équation donnée sous la forme $Q(\vec{u}) = 1$ où Q est une forme quadratique à 2 ou 3 variables. Application numérique :

$$x^2 + 3y^2 + 10z^2 + 2xy - 2xz - 10yz = 1$$

L'aire de l'ellipse d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ est $S = \pi ab$.

Le volume de l'ellipsoïde d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ est

$$V = \frac{4}{3}\pi abc$$

$Q(\vec{u}) = 1$ représente une ellipse ou un ellipsoïde si Q est définie positive, l'équation réduite est

$$\lambda_1 x^2 + \lambda_2 y^2 = 1$$

ou

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 = 1$$

les λ_i étant les valeurs propres, qu'il n'est pas nécessaire de calculer ; on aura

$$S = \frac{\pi}{\sqrt{\Delta}}, \qquad V = \frac{4}{3}\frac{\pi}{\sqrt{\Delta}}.$$

Δ étant le déterminant de la matrice A de Q.

Dans l'application numérique, il suffit de décomposer en carrés la forme quadratique, par exemple sous la forme

$$(x + y - z)^2 + 2(y - 2z)^2 + z^2$$

ce qui montre qu'il s'agit bien d'un ellipsoïde, et de calculer son déterminant.

Le volume est $V = \dfrac{2\pi\sqrt{2}}{3}$.

■ **12.** Calculer, le repère étant orthonormé et en supposant $b^2 - ac < 0$, l'intégrale double

$$I = \iint_{\mathbb{R}^2} e^{-ax^2-2bxy-cy^2}\,\mathrm{d}x\mathrm{d}y$$

Par changement de repère orthonormé et passage en coordonnées polaires, on obtient

$$I = \frac{\pi}{\sqrt{ac-b^2}}$$

■ **13.** Établir que la matrice $M = \dfrac{1}{3}\begin{pmatrix} -2 & 1 & -2 \\ 2 & 2 & -1 \\ 1 & -2 & -2 \end{pmatrix}$ est celle d'une rotation, dont on déterminera l'angle et l'axe.

M est orthogonale et son déterminant vaut 1, elle définit bien une rotation R. Son angle θ est donné par

$$1 + 2\cos\theta = \operatorname{tr} M = -\frac{2}{3}, \quad \cos\theta = -\frac{5}{6}$$

Prenons $\theta = \operatorname{Arc\,cos}\left(-\dfrac{5}{6}\right)$, soit $\theta \simeq 146{,}44°$.

L'axe de rotation, au sens près, est la droite ensemble des vecteurs invariants par R, sa base est $\vec{u}\ (1, 3, -1)$. Le sens de l'axe est celui du vecteur $\vec{v} \wedge R(\vec{v})$, où $\vec{v}$ est un vecteur orthogonal à $\vec{u}$, par exemple $\vec{v}\ (3, -1, 0)$. On obtient $\vec{v} \wedge R(\vec{v}) = \vec{w}\left(-\dfrac{5}{3}, -5, \dfrac{5}{3}\right)$, l'axe $\vec{D}$ a donc pour vecteur directeur $-\vec{u}$.

■ **14.** Montrer que si A est une matrice orthogonale n'admettant pas -1 pour valeur propre, la relation $A = (I+B)^{-1}(I-B)$ définit une matrice B antisymétrique.

Il suffit de résoudre la relation par rapport à B et de vérifier ${}^tB = -B$. On aura à établir que $I - A$ et $(I+A)^{-1}$ commutent, ce qui est facile en écrivant $I - A = 2I - (I+A)$.

■ **15.** a, b, c, d étant quatre réels tels que $a^2 + b^2 + c^2 + d^2 = 1$, réduire dans un espace euclidien de dimension 4 l'endomorphisme φ de matrice

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

φ est une isométrie directe. Ses valeurs propres sont données par

$$\text{dét}\,(M - \lambda I) = [(a - \lambda)^2 + b^2 + c^2 + d^2]^2 = 0$$

soit deux valeurs propres doubles complexes

$$\lambda_1 = a + i\sqrt{b^2 + c^2 + d^2} \qquad \lambda_2 = a - i\sqrt{b^2 + c^2 + d^2}$$

En posant $a = \cos\theta$, $\sqrt{b^2 + c^2 + d^2} = \sin\theta$, φ est diagonalisable dans une base $(\vec{u_1}, \vec{u_2}, \vec{u_3}, \vec{u_4})$

$$D = \begin{pmatrix} e^{i\theta} & 0 & 0 & 0 \\ 0 & e^{-i\theta} & 0 & 0 \\ 0 & 0 & e^{i\theta} & 0 \\ 0 & 0 & 0 & e^{-i\theta} \end{pmatrix}$$

et par changement de base dans chacun des sous-espaces $(\vec{u_1}, \vec{u_2})$ et $(\vec{u_3}, \vec{u_4})$, φ peut aussi être défini par la matrice

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}$$

CHAPITRE 10

FORCES HERMITIENNES. MATRICES HERMITIENNES

Dans ce chapitre, nous introduisons brièvement des notions comparables à celles du chapitre précédent, mais en raisonnant cette fois systématiquement dans un espace vectoriel sur $\mathbb{C}$: les matrices considérées seront, sauf exception, à éléments complexes. Ce chapitre étant très bref, nous n'en donnerons pas de résumé.

1. FORMES SESQUILINÉAIRES. FORMES HERMITIENNES

a) E étant un espace vectoriel sur $\mathbb{C}$, une **forme sesquilinéaire** de E est une application $(\vec{u}, \vec{v}) \to \varphi(\vec{u}, \vec{v})$ satisfaisant aux conditions suivantes, où $\vec{u}$, $\vec{v}$, $\vec{w}$ sont des vecteurs quelconques et λ un scalaire quelconque :

$$(1)\quad \begin{cases} \varphi(\vec{u} + \vec{v}, \vec{w}) = \varphi(\vec{u}, \vec{w}) + \varphi(\vec{v}, \vec{w}) \\ \varphi(\vec{u}, \vec{v} + \vec{w}) = \varphi(\vec{u}, \vec{v}) + \varphi(\vec{u}, \vec{w}) \\ \quad \varphi(\lambda\vec{u}, \vec{v}) = \lambda\varphi(\vec{u}, \vec{v}) & \text{(linéarité à gauche)} \\ \quad \varphi(\vec{u}, \lambda\vec{v}) = \bar{\lambda}\varphi(\vec{u}, \vec{v}) & \text{(semi-linéarité à droite)} \end{cases}$$

C'est ce caractère « une fois et demie » linéaire qu'évoque, étymologiquement, le mot « sesquilinéaire ».

b) Une forme sesquilinéaire est dite **hermitienne** si elle satisfait en outre à la condition suivante, pour tous $\vec{u}, \vec{v}$:

$$\boxed{\varphi(\vec{v}, \vec{u}) = \overline{\varphi(\vec{u}, \vec{v})}} \qquad (2)$$

On voit que les formes bilinéaires symétriques sont des cas particuliers des formes hermitiennes.

Exemple de forme hermitienne.

Sur l'espace des fonctions complexes de la variable réelle t, intégrables sur $[0, 1]$, l'application

$$(f, g) \to \int_0^1 f(t)\,\overline{g(t)}\,\mathrm{d}t$$

est une forme hermitienne.

2. MATRICE D'UNE FORME HERMITIENNE EN DIMENSION FINIE

L'application des propriétés (1) et (2) montre que, dans une base $(\vec{e_1}, \ldots, \vec{e_n})$, une forme hermitienne s'écrit :

$$\varphi(\vec{u}, \vec{v}) = \sum_{i,j} x_i \bar{y}_j \varphi(\vec{e_i}, \vec{e_j})$$

si $$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

En posant $\varphi(e_i, e_j) = a_{ij}$ et en appelant A la matrice des a_{ij}, on a donc

$$\boxed{\varphi(\vec{u}, \vec{v}) = {}^tXA\bar{Y}} \tag{3}$$

A est la matrice de φ.

Le fait que $\varphi(\vec{e_j}, \vec{e_i}) = \overline{\varphi(\vec{e_i}, \vec{e_j})}$ montre que $a_{ji} = \overline{a_{ij}}$ donc

$$\boxed{{}^tA = \bar{A}} \tag{4}$$

Une matrice A vérifiant cette relation est une **matrice hermitienne.** Comme $a_{ii} = \overline{a_{ii}}$ les éléments diagonaux de A sont réels, par exemple

$$A = \begin{pmatrix} 3 & i & 1+i \\ -i & -1 & 1 \\ 1-i & 1 & 2 \end{pmatrix}$$

est une matrice hermitienne d'ordre 3.

Une matrice symétrique réelle est un cas particulier d'une matrice hermitienne. Toute matrice hermitienne définit, par la formule (3), une forme hermitienne. On vérifie aisément qu'un changement de base conduit à la nouvelle matrice

$$\boxed{A' = {}^tPA\bar{P}}$$

Remarquons que, sauf si les y_i sont réels, $\varphi(\vec{u}, \vec{v})$ n'est pas un polynôme des variables $x_1, \ldots, x_n, y_1, \ldots, y_n$. Enfin $\varphi(\vec{u}, \vec{u}) = \overline{\varphi(\vec{u}, \vec{u})}$ donc $\varphi(\vec{u}, \vec{u})$ est réel. Bien que ce ne soit pas un polynôme en $x_1, x_2, \ldots, x_n$, on l'appelle souvent forme quadratique hermitienne associée à φ.

En posant

$$\varphi(\vec{u}, \vec{u}) = Q(\vec{u})$$

on a

$$Q(\vec{u}) = \sum_{i,j} a_{ij}\, x_i\, \overline{x_j} = {}^tXA\overline{X}$$

et on pourra facilement vérifier que

$$\varphi(\vec{u}, \vec{v}) = \frac{1}{4}[Q(\vec{u} + \vec{v}) - Q(\vec{u} - \vec{v}) + iQ(\vec{u} + i\vec{v}) - iQ(\vec{u} - i\vec{v})]$$

3. PRODUIT SCALAIRE HERMITIEN. ORTHOGONALITÉ

a) Une forme hermitienne est dite définie positive si

$$\boxed{\vec{u} \neq \vec{0} \Rightarrow Q(\vec{u}) = \varphi(\vec{u}, \vec{u}) > 0}$$

Un espace sur $\mathbb{C}$, muni d'une forme hermitienne définie positive, est dit **préhilbertien** (ou **hermitien**), et la forme hermitienne associée est dite produit scalaire hermitien (ou simplement, produit scalaire) : on le note $(\vec{u} \mid \vec{v})$ ou $\vec{u} \cdot \vec{v}$.

b) Sur un espace préhilbertien, deux vecteurs $\vec{u}$, $\vec{v}$ sont orthogonaux si $\vec{u} \cdot \vec{v} = 0$. Comme dans un espace euclidien (qui n'est autre que le « cas réel » d'un espace préhilbertien), à tout sous-espace F de E, on peut associer le sous-espace orthogonal $F^\perp$ (défini comme dans un espace euclidien) et, en dimension finie, $F^\perp$ est un supplémentaire de F.

c) On démontre, et nous admettrons, qu'il existe dans un espace préhilbertien au moins une **base orthonormée** $(\vec{e_1}, \vec{e_2}, \ldots, \vec{e_n})$ telle que que

$$\vec{e_i} \cdot \vec{e_j} = \delta_{ij}$$

symbole de Kronecker.

Exprimé dans une telle base, le produit scalaire de $\vec{u}$ $(x_1, \ldots, x_n)$ et $\vec{v}$ $(y_1, \ldots, y_n)$ est

$$\boxed{\vec{u} \cdot \vec{v} = \sum_1^n x_i\, \overline{y_i} \quad = {}^tX\overline{Y}}$$

d) Enfin $N(\vec{u}) = \sqrt{\vec{u}\,.\,\vec{u}}$ définit une norme sur l'espace préhilbertien E. En effet

$$\vec{u}\,.\,\vec{u} = \sum_1^n x_i \overline{x_i} = \sum_1^n |x_i|^2 \geqslant 0$$

donc $N(\vec{u})$ est un réel positif, avec $N(\vec{u}) = 0 \Rightarrow \vec{u} = \vec{0}$

$$[N(\lambda\vec{u})] = \sum_1^n \lambda\bar{\lambda}\,|x_i|^2 = |\lambda|^2 \sum_1^n |x_i|^2$$

donc
$$N(\lambda\vec{u}) = |\lambda|\,N(\vec{u})$$

Nous admettrons l'inégalité triangulaire $N(\vec{u} + \vec{v}) \leqslant N(\vec{u}) + N(\vec{v})$. La norme ainsi définie est dite norme hermitienne.

4. MATRICE UNITAIRE, ENDOMORPHISME UNITAIRE

a) La matrice P d'une base orthonormée par rapport à une autre base orthonormée est telle que $\sum_{i=1}^{i=n} a_{ij}\,\overline{a_{ik}} = \delta_{jk}$ donc

$${}^tP\,.\,\overline{P} = I \qquad \text{et} \qquad \boxed{P^{-1} = {}^t\overline{P}}$$

Une telle matrice, qui admet pour inverse sa « transconjuguée » ou « adjointe » ${}^t\overline{P}$ est dite unitaire. Comme dét $({}^tP)$ dét $(\overline{P}) = 1$, on a $|\text{dét } P| = 1$.

b) Un endomorphisme f, dont la matrice dans une base orthonormée de E est unitaire, est appelée endomorphisme unitaire. Cette définition équivaut à la suivante : un endomorphisme unitaire conserve le produit scalaire hermitien, donc aussi la norme hermitienne.

On en déduit aisément que les valeurs propres d'un endomorphisme unitaire, donc aussi d'une matrice unitaire, ont pour module 1. De plus on démontre, et nous admettrons, qu'un endomorphisme unitaire est diagonalisable et qu'il existe une base orthonormée de vecteurs propres. Ces résultats renferment, comme cas particuliers, ceux concernant les matrices orthogonales.

5. RÉDUCTION D'UNE MATRICE HERMITIENNE

Soit λ une valeur propre de la matrice hermitienne A : il existe une matrice-colonne X telle que

$$AX = \lambda X \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

donc $$\bar{A}\bar{X} = \bar{\lambda}\bar{X}$$

Soit la matrice $M = {}^t\bar{X}AX$ réduite à un seul élément α. On a

$$M = {}^t\bar{X}\lambda X = \lambda{}^t\bar{X}X$$

or $${}^t\bar{M} = {}^t({}^tX\bar{A}\bar{X}) = {}^t\bar{X}\,{}^t\bar{A}X = {}^t\bar{X}AX = M$$

(car ${}^t\bar{A} = A$). M, réduite à l'élément α, est donc réelle. α l'est aussi, et comme ${}^t\bar{X}X = \sum_1^n x_i\bar{x_i} = \sum_1^n |x_i|^2 = \beta$, réel positif, on a $\lambda = \dfrac{\alpha}{\beta}$, nombre réel.

Il en résulte que **les valeurs propres d'une matrice hermitienne sont réelles.** En outre, nous admettrons qu'il existe pour A une base orthonormée de vecteurs propres, donc **une matrice hermitienne est diagonalisable.**

EXERCICES

■ **1.** E étant un espace hermitien de dimension finie, u un endomorphisme unitaire de E, e l'identité, on pose $v = e - u$.

a) Établir $\ker v = (\operatorname{Im} v)^\perp$

b) Montrer que si φ est le projecteur orthogonal de E sur $\ker v$, on a, pour tout vecteur $\vec{x}$,

$$\varphi(\vec{x}) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{k=n-1} u^k(\vec{x})$$

---

a) $$\vec{x} \in \ker v \Leftrightarrow u(\vec{x}) = \vec{x}$$

$$\vec{y} \in \operatorname{Im} v \Leftrightarrow \exists\, \vec{z} \quad \text{tel que} \quad \vec{y} = v(\vec{z}) = \vec{z} - u(\vec{z})$$

$$\vec{x} \cdot \vec{y} = u(\vec{x}) \cdot \vec{z} - u(\vec{x}) \cdot u(\vec{z}) = 0$$

car, u conservant le produit scalaire, on a $u(\vec{x}) \cdot u(\vec{z}) = \vec{x} \cdot \vec{z}$.

Tout vecteur de $\ker v$ est donc orthogonal à tout vecteur de $\operatorname{Im} v$, donc $\ker v \subset (\operatorname{Im} v)^\perp$.

Mais $(\operatorname{Im} v)^\perp$ est un supplémentaire de $\operatorname{Im} v$, donc

$$\dim (\operatorname{Im} v)^\perp = \dim E - \dim \operatorname{Im} v = \dim \ker v$$

et on a bien $$\ker v = (\operatorname{Im} v)^\perp$$

b) $\vec{x}$ étant cette fois un vecteur quelconque de E, on a

$$\vec{x} = \varphi(\vec{x}) + \vec{z}$$

où $\varphi(\vec{x}) \in \ker v \qquad \vec{z} \in \operatorname{Im} v$

Donc $\exists\, \vec{y}$ tel que $\vec{z} = v(\vec{y}) = \vec{y} - u(\vec{y})$

et $\varphi(\vec{x}) = \vec{x} + u(\vec{y}) - \vec{y} \qquad (1)$

Comme $\varphi(\vec{x}) \in \ker v \qquad u\,[\varphi(\vec{x})] = \varphi(\vec{x})$

donc $(1) \Rightarrow \varphi(\vec{x}) = u(\vec{x}) + u^2(\vec{y}) - u(\vec{y}) \qquad (2)$

De même $\varphi(\vec{x}) = u^2(\vec{x}) + u^3(\vec{y}) - u^2(\vec{y}) \qquad (3)$

et ainsi de suite $\varphi(\vec{x}) = u^3(\vec{x}) + u^4(\vec{y}) - u^3(\vec{y}) \qquad (4)$

. . . .

$$\varphi(\vec{x}) = u^{n-1}(\vec{x}) + u^n(\vec{y}) - u^{n-1}(\vec{y}) \qquad (n)$$

En ajoutant membre à membre les relations (1), (2), . . ., (n), on obtient :

$$n\varphi(\vec{x}) = \vec{x} + u(\vec{x}) + u^2(\vec{x}) + \ldots + u^{n-1}(\vec{x}) + u^n(\vec{y}) - \vec{y}$$

$$\varphi(\vec{x}) = \frac{1}{n}\sum_0^{n-1} u^k(\vec{x}) + \frac{1}{n}[u^n(\vec{y}) - \vec{y}]$$

Or u est unitaire, donc conserve la norme d'un vecteur, donc

$$N\,(u^n(\vec{y})) = N(\vec{y})$$

et l'inégalité triangulaire implique $N\,[u^n(\vec{y}) - \vec{y}] \leqslant 2\,N(\vec{y})$

$$0 \leqslant N\left[\frac{u^n(\vec{y}) - \vec{y}}{n}\right] \leqslant \frac{2}{n}N(\vec{y})$$

expression qui tend vers 0 pour n infini.

On a donc bien

$$\varphi(\vec{x}) = \lim_{n\to+\infty}\frac{1}{n}\sum_0^{n-1} u^k(\vec{x})$$

■ **2.** A étant une matrice hermitienne de dimension n, montrer que la matrice

$$P = (A + iI)(A - iI)^{-1}$$

existe et est unitaire, et exprimer les valeurs propres de A en fonction de celles de P.

---

a) A n'a que des valeurs propres réelles, donc $A - iI$ est inversible et P existe.

Exprimons ${}^t\bar{P}$

$$\bar{P} = (\bar{A} - iI)(\bar{A} + iI)^{-1}$$

$$ {}^t\bar{P} = ({}^t\bar{A} + iI)^{-1}({}^t\bar{A} - iI) = (A + iI)^{-1}(A - iI) $$

(car ${}^t\bar{A} = A$ et ${}^tI = I$).

Comme $P^{-1} = (A - iI)(A + iI)^{-1}$ (elle existe car $-i$ n'est pas valeur propre de A), il suffit de montrer que $A - iI$ et $(A + iI)^{-1}$ commutent; or

$$ \begin{aligned} (A - iI)(A + iI)^{-1} &= (A + iI - 2\,iI)(A + iI)^{-1} \\ I - 2\,i(A + iI)^{-1} &= (A + iI)^{-1}(A + iI) - 2\,i(A + iI)^{-1} \\ &= (A + iI)^{-1}(A - iI) \end{aligned} $$

On a donc bien ${}^t\bar{P} = P^{-1}$ et P est unitaire.

b) Soit X la matrice-colonne d'un vecteur propre de P, associé à une valeur propre λ, on a $PX = \lambda X$ donc $(A + iI)(A - iI)^{-1}X = \lambda X$ ou

$$ \begin{aligned} (A - iI)^{-1}(A + iI)X &= \lambda X \\ (A + iI)X &= (A - iI)\lambda X \\ AX + iX &= \lambda AX - \lambda iX \\ (\lambda - 1)AX &= i(\lambda + 1)X \end{aligned} $$

Ceci est impossible pour $\lambda = 1$ car $X \neq 0$ donc 1 n'est pas valeur propre de P, et

$$ AX = i\,\frac{\lambda + 1}{\lambda - 1}\,X $$

X est donc vecteur propre de A, et la valeur propre associée est

$$ \mu = i\,\frac{\lambda + 1}{\lambda - 1} $$

Comme A et P sont diagonalisables et qu'ainsi les dimensions des sous-espaces propres de chacune des matrices sont égales aux ordres de multiplicité des valeurs propres correspondantes, A et P ont mêmes sous-espaces propres et toutes les valeurs propres de A sont données par $\mu = i\,\dfrac{\lambda + 1}{\lambda - 1}$ les λ étant les valeurs propres de P.

Rappelons que $|\lambda| = 1$ donc $\lambda = e^{i\theta}$ avec $\theta \neq 2\,k\pi$ puisque $\lambda \neq 1$, donc

$$ \mu = i\,\frac{e^{i\theta} + 1}{e^{i\theta} - 1} = i\,\frac{2\cos\dfrac{\theta}{2}}{2\,i\sin\dfrac{\theta}{2}} = \operatorname{cotg}\frac{\theta}{2} $$

Ces valeurs propres sont réelles, comme prévu.