André et Germaine REVUZ

LE COURS DE L'A.P.M.

II - Espaces vectoriels

Les brochures de l'A.P.M.

8

Le Cours de l'A.P.M. - II. Espaces vectoriels — A. G. REVUZ

8

A.P.M.

Les brochures de l'A.P.M.

Collection de monographies utiles à l'enseignement des mathématiques.

1. **Le langage simple et précis des Mathématiques modernes,** par André REVUZ et Léonce LESIEUR (épuisé).
2. **Congruences paratactiques de cycles,** par Paul ROBERT (64 pages, 3 F, franco 3,50 F).
3. **Recherche d'une axiomatique commode pour le premier enseignement de la géométrie élémentaire,** par Gustave CHOQUET. (40 pages, 2 F, franco 2,50 F).
4. **Le calcul des probabilités et l'enseignement,** par A. HUISMAN, R. FORTET, E. MOURIER, A. FUCHS, D. DUGUE, G.-T. GUILBAUD, J. BOUZITAT, J. VILLE, F. GENUYS (épuisé).
5. **L'enseignement de la mécanique,** par P. GERMAIN, R. MAZET, J. KAMPE de FERIET (épuisé).
6. **Le Cours de l'A.P.M. I. — Groupes, anneaux, corps,** par André et Germaine REVUZ (160 pages, 20 F, franco 22 F).
7. **Etude commentée d'une méta-démonstration de Gödel,** par Jean BALIBAR (40 pages, 2 F, franco 2,50 F).

En préparation.

- **Les Mathématiques « modernes » dans l'enseignement du second degré,** par André HUISMAN.
- **Le cours de l'A.P.M. III, Eléments de topologie,** par André et Germaine REVUZ.

Pour se procurer les brochures de l'A.P.M.,

Adresser commandes et virements postaux à l'Association des Professeurs de Mathématiques de l'Enseignement Public, C.C.P. Paris 5708-21.
Préciser au dos du virement les brochures commandées.

Voir page 3 de la couverture les conditions d'adhésion à l'A.P.M.E.P.

André REVUZ
Professeur à la Faculté des Sciences de Poitiers

Germaine REVUZ
Maître-Assistant à la Faculté des Sciences de Poitiers

LE COURS DE L'A.P.M.

II

Espaces vectoriels

Association des Professeurs
de Mathématiques de l'Enseignement Public
PARIS - 1963

« *Afin que, ... aux dépens d'autrui*
Sage, je m'enseignasse. »

REGNIER.

« *C'est assez désagréable... de ne pouvoir plus rien apprendre pour toute la vie ! Nos aïeux s'en tenaient aux enseignements qu'ils avaient reçus dans leur jeunesse : mais, nous, il nous faut recommencer tous les cinq ans, si nous ne voulons pas être complètement démodés.* »

GŒTHE (Les affinités électives).

Les brochures de l'A.P.M. mettent à la disposition des professeurs des textes utiles à l'enseignement.

Ou bien ces textes sont inédits, ou bien ils ont déjà paru, soit dans le Bulletin de l'A.P.M., soit ailleurs. Dans tous les cas, il a paru intéressant de regrouper des écrits sous une forme commode pour les maîtres qui auront à s'en servir.

Brochures parues :

1. Le langage simple et précis des mathématiques modernes, par A. REVUZ et L. LESIEUR, Professeurs à la Faculté des Sciences de Poitiers (avril 1960) **(épuisé).**
2. Congruences Paratactiques de cycles, par Paul ROBERT, Inspecteur général honoraire de l'Instruction Publique (avril 1960).
3. Recherche d'une axiomatique commode pour le premier enseignement de la géométrie élémentaire, par Gustave CHOQUET, Professeur à la Sorbonne (février 1961).
4. Le calcul des probabilités et l'enseignement, par A. HUISMAN, R. FORTET E. MOURIER, A. FUCHS, D. DUGUE, G.-T. GUILBAUD, J. BOUZITAT, J. VILLE et F. GENUYS (novembre 1961) **(épuisé).**
5. L'enseignement de la mécanique, par P. GERMAIN, J. KAMPE DE FERIET et R. MAZET (novembre 1961) **(épuisé).**
6. Le Cours de l'A.P.M. — I. Groupes, anneaux, corps, par André et Germaine REVUZ (novembre 1962).
7. Lecture commentée d'une méta-démonstration de Gödel, par J. BALIBAR (mai 1962).
8. Le Cours de l'A.P.M. — II. Espaces vectoriels, par André et Germaine REVUZ (1963).

Brochures en préparation :

- Les Mathématiques « modernes » dans l'enseignement du second degré, par A. HUISMAN.

AVANT-PROPOS

Ce deuxième tome du « Cours de l'A.P.M. » a été élaboré dans les mêmes conditions que le premier : personne n'a relâché son effort et il est réconfortant de constater la continuité du travail de l'Association.

Mis à part le premier chapitre (nombres cardinaux), dont la place logique aurait été le début du premier tome, mais qui y aurait inutilement effarouché certains lecteurs, l'essentiel de ce volume est consacré à l'étude de la structure d'espace vectoriel sur un corps commutatif et à ses morphismes, les applications linéaires, c'est-à-dire à la partie la plus simple de l'algèbre linéaire, dont l'objet véritable est la structure plus générale de module sur un anneau.

Il n'y a pas de branche des mathématiques pures ou appliquées qui ne fasse une grande consommation d'algèbre linéaire : son importance est donc considérable et elle est l'élément indispensable de toute culture mathématique. Dans l'état actuel de la diffusion de la science mathématique, un étudiant de Propédeutique devrait connaître l'essentiel de ce qui est traité dans ce volume, un étudiant de Mathématiques I devrait l'avoir parfaitement assimilé.

Aucune des notions d'algèbre linéaire n'est très compliquée : elles semblent même si naturelles et la plupart des démonstrations si évidentes que l'on soupçonne mal, au premier abord, de quel puissant outil on est en train de se munir. Mais il faut croire que cette simplicité est trompeuse, si l'on en juge par les fautes commises par les débutants : l'étude de ces fautes m'amène à penser que la plupart ont leur origine dans une intuition confuse qui attribue inconsciemment au sous-espace vectoriel engendré par une partie les propriétés d'une réunion, d'où le lapsus : « le » supplémentaire d'un sous-espace, qui chez beaucoup trahit une confusion pas seulement verbale avec le complémentaire ; d'où l'utilisation, explicite ou non, du théorème faux : l'intersection est distributive par rapport à la somme (directe ou non) de sous-espaces ; d'où encore l'oubli de ce fait si évident que deux bases d'un même espace peuvent être des ensembles disjoints. Savoir penser linéairement, c'est sans doute, en premier lieu, s'être débarrassé de ces confusions (surprenantes aux yeux de qui a compris, mais dont l'expérience montre qu'elles renaissent avec chaque nouvelle génération d'étudiants), et avoir une intuition correcte de la notion de somme directe. C'est, en second lieu, avoir bien compris le sens et l'importance du théorème fondamental (III, 1) relatif à la factorisation des applications linéaires, clé de presque tous les autres résultats.

Le libellé de certains programmes, la rédaction de certains livres feraient croire à l'égalité : algèbre linéaire = calcul matriciel. Il s'agit, en fait, d'une inclusion : le calcul matriciel est une partie de l'algèbre linéaire, partie importante pour les applications, en particulier les applications numériques, mais partie mineure pour l'intelligence de la théorie. Qui a bien compris l'algèbre linéaire manipulera les matrices avec aisance, qui ne connaît que les matrices ou ne veut rien exprimer sans leur inter-

médiaire risque — et l'expérience montre que ce risque n'est pas imaginaire — de masquer des idées très simples par un appareil analytique souvent lourd. A ce propos n'est-il pas étrange de constater que, souvent, de soi-disant défenseurs de la géométrie ne voient pas que l'esprit géométrique, c'est-à-dire le recours à des idées simples et puissantes de préférence à des calculs, anime l'algèbre linéaire lorsqu'on évite de faire intervenir trop vite la représentation matricielle ?

Il a été dit que ce Cours de l'A.P.M. s'adressait spécialement aux professeurs : le cours oral a eu des professeurs pour auditeurs, le cours écrit a surtout des professeurs pour souscripteurs, mais il est clair qu'il s'adresse à tous ceux qu'intéressent les Mathématiques, car il se veut avant tout un instrument de culture. S'il est destiné d'abord à des professeurs, c'est que nous estimons que la culture chez le professeur et ses efforts personnels pour l'approfondir et l'accroître ne sont pas un luxe, mais la garantie première de la valeur de son enseignement — ce dernier se plaçât-il à un niveau très élémentaire. Un professeur n'est pas une machine à enseigner, son enseignement ne porte que s'il a parfaitement assimilé ce qu'il enseigne et s'il est capable de sortir du monologue dogmatique pour provoquer le dialogue avec ses élèves. Son enseignement ne sera fécond que si, non content de bien connaître ce qu'il enseigne, le professeur sait aussi bien ce que l'élève apprendra à un stade ultérieur : la formation des élèves est une course de relais chez les professeurs. Est-on sûr que le « témoin » soit toujours bien passé ?

Ceci dit, ce Cours n'est ni un manuel déguisé, ni un modèle de manuel destiné à des élèves : la matière y a été choisie sans référence à aucun programme. Je ne pense pas, pour ma part, que la théorie des nombres cardinaux ait à figurer à quelque niveau que ce soit de l'enseignement du Second Degré, que les espaces vectoriels de dimension infinie ou la réduction de Jordan aient à être étudiés en Propédeutique, mais il me paraît important que les professeurs qui traiteront des nombres entiers, des espaces vectoriels de dimension finie ou de la mise des matrices sous forme triangulaire sachent aller au-delà et sachent ce qui, dans les situations plus générales, se conserve ou ne se conserve pas : satisfaire la curiosité des professeurs, ou l'exciter, n'est-il pas légitime ? Inversement, des questions classiques et très connues : utilisation des déterminants pour la résolution des systèmes linéaires scalaires, par exemple, n'ont pas été reprises.

Nous espérons que le lecteur, qui aura eu le courage de travailler le présent Cours, sera convaincu que l'esprit moderne a su trouver dans de très vieilles — et assez rébarbatives — questions les lignes d'action qui en rendent l'étude aussi attrayante que celle de la plus belle géométrie. Alors il aura compris que la modernisation des Mathématiques ne se réduit pas à l'utilisation de quelques symboles nouveaux, ou à la transformation du vocabulaire, mais qu'elle est avant tout une manière plus féconde de concevoir les êtres mathématiques et leurs relations. Peut-être peut-on espérer aussi que l'ère des polémiques puériles et des étalages d'ignorance agressive est passée et que la diffusion de l'esprit moderne se poursuivra désormais sans précipitation, mais sans timidité ?

André REVUZ.

TABLE DES MATIÈRES

CHAP. III. — APPLICATIONS LINEAIRES.

CHAP. IV. — DUALITE.

CHAP. V. — ETUDE DES ENDOMORPHISMES DES ESPACES VECTORIELS DE DIMENSION FINIE. FORME REDUITE DES MATRICES CARREES.

CHAP. VI. — DETERMINANTS.

CHAP. VII. — STRUCTURE AFFINE.

CHAPITRE I

PUISSANCE DES ENSEMBLES NOMBRES CARDINAUX

§ 1. POSITION DU PROBLEME. PREMIERES DEFINITIONS

La Mathématique la plus fruste, celle qui se trouve attestée dans la langue des sociétés humaines les moins évoluées, consiste à donner un sens aux questions suivantes et un moyen commode d'y répondre :

« Deux collections d'objets étant données,

a) ont-elles autant d'éléments ?

b) l'une d'elles a-t-elle plus d'éléments que l'autre ? »

La théorie élémentaire correspondante, celle des « entiers naturels », s'occupe des collections finies. Mais on peut se poser les mêmes questions pour des ensembles infinis. Dans une théorie axiomatique des ensembles, on la pose pour tous les ensembles et on définit alors les ensembles finis, et par suite les entiers, par leur comportement particulier à son égard. Ici, nous supposerons que les entiers sont connus.

1. Nombres cardinaux. Egalité.

La formulation mathématique de l'énoncé : « deux ensembles E et F ont autant d'éléments l'un que l'autre », sera : *il existe une bijection de* E *sur* F. On dira *que ces deux ensembles ont même puissance,* ou *que* E *est équipotent à* F, ou *qu'ils ont même nombre cardinal,* et on écrira :

$$\text{card } E = \text{card } F$$

Il est clair qu'il existe toujours une bijection de E sur E (l'identité), que s'il existe une bijection f de E sur F, $\overset{-1}{f}$ fournit une bijection de F sur E, et que si f et g sont respectivement des bijections de E sur F et de F sur G, $g \circ f$ est une bijection de E sur G. Nous retrouvons donc les propriétés de réflexivité, symétrie et transitivité des relations d'équivalence et si nous ne considérons comme ensemble E, F ... que les parties d'un ensemble X donné *a priori,* « avoir même puissance » est une relation d'équivalence sur l'ensemble $\mathcal{P}(X)$.

Si on ne veut pas se restreindre aux parties d'un ensemble donné, la relation d'équivalence « avoir même puissance » n'est plus définie sur un ensemble, car la considération d'un ensemble dont tout ensemble serait élément (« l'ensemble de tous les ensembles ») conduit à une contradiction. Certains formalismes introduisent la notion de classe, dont celle d'ensemble serait une spécialisation. Dans ce cas, la relation « E a même puissance que F » est une relation d'équivalence sur la classe de tous les ensembles et card E est, relativement à cette relation, une classe d'équivalence (qui, elle non plus, n'est pas un ensemble).

Dans d'autres formalismes (Bourbaki), card E est à interpréter pour l'intuition comme un représentant de cette classe d'équivalence. C'est alors un ensemble qui est équipotent à E.

Remarque : Contrairement à ce que pourrait suggérer une première intuition, un ensemble peut être une partie d'un autre et avoir pourtant autant d'éléments que cet autre au sens que nous venons d'indiquer, c'est-à-dire que nous pouvons avoir :

$$E \subset F \text{ et } E \neq F \qquad \text{avec card } E = \text{card } F.$$

par exemple **N** et $2\mathbf{N} = \mathbf{N}'$ (ensemble des nombres pairs) se correspondent dans la bijection :

$$n \in \mathbf{N} \longrightarrow 2n \in \mathbf{N}'$$

Un ensemble fini, au contraire, ne peut avoir même puissance qu'une de ses vraies parties. C'est cette propriété qui est souvent prise comme définition des ensembles finis dans les théories axiomatiques des ensembles.

2. Comparaison de deux nombres cardinaux.

On dira que card E $\leqslant$ *card* F *si* E *a même puissance qu'une partie de* F, *ou encore s'il existe une injection de* E *dans* F.

Remarquons d'abord que :

$$\text{card } E = \text{card } E', \quad \text{card } E \leqslant \text{card } F \implies \text{card } E' \leqslant \text{card } F$$

Examinons si cette relation jouit des propriétés d'une relation d'ordre.

La réflexivité est évidente. La transitivité s'établit immédiatement : f et g étant des injections :

$$\exists f : E \longrightarrow F, \quad g : F \longrightarrow G \implies \exists\, g\circ f : E \longrightarrow G$$

$g\circ f$ étant encore une injection ; donc :

$$\text{card } E \leqslant \text{card } F, \quad \text{card } F \leqslant \text{card } G \implies \text{card } E \leqslant \text{card } G$$

La propriété d'antisymétrie, beaucoup moins immédiate, fait l'objet du *théorème de Cantor-Bernstein :*

$$\left.\begin{matrix}\text{card } E \leqslant \text{card } F\\ \text{card } F \leqslant \text{card } E\end{matrix}\right\} \implies \text{card } E = \text{card } F$$

ce qui peut encore s'énoncer : *s'il existe une injection de* E *dans* F *et une injection de* F *dans* E, *alors il existe une bijection de* E *sur* F.

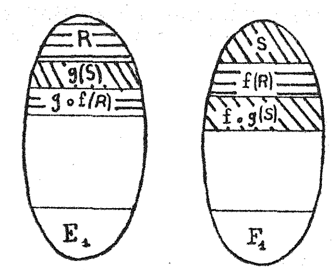
Démonstration : Nous supposerons que :

$$f : E \longrightarrow F \text{ et } g : F \longrightarrow E$$

sont des injections qui ne sont pas bijectives, sans quoi le théorème serait démontré.

$f(E)$ est donc différent de F ; posons $S = \complement_F f(E)$. $g(F)$ est différent de E ; posons $R = \complement_E g(F)$. Considérons alors $f(R)$ qui est une partie de F disjointe de S d'après la définition même de S, et considérons de même $g(S)$ disjointe de R. Puis, considérons les images successives $g\circ f(R) \subset E$, $f\circ g(S) \subset F$, $f\circ g\circ f(R) \subset F$, $g\circ f\circ g(S) \subset E$, ..., etc...

Observons qu'une injection donne toujours de deux ensembles disjoints des images disjointes. Le fait que S et $f(R)$ sont disjoints entraîne donc que $g(S)$ et $g\circ f(R)$ le soient et ceci entraînera que $f\circ g(S)$ et $f\circ g\circ f(R)$ le soient, etc..., ainsi de suite. Nous déterminerons dans nos ensembles E et F une succession de sous-ensembles deux à deux disjoints. En



E F

Fig. 1

continuant ce processus indéfiniment, il peut se faire que nous n'épuisions pas les éléments de E et de F ; soient E_1 et F_1 les sous-ensembles respectivement constitués des éléments de E et de F qui n'appartiennent à aucune des images considérées précédemment. Nous avons ainsi établi une partition de E et de F en ensembles disjoints :

$$E \begin{cases} R \\ E_2 = \bigcup_{n=1}^{\infty} (g\circ f)^n(R) \\ E_3 = \bigcup_{n=0}^{\infty} g\circ(f\circ g)^n(S) \\ E_1 \end{cases} \qquad F \begin{cases} S \\ F_2 = \bigcup_{n=1}^{\infty} (f\circ g)^n(S) \\ F_3 = \bigcup_{n=0}^{\infty} f\circ(g\circ f)^n(R) \\ F_1 \end{cases}$$

et nos définitions entraînent que :

$$f(R \cup E_2) = F_3, \quad f(E_3) = F_2.$$

Mais alors E_1 doit avoir une image $f(E_1)$ disjointe de F_2 et de F_3 et telle que :

$$f(E_1) \cup F_2 \cup F_3 = f(E)$$

ceci exige que $f(E_1) = F_1$. La restriction de f à E_1 est donc une bijection de E_1 sur F_1. Mais alors la restriction de f à $R \cup E_2 \cup E_1$ est une bijection de cet ensemble sur $F_3 \cup F_1$.

Considérons maintenant g qui est telle que $g(S \cup F_2) = E_3$, et dont la restriction à $S \cup F_2$ est de même une bijection de $S \cup F_2$ sur E_3, qui admet une bijection réciproque, restriction de $\overset{-1}{g}$ à E_3 :

$$\overset{-1}{g}(E_3) = S \cup F_2$$

Nous pouvons alors définir une bijection h de E sur F de la façon suivante : avoir même restriction que f sur $R \cup E_2 \cup E_1$, avoir même restriction que $\overset{-1}{g}$ sur E_3.

$$h : \begin{cases} R \cup E_2 \cup E_1 \xrightarrow{f} F_3 \cup F_1 \\ E_3 \xrightarrow[\overset{-1}{g}]{} S \cup F_2 \end{cases}$$

Le théorème est démontré.

Remarque : On aurait aussi bien pu définir une bijection h' de F sur E par la condition d'avoir même restriction que g sur $S \cup F_2 \cup F_1$ et même restriction que $\overset{-1}{f}$ sur F_3 :

$$h' : \begin{cases} S \cup F_2 \cup F_1 \xrightarrow{g} E_3 \cup E_1 \\ F_3 \xrightarrow[\overset{-1}{f}]{} R \cup E_2 \end{cases}$$

Remarquons que sur une partie des ensembles E et F les bijections h et h' ont des restrictions réciproques :

$$R \cup E_2 \underset{h' = \overset{-1}{f}}{\overset{h=f}{\rightleftarrows}} F_3$$

$$E_3 \underset{h' = g}{\overset{h=\overset{-1}{g}}{\rightleftarrows}} S \cup F_2$$

Mais les bijections

$$E_1 \underset{h' = g}{\overset{h=f}{\rightleftarrows}} F_1$$

n'ont pas de raison d'être réciproques l'une de l'autre.

3. Application à l'ensemble des rationnels.

Nous dirons d'un ensemble qui a même puissance que **N** qu'il est dénombrable (*). Et nous appellerons $\aleph_0$ (aleph zéro) son nombre cardinal.

Nous démontrerons d'abord que *le produit cartésien* $\mathbf{N}^2$ *est dénombrable.* Les éléments de $\mathbf{N}^2$ étant les points à coordonnées entières du premier quadrant, il suffit de prendre successivement l'origine, puis les points de la droite $x + y = 1$ dans l'ordre des ordonnées croissantes, puis ceux de la droite $x + y = 2$ dans ce même ordre, puis ceux de $x + y = 3$ ... et ainsi de suite ... et d'attribuer à chaque point son numéro d'ordre dans la liste ainsi dressée pour obtenir une bijection :

$\mathbf{N} \longrightarrow$	$\mathbf{N} \times \mathbf{N}$
0	(0,0)
1	(1,0)
2	(0,1)
3	(2,0)
4	(1,1)
5	(0,2)
6	(3,0)

et ainsi de suite.

(*) A noter que certains auteurs utilisent le terme « dénombrable » pour ce que nous appellerions ici « fini ou dénombrable ». Cet usage se justifie par le fait que dans la pratique mathématique la plupart des énoncés où intervient le dénombrable concernent le cas « fini ou dénombrable ».

On établira sans peine qu'avec ces conventions au point (x, y) correspond l'entier :

$$\frac{(x+y)(x+y+1)}{2} + y$$

Venons-en maintenant à l'ensemble $\mathbf{Q}^+$ des rationnels positifs. On a $\mathbf{N} \subset \mathbf{Q}^+$ et $\mathbf{Q}^+$ a même puissance que l'ensemble des fractions irréductibles, donc aussi même puissance qu'une partie de $\mathbf{N}^2$. On a donc,
d'une part : $\operatorname{card} \mathbf{N} \leqslant \operatorname{card} \mathbf{Q}^+$,
d'autre part : $\operatorname{card} \mathbf{Q}^+ \leqslant \operatorname{card} \mathbf{N}^2 = \operatorname{card} \mathbf{N}$,
d'où, en vertu du théorème de Cantor-Bernstein : $\operatorname{card} \mathbf{N} = \operatorname{card} \mathbf{Q}^+$.
L'ensemble des rationnels positifs est dénombrable.

Passons maintenant aux rationnels des deux signes ou plus généralement à la réunion de deux ensembles dénombrables disjoints E_1 et E_2. A l'élément x_n de E_1 qui est l'image de n dans $\mathbf{N} \to E_1$ on fait correspondre $2n$. A l'élément y_n de E_2 qui est l'image de n dans $\mathbf{N} \to E_2$ on fait correspondre $2n + 1$. On a ainsi défini une bijection $E_1 \cup E_2 \longrightarrow \mathbf{N}$ et $E_1 \cup E_2$ est dénombrable.

En particulier Q *est dénombrable.*

Le résultat précédent peut se généraliser. $\mathbf{N}^3 = \mathbf{N}^2 \times \mathbf{N}$. $\mathbf{N}^2$ étant dénombrable, à chacun de ses éléments on fait correspondre bijectivement un entier n et au couple (n, z) où $z \in \mathbf{N}$ et où n est l'image d'un élément $(x, y) \in \mathbf{N}^2$, on fait correspondre un nouvel entier n' ... et ainsi de suite... Par récurrence, on montrera que, pour tout k, $\mathbf{N}^k$ est dénombrable.

4. Cardinal de l'ensemble des parties d'un ensemble.

Devant un tel résultat on pourrait être amené à se demander si tous les ensembles ne sont pas dénombrables. Le théorème suivant montre qu'il n'en est rien.

L'ensemble $\mathcal{P}$(E) *des parties d'un ensemble ne peut avoir même puissance que* E.

Nous le démontrerons par l'absurde. Supposons qu'il existe une bijection φ :

$$\varphi : E \longrightarrow \mathcal{P}(E)$$

qui, à tout $a \in E$, fasse correspondre une partie $\varphi(a) \subset E$; un élément a peut appartenir à $\varphi(a)$ ou ne pas lui appartenir. Considérons l'ensemble des éléments qui n'appartiennent pas à leur image par φ, soit :

$$B = \{ a \,;\, a \in E, \quad a \notin \varphi(a) \}.$$

B est une partie de E ; donc :

$$\exists\, b \in E \quad B = \varphi(b).$$

Nous allons montrer que la considération de b nous conduit à une contradiction. En effet :

ou bien $b \in \varphi(b) \iff b \in B$; or, d'après la définition de B, $b \notin \varphi(b)$,
ou bien $b \notin \varphi(b) \iff b \notin B$; or, d'après la définition de B, $b \in \varphi(b)$.

Remarque : Une réaction fréquente au raisonnement précédent s'exprime par : « Et si B n'existait pas ? ». Il faut remarquer que B existe toujours ; il peut seulement être l'ensemble vide, mais ce cas n'a rien de particulier car $\varnothing \in \mathcal{P}(E)$ et est par conséquent l'image par φ d'un élément b de E.

Il ne peut donc exister de bijection entre E et $\mathcal{P}(E)$. Comme $\mathcal{P}(E)$ contient l'ensemble $\tilde{E}$ constitué des parties réduites à un seul élément et comme E et $\tilde{E}$ ont évidemment même cardinal, on écrit :

$$\text{card } E < \text{card } \mathcal{P}(E),$$

ce qui se lit : « card E est strictement inférieur à card $\mathcal{P}(E)$ » et signifie qu'il existe une injection de E dans $\mathcal{P}(E)$, et qu'il n'existe aucune bijection de E sur $\mathcal{P}(E)$.

On a de même :

$$\text{card } \mathcal{P}(E) < \text{card } \mathcal{P}\mathcal{P}(E).$$

On voit donc qu'on peut par ce procédé, à partir de n'importe quel cardinal, construire une chaîne infinie strictement croissante de nombres cardinaux.

§ 2. AXIOME DE ZERMELO

L'ordre sur les cardinaux est-il total ?

Ayant défini un ordre pour les nombres cardinaux, il est naturel de se demander s'il s'agit d'un ordre total. La réponse à cette question ne peut être donnée sans avoir recours à un axiome dont l'adoption a soulevé de nombreuses controverses. A l'heure actuelle, le point de vue presque unanimement adopté est le suivant : si l'application de cet axiome à certaines familles conduit à des contradictions, ce n'est pas l'axiome qui doit être mis en cause mais la nature de ces familles ; celles-ci ne doivent pas être considérées comme des ensembles.

Nous donnerons, de l'axiome de Zermelo, une liste non exhaustive d'énoncés que l'on rencontre dans la littérature et qui sont équivalents. Un premier groupe comprend les énoncés suivants, très voisins les uns des autres :

I. — *Etant donnée une famille $\mathcal{E}$ d'ensembles E_i deux à deux disjoints, il existe un ensemble contenant exactement un élément de chacun des E_i.* En fait, dans cet énoncé que nous venons de donner sous sa forme historique, l'hypothèse « disjoints » peut être abandonnée ; et nous arrivons à l'énoncé suivant.

II. — Soit une famille d'ensembles E_i disjoints ou non. On appelle produit cartésien (infini) de ces ensembles, l'ensemble dont les éléments sont les familles (x_i) où i décrit l'ensemble I des indices et où chaque x_i appartient à l'ensemble E_i de même indice. Soit $\prod_{i \in I} E_i$ ce produit. On énonce alors :

$$\forall i \in I,\ E_i \neq \varnothing \implies \prod_{i \in I} E_i \neq \varnothing.$$

III. — Soit R une relation binaire entre éléments $x \in E$ et $y \in F$ (la donnée de R est celle d'un sous-ensemble de $E \times F$). On énonce :

$$\forall x \in E \quad E_x = \{ y \,;\, y \in F, xRy \} \neq \varnothing \implies \exists f : E \to F \,;\, \forall x \in E, xRf(x),$$

ce qui revient à dire intuitivement que l'on peut « choisir » un élément dans chacun des ensembles non vides E_x.

IV. — Pour tout ensemble E, il existe une application :

$$f : \mathcal{P}(E) - \{\varnothing\} \longrightarrow E \,; \quad \forall A \in \mathcal{P}(E) - \{\varnothing\},\ f(A) \in A,$$

c'est-à-dire une application qui, à toute partie non vide de E, fait correspondre un élément de cette partie.

Un second groupe comprend deux énoncés. Leur équivalence, entre eux et avec les précédents, est plus difficile à établir.

I. — *L'ordre sur les cardinaux est total.*

II. — (Souvent appelé théorème de Zermelo). *Tout ensemble peut être bien ordonné.*

On dit qu'un ensemble est bien ordonné si, pour l'ordre considéré, toute partie a un plus petit élément (ceci entraîne, évidemment, que pour cet ordre l'ensemble soit totalement ordonné).

Un ensemble bien ordonné a un plus petit élément que nous pouvons noter a_1. Le complémentaire de $\{a_1\}$ a un plus petit élément que nous pouvons noter a_2 ; le complémentaire de $\{a_1, a_2\}$ a un plus petit élément que nous pouvons noter a_3... Si le complémentaire de $\{a_1, a_2, \dots a_n, \dots\}$ $n \in \mathbf{N}$ n'est pas vide, son plus petit élément est noté a_ω, le suivant s'il existe $a_{\omega+1}$, etc... Par exemple, si nous considérons le bon ordre des entiers qui consiste à prendre d'abord tous les nombres pairs dans leur ordre naturel, puis tous les nombres impairs également dans leur ordre naturel, le nombre d'indice ω sera 1, le nombre d'indice $\omega + 1$ sera 3, etc.

Le théorème essentiel sur les ensembles bien ordonnés est que deux ensembles bien ordonnés E et F étant donnés, il existe toujours une bijection croissante de l'un, E par exemple, sur une section commençante de l'autre. Il en résulte que II implique I, car l'existence de cette bijection se traduit par card E $\leqslant$ card F.

Nous donnerons enfin l'énoncé suivant, très commode dans de nombreuses applications, dont on démontre qu'il est équivalent aux précédents.

Axiome (ou théorème) de Zorn : Si un ensemble ordonné est tel que chacune de ses chaînes (c'est-à-dire chacun de ses sous-ensembles totalement ordonnés) soit majorée, alors cet ensemble possède au moins un élément maximal.

Un ensemble satisfaisant aux hypothèses du théorème de Zorn est appelé *inductif*. Avec cette terminologie, le théorème s'énonce : Un ensemble inductif possède au moins un élément maximal.

Nous comprendrons la signification de cet énoncé en l'appliquant à un exemple : celui des idéaux bilatères propres (c'est-à-dire $\neq$ A) d'un anneau A unitaire, qui contiennent un idéal donné I. Tout sous-ensemble totalement ordonné (ou chaîne) d'idéaux propres est majoré par la réunion de ces idéaux. En effet, cette réunion $\mathcal{I}$ est un idéal (car si $a \in I_1$, $b \in I_2$, $a - b$ appartient à celui des idéaux I_1 et I_2 qui contient l'autre donc à la réunion ; si $a \in I_1$, $\forall x \in A$ $ax \in I_1$ donc $ax \in \mathcal{I}$), et cet idéal contient tous les idéaux de la chaîne. D'autre part, c'est un idéal propre car s'il était A il contiendrait l'unité et il ne peut la contenir sans qu'un des idéaux de la chaîne la contienne, ce qui impliquerait que cet idéal ne soit pas propre. Nous sommes donc dans les hypothèses du théorème de Zorn et pouvons conclure que la famille considérée possède au moins un élément maximal, c'est-à-dire un idéal propre tel qu'il n'en existe aucun plus grand, c'est-à-dire un idéal maximal. Il en résulte que l'idéal I est contenu dans un idéal maximal.

Le théorème de Zorn a donc permis de démontrer : *dans un anneau unitaire, tout idéal bilatère propre est contenu dans un idéal maximal.*

§ 3. OPERATION SUR LES CARDINAUX

1. Somme.

Nous poserons :

$$\text{card } E + \text{card } F = \text{card } E \cup F, \text{ avec } E \cap F = \varnothing.$$

Cette définition n'a de sens que si le résultat n'est pas modifié par le remplacement d'un des ensembles par un ensemble équipotent. Il est clair qu'il en est bien ainsi, car l'égalité des cardinaux de E et E' entraîne l'existence d'une bijection f de E sur E'. Puisque $E \cap F = E \cap F' = \varnothing$, on peut définir une bijection :

$$E \cup F \longrightarrow E' \cup F$$

dont la restriction à E est f et la restriction à F est la bijection identique.

On vérifie aussi immédiatement qu'associativité et commutativité de l'opération réunion entraînent l'associativité et la commutativité de cette somme.

Enfin on vérifie, en appliquant cette définition à des ensembles finis, qu'on obtient, comme cas particulier, l'addition sur **N** puisque si E et F disjoints ont respectivement n et p éléments :

$$\text{card } E \cup F = n + p.$$

Mais si a ou b est infini, nous démontrerons plus loin que :

$$a + b = \sup(a, b).$$

Nous allons cependant le démontrer tout de suite dans le cas particulier, important pour les applications, où le plus petit des deux cardinaux est $\aleph_0$.

Soient donc un ensemble E infini, de puissance $a \geqslant \aleph_0$, et un ensemble dénombrable D. Il existe une bijection entre D et un ensemble dénombrable $D' \subset E$. Et nous voulons montrer qu'il en existe une entre E et $E \cup D$. Coupons E en $\complement_E D'$ et D' disjoints ; coupons $E \cup D$ en $\complement_E D'$ et $D \cup D'$ disjoints. $D \cup D'$, réunion de deux ensembles dénombrables, est dénombrable. Il existe donc une bijection $f : D' \longrightarrow D \cup D'$. Il nous suffit alors de considérer la bijection dont la restriction à D' est f et la restriction à $\complement_E D'$ est la bijection identique pour avoir la bijection cherchée.

Nous pouvons donc énoncer : $\forall a \geqslant \aleph_0 \quad a + \aleph_0 = a.$

2. Produit.

On pose :

$$\text{card } E \times \text{card } F = \text{card } E \times F.$$

On vérifie immédiatement que cette définition a un sens $\Big($ si card E = card E', c'est qu'il existe une bijection $\varphi : x \in E \longrightarrow \varphi(x) \in E'$; il existe alors une bijection $(x, y) \in E \times F \longrightarrow (\varphi(x), y) \in E' \times F \Big)$; que l'opération ainsi définie est associative et commutative ; que sur l'ensemble des entiers naturels, elle est la multiplication ordinaire, car si card $E = n$ et card $F = p$, on a card $E \times F = np$.

Nous montrerons plus loin que si a ou b est infini :

$$a\, b = \sup(a, b).$$

3. Exponentiation.

On pose :

$$\text{card } F^{\text{card } E} = \text{card } \mathcal{F}(E, F)$$

$\mathcal{F}$ représentant l'ensemble des applications de E dans F. Il est évident que cette définition a un sens. Si card E = card E', c'est qu'il existe une bijection $\varphi : E' \longrightarrow E$, et à toute application $f : E \longrightarrow F$ correspond une application $f \circ \varphi : E' \longrightarrow F$ et une seule. De même, si card F = card F', c'est qu'il existe $\psi : F \longrightarrow F'$, et à toute application $f : E \longrightarrow F$ correspond une application $\psi \circ f : E \longrightarrow F'$ et une seule. On constate aussi qu'elle donne sur **N** l'exponentiation habituelle, car si F a n éléments et si E a p éléments, il existe n^p applications de E dans F.

Considérons alors le nombre cardinal $2^{\aleph_0}$. C'est la puissance de l'ensemble $\mathcal{F}$ des applications d'un ensemble dénombrable E sur l'ensemble F constitué de 2 éléments. Pour étudier cette puissance, prenons **N** pour E et $\{0, 1\}$ pour F. Une application de **N** dans F est donnée par $\varphi(n) = \alpha_n$, α_n valant 0 ou 1 pour tout n. Elle peut être donnée sous la forme : $0, \alpha_1 \alpha_2 \ldots \alpha_n \ldots$

Or, ce symbole est le développement illimité en numération binaire d'un nombre réel de $[0, 1]$. A toute application $f \in \mathcal{F}$ on fait donc correspondre $x \in [0, 1]$. Réciproquement, tout x appartenant à $[0, 1]$ a un développement binaire illimité unique si $x \neq \frac{k}{2^n}$ (n entier et k entier) et deux développements s'il est de la forme $\frac{k}{2^n}$ $\Big[\frac{9}{16}$ a les deux développements $0{,}100100 \ldots 0 \ldots$ et $0{,}100011 \ldots 1 \ldots \Big]$. Mais l'ensemble des nombres de la forme $\frac{k}{2^n}$ dépendant des deux entiers k et n a la puissance d'une partie infinie de $\mathbf{N} \times \mathbf{N}$, donc est dénombrable. Retirer de l'ensemble des développements binaires illimités un des deux ensembles qui font double emploi (par exemple celui des éléments où tous les chiffres sont des 1 à partir d'un certain rang) ne change pas la puissance de cet ensemble car, si la puissance de l'ensemble restant, qui est infini, est a, celle de l'ensemble primitif est $a + \aleph_0 = a$. L'ensemble restant correspond alors bijectivement à l'ensemble des réels de $[0, 1]$. On peut donc conclure

$$\boxed{2^{\aleph_0} = c}$$

c, appelé *puissance du continu,* désignant la puissance de l'ensemble des nombres de $[0, 1]$.

Il résulte de l'égalité précédente que c est strictement supérieur à $\aleph_0$. En effet, nous avons montré que l'ensemble des parties d'un ensemble avait une puissance strictement supérieure à celle de cet ensemble. Or, choisir une partie dans un ensemble E c'est dire, pour chaque élément de l'ensemble, si on le met ou non dans la partie, c'est donc mettre en face de chaque élément de l'ensemble un « oui » ou un « non » ; c'est donc définir une application de E dans l'ensemble à 2 éléments $\{$oui, non$\}$; et l'ensemble des parties de E a la puissance de l'ensemble des applications de E dans $\{$oui, non$\}$. Sous une autre forme, remplaçant oui par 1, et non par 0, on voit qu'il existe une bijection de $\mathcal{P}(E)$ sur $\mathcal{F}[E, \{0, 1\}]$ qui, à toute partie A de E, fait correspondre l'application χ_A, appelée *fonction caractéristique* de A, définie par :

$$\chi_A(x) = 1 \text{ si } x \in A \qquad \chi_A(x) = 0 \text{ si } x \notin A.$$

Si donc card $E = a$, on a : card $\mathcal{P}(E) = 2^a$ et $2^a > a$.

En particulier, si $a = \aleph_0$, il vient $c = 2^{\aleph_0} > \aleph_0$.

La question qui se pose alors est de savoir s'il existe des puissances intermédiaires entre $\aleph_0$ et c. La non-existence de telles puissances a été conjecturée par Cantor (hypothèse du continu). Les travaux de Gödel prouveraient que cette hypothèse n'est pas incompatible avec les axiomes de la théorie des ensembles. Il n'y a pas d'ensemble dont tout cardinal soit élément, mais on peut établir que sur la classe des cardinaux, l'ordre que nous avons introduit est un bon ordre. $\aleph_0$ est le plus petit cardinal infini. Il existe un cardinal $\aleph_1$ immédiatement supérieur à $\aleph_0$. L'hypothèse du continu s'écrit alors $2^{\aleph_0} = \aleph_1$.

4. Théorème. Pour tout cardinal infini a, $a^2 = a$.

Montrons d'abord que $na = \underbrace{a + a + \dots + a}_{n \text{ termes}}$; na est la puissance du produit cartésien de l'ensemble fini $1, 2, \dots, n$ par l'ensemble E de puissance a. Ce produit est l'ensemble des couples :

$$\{ (p, x) \; ; \; 1 \leqslant p \leqslant n, \quad x \in E \} \qquad (1)$$

$a + a + \dots + a$ est la puissance de la réunion $\bigcup_{p=1}^{n} E_p$ de n ensembles disjoints de même puissance $a, E_1, E_2 \dots E_n$ qui sont tous tels qu'il existe une bijection f_p de E sur E_p. Soit x_p l'image de x dans la bijection f_p.

$\bigcup_{p=1}^{n} E_p$ est l'ensemble :

$$\{ x_p \; ; \quad 1 \leqslant p \leqslant n \; ; \quad x \in E \} \qquad (2)$$

Il existe une bijection évidente entre les ensembles (1) et (2) : celle qui à (p, x) fait correspondre x_p.

Démontrons maintenant que, pour *tout cardinal* a *infini*, $a^2 = a$. Nous avons déjà démontré ce théorème dans le cas où $a = \aleph_0$. Soit maintenant un ensemble E de cardinal $a > \aleph_0$. Nous voulons prouver qu'il existe une bijection de E^2 sur E. Or il existe certainement des parties X de E telles qu'il existe une bijection de X^2 sur X. En effet, $a > \aleph_0$ signifie qu'il existe des parties de E qui sont dénombrables d'après la définition de la relation d'ordre sur les cardinaux, et ces parties peuvent être mises en correspondance biunivoque avec leur carré cartésien.

Nous allons considérer l'ensemble $\mathcal{M}$ des couples (X, f) où X représente une de ces parties et f la bijection de X^2 sur X qui lui est associée. Sur cet ensemble, on peut définir un ordre par :

$$(X_1, f_1) \prec (X_2, f_2) \iff \begin{cases} X_1 \subset X_2 \\ f_1 \text{ est la restriction de } f_2 \text{ à } X_1^2 \end{cases}$$

On vérifie immédiatement que cette relation est bien réflexive, transitive et antisymétrique. Nous allons montrer que cet ensemble ordonné satisfait aux hypothèses de l'axiome de Zorn, c'est-à-dire qu'il est inductif. Considérons une chaîne de (X_i, f_i) ; elle est majorée par le couple (X, f) où X est la réunion des X_i associée à une bijection f que nous allons définir. Soit $(x, y) \in X^2$. Ceci veut dire : $\exists\, i \quad x \in X_i, \; \exists\, j \quad y \in X_j$. Mais les X_i formant une chaîne, un des deux ensembles X_i, X_j contient l'autre. Il existe donc au moins un indice k tel que $(x, y) \in X_k^2$. Mais s'il en existe d'autres et si l est l'un d'eux, un des deux ensembles X_l^2 et X_k^2 est contenu dans l'autre, et une des deux applications f_k et f_l est la restriction de l'autre. On en déduit que, pour tout indice k tel que $(x, y) \in X_k^2$, $f_k(x, y)$ a la même valeur.

En posant $f(x, y) = f_k(x, y)$, on définit donc une application f de X^2 sur X.

Reste à montrer que cette application est bijective. Elle est injective, car étant donnés $(x, y) \neq (x', y')$, il existe un X_k qui contient x, y, x', y', donc :

$$(x, y) \in X_k^2 \qquad f(x, y) = f_k(x, y)$$
$$(x', y') \in X_k^2 \qquad f(x', y') = f_k(x', y')$$

f_k étant injective, $f(x, y) \neq f(x', y')$ et f est injective. Elle est surjective car si $u \in X$, $\exists\, i \;\; u \in X_i$; f_i étant surjective :

$$\exists\, (x, y) \in X_i^2 \qquad f_i(x, y) = u = f(x, y).$$

L'ensemble $\mathcal{M}$ est donc inductif et possède un élément maximal, c'est-à-dire qu'il existe un ensemble $F \subset E$ et une bijection $f : F^2 \longrightarrow F$ et qu'il est impossible de trouver $G \supset F$, $G \neq F$ avec une bijection $g : G^2 \longrightarrow G$ dont f soit la restriction à F^2.

Si card $F = a$, le théorème est démontré.

Supposons que card $F = b < a$. Nous avons démontré que $b^2 = b$.

Soit $Y = \complement_E F$. Si on avait card $Y \leqslant b$, on aurait :

$$\text{card } E = \text{card } F + \text{card } Y \leqslant b + b = 2b \qquad (1)$$

Or, $2b \leqslant b^2$ (ceci se vérifie immédiatement en revenant aux définitions). Donc, card $E \leqslant b^2 = b$, ce qui est en contradiction avec l'hypothèse $b < a$. Donc, card $Y > b$. Mais alors :

$$\exists\, Z \subset Y \qquad \text{avec card } Z = b.$$

Posons : $\quad G = F \cup Z \quad$ et formons :

$$G \times G = (F \times F) \cup (F \times Z) \cup (Z \times F) \cup (Z \times Z)$$

les quatre ensembles considérés étant deux à deux disjoints. Les trois derniers ont pour cardinal $b \times b = b^2 = b$; leur union a pour cardinal $3b$. Mais en (1) nous avons montré que :

$$b \leqslant 2b \leqslant b \quad \text{donc : } 2b = b \text{ et } 3b = 2b + b = b + b = b$$

$G \times G$ se décompose alors en : 1° $F \times F$, que f applique bijectivement sur F ; 2° $(F \times Z) \cup (Z \times F) \cup (Z \times Z)$ de cardinal b qui peut être appliqué bijectivement sur Z. Donc $G \times G$ peut être appliqué bijectivement sur $F \cup Z = G$; (G, g) appartient à $\mathcal{M}$ et majore strictement (F, f) qui ne serait donc pas un élément maximal de $\mathcal{M}$. L'hypothèse $b < a$ a donc abouti à une contradiction et card $F = a$, ce qui démontre que $a^2 = a$.

Conséquences : Nous pouvons en déduire les propriétés :

$$\forall a, b \text{ infinis} \qquad ab = \sup(a, b)$$
$$a + b = \sup(a, b).$$

Pour la première, il suffit de remarquer que :

$$a \leqslant b \implies ab \leqslant b^2 = b \; ;$$

comme d'autre part $ab \geqslant b$, il en résulte $ab = b$.

Pour la deuxième, on remarque de même que :

$$a \leqslant b \implies a + b \leqslant 2b = b \; ;$$

comme d'autre part $a + b \geqslant b$, il en résulte $a + b = b$.

Exercice 1. — Démontrer que $c^2 = c$ sans utiliser le théorème général. Pour cela on cherche à définir une bijection entre le carré $(0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1)$ et le segment $[0,\ 1]$.

A tout nombre t de $[0,\ 1]$ donné par son développement décimal illimité, on pense d'abord à faire correspondre $(x,\ y)$, x (resp. y) ayant pour développement décimal illimité le développement constitué des décimales de rang impair (resp. pair) de t, prises dans leur ordre. Toutefois, le fait que certains réels admettent deux développements décimaux distincts soulève des difficultés. Préciser ces difficultés et montrer qu'on peut y pallier en modifiant ainsi l'application : le développement de t sera coupé en tranches de chiffres dont le premier à partir de la gauche sera obligatoirement différent de 9. Par exemple : $t = 0{,}09\ \ 7\ \ 8\ \ 6\ \ 5999\ \ 29\ \ 399 \ldots$ sera coupé comme il est indiqué et les développements de x (resp. y) seront constitués par la suite des tranches de chiffres d'ordre impair (resp. pair) ; dans l'exemple ci-dessus :

$$x = 0{,}0985999399\ldots.$$
$$y = 0,\ 7629\ldots\ldots$$

Montrer que dans ces conditions on a bien défini une bijection.

$$\{ t,\ 0 \leqslant t \leqslant 1 \} \longrightarrow \{ (x,\ y)\ ;\ 0 \leqslant x,\ 1\ ;\ 0 \leqslant y \leqslant 1 \}$$

Montrer que cette bijection et son inverse sont discontinues et préciser leurs points de discontinuité.

Montrer qu'une application continue du carré sur le segment $[0,\ 1]$ n'est jamais bijective.

Exercice 2. — Montrer $\aleph_0{}^{\aleph_0} = c$. On montrera que $\mathcal{F}$ $(\mathbf{N}^*,\ \mathbf{N}^*)$ a la puissance du continu en observant que la donnée du développement en fraction continue d'un irrationnel de $[0,\ 1]$ est la donnée d'une application $\mathbf{N}^* \longrightarrow \mathbf{N}^*$. $(\mathbf{N}^* = \mathbf{N} - \{ 0 \})$.

Exercice 3. — 1° Montrer que $c^{\aleph_0} = c$ en considérant l'ensemble des suites de réels $\{ x_n \} = x_1, \ldots x_n \ldots$ et en considérant l'application définie de la façon suivante :

Soit $x_n = 0,\ \alpha_n^1\ \alpha_n^2 \ldots \alpha_n^n \ldots$, les α représentant des tranches de décimales définies comme dans l'exercice 1, l'indice inférieur étant relatif au rang du terme dans la suite, l'indice supérieur étant relatif au rang de la tranche de décimales. A la suite $\{ x_n \}$ on fait correspondre :

$$x = 0,\ \alpha_1^1\ \alpha_1^2\ \alpha_2^1\ \alpha_1^3\ \alpha_2^2\ \alpha_3^1\ \alpha_1^4\ \alpha_2^3 \ldots$$

c'est-à-dire un réel dont le développement décimal est constitué par juxtaposition des tranches α de décimales des x_n rangées dans un ordre tel que la somme des indices des α soit non décroissante et qu'à l'intérieur d'un groupe où cette somme est constante l'indice inférieur aille en croissant.

2° En déduire que l'ensemble $\mathcal{C}$ $([0,\ 1],\ \mathbf{R})$ que l'on notera plus brièvement $\mathcal{C}$, des applications continues de $[0,\ 1]$ dans $\mathbf{R}$ a la puissance c. Comparer cette puissance à celle de l'ensemble $\mathcal{F}$ de toutes les applications de $[0,\ 1]$ dans $\mathbf{R}$.

Remarque. — $\mathbf{R}$ a pour puissance c. On peut le montrer soit en faisant la bijection $x \in]0,\ 1[\longrightarrow \operatorname{tg} \pi \left(x - \frac{1}{2}\right) \in] -\infty,\ +\infty[$, soit en faisant la bijection de $\mathbf{N} \times [0,\ 1[$ sur $\mathbf{R}$ qui consiste à faire correspondre au couple formé d'un entier et d'un nombre $x \in [0,\ 1[$, le réel admettant n pour partie entière et x pour partie décimale, bijection qui prouve que

$$\text{card. } \mathbf{R} = \text{card. } [0,\ 1] \times \aleph_0 = \text{card. } [0,\ 1]$$

d'après la règle de multiplication des cardinaux.

CHAPITRE II

ESPACES VECTORIELS NOTIONS FONDAMENTALES

§ 1. DEFINITIONS

1. Axiomes de la structure d'espace vectoriel sur un corps.

Nous rappelons que l'on dit qu'un ensemble E a une structure d'espace vectoriel sur un corps K appelé *corps des opérateurs* ou *corps des scalaires* (corps que, dans ce qui suit, nous supposerons commutatif), si :

1) E est un groupe abélien, pour une opération notée $+$,

2) il existe une opération externe :

$$\begin{array}{ccc} K \times E & \longrightarrow & E \\ (\lambda,\ x) & & \lambda x \end{array}$$

qui satisfait aux axiomes suivants (*) :

$$\lambda(x + y) = \lambda x + \lambda y$$
$$\lambda(\mu x) = (\lambda\mu)x$$
$$(\lambda + \mu)x = \lambda x + \mu x$$
$$1x = x$$

ces axiomes entraînent :

$$\forall x \in E \qquad 0x = 0$$
$$\forall \lambda \in K \qquad \lambda 0 = 0$$

Remarque : Les notations utilisées ci-dessus sont classiques. Il importe cependant de remarquer leur ambiguïté : $+$ est suivant les cas le symbole de l'opération interne de E ou le symbole de l'addition de K ; la juxtaposition de deux lettres désigne suivant les cas la multiplication dans K ou l'opération externe de E ; 0 désigne suivant les cas le zéro du groupe additif de E ou le zéro de K. Aucun inconvénient pratique ne résulte de ces confusions, mais il est peut-être avantageux, pour bien comprendre le sens des axiomes et en déduire les premières conséquences, de recourir à un double jeu de notations tel que le suivant :

T opération interne de E,

Θ élément neutre de cette opération,

(*) Nous rappelons qu'un *module* a la même définition qu'un espace vectoriel, à ceci près que l'ensemble des opérateurs n'est plus un corps K, mais un anneau. Si en outre, l'anneau n'est pas unitaire, le quatrième axiome énoncé pour la loi externe disparaît.

. opération externe,
$+$ addition de K,
0 zéro de K
juxtaposition : multiplication de K.

Les axiomes s'écrivent alors :

$$\lambda.(\mu.x) = \lambda\mu.x,$$
$$(\lambda + \mu).x = (\lambda.x) \mathbin{\top} (\mu.x),$$
$$\lambda.(x \mathbin{\top} y) = (\lambda.x) \mathbin{\top} (\lambda.x),$$
$$1.x = x$$

De ces axiomes, on déduit :

$$\lambda.x = (\lambda + 0).x = (\lambda.x) \mathbin{\top} (0.x)$$

d'où $\quad 0.x = \Theta$

et $\quad \lambda.x = \lambda.(x \mathbin{\top} \Theta) = (\lambda.x) \mathbin{\top} (\lambda.\Theta)$

d'où $\quad \lambda.\Theta = \Theta.$

Remarquons encore que les axiomes donnés, quoique classiques, sont surabondants : si l'on suppose que l'opération interne de E est associative, commutative, simplifiable et possède un élément neutre Θ, les axiomes de l'opération externe entraînent d'abord que :

$$x = 1.x = (1 + 0).x = x \mathbin{\top} (0.x)$$

d'où, puisqu'il existe un élément neutre Θ et que l'opération est simplifiable, $0.x = \Theta$; ensuite que

$$(1.x) \mathbin{\top} [(-1).x] = [1 + (-1)].x = 0.x = \Theta$$

donc que x a $(-1.)x$ pour opposé, ce qui entraîne que E est bien un groupe pour l'opération interne.

Les espaces vectoriels les plus simples et les premiers à avoir été considérés sont ceux qui sont liés à la géométrie euclidienne dans laquelle ils jouent un rôle essentiel (le théorème de Thalès, par exemple, exprime la distributivité de l'opération externe par rapport à l'opération interne : la somme « vectorielle »). Les éléments de ces espaces sont appelés *vecteurs* (*), d'où vient le terme d'*espace vectoriel.* Mais la structure d'espace vectoriel joue un rôle peut-être plus important encore en analyse, où la plupart des espaces fonctionnels qu'on considère sont des espaces vectoriels sur **R** ou **C**. Cela tient d'une part à ce que (comme nous l'avons vu, exercice 37, *Cours A.P.M.* 1), l'ensemble $\mathcal{F}(A, E)$ des applications d'un ensemble A quelconque dans un espace vectoriel sur un corps K a une structure naturelle d'espace vectoriel sur K, et d'autre part à ce que la plupart des sous-ensembles intéressants de $\mathcal{F}(A, E)$, caractérisés par des propriétés relatives à d'autres structures de A et de E (applications continues, dérivables, linéaires...), en sont des sous-espaces vectoriels.

Dans le cas particulier où A est l'ensemble fini $\{1, ..., n\}$ et où E est le corps K lui-même, une application de A dans K est la donnée de n éléments distincts de K, c'est-à-dire un élément de K^n. K^n peut donc, en tant qu'espace du type $\mathcal{F}(A, K)$, recevoir une structure d'espace vectoriel sur K, dont les opérations sont définies par :

$$(\lambda_1, ..., \lambda_n) + (\mu_1, ..., \mu_n) = (\lambda_1 + \mu_1, ..., \lambda_n + \mu_n)$$
$$\lambda(\lambda_1, ..., \lambda_n) = (\lambda\lambda_1, ..., \lambda\lambda_n).$$

On retrouve les espaces vectoriels liés à la géométrie euclidienne, le plan rapporté à 2 axes de coordonnées comme espace $\mathbf{R}^2$, l'espace usuel rapporté à 3 axes comme espace $\mathbf{R}^3$.

(*) Le terme « vecteur » peut être utilisé, sans que ce soit obligatoire, pour désigner un élément d'un espace vectoriel quelconque.

2. Homomorphisme d'espaces vectoriels ou applications linéaires.

E et F étant deux espaces vectoriels sur le même corps K, on dira que l'application f de E dans F est une application linéaire si elle est un homomorphisme pour la structure d'espace vectoriel, ce qui équivaut à :

$$\forall (x, y) \in E^2 \quad f(x + y) = f(x) + f(y)$$
$$\forall \lambda \in K \quad \forall x \in E \quad f(\lambda x) = \lambda f(x).$$

Il faut remarquer ici que, s'il ne s'agit pas d'un endomorphisme, c'est-à-dire si $E \neq F$, les signes $+$ du premier et du deuxième membre des égalités ci-dessus ne sont pas relatifs à la même opération, de même que les multiplications λx et $\lambda f(x)$.

3. Sous-espaces vectoriels.

Soit une partie H de E, stable pour les opérations définissant la structure d'espace vectoriel, c'est-à-dire telle que :

$$H + H \subset H \qquad (1)$$
$$\forall \lambda \in K \quad \lambda H \subset H \qquad (2)$$

Il résulte de cette définition que la structure d'espace vectoriel de E induit sur H une structure d'espace vectoriel. En effet, pour que H jouisse de cette propriété, il faut qu'il soit un sous-groupe. Nous avons vu (*Cours de l'A.P.M.* 1, II, 2, 5) que la condition nécessaire et suffisante pour qu'il en soit ainsi était (en passant de la notation multiplicative, utilisée alors, à la notation additive) que :

$$H + H \subset H \qquad (-H) \subset H\ ;$$

mais, puisque,

$$\forall x \qquad -x = (-1)x$$
$$-H = (-1)H$$

et la condition $(-H) \subset H$ est incluse dans la condition (2). Les conditions (1) et (2) sont donc bien suffisantes pour que la structure induite sur H soit celle d'un espace vectoriel sur K (*). *On dira que* H *est un sous-espace vectoriel de* E.

Remarque : 0 appartenant à H, $H + H \supset H + 0 = H$, donc :

$$H + H = H.$$

D'autre part, $\lambda^{-1}H \subset H \implies H \subset \lambda H$, donc $H = \lambda H$.

Exercice 4. — Dans ce qui suit toutes les lois de composition sont supposées commutatives. Répondre aux questions suivantes : en cas de réponse affirmative justifier la réponse ; en cas de réponse négative, donner un contre-exemple.

1° Tout groupe abélien est un module sur **Z**.
Un sous-groupe est-il un sous-module ?
Un sous-module est-il un sous-groupe ?

2° Tout anneau est un module sur lui-même.
Un sous-anneau est-il un sous-module ? un idéal ?
Un sous-module est-il un sous-anneau ? un idéal ?
Un idéal est-il un sous-anneau ? un sous-module ?

3° Un corps K est un espace vectoriel sur un de ses sous-corps k.
Un autre sous-corps de K est-il un sous-espace vectoriel sur k ?
Un sous-espace vectoriel de K est-il un sous-corps de K ?

(*) Nous avons donc établi qu'une partie stable d'un espace vectoriel est un sous-espace vectoriel. Rappelons qu'au contraire, être une partie stable d'un groupe (respectivement d'un anneau ou d'un corps) n'est pas suffisant pour être un sous-groupe (resp. sous-anneau ou sous-corps).

Théorème : L'intersection d'une famille de sous-espaces vectoriels est un sous-espace vectoriel.

Soit H_i ($i \in I$) une famille de sous-espaces vectoriels de E ; $H = \bigcap_{i \in I} H_i$ est un sous-espace vectoriel de E comme il est immédiat de le vérifier.

4. Sous-espace vectoriel engendré par une partie A de E.

Soit A une partie quelconque de E. La famille des sous-espaces vectoriels qui contiennent A n'est pas vide, — elle comprend au moins E. Leur intersection H est un sous-espace vectoriel qui contient A et est inclus dans les autres. C'est le plus petit sous-espace vectoriel qui contient A ; on dit qu'il est *engendré par* A, *ou que* A *est un système de générateurs de* H.

CONSTRUCTION DE H.

a) *Combinaisons linéaires finies :* Les conditions (1) et (2) entraînent :

$$\forall (x, y) \in A \times A \qquad x + y \in H$$
$$\forall \lambda \in K \quad \forall x \in A \qquad \lambda x \in H.$$

Il en résulte que toute combinaison linéaire finie d'éléments de A appartient à H, c'est-à-dire que, pour tout entier positif n, pour toute famille finie $x_1, ..., x_n$, d'éléments de A, pour toute famille finie $\lambda_1, ..., \lambda_n$, d'éléments de K, l'élément de E égal à :

$$\lambda_1 x_1 + \ldots\ldots + \lambda_n x_n$$

est un élément de H.

Une manière commode d'écrire les combinaisons linéaires finies d'éléments de A consiste à attribuer à chaque élément x de A un élément λ_x de K, en imposant que les λ_x soient nuls, sauf pour un nombre fini d'indices, et de désigner la combinaison linéaire finie correspondante par :

$$\sum_{x \in A} \lambda_x x \qquad \text{ou} \qquad \sum \{ \lambda_x x \,;\, x \in A \}$$

Une autre écriture consiste à indexer les éléments de A en utilisant un ensemble d'indice I. Mais il est alors avantageux de supposer que chaque élément est caractérisé par son indice, autrement dit que l'application de I dans A est bijective. La notation des combinaisons linéaires finies d'éléments de A sera dans ce cas :

$$\sum_{i \in I} \lambda_i x_i \qquad \text{ou} \qquad \sum \{ \lambda_i x_i \,;\, i \in I \}$$

avec toujours la convention fondamentale qu'un nombre fini seulement de λ_i sont différents de zéro.

b) H est l'ensemble des combinaisons linéaires finies d'éléments de A. En effet, nous venons de voir que, si H contient A, il contient cet ensemble, mais comme cet ensemble est un sous-espace vectoriel de E, en vertu de :

$$\sum \{ \lambda_x x \,;\, x \in A \} + \sum \{ \mu_x x \,;\, x \in A \} = \sum \{ (\lambda_x + \mu_x) x \,;\, x \in A \}$$
$$\text{et} \qquad \lambda \sum \{ \lambda_x x \,;\, x \in A \} = \sum \{ \lambda \lambda_x x \,;\, x \in A \}$$

H qui est le plus petit sous-espace vectoriel contenant A lui est identique.

§ 2. GENERATEURS. PARTIES LIBRES. BASES

1. Générateurs et parties libres.

Soit S un système de générateurs d'un espace vectoriel E. Il est clair que toute partie contenant S est un système de générateurs de E ; mais il peut arriver que S soit inutilement grand (si, par exemple, S contient x_1, x_2 et $x_1 + x_2$, E serait le même si on retirait de S un de ces trois éléments). Il est donc naturel de chercher un *système minimal de générateurs* d'un espace E donné, c'est-à-dire un système tel que, si on en retranche un seul générateur, l'espace E ne soit plus engendré. Supposons donc que nous ayons un tel système minimal :

$$S = \{ x_i \,;\, i \in I \}$$

et que nous en retranchions un élément x_{i_0}. Nous considérons le système restant :

$$S' = \{ x_i \,;\, i \in I, \quad i \neq i_0 \}.$$

Si on avait :

(1) $\quad x_{i_0} = \sum \lambda_i x_i \quad$ ($\lambda_i = 0$, sauf pour un nombre fini d'indices),

tout élément de E, combinaison linéaire des éléments de S, serait aussi combinaison linéaire des éléments de S' et S' serait un système de générateurs de E, contrairement à l'hypothèse. Une égalité du type (1) est donc impossible pour un élément x_{i_0} quelconque appartenant au système minimal de générateurs. Il en résulte qu'une égalité du type :

(2) $$\sum \lambda_i x_i = 0$$

où au moins un des λ_i ne serait pas nul est impossible, car soit i_0 l'indice correspondant, en multipliant l'égalité par $\lambda_{i_0}^{-1}$, on obtiendrait :

$$x_{i_0} = - \sum \lambda_i \lambda_{i_0}^{-1} x_i = - \sum \mu_i x_i$$

ce qui, nous venons de le montrer, est impossible.

Autrement dit, si les x_i forment un système minimal de générateurs de E, dans une égalité du type (2), il est impossible qu'un quelconque des λ_i ne soit pas nul. Ou encore :

$$\sum \lambda_i x_i = 0 \implies \forall i, \lambda_i = 0.$$

Une partie d'un espace vectoriel telle qu'aucune combinaison linéaire de ses éléments ne puisse être nulle sans que tous les coefficients de cette combinaison le soient est appelée une partie libre de l'espace vectoriel. On dit encore que *ses éléments sont linéairement indépendants.* Nous pouvons énoncer :

Un système minimal de générateurs est une partie libre.

En particulier un élément différent de 0 est une partie libre car $\lambda x = 0$, avec $\lambda \neq 0$, entraînerait $\lambda\lambda^{-1} x = 0$, donc $x = 0$ (*).

(*) Observons que c'est dans cette question que la *théorie des espaces vectoriels se sépare de celles des modules*, les raisonnements précédents ayant supposé l'existence de $\lambda_{i_0}^{-1}$
En particulier, dans un module, on peut avoir $\lambda x = 0$, sans avoir ni $x = 0$, ni $\lambda = 0$; exemple : soit l'anneau $\mathbb{Z}/6\,\mathbb{Z}$ considéré comme module sur $\mathbb{Z}$:

$$3 \times \dot{2} = \dot{0}$$

De même, un système minimal de générateurs pourra ne pas être une partie libre : exemple, dans $\mathbb{Z}$ considéré comme module sur lui-même, deux nombres premiers entre eux, a et b (tous deux différents de 1) constituent un système de générateurs de $\mathbb{Z}$ tout entier en vertu de l'identité de Bezout ; et ce système est minimal car a (resp. b) n'engendrerait que l'ensemble des multiples de a (resp. de b). Or, on a l'égalité $ab + (-b)\,a = 0$, donc $\{ a, b \}$ n'est pas une partie libre.

2. Bases.

Cherchons à caractériser une *partie libre maximale,* c'est-à-dire telle que l'adjonction d'un seul élément lui fasse perdre sa qualité de partie libre. Soit $A = \{ x_i \,;\, i \in I \}$ une telle partie. Soit $x \in E$, $x \notin A$, et considérons $A \cup \{ x \}$. Par hypothèse, ce n'est pas une partie libre, donc il existe des λ non tous nuls tels que :

$$\lambda x + \sum_{i \in I} \lambda_i x_i = 0.$$

Parmi les λ non nuls, il y a certainement le premier, sans quoi A ne serait pas libre. En multipliant par λ^{-1}, on tire de cette égalité :

$$x = \sum_{i \in I} \mu_i x_i.$$

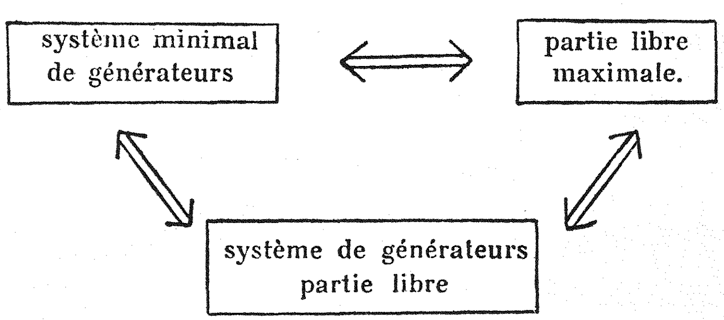
x étant un élément arbitraire de E, on a démontré que tout $x \in E$ appartient au sous-espace vectoriel engendré par A ; donc que A est un système de générateurs. Nous énoncerons : *une partie libre maximale est un système de générateurs.*

Réciproquement, on peut affirmer :

1) Une partie libre qui est un système de générateurs est un système minimal, car s'il n'était pas minimal, c'est qu'un au moins des éléments serait engendré par les autres, donc en serait une combinaison linéaire.

2) Un système de générateurs qui est une partie libre est une partie libre maximale puisque tout élément qu'on pourrait lui ajouter est engendré par les siens, donc en est une combinaison linéaire.

Finalement, nous avons les équivalences suivantes :



On appelle *base d'un espace vectoriel,* un système d'éléments qui vérifie ces différentes propriétés et qui est donc à la fois : 1° partie libre maximale, 2° système de générateurs minimal ; et nous arrivons à la propriété fondamentale suivante :

Tout élément d'un espace vectoriel peut s'exprimer de manière unique comme combinaison linéaire d'éléments d'une base.

Il s'exprime, en effet, comme combinaison d'éléments de la base, puisque celle-ci est un système de générateurs ; cette expression est unique, car si on a :

$$x = \sum_{\in I} \lambda_i x_i = \sum_{i \in I} \mu_i x_i$$

(les indices des λ et μ non nuls n'étant pas nécessairement les mêmes), on en déduit :

$$\sum_{i \in I} (\lambda_i - \mu_i)\, x_i = 0$$

d'où, pour tout i, $\lambda_i = \mu_i$, puisque les x_i constituent une partie libre.

Cette propriété est caractéristique d'une base.

Soit en effet $S = \{ x_i \,;\, i \in I \}$ un système d'éléments qui y satisfait. Puisque tout x peut s'exprimer comme combinaison linéaire des éléments de S, c'est que S est un système de générateurs. D'autre part, c'est une partie libre car si on avait :

$$\sum \lambda_i x_i = 0$$

avec des λ non tous nuls, on pourrait donner de tout élément x deux expressions distinctes :

$$x = \sum \mu_i x_i \qquad x = \sum (\mu_i + \lambda_i) x_i$$

ce qui est contraire à l'hypothèse.

3. Existence des bases.

Nous allons maintenant montrer que *tout espace vectoriel possède une base* (*).

D'abord, un espace vectoriel possède toujours des parties libres (puisque tout vecteur différent de zéro est une partie libre). Considérons alors l'ensemble $\mathcal{L}$ de toutes les parties libres d'un espace vectoriel E ; il est ordonné par inclusion. Il satisfait aux hypothèses du théorème de Zorn. En effet, soit :

$$\Lambda \subset \mathcal{L}$$

une famille totalement ordonnée de parties libres L.

Si $M = \bigcup \{ L \,;\, L \in \Lambda \}$ est une partie libre, elle est un majorant de l'ensemble Λ dans $\mathcal{L}$. Vérifier que c'est une partie libre, c'est vérifier que toute combinaison linéaire finie :

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$$

d'éléments $x_1, x_2 \dots x_n$ de M ne peut être nulle que si les λ_i le sont. Or :

$$x_i \in M \implies \exists\, L_i \in \Lambda \quad x_i \in L_i$$

$L_1, L_2, \dots L_n$ sont en nombre fini et puisqu'elles appartiennent à une chaîne, l'une d'entre elles, soit L_k, contient toutes les autres et contient donc tous les x_n. Mais L_k étant une partie libre,

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0 \implies \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

Donc, M est une partie libre et $\mathcal{L}$ satisfait aux hypothèses du théorème de Zorn. Donc, $\mathcal{L}$ possède un élément maximal qui est une partie libre maximale, c'est-à-dire une base de E. On peut, en particulier, considérer toutes les parties libres qui contiennent une partie libre donnée L. Leur ensemble possédera un élément maximal (il est aisé de voir qu'il satisfait encore aux hypothèses du théorème de Zorn) qui sera une base. De façon plus précise, donnons-nous un système de générateurs S et une partie libre $L \subset S$. On peut considérer $\mathcal{L}$ (L, S) ensemble des parties libres incluant L et incluses dans S. Cet ensemble satisfait encore aux hypothèses du théorème de Zorn. [L'union des parties libres d'une chaîne sera

(*) Il est essentiel de remarquer que ce théorème d'existence d'une base est faux pour les modules : un module peut posséder une base (il est alors dit « module libre »), mais peut ne pas en posséder.

encore une partie libre incluant L et incluse dans S]. Donc, $\mathcal{L}$ (L, S) possède un élément maximal L_0 :

$$L \subset L_0 \subset S.$$

Cet élément est une partie libre maximale dans $\mathcal{L}$ (L, S). *A priori*, il n'est pas sûr qu'elle soit maximale dans $\mathcal{L}$. Mais nous pouvons voir directement qu'elle constitue une base de E. En effet, tout élément de S peut s'exprimer comme combinaison linéaire d'éléments de L_0, sans quoi un élément de S, au moins, pourrait être ajouté à L_0 qui ne serait plus maximal ; S étant un système de générateurs, tout élément de E pourra s'exprimer comme combinaison linéaire d'éléments de S, donc comme combinaison linéaire d'éléments de L_0, et nous pouvons énoncer :

Etant donnée une partie libre L *incluse dans un système de générateurs* S, *il existe une base* B *telle que :*

$$L \subset B \subset S.$$

Corollaire : Soit une partie libre L et un système quelconque de générateurs S ; S ∪ L est, *à fortiori*, un système de générateurs qui contient la partie libre L. On peut donc leur appliquer le théorème précédent. Il existe une base B telle que :

$$L \subset B \subset S \cup L,$$

les éléments de B sont les éléments de L et une partie des éléments de S. *Etant donnés une partie libre et un système de générateurs, on peut compléter la partie libre par des éléments du système de générateurs pour obtenir une base.* Cette proposition est souvent appelée *théorème d'échange.*

Remarque importante : La démonstration générale de l'existence d'une base que nous venons de donner fait appel au théorème de Zorn. Dans des cas particuliers importants, cette référence est inutile :

1) E *possède un système fini de générateurs* $\{ x_1, x_2 \dots x_n \}$. $\{ x_1 \}$ est une partie libre ; si x_2 est indépendant de x_1, on considère la partie libre $\{ x_1, x_2 \}$; sinon, on néglige x_2. Si x_3 est indépendant de la partie libre déjà considérée, on l'ajoute à cette partie ; sinon, on le néglige... et ainsi de suite jusqu'à épuisement des x_i dont on extrait ainsi un système de vecteurs indépendants, tels que tous les autres vecteurs du système de générateurs soient engendrés par eux, autrement dit une base.

2) E *possède un système dénombrable de générateurs.* On peut opérer comme ci-dessus. Pour tout rang *n*, l'élément x_n correspondant sera adjoint à la base en construction ou rejeté suivant qu'il est indépendant des vecteurs de rang inférieur ou qu'il est engendré par eux. On constituera ainsi une base qui pourra être finie ou dénombrable.

§ 3. MISE D'UN ESPACE VECTORIEL SOUS FORME DE SOMME DIRECTE

1. Considérons une partition d'une base B en deux ensembles B_1 et B_2 et les deux sous-espaces V_1 et V_2 engendrés par B_1 et B_2. Il est clair que tout élément x de E peut s'écrire :

$$x = \sum \{ \lambda_a a \, ; \, a \in B \}$$

et peut aussi s'écrire :

$$x = \sum \{ \lambda_a a \, ; \, a \in B_1 \} + \sum \{ \lambda_a a \, ; \, a \in B_2 \}.$$

Sous cette forme, on voit que l'on a $x = x_1 + x_2$ avec $x_1 \in V_1$, $x_2 \in V_2$. Tout élément x de E est la somme d'un élément de V_1 et d'un élément de V_2, ce qui se traduit par :

$$E = V_1 + V_2.$$

On sait que, pour qu'un groupe commutatif, somme de deux de ses sous-groupes, en soit somme directe, il suffit que l'intersection de ces sous-groupes se réduise à l'élément neutre. C'est le cas ici pour V_1 et V_2, car :

$$x = \sum \{ \lambda_a a \, ; \, a \in B_1 \} = \sum \{ \lambda_a a \, ; \, a \in B_2 \}$$

entraînerait $\sum \{ \mu_a a \, ; \, a \in B \} = 0$

avec $\mu_a = \lambda_a$ **si** $a \in B_1$ $\mu_a = -\lambda_a$ si $a \in B_2$.

Comme B est une partie libre, cela entraîne $\forall a \in B$ $\mu_a = 0$; d'où $x = 0$. On a donc :

$$E = V_1 \oplus V_2.$$

2. Généralisation. Somme directe infinie.

La partition de la base B peut être effectuée en plus de deux parties et même en une infinité de parties. Soit $\{ B_j, \, j \in J \}$ l'ensemble de ces parties, indexé bijectivement par l'ensemble J. On a :

$$B = \bigcup \{ B_j \, ; \, j \in J \}$$

$$\text{et } j \neq j' \implies B_j \cap B_{j'} = \varnothing.$$

Soit V_j le sous-espace vectoriel engendré par B_j. Soit $x \in E$, et la famille des scalaires λ_a tels que :

$$x = \sum \{ \lambda_a a \, ; \, a \in B \}.$$

L'ensemble des λ_a non nuls est fini. Il en est de même, *à fortiori*, de l'ensemble des λ_a non nuls, avec $a \in B_j$. On peut poser :

$$x_j = \sum \{ \lambda_a a \, ; \, a \in B_j \}, \text{ on a } x_j \in V_j$$

et le nombre d'indices j pour lesquels x_j est différent de 0 est fini : x est la somme de ces x_j non nuls. On écrira :

$$x = \sum \{ x_j \, ; \, j \in J \}$$

étant entendu que, dans cette somme, un nombre fini seulement de x_j sont différents de zéro.

On voit donc que tout vecteur de E est somme d'un nombre fini de vecteurs non nuls appartenant respectivement à un nombre fini d'espaces V_j. Il est immédiat que cette décomposition est unique. On dit alors que E est somme directe de la famille des sous-espaces V_j, et on écrit :

$$E = \bigoplus_{j \in J} V_j$$

3. Sous-espaces supplémentaires.

Revenons au cas où les B_j sont au nombre de deux :

$$V = V_1 \oplus V_2.$$

On dit que V_1 *et* V_2 *sont supplémentaires.*

Théorème : Etant donné un sous-espace V_1 *de* E, *il existe toujours un sous-espace* V_2 *de* E *tel que :* $V = V_1 \oplus V_2$.

En effet, une base B_1 de V_1 est une partie libre de E. On peut la compléter par un ensemble B_2 d'éléments (pris dans un système arbitraire de générateurs) pour obtenir une base B :

$$B = B_1 \cup B_2.$$

Le sous-espace engendré par B_2 est un supplémentaire de V_1.

Cette propriété essentielle des espaces vectoriels n'est pas vraie dans les modules (*). Elle joue un rôle capital dans l'algèbre des espaces vectoriels.

Le sous-espace V_2 que nous venons de construire n'est, en général, pas unique (**). Un sous-espace V_1 admet généralement plusieurs sous-espaces supplémentaires, et même une infinité si K est infini, ce que nous allons voir dans l'exemple et l'exercice suivants.

Exemple : E est le plan $\mathbf{R}^2$. V_1 est une droite (sous-espace engendré par un vecteur x_1) ; V_1 admet pour supplémentaire toute droite issue de l'origine et distincte de V_1, car tout $x \in \mathbf{R}^2$ peut se décomposer en un vecteur appartenant à V_1 et un vecteur appartenant à cette droite.

Exercice 5. — Un espace V_1 admet une infinité de supplémentaires (au moins, si le corps K est infini). On montrera en effet que si $E = V_1 \oplus V_2$, $\{a_i\}$ constituant une base de V_1 et $\{b_j\}$ est une base de V_2, si a est un élément différent de zéro de V_1 et si on pose :

$$b'_j = a + b_j$$

$\{a_i\} \cup \{b'_j\}$ constitue une base de E et $\{b'_j\}$ engendre un supplémentaire de V_1, soit V'_2, distinct de V_2.

Montrer en outre que deux éléments distincts a et $\bar{a}$ de V_1 permettent de définir par le procédé ci-dessus deux sous-espaces supplémentaires de V_1, soient V'_2 et $\overline{V}'_2$, distincts.

Les sous-espaces supplémentaires d'un même sous-espace jouissent de la propriété fondamentale suivante : *tous les sous-espaces supplémentaires d'un même sous-espace vectoriel sont isomorphes entre eux.*

Soient : $E = V_1 \oplus V_2 \qquad E = V_1 \oplus V'_2$.

Nous allons montrer qu'il existe une bijection entre V_2 et V'_2 qui est un isomorphisme. Tout x_2 appartenant à V_2 peut, puisqu'il est élément de E, se mettre de manière unique sous la forme :

$$x_2 = \xi_1 + x'_2 \qquad \xi_1 \in V_1 \qquad x'_2 \in V'_2.$$

Ceci définit donc une application : $x_2 \longrightarrow x'_2$; cette application est surjective car :

$$\forall x'_2 \in V'_2 \quad \exists \xi_1 \in V_1 \quad \exists x_2 \in V_2 \quad x'_2 = \xi_1 + x_2$$

d'où $x_2 = -\xi_1 + x'_2$, et x'_2 est l'image de l'élément x_2 de V_2.

(*) Il peut même arriver qu'il soit toujours impossible, un sous-module étant donné, d'en trouver un supplémentaire. C'est le cas de $\mathbf{Z}$ (considéré comme module sur lui-même), dont nous avons montré qu'il ne pouvait jamais être somme directe de deux de ses sous-groupes non triviaux (A.P.M. I, II, 3, 1).

(**) Il faut se garder du lapsus : « le » supplémentaire d'un sous-espace, qui dégénère facilement en faute de raisonnement.

Elle est injective car :

$$x_2 = \xi_1 + x'_2 \qquad y_2 = \eta_1 + x'_2$$

entraînerait

$$x_2 - y_2 = \xi_1 - \eta_1,$$

ce qui, $x_2 - y_2$ et $\xi_1 - \eta_1$ étant respectivement éléments de V_2 et V_1 dont l'intersection est réduite à zéro, entraînerait $x_2 = y_2$. Enfin, cette bijection est un isomorphisme car elle est linéaire :

$$x_2 = \xi_1 + x'_2 \qquad y_2 = \eta_1 + y'_2$$

$$x_2 + y_2 = (\xi_1 + \eta_1) + (x'_2 + y'_2)$$

$x'_2 + y'_2$, image de $x_2 + y_2$, est la somme des images ; et de même :

$$x_2 = \xi_1 + x'_2 \qquad \lambda x_2 = \lambda\xi_1 + \lambda x'_2,$$

$\lambda x'_2$ est l'image de λx_2.

L'isomorphisme de V_2 et V'_2 est donc établi.

Exercice 6. — Soit $H_1, H_2 \ldots H_n$ une famille finie de sous-espaces vectoriels d'un espace E telle que :

$$E = H_1 + H_2 + \ldots + H_n$$

Montrer que la condition nécessaire et suffisante pour que E soit somme directe des H_i est que :

$$\forall i \in \{1, 2, \ldots, n-1\} \qquad (H_1 + \ldots H_i) \cap H_{i+1} = \{0\}$$

La condition $i \neq j \implies H_i \cap H_j = \{0\}$ est nécessaire mais non suffisante.

Exercice 7. — Soit A un ensemble et K un corps, on considère l'ensemble $\mathcal{H}$ des applications φ de A dans K qui ne prennent de valeurs différentes de 0 que pour un nombre fini d'éléments de A.

1° Montrer que $\mathcal{H}$ est un sous-espace vectoriel de $\mathcal{F}$ (A, K). On l'appelle l'espace des combinaisons linéaires formelles d'éléments de A à coefficients dans K.

On considère, pour tout élément a de A, l'application φ_a définie par :

$$\varphi_a(a) = 1 \qquad \varphi_a(b) = 0 \qquad \text{pour } b \neq a$$

Montrer que leur ensemble est une base de $\mathcal{H}$.

2° On suppose que A est une partie d'un ensemble vectoriel E sur K. Soit H le sous-espace engendré par A dans E. Montrer qu'il existe une application linéaire de $\mathcal{H}$ sur H.

Dans quel cas est-ce un isomorphisme ?

Exercice 8. — Soit une famille d'espaces vectoriels E_i ($i \in I$) sur un même corps K et $E = \prod E_i$ leur produit cartésien.

1° Montrer qu'on peut attribuer à E une structure d'espace vectoriel liée de façon simple à celle des E_i, et telle qu'il existe alors dans E des sous-espaces vectoriels H_i respectivement isomorphes aux E_i.

2° On peut, dans E, déterminer le sous-espace vectoriel E_0 engendré par la réunion des H_i. Caractériser ce sous-espace. Montrer qu'il est somme directe des H_i.

3° Soit un nombre cardinal, a, inférieur à card I. A tout

$$x = \{x_1, x_2 \ldots x_i \ldots\}$$

on associe le cardinal c_x de l'ensemble des indices $i \in I$, pour lesquels x_i est différent de zéro ; on considère les deux ensembles

$E_a = \{x ; c_x < a\}$, $F_a = \{x : c_x \leqslant a\}$; on précise que ce cardinal c_x est infini.

Montrez que ce sont deux sous-espaces vectoriels et que l'un d'eux n'est autre que E_0, pour une valeur de a que l'on indiquera.

§ 4. EGALITE DES NOMBRES CARDINAUX DES BASES DIMENSION D'UN ESPACE VECTORIEL

L'égalité des nombres cardinaux des différentes bases va être démontrée par récurrence dans le cas où ce nombre est fini. Ce théorème est vrai pour $n = 1$, car si un espace vectoriel admet $\{ a_1 \}$ comme base, c'est que tout x de l'espace est de la forme λa_1. Un système de plus de deux vecteurs $\{ a_1, \lambda_1 a_1 \}$ ne peut être une partie libre puisque :

$$\lambda_1(\lambda a_1) - \lambda(\lambda_1 a_1) = 0$$

(inversement, tout vecteur $a_2 = \lambda_2 a_1$, avec $\lambda_2 \neq 0$, est une base car $x = \frac{\lambda}{\lambda_2} a_2$).

Admettons donc que le théorème soit vrai pour $n - 1$, c'est-à-dire que si un espace vectoriel a une base de $n - 1$ éléments, toute autre base aura $n - 1$ éléments. Soit alors un espace E possédant une base B_2 ayant $n_2 = n$ éléments et une autre base B_1 de n_1 éléments. Si n_1 était inférieur à n_2, l'hypothèse de récurrence s'appliquerait à E et à la base B_1, ce qui conduirait à une contradiction. Nous avons donc $n_1 \geqslant n_2 = n$. Soit alors $a_1 \in B_1$. D'après le théorème d'échange, nous pouvons construire une base B' telle que :

$$\{ a_1 \} \subset B' \subset \{ a_1 \} \cup B_2.$$

B' contient a_1 et des éléments de B_2 ; mais elle ne peut les contenir tous sans quoi elle ne serait plus une partie libre. B' contient donc au plus $n - 1$ éléments de B_2.

Décomposons alors la base B_1 en $\{ a_1 \}$ et $B_1 - \{ a_1 \}$ et décomposons la base B' en $\{ a_1 \}$ et $B' - \{ a_1 \}$; ces décompositions de base conduisent à deux décompositions de E en somme directe :

$$E = Ka_1 \oplus V_1$$
$$E = Ka_1 \oplus V',$$

Ka_1, V_1 et V' étant les sous-espaces vectoriels engendrés respectivement par a_1, $B_1 - \{ a_1 \}$ et $B' - \{ a_1 \}$; $V_1 \approx V'$ comme supplémentaires d'un même sous-espace. Mais $B' - \{ a_1 \}$ a un nombre d'éléments p au plus égal à $n - 1$. Toutes les bases de V' ont p éléments (hypothèse de récurrence) et l'isomorphisme exige que les bases de V_1 aient aussi p éléments. Donc :

$$n_1 = p + 1 \leqslant n \text{ d'où } n_1 = n_2.$$

La propriété est encore vraie si le cardinal des bases est infini. Nous ne le démontrerons pas.

Ce nombre, cardinal commun à toutes les bases d'un espace vectoriel, est appelé *dimension de l'espace ;* il est noté « dim E ».

Remarque 1 : Lorsque le même ensemble peut être considéré comme espace vectoriel sur plusieurs corps (en particulier, si E a une structure d'espace vectoriel sur K, il a aussi une structure d'espace vectoriel sur tout sous-corps k de K), les dimensions de ces diverses structures sont en général différentes. Exemples : 1° un espace de dimension n sur **C** est de dimension $2n$ sur **R** ; — 2° **R** est de dimension 1 quand on le considère comme espace vectoriel sur lui-même et est de dimension infinie quand on le considère comme espace vectoriel sur **Q** ; — 3° si K est un corps, E une extension algébrique simple de degré n de K, et F une extension algébrique simple de degré m de E, F est un espace vectoriel de dimension m sur E et un espace vectoriel de dimension $m.n$ sur K.

Remarque 2 : Dans un espace vectoriel de dimension n, toutes les bases ont n éléments. Il en résulte que, dans un tel espace, *tout système de n générateurs est une base* (puisque la base qu'on sait pouvoir en extraire doit avoir n éléments), *toute partie libre à n éléments est une base* (puisque la base qu'on peut obtenir en lui adjoignant éventuellement d'autres éléments doit avoir n éléments). Cette remarque est pratiquement très utile.

Remarque 3 : Il est clair aussi que si un espace vectoriel est somme directe de deux sous-espaces vectoriels, sa dimension est la somme de celles des sous-espaces. Réciproquement, si deux sous-espaces ont une intersection réduite à $\{ 0 \}$ et si la somme de leurs dimensions est celle de l'espace, la réunion d'une base de l'un et d'une base de l'autre constitue une base de l'espace : ce sont deux sous-espaces supplémentaires et l'espace en est somme directe.

La dimension des sous-espaces supplémentaires d'un sous-espace V d'un espace E est appelée la *codimension* de V. On a :

$$\text{codim } V = \dim E - \dim V.$$

Remarque 4 : Deux espaces E et F de même dimension (finie ou non) sont évidemment isomorphes. L'égalité des dimensions, c'est-à-dire des cardinaux des bases, signifie l'existence d'une bijection φ d'une base de E sur une base de F. Il est clair alors que l'application qui à $\Sigma \lambda_a a$ fait correspondre $\Sigma \lambda_a \varphi(a)$ est un isomorphisme de E sur F.

Plus spécialement, tous les espaces de dimension n sur un même corps K sont isomorphes entre eux, et en particulier à K^n.

Exercice 9. — On considère l'espace vectoriel $P_n \subset \mathbf{R}[x, y]$ des polynômes à deux indéterminées, à coefficients réels, de degré inférieur ou égal à n. En indiquer une base et la dimension. Dire si les sous-ensembles suivants sont des sous-espaces vectoriels. Si oui, en trouver un supplémentaire ; indiquer une base de chacun de ces sous-espaces. Vérifier que la dimension de P_n soit bien la somme des dimensions des sous-espaces supplémentaires :

- — Polynômes de degré fixé $n_0 < n$.
- — Polynômes homogènes.
- — Polynômes homogènes de degré fixé $n_0 < n$.
- — Polynômes de degré inférieur ou égal à n à une indéterminée.
- — Polynômes pairs $[P(x, y) = P(-x, -y)]$.
- — Polynômes symétriques $[P(x, y) = P(y, x)]$.
- — Polynômes de degré inférieur ou égal à $p < n$.
- — Polynômes s'annulant pour $x = a$, $y = b$.
- — Polynômes tels que $P(x, 1) = P(x, 0) + 1$.

Exercice 10. — On rappelle qu'étant donnée une fonction réelle f deux fois dérivable de deux variables réelles x et y, on nomme laplacien de cette fonction la fonction :

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

et que la foncton est dite harmonique si son laplacien est nul.

P étant un polynôme homogène de degré n, on demande de montrer que si la fonction $(x^2 + y^2)^\lambda P(x, y)$ est harmonique, $P(x, y)$ est divisible par $x^2 + y^2$ et d'en déduire que $P = 0$.

2° Soit $\mathcal{P}_n$ l'espace vectoriel des polynômes homogènes de degré n en x et y ; $\mathcal{H}_n$ le sous-espace des polynômes harmoniques de degré n ; $\mathcal{A}_n$ le sous-espace des polynômes de degré n divisibles par $x^2 + y^2$.

Montrer que $\mathcal{P}_n$ est somme directe de $\mathcal{H}_n$ et $\mathcal{A}_n$ (on commencera par déterminer les dimensions respectives de $\mathcal{H}_n$ et $\mathcal{A}_n$).

3° En déduire que tout polynôme Q en x et y, homogène ou non, peut s'écrire d'une façon et d'une seule, sous la forme d'une somme finie.

$$Q = H_0 + AH_1 + A^2H_2 + \ldots + A^pH_p$$

où les H_p sont des polynômes harmoniques, et où A désigne le polynôme $x^2 + y^2$ (commencer par le cas où Q est homogène).

4° Soit E la circonférence $x^2 + y^2 = 1$; appelons « fonction polynominale sur E », la restriction à E d'un polynôme Q (x, y).

Montrer que toute fonction polynominale sur E se prolonge dans tout le plan par un polynôme harmonique (unique).

5° (Question faisant appel à des connaissances étrangères à ce cours). Montrer que pour toute fonction continue φ sur E, et pour tout $\varepsilon > 0$, il existe un polynôme harmonique qui, sur E, approxime φ à ε près.

En déduire l'existence d'une fonction continue dans le disque fermé $x^2 + y^2 \leqslant 1$, qui prolonge φ et qui est harmonique à l'intérieur du disque.

Exercice 11. — Soit un espace vectoriel E et soit $\mathcal{V}(E)$ l'ensemble de tous les sous-espaces vectoriels de E. On ordonne cet ensemble par inclusion. Vérifier que $\mathcal{V}(E)$ est un treillis, le sup de deux éléments étant leur somme (qui n'est pas, en général, une somme directe), et que ce treillis n'est pas distributif, c'est-à-dire que les deux opérations sup et inf ne sont pas distributives l'une par rapport à l'autre.

Démontrer que :

$$\dim (V_1 + V_2) + \dim (V_1 \cap V_2) = \dim V_1 + \dim V_2$$

CHAPITRE III

APPLICATIONS LINÉAIRES

§ 1. FACTORISATION CANONIQUE

Nous avons donné (II, 1, 2) la définition des applications linéaires. On peut toutefois en donner une définition un peu plus générale en admettant que E et F sont deux espaces vectoriels sur deux corps K et K' différents tels que K' soit isomorphe à K dans un isomorphisme φ (*). Les conditions de linéarité de l'application f seraient alors :

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

Mais, dans ce qui suit, nous nous bornerons au cas où $K = K'$ et où φ est l'application identique.

A l'homomorphisme f, nous allons appliquer les résultats généraux relatifs aux homomorphismes (*Cours A.P.M. I*, III, 1, 6), et d'abord le fait que $f(E) \subset F$ est un espace vectoriel et que $f : E \longrightarrow f(E)$ est un homomorphisme surjectif (**).

Nous savons alors que si R est la relation d'équivalence définie sur E par :

$$x R y \Longleftrightarrow f(x) = f(y),$$

si φ est l'application canonique de E sur E/R, g l'application de E/R sur $f(E)$ telle que :

$$f = g \circ \varphi,$$

il est possible de donner à E/R une structure d'espace vectoriel telle que φ soit un homomorphisme et g un isomorphisme. La structure intervenant ici étant celle d'un espace vectoriel est d'abord une structure de groupe commutatif. La relation $f(x) = f(y)$ s'écrit :

$$x - y \in \overset{-1}{f}(0)$$

où $\overset{-1}{f}(0)$ est le *noyau* V de l'homomorphisme et est un sous-groupe de E.

Mais ici V est en outre un sous-espace vectoriel. En effet :

$$x \in V \Longrightarrow f(x) = 0 \Longrightarrow \lambda f(x) = 0 \Longrightarrow f(\lambda x) = 0 \Longrightarrow \lambda x \in V$$

(*) On sait que tout homomorphisme surjectif de corps est un isomorphisme.

(**) Dans les pages citées, nous nous étions d'ailleurs placés dans des hypothèses un peu plus générales. Ne supposant pas *à priori* que F est un espace vectoriel, mais seulement que c'était un ensemble muni d'une loi de composition interne et d'une loi externe (avec K comme ensemble d'opérateurs), nous avions montré que $f(E) \subset F$ recevait une structure d'espace vectoriel sur K.

On notera E/V l'espace vectoriel-quotient et on obtient le schéma suivant :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[dr, "\varphi"'] & f(E) \subset F \\ & E/V \arrow[u, "g"'] \end{tikzcd}$$

où f, φ, g sont trois applications linéaires, E, E/V, f(E) trois espaces vectoriels.

Réciproquement, si on se donne un sous-espace V de E, la relation $x - y \in V$ est une relation d'équivalence R et l'ensemble quotient E/R noté E/V est image homomorphe de E dans l'application canonique $x \longrightarrow \dot{x}$.

L'homomorphisme pour la somme résulte des théorèmes généraux relatifs aux groupes puisque V est sous-groupe (le groupe étant commutatif, la question de l'invariance ne se pose pas).

Quant à l'homomorphisme pour la loi externe, il s'obtient en remarquant que $\dot{x} = x + V$ et que si on pose :

$$\lambda\dot{x} = \{\,\lambda x : x \in \dot{x}\,\} \qquad \lambda\dot{x} = \lambda x + \lambda V$$

mais $\lambda V = V$, car V est sous-espace vectoriel, donc $\lambda\dot{x} = \dot{\widehat{\lambda x}}$.

Considérons alors un sous-espace supplémentaire de V, soit W. Nous avons vu (*Cours A.P.M. I*, II, 3, 1) que si un groupe G est produit direct de deux sous-groupes, chacun est isomorphe au groupe quotient de G par l'autre. Appliquée au groupe additif de E, somme directe de V et W, cette propriété montre que W est, en tant que groupe additif, isomorphe au groupe additif E/V dans l'application :

$$\varphi : x \in W \longrightarrow \dot{x} = x + V \in E/V \qquad (*)$$

D'autre part, nous venons de montrer que $\lambda\dot{x} = \dot{\widehat{\lambda x}}$, donc φ qui, à $\lambda x \in W$ fait correspondre $\dot{\widehat{\lambda x}} = \lambda\dot{x}$, est un homomorphisme pour l'opération externe. L'isomorphisme complet est donc établi entre W et E/V :

$$W \approx E/V.$$

L'espace quotient d'un espace vectoriel modulo un sous-espace est isomorphe à tout supplémentaire de ce sous-espace.

Exemple dans $\mathbf{R}^2$ (fig. 2). Soient V et W deux sous-espaces supplémentaires de dimension 1, x (origine O, extrémité m) un élément de W. Sa classe modulo V, $\dot{x}$ est l'ensemble des vecteurs d'origine O et ayant leur

---

(*) On peut redémontrer directement cette propriété ici ; d'abord :

$$\varphi : x \in W \longrightarrow \dot{x} = x + V \in E/V$$

est un homomorphisme pour la somme car :

$$\dot{x} + \dot{y} = x + V + y + V = x + y + V = \dot{\widehat{x + y}}\ ;$$

ensuite φ est injective, car si $x \in W$, $y \in W$ pour que x et y aient la même image, il faudrait que $x\text{-}y \in V$, mais comme $x\text{-}y \in W$ et que $V \cap W = \{\,0\,\}$, il en résulte que $x = y$. Enfin, φ est surjective, car si $\dot{\xi} \in E/V$, c'est qu'il existe $\xi \in E$ tel que $\dot{\xi} = \xi + V$, mais puisque $E = V \oplus W$, $\xi = x + y$, $x \in W$, $y \in V$ donc $\dot{\xi} = \dot{x}$ et $\dot{\xi}$ est image de x par φ.

φ, homomorphisme bijectif est bien un isomorphisme.

extrémité sur la droite D parallèle à V passant par m. L'espace de ces classes, muni de la structure quotient, est isomorphe à W ; ce qui signifie, pour l'opération interne, rappelons-le, que si x_1 et x_2 sont deux vecteurs de W, la somme d'un vecteur $y_1 \in \dot{x}_1$ et d'un vecteur $y_2 \in \dot{x}_2$ appartient à la classe de $x_1 + x_2$, c'est-à-dire a son extrémité sur la droite parallèle à W issue de l'extrémité μ de $x_1 + x_2$.

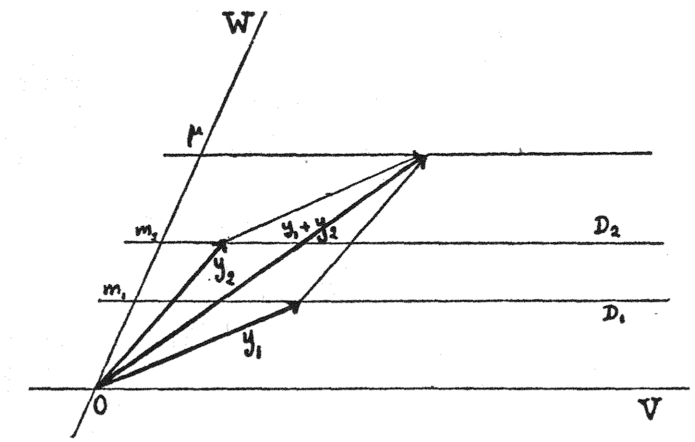


FIG. 2

EN ?

Reprenons l'étude générale. On sait qu'il existe un isomorphisme : $E/W \approx f(E)$. Nous pouvons donc conclure :

L'image d'un espace vectoriel par une application linéaire est isomorphe à tout supplémentaire du noyau de l'application.

Considérons la restriction de f à W, soit g. $\overset{-1}{g}(0)$ est réduit à zéro, d'après la définition de W, donc g est bijective.

Théorème fondamental : La restriction d'une application linéaire à tout supplémentaire de son noyau en est un isomorphisme sur l'image.

Si E *est de dimension finie*, nous avons :

$$\dim E = \dim V + \dim W \qquad \text{avec} \quad W \approx f(E)$$

d'où

$$\dim E = \dim \overset{-1}{f}(0) + \dim f(E)$$

ou

$$\operatorname{codim} \overset{-1}{f}(0) = \dim f(E).$$

Dans le cas d'un isomorphisme, $\dim f(E) = \dim E$. Et, réciproquement, si f est une application de E dans F telle que f(E) ait même dimension que E, il en résulte que la dimension du noyau est zéro.

La dimension de f(E) est appelée *rang de l'application linéaire*.

Exercice 12. — Pour chacune des applications linéaires suivantes trouver le noyau, un supplémentaire du noyau et vérifier que la restriction de l'application à ce supplémentaire est un isomorphisme sur l'image.

Espace de départ	Application	Espace d'arrivée
Fonctions d'une variable x définies sur [0,1] et dérivables.	$f \longrightarrow \frac{df}{dx}$	Fonctions d'une variable x définies sur [0, 1].
Fonctions d'une variable x définies sur [0,1] et p fois dérivables.	$f \longrightarrow \frac{d^pf}{dx^p}$	*Idem*
Fonctions d'une variable x continues sur [0,1].	$f \longrightarrow g$ avec $g(x) = \int_0^x f(t)dt$	Fonctions d'une variable x définies sur [0, 1] et dérivables.
Idem	$f \longrightarrow \int_0^1 f(x)dx$	**R**
Suites infinies indexées par **Z**	$\{u_n\} \to \{v_n = u_n - u_{n-1}\}$	Suites infinies indexées par **Z**.

Exercice 13. — 1° E et F étant deux espaces vectoriels de même dimension finie n, montrez que les quatre affirmations suivantes relatives à une application linéaire f de E dans F sont équivalentes :

— f est un isomorphisme de E sur F ;
— f est surjective ;
— f est injective ;
— f est de rang n.

2° Prouver par des contre-exemples que, dans le cas d'espaces de dimensions infinies, une application linéaire,

a) peut être injective sans être un isomorphisme ;
b) peut être surjective sans être un isomorphisme.

§ 2. ESPACES D'APPLICATIONS LINEAIRES

1. Structure d'espace vectoriel de $\mathcal{L}$ (E, F).

Soient E et F deux espaces vectoriels sur le même corps K et l'ensemble $\mathcal{L}$(E, F) de toutes les applications linéaires de E dans F qui est une partie de $\mathcal{F}$ (E, F), ensemble des applications de E dans F. On sait (exercice 37, *Cours A.P.M. I*) que $\mathcal{F}$ a une structure d'espace vectoriel sur K, la somme $f + g$ de deux applications f et g et le produit d'une application f par un scalaire λ de K étant respectivement définis par :

$$\forall x \in E, \quad (f+g)(x) = f(x) + g(x) \qquad (\lambda f)(x) = \lambda f(x).$$

Nous allons chercher si $\mathcal{L}$ est un sous-espace vectoriel de $\mathcal{F}$, ce qui se réduit à chercher si c'est une partie stable pour l'opération interne et l'opération externe, autrement dit à voir si, f et g étant linéaires, $f + g$ et λf le sont.

$\forall (x, y) \in E \times E$

$(f+g)(x+y) = f(x+y) + g(x+y)$	définition de la somme sur $\mathcal{F}$
$= f(x) + f(y) + g(x) + g(y)$	linéarité de f et g
$= f(x) + g(x) + f(y) + g(y)$	commutativité de l'opération interne sur $\mathcal{F}$
$= (f+g)(x) + (f+g)(y)$	définition de la somme sur $\mathcal{F}$

$\forall x \in E \quad \forall \mu \in K$

$(f+g)(\mu x) = f(\mu x) + g(\mu x)$	définition de la somme sur $\mathcal{F}$
$= \mu f(x) + \mu g(x)$	linéarité de f et g
$= \mu[f(x) + g(x)]$	distributivité
$= \mu[(f+g)(x)]$	définition de la somme sur $\mathcal{F}$

Nous avons de même : $\forall (x, y) \in E \times E \quad \forall \lambda \in K.$

$(\lambda f)(x+y) = \lambda f(x+y)$	définition de l'opération externe sur $\mathcal{F}$
$= \lambda[f(x) + f(y)]$	linéarité de f
$= \lambda f(x) + \lambda f(y)$	distributivité
$= (\lambda f)(x) + (\lambda f)\ (y)$	définition de l'opération externe sur $\mathcal{F}$

Enfin, $\forall x \in E \quad \forall \mu \in K.$

$(\lambda f)(\mu x) = \lambda f(\mu x)$	définition de l'opération externe sur $\mathcal{F}$
$= \lambda \mu\, f(x)$	linéarité de f
$= (\lambda \mu) f(x)$	associativité de l'opération externe sur $\mathcal{F}$

Nous voulons comparer cette expression à $\mu(\lambda f)(x)$ qui vaut $\mu\lambda f(x) = (\mu\lambda) f(x)$. On voit que nous ne pouvons conclure à la linéarité de l'application λf que si $\lambda\mu = \mu\lambda$, donc si le corps K est commutatif (*). C'est une des raisons pour laquelle nous avons choisi de nous placer toujours dans l'hypothèse où le corps K est commutatif (**R** et **C** étant commutatifs, presque toutes les applications de l'algèbre linéaire à l'analyse et à la géométrie s'accommoderont de cette restriction).

Moyennant cete hypothèse, nous pouvons conclure de ce qui précède que $\mathcal{L}$ (E, F) *est un espace vectoriel sur* K, sous-espace de l'espace $\mathcal{F}$.

(*) Dans le cas où le corps K n'est pas commutatif, il faut, pour pouvoir conclure que λ (multiplicateur de f), commute avec μ pour tout μ appartenant à K ; λ doit donc appartenir au centre du groupe multiplicatif du corps (cf. exercice 21, Cours A.P.M. I) qui constitue un sous-corps de K, et on pourra seulement dire que l'ensemble $\mathcal{L}$ est un espace vectoriel sur ce centre. Ce n'est pas un sous-espace vectoriel de $\mathcal{F}$.

2. Composition des applications linéaires.

Soient $E \xrightarrow{f} F \xrightarrow{g} G$, trois espaces vectoriels sur un même corps K, f et g deux applications linéaires et $g \circ f$ l'application composée dont on vérifie immédiatement qu'elle est linéaire :

$$g \circ f \in \mathcal{L}(L, G).$$

La composition des applications définit donc une application :

$$\begin{array}{ccc} \mathcal{L}(E, F) \times \mathcal{L}(F, G) & \longrightarrow & \mathcal{L}(E, G) \\ (f, g) & & (g \circ f) \end{array} \qquad (1)$$

du produit cartésien $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$ dans $\mathcal{L}(E, G)$. On vérifie que :

$$\begin{gathered} g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2 \\ (g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f \\ (\lambda g) \circ f = \lambda(g \circ f) \\ g \circ (\lambda f) = \lambda(g \circ f). \end{gathered}$$

Une application $\varphi : E_1 \times E_2 \longrightarrow F$ qui, comme la précédente, vérifie :

$$\left.\begin{array}{l} \forall x_1, y_1 \in E_1 \\ \forall x_2, y_2 \in E_2 \\ \forall \lambda \in K \end{array}\right\} \begin{array}{l} \varphi(x_1 + y_1, x_2) = \varphi(x_1, x_2) + \varphi(y_1, x_2) \\ \varphi(\lambda x_1, x_2) = \lambda\varphi(x_1, x_2) \\ \varphi(x_1, x_2 + y_2) = \varphi(x_1, x_2) + \varphi(x_1, y_2) \\ \varphi(x_1, \lambda x_2) = \lambda\varphi(x_1, x_2) \end{array}$$

est dite une application *bilinéaire* du produit cartésien $E_1 \times E_2$ dans F. Ceci signifie que c'est une application linéaire dans F de l'espace vectoriel obtenu en fixant x_2 et en faisant décrire E_1 à x_1 :

$$\{ (x, x_2) \; ; \; x \in E_1 \}$$

ainsi qu'une application linéaire dans F de l'espace :

$$\{ (x_1, x) \; ; \; x \in E_2 \}$$

ces deux sous-espaces étant d'ailleurs respectivement isomorphes à E_1 et à E_2.

L'application (1) définie par la composition des applications est donc bilinéaire.

3. Endomorphismes.

Plaçons-nous maintenant dans le cas particulièrement important où $E = F = G$. Ces applications linéaires, qui sont des homomorphismes d'un ensemble dans lui-même, sont alors des *endomorphismes*, et, dans le cas où elles sont des isomorphismes, ce sont des automorphismes.

L'application (1) définit alors une loi de composition interne sur $\mathcal{L}(E, E)$ qui est distributive par rapport à l'addition. $\mathcal{L}(E, E)$ a donc une structure d'anneau, mais comme, en outre, il possède une loi externe vérifiant par rapport aux lois internes les axiomes convenables, $\mathcal{L}(E, E)$ a une structure d'algèbre (non commutative). C'est *l'algèbre des endomorphismes d'un espace vectoriel.*

On peut étudier différents sous-ensembles de cette algèbre ; en particulier celui des automorphismes, mais celui-ci ne constitue pas une sous-algèbre car ce n'est pas une partie stable pour l'addition (donnons le contre-exemple suivant : l'automorphisme $x \to (-x)$ et l'automorphisme identique ont pour somme l'application qui, à tout $x \in E$, fait correspondre 0, qui n'est évidemment pas un automorphisme). Cet ensemble des automorphismes est, inversement, stable pour la composition des applications. D'autre part, chaque automorphisme admet pour cette loi un inverse (automorphisme réciproque), donc l'ensemble des automorphismes forme un groupe que l'on nomme *groupe linéaire de l'espace vectoriel* **E.**

Exercice 14. — E, E', F, F' étant quatre espaces vectoriels sur un même corps K montrer que :

$$\left.\begin{array}{l} E \approx E' \\ F \approx F' \end{array}\right\} \Longrightarrow \mathcal{L}(E, F) \approx \mathcal{L}(E', F')$$

Pour cela si φ et ψ sont respectivement les isomorphismes

$$\varphi : E \longrightarrow E' \qquad \psi : F \longrightarrow F'$$

à toute application $f : E \longrightarrow F$, on fera correspondre

$$f' = \psi \circ f \circ \overset{-1}{\varphi}$$

c'est-à-dire l'application choisie de façon que le diagramme

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\varphi"'] & F \arrow[d, "\psi"] \\ E' \arrow[r, "f'"'] & F' \end{tikzcd}$$

soit commutatif et on montrera que $f \longrightarrow f'$ est un isomorphisme.

2) Dans le cas où $E = F$, $E' = F'$ montrer que l'isomorphisme

$$\mathcal{L}(E, E) \approx \mathcal{L}(E', E')$$

s'étend à la structure d'algèbre.

Dans l'isomorphisme envisagé dans cet exercice, à un automorphisme $E \to E$ correspond un automorphisme $E' \to E'$. En effet, un automorphisme est un élément inversible (pour l'opération : composition) de l'ensemble $\mathcal{L}(E, E)$ des endomorphismes et dans l'isomorphisme $\mathcal{L}(E, E) \to \mathcal{L}(E', E')$ l'homologue d'un élément inversible est un élément inversible. Les groupes linéaires de E et E' sont donc isomorphes. En particulier, si E et E' sont des espaces vectoriels ayant la même dimension finie sur le corps K, la structure de leur groupe linéaire est la même. On désigne par :

$$GL_n(K)$$

cette structure de groupe qui est déterminée par n et K. Ces groupes sont appelés *groupes classiques*.

§ 3. DETERMINATION DES APPLICATIONS LINEAIRES. MATRICES

1. Nous allons nous occuper maintenant de la façon dont on peut déterminer une application linéaire $f : E \longrightarrow F$ (E et F, espaces vectoriels sur un même corps). Remarquons d'abord que *l'image d'un système de générateurs est un système de générateurs de* $f(E)$. En effet, si $\{ x_i \}$ est un système de générateurs de E :

$$\forall x \in E \quad x = \Sigma \lambda_i x_i,$$

ce qui entraîne, du fait de la linéarité,

$$f(x) = \Sigma \lambda_i f(x_i).$$

Les $f(x_i)$ constituent donc un système de générateurs. En particulier, l'image d'une base de E est un système de générateurs, mais elle n'est généralement pas une base car *l'image d'une partie libre n'est pas nécessairement une partie libre.* Mais inversement,

$$f(A) \text{ partie libre} \Longrightarrow A \text{ partie libre.}$$

Car si A n'était pas libre, il existerait entre les éléments a_i de A une relation $\Sigma\lambda_i a_i = 0$ (λ_i non tous nuls), ce qui entraînerait une relation $\Sigma\lambda_i f(a_i) = 0$ et $f(A)$ ne serait pas une partie libre.

Ceci étant, pour définir l'application f, nous partirons d'une base de E. *Les images des éléments de la base peuvent être choisies arbitrairement et il est clair que l'application est parfaitement déterminée par la donnée de ces images.* Autrement dit, une base B de E étant choisie, il y a une bijection naturelle entre l'ensemble $\mathcal{L}(E, F)$ des applications linéaires de E dans F et l'ensemble $\mathcal{F}(B, F)$ des applications de B dans F. On voit d'ailleurs aisément que cette bijection est un isomorphisme pour les structures naturelles d'espace vectoriel de $\mathcal{L}(E, F)$ et $\mathcal{F}(B, F)$. On voit en même temps que deux espaces E et F de même dimension sont isomorphes d'une infinité de manières, tout isomorphisme s'obtenant à partir d'un élément de $\mathcal{F}(B, F)$ qui envoie B sur une base de F, ce qui précise la remarque 4 de II, 4.

Exercice 15. — Soit E un espace vectoriel de dimension finie. Si f et g sont deux endomorphismes de E et si gof est l'identité, montrer que f et g possèdent des inverses. Montrer par un exemple que cette conclusion ne s'étend pas au cas où E est de dimension infinie.

Exercice 16. — Soit E un espace vectoriel sur un corps K. Soit u une application de E dans E qui permute avec tous les automorphismes f de E (c'est-à-dire que $uof = fou$).

Montrer que u est une homothétie ($x \longrightarrow \lambda x$ où $\lambda \in K$). [Ecrire que u est, en particulier, permutable avec tout automorphisme qui laisse invariant un élément $x \in E$ et en déduire que $u(x) = \lambda(x)x$ où $\lambda(x) \in K$].

Exercice 17. — Pour abréger l'écriture, on appelle ici « espaces » les espaces vectoriels sur un corps donné (fixé une fois pour toutes) et « applications » les applications linéaires de ces espaces les uns dans les autres.

Soit $\varphi : E \longrightarrow E'$ une application. La dimension du noyau de φ sera notée $n(\varphi)$. D'autre part, on notera $c(\varphi)$ la codimension dans E' de l'image $\varphi(E)$ de E par φ ; autrement dit : $c(\varphi) = \dim (E'/\varphi(E))$. Dans le cas général $n(\varphi)$ et $c(\varphi)$ peuvent être infinis. Lorsqu'ils sont tous deux finis, on pose :

$$d(\varphi) = c(\varphi) - n(\varphi).$$

1) On se donne les espaces E et E', de dimensions *finies* e et e' respectivement. Quelles relations doivent vérifier les entiers a et b pour qu'il existe une application $\varphi : E \longrightarrow E'$ telle que

$$n(\varphi) = a \text{ et } c(\varphi) = b\,?$$

Calculer $d(\varphi)$ en fonction de e et e'.

2) Soient E, E', E'' trois espaces de dimensions finies et $\varphi : E \longrightarrow E'$, $\varphi' : E' \longrightarrow E''$ deux applications. On considère l'application composée $\varphi'' = \varphi' o \varphi$. Démontrer la relation :

$$d(\varphi'') = d(\varphi) + d(\varphi').$$

3) On considère deux espaces E et E' ayant chacun une base infinie dénombrable. Démontrer que, quels que soient les entiers a et b, il existe une application $\varphi : E \longrightarrow E'$ telle que $n(\varphi) = a$ et $c(\varphi) = b$.

4) On considère, comme au 2) trois espaces E, E', E'' et trois applications φ, φ', φ'', avec $\varphi'' = \varphi' o \varphi$, mais on n'exige plus que les espaces soient de dimensions finies. Démontrer que si $n(\varphi)$ et $n(\varphi')$ [respectivement $c(\varphi)$ et $c(\varphi')$] sont finis, alors $n(\varphi'')$ [respectivement $c(\varphi'')$] est fini.

5) On conserve les notations du 4), et on suppose que $n(\varphi)$, $n(\varphi')$, $c(\varphi)$, $c(\varphi')$ sont tous les quatre finis. Démontrer la relation :

$$d(\varphi'') = d(\varphi) + d(\varphi').$$

Convention relative aux notations.

Les éléments de la base de E sont désignés par a_i (i décrivant un ensemble I d'indices). Les coordonnées d'un vecteur x par rapport aux éléments de cette base seront désignées par x^i (*). On écrira donc :

$$x = \sum_{i \in I} x^i a_i,$$

ce qui entraîne :

$$f(x) = \sum_{i \in I} x^i f(a_i).$$

L'application est, comme nous venons de le dire, définie par la donnée des $f(a_i)$. Soit $\{ b_j \}$ une base de F. Nous nous donnerons $f(a_i)$ sous la forme :

$$f(a_i) = \sum \alpha_i^j b_j \tag{1}$$

Les α_i^j sont des éléments du corps dont la donnée définit f.

2. Cas des espaces à nombre fini de dimensions. Matrices.

Plaçons-nous dans le cas où $\dim E = n$ et $\dim F = m$. Les α_i^j au nombre de mn peuvent être rangés dans un tableau rectangulaire appelé *matrice* et la donnée de la matrice équivaudra à celle de l'application f.

Nous convenons d'écrire la matrice en plaçant l'élément α_i^j dans la $i^{\text{ème}}$ colonne et la $j^{\text{ème}}$ ligne. La matrice a alors n colonnes et m lignes *Les éléments de la $i^{\text{ème}}$ colonne sont les coordonnées de $f(a_i)$ dans la base* $\{ b_j \}$.

$$\begin{array}{cc}
\begin{pmatrix}
\alpha_1^1 & \cdots & \alpha_i^1 & \cdots & \alpha_n^1 \\
\vdots & & \vdots & & \vdots \\
\alpha_1^j & \cdots & \alpha_i^j & \cdots & \alpha_n^j \\
\vdots & & \vdots & & \vdots \\
\alpha_1^m & \cdots & \alpha_i^m & \cdots & \alpha_n^m
\end{pmatrix} & m \\
n &
\end{array}$$

(*) La place des indices est choisie de telle sorte que les sommations qui interviennent se fassent toujours sur un indice qui apparaît dans chacun des termes de la somme, une fois en haut et une fois en bas. Le nom de l'indice par rapport auquel on somme ne figure évidemment pas dans la somme et est appelé indice muet par opposition aux autres indices que l'on retrouvera dans la somme à la place où ils se trouvaient dans les termes. L'intérêt de cette convention d'écriture est de permettre un certain contrôle des calculs. Certains, pour alléger l'écriture, font la convention supplémentaire de supprimer le signe Σ et d'interpréter systématiquement toute expression du type $u^i v_i$, comme étant $\sum_{i \in I} u^i v_i$, quitte à signaler explicitement le cas, extrêmement rare, où le produit $u^i v_i$, pour une valeur donnée de i, interviendrait seul.

3. Opérations sur les matrices.

Les opérations sur les applications linéaires définissent des opérations sur les matrices.

— *Somme.* Soient f et g deux applications linéaires de E dans F :

f représenté par la matrice $A(\alpha_i^j)$

g représenté par la matrice $B(\beta_i^j)$

A + B *est par définition la matrice qui représente l'application* $f+g$.

Puisque par définition de la somme :

$$(f+g)(a_i) = f(a_i) + g(a_i),$$

et puisque la décomposition d'un vecteur dans la base $\{b_j\}$ est unique, il vient immédiatement que tout coefficient de la matrice somme est la somme des coefficients correspondants des deux matrices :

$$\forall i,\ \forall j,\ \gamma_i^j = \alpha_i^j + \beta_i^j.$$

— *Multiplication externe.* Par définition, si $\lambda \in K$, la matrice λA *représente l'application* λf et tous ses coefficients sont obtenus en multipliant par λ ceux de la matrice A.

Il résulte de ces définitions que *l'ensemble des matrices représentant les applications linéaires de* E *dans* F, *c'est-à-dire l'ensemble des matrices à n colonnes et m lignes (ou (n, m)-matrices), forment un espace vectoriel sur* K comme $\mathcal{L}(E, F)$ lui-même. Nous allons voir qu'on peut définir une base de cet espace vectoriel liée simplement à celles qui ont été choisies dans E et F. Considérons l'application linéaire φ_j^i dont la matrice a tous ses coefficients nuls sauf son coefficient d'indices i, j (i en bas, j en haut) qui vaut 1.

$$\forall k \neq i \quad \varphi_j^i(a_k) = 0$$
$$\varphi_j^i(a_i) = b_j$$

On voit qu'on peut écrire f sous la forme :

$$f = \sum_j \sum_i \alpha_i^j \varphi_j^i$$

En effet, $\forall k \quad \sum_j \sum_i \alpha_i^j \varphi_j^i(a_k) = \sum_j \alpha_k^j b_j = f(a_k)$.

Les applications φ_j^i forment donc un système de générateurs de $\mathcal{L}(E, F)$. Nous allons montrer qu'elles sont linéairement indépendantes, donc constituent une base.

Supposons qu'on ait une égalité :

$$\sum_{i,j} \lambda_i^j \varphi_j^i = 0,$$

ce qui signifierait que l'image d'un vecteur quelconque par cette application est nulle. Prenons pour ce vecteur a_k. Il vient :

$$\sum_{i,j} \lambda_i^j \varphi_j^i(a_k) = 0,$$

ce qui, en tenant encore compte des valeurs de φ_j^i pour a_k, donne en calculant comme ci-dessus :

$$\sum_j \lambda_k^j b_j = 0.$$

Mais $\{b_j\}$ constituant une base, cette égalité entraînerait :

$$\forall j \qquad \lambda_k^j = 0$$

et k étant quelconque dans I, on arriverait à :

$$\forall k, j \quad \lambda_k^j = 0,$$

c'est-à-dire que les φ_j^i forment une base. En langage de matrice, nous dirons que les (n, m)-matrices dont seul le coefficient d'indices (i, j) vaut 1 tous les autres étant nuls, constituent une base de l'espace vectoriel, des (n, m)-matrices.

— *Calcul des coordonnées de $f(x)$ en fonction des coordonnées de x.* Soit un vecteur x de coordonnées x^i dans la base $\{a_i\}$:

$$f(x) = \sum_i x^i f(a_i)$$

$$\text{avec} \quad f(a_i) = \sum_j \alpha_i^j b_j$$

$$\text{donc} \quad f(x) = \sum_{i,j} x^i \alpha_i^j b_j$$

D'autre part, dans la base $\{b_j\}$, $f(x)$ s'exprime par :

$$\sum_j y^j b_j\,;$$

donc $$\forall j, \quad y^j = \sum_i x^i \alpha_i^j$$

— *Produit de deux matrices.* E, F, G étant trois espaces vectoriels, considérons deux applications linéaires f et g, respectivement de E dans F et de F dans G.

	$E \xrightarrow{\ f\ }$	$F \xrightarrow{\ g\ }$	G
Soient respectivement	$\{a_i\}$	$\{b_j\}$	$\{c_k\}$
les bases de ces espaces et	n	m	p
leurs dimensions. Soient	$A = (\alpha_i^j)$	$B = (\beta_j^k)$	

les matrices de f et g respectivement à n colonnes et m lignes, à m colonnes et p lignes.

Considérons *l'application* $g \circ f$. *Sa matrice* C est par définition *le produit des matrices* B *et* A *associées à* g *et* f. Cherchons la forme de cette matrice.

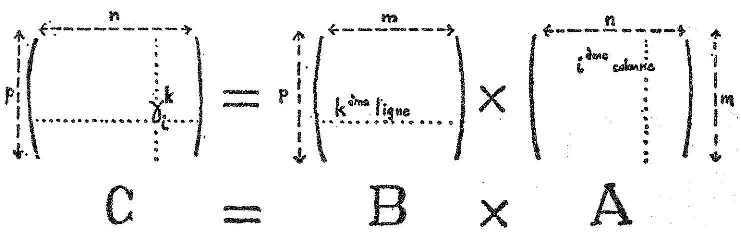
$$(g \circ f)(a_i) = g(f(a_i)) = g\left[\sum_j \alpha_i^j b_j\right] = \sum_j \alpha_i^j g(b_j)$$

$$= \sum_j \alpha_i^j \sum_k \beta_j^k c_k = \sum\sum \alpha_i^j \beta_j^k c_k \qquad (1)$$

la $k^{\text{ème}}$ coordonnée de $(g \circ f)(a_i)$ est, d'après nos conventions, le coefficient γ_i^k de la $i^{\text{ème}}$ colonne et de la $k^{\text{ème}}$ ligne ; on tire sa valeur de (1) :

$$\gamma_i^k = \sum_j \beta_j^k \alpha_i^j$$

Cette règle peut être schématisée ainsi :
L'élément de la $k^{\text{ème}}$ ligne et de la $i^{\text{ème}}$ colonne est obtenu à l'aide de la



$k^{\text{ème}}$ ligne de la matrice de gauche (celle qui correspond à la deuxième application) et de la $i^{\text{ème}}$ colonne de la matrice de droite (*).

Il importe de remarquer que le produit de deux matrices rectangulaires n'est défini que si le nombre de lignes de A est égal au nombre de colonnes de B (ce nombre étant la dimension de l'espace vectoriel intermédiaire).

Remarque. — La formule $y^j = \sum_i \alpha_i^j x^i$ donnant les coordonnées du vecteur $y = f(x)$ en fonction des coordonnées du vecteur x peut être considérée comme représentant un produit de matrices. Si X est la matrice à une colonne et n lignes dont les éléments sont les coordonnées de x par rapport à la base $\{a_i\}$ et Y la matrice à une colonne et m lignes dont les éléments sont les coordonnées de y par rapport à la base $\{b_j\}$, et A la matrice de l'application linéaire f par rapport à ces deux bases, la formule

$$y^j = \sum_i \alpha_i^j x^i$$

exprime l'égalité suivante entre matrices :

$$Y = AX.$$

Les matrices X et Y sont souvent appelées « matrices-colonne » ou encore « vecteurs-colonne ».

Exercice 18. — Soit E un espace vectoriel sur un corps K commutatif et $E = \bigoplus_{i=1}^{m} E_i$ une décomposition de E en une somme directe finie. Tout vecteur x de E admet donc une décomposition unique de la forme

$$x = \sum_{i=1}^{m} x_i \quad \text{ou} \quad x_i \in E_i$$

1) On désigne par pr_i l'application de E dans E_i qui à x fait correspondre x_i. Montrer que pr_i est linéaire et que l'on a

$$\sum_{i=1}^{m} pr_i = e_E \qquad \text{(application identique)}$$

(*) Les conventions d'écriture, que nous avons indiquées, ne sont pas les seules possibles, ni les seules effectivement utilisées. Il importe pourtant de remarquer qu'un changement de convention en entraîne d'autres. Par exemple, si on intervertit le rôle des lignes et des colonnes, la règle du produit ne sera plus la même.

et $\qquad pr_i \circ pr_{i'} = 0 \qquad$ si $i \neq i'$

$\qquad pr_i \circ pr_i = pr_i$

pr_i peut être appelée projection sur E_i parallèlement à E'_i si $E'_i = \bigoplus_{j \neq i} E_j$.

2) F étant un second espace vectoriel sur le même corps K et $F = \bigoplus_{j=1}^{n} F_j$ une décomposition de F en somme directe finie, montrer que tout $f \in \mathcal{L}(E, F)$ peut s'écrire

$$f = \sum_{i,j} pr_j^F \circ f \circ pr_i^E$$

où pr_i^E sont les projections relatives à la décomposition de E, et pr_j^F celles relatives à la décomposition de F.

En déduire

$$\mathcal{L}(E, F) \approx \bigoplus_{i,j} \mathcal{L}(E_i, F_j)$$

3) Montrer que le résultat précédent ne s'étend pas au cas des sommes directes infinies.

4) G étant un troisième espace vectoriel sur K et $G = \bigoplus_{k=1}^{p} G_k$ une décomposition de G en somme directe finie, montrer que si $g \in \mathcal{L}(F, G)$ on a

$$g \circ f = \sum_{i,j,k} pr_k \circ g \circ pr_j \circ f \circ pr_i$$

Supposant que E, F, G sont de dimension finie, en déduire la règle de multiplication des matrices dite « par blocs ».

Exercice 19. — L'espace géométrique usuel est rapporté à trois axes Ox, Oy, Oz et identifié à $\mathbf{R}^3$. On donne un plan P d'équation

$$ux + vy + wz = 0$$

et une droite D d'équations

$$\frac{x}{\alpha} = \frac{y}{\beta} = \frac{z}{\gamma} \quad \text{avec } u\alpha + v\beta + w\gamma \neq 0$$

On considère les deux applications linéaires de $\mathbf{R}^3$ dans lui-même correspondant à la projection sur P parallèlement à D et à la projection sur D parallèlement à P. Ecrire les matrices A et B de ces deux applications. Vérifier que les produits AB et BA ont bien la valeur prévisible *à priori*.

4. Matrices des endomorphismes d'un espace vectoriel.

Nous venons de voir qu'à toute application linéaire de l'espace E (de dimension n) dans l'espace F de dimension m correspond biunivoquement une (n, m)-matrice (n colonnes, m lignes) quand une base a été choisie dans chacun des espaces. Quand E et F ont le même nombre de dimensions, la matrice devient une matrice carrée. Un cas particulièrement intéressant où il en est ainsi est le cas où $E = F$, E étant un espace vectoriel de dimension n sur K, c'est-à-dire le cas où les matrices considérées sont celles des endomorphismes de $\mathcal{L}(E, E)$, E étant rapporté à la même base, qu'il soit considéré comme espace de départ ou espace d'arrivée.

Les matrices carrées d'ordre n ou (n, n)-matrices forment évidemment une algèbre sur K isomorphe à $\mathcal{L}(E, E)$ puisque les opérations sur les matrices ont justement été définies de façon à assurer cet isomorphisme. Parmi celles-ci figure la *matrice unité* I_n qui représente l'automorphisme identique et dont tous les éléments sont nuls, sauf ceux de la diagonale principale qui valent 1.

Parmi les matrices carrées d'ordre n figurent aussi les matrices qui représentent les automorphismes et dont l'ensemble forme, pour le produit, un groupe isomorphe à $GL_n(K)$. Elles sont *inversibles* et sont les seules à l'être.

§ 4. RELATION ENTRE APPLICATIONS LINEAIRES ET MATRICES CHANGEMENTS DE BASES

Nous avons jusqu'ici supposé les bases fixées dans E et dans F. Mais si on change la base de E et la base de F, c'est une matrice différente qui représentera la même application. Et c'est à une application différente que correspondra la même matrice. On est donc amené à se poser le problème de caractériser les applications susceptibles d'être représentées par une même matrice M grâce à un choix convenable des deux bases et celui de caractériser les matrices qui représentent une même application.

1. Applications représentées par une même matrice. Notion de rang d'une matrice.

Soit d'abord une matrice M qui représente l'application $f : E \longrightarrow F$ relativement à la base (a) de E et à la base (b) de F. Cherchons quelle application g représente la matrice M quand E est rapporté à une nouvelle base (c) et F à une nouvelle base (d). (Les éléments de la ième colonne deviennent les coordonnées de $g(c_i)$ dans la base (d)).

L'application g peut être décomposée de la façon suivante : considérons l'automorphisme u de E défini par :

$$\forall i \quad c_i \xrightarrow{u} a_i$$

et l'automorphisme v de F défini par :

$$\forall i \quad b_i \xrightarrow{v} d_i$$

on a : $g = v \circ f \circ u$

$$\begin{tikzcd} \underset{(a)}{E} \arrow[r, "f"] & \underset{(b)}{F} \arrow[d, "v"] \\ (c) \arrow[u, "u"] \arrow[r, "g"] & (d) \end{tikzcd}$$

puisque $\quad \forall i \qquad (v \circ f \circ u)(c_i) = g(c_i)$
et qu'une application linéaire est déterminée par les images des éléments de la base.

Cherchons le rang de g, c'est-à-dire la dimension de $g(E)$. $u(E) = E$, donc $(f \circ u)(E) = f(E)$, donc $g(E) = (v \circ f)(E)$, v étant un automorphisme de F, $f(E)$ et $v(f(E))$ ont même dimension.

Toutes les applications susceptibles d'être représentées par M *ont donc même rang.* Ce nombre est appelé le *rang de la matrice.*

Réciproquement, si f et g sont deux applications linéaires de même rang, il existe deux automorphismes v de F et u de E tels que :

$$g = v \circ f \circ u.$$

On a, en effet, par hypothèse $\dim f(E) = \dim g(E)$. Si H et K sont respectivement des supplémentaires de $f(E)$ et $g(E)$ par rapport à F, on a aussi $\dim H = \dim K$. On peut donc construire un automorphisme v de F qui envoie isomorphiquement $f(E)$ sur $g(E)$ et H sur K ; (il suffit de choisir des bases dans $f(E)$, $g(E)$, H et K et d'envoyer la base de $f(E)$ sur celle de $g(E)$, la base de H sur celle de K). Posons

$$\varphi = v \circ f$$

On a

$$\varphi(E) = g(E)$$

Soit N le noyau de φ (qui est aussi celui de f) et L celui de g, V et W respectivement des supplémentaires de N et L par rapport à E. En vertu du théorème fondamental (III, 1), la restriction ψ de φ à V en est un isomorphisme sur $\varphi(E) = g(E)$, la restriction h de g à W en est un isomorphisme sur $g(E)$. $\psi^{-1} \circ h$ est donc un isomorphisme de W sur V.

N et L ont même dimension. Il existe donc un isomorphisme u de E dont la restriction à W est $\psi^{-1} \circ h$ et qui envoie L sur N. L'application $\varphi \circ u$ a donc, comme g, L pour noyau et a même restriction que g à W ($\psi \circ \psi^{-1} \circ h = h$) ; elle coïncide avec $g : g = \varphi \circ u$, c'est-à-dire

$$g = v \circ f \circ u.$$

Par suite, si f est l'application de rang r déterminée par la matrice M relativement à une base (a) de E et une base (b) de F, g sera représentée par la même matrice relativement à la base (a') de E, image de (a) par u^{-1}, et à la base (b') de F, image de (b) par v. Autrement dit, *une matrice de rang donné est susceptible de représenter toutes les applications linéaires de ce rang* (et ne représente que celles-là).

En particulier, si nous revenons aux matrices carrées d'ordre n, nous voyons que celles qui sont inversibles sont celles qui sont de rang égal à leur ordre. *Dans l'anneau des matrices carrées d'ordre n, le groupe des éléments inversibles est constitué par l'ensemble des matrices de rang n.* On les appelle aussi *matrices régulières.*

Parmi les automorphismes qu'une matrice de rang n est susceptible de représenter, il y a l'automorphisme identique. Si une matrice M représente l'automorphisme identique d'un espace E rapporté à une base (a_i) sur E rapporté à une base (a'_j), la $i^{\text{ème}}$ colonne de M est constituée des coordonnées du vecteur a_i dans la base (a'_j) (application de la règle générale puisqu'ici $f(a_i) = a_i$). On dit alors que M est la *matrice de passage* de la base (a_i) à la base (a'_j).

2. Matrices représentant une même application.

On se donne maintenant f et on cherche toutes les matrices qui représentent f par rapport à tous les couples de bases possibles.

Soit encore M la matrice de f relative aux bases (a) de E et (b) de F ; (a') et (b') d'autres bases de E et F respectivement, par rapport auxquelles on va chercher la matrice M' de f. On peut écrire :

$$f = I_F \circ f \circ I_E$$

où I_E représente l'automorphisme identique de E et I_F l'automorphisme identique de F. Cette égalité se traduit par l'égalité entre matrices $M' = B^{-1}MA$, où A et B représentent les automorphismes identiques et sont les matrices de passage des nouvelles bases aux anciennes (*).

$$\begin{tikzcd}
(a) \arrow[r, "M"] & (b) \arrow[d, shift left, "B^{-1}"'] \\
(a') \arrow[u, "A"] \arrow[r, "M'"'] & (b') \arrow[u, shift right, "B"']
\end{tikzcd}$$

(fig. 3)

Deux matrices telles que M et M' sont dites *équivalentes*. Cette définition qui ne s'applique évidemment qu'à des matrices de même type se traduit donc par :

Deux (n, m)-matrices sont dites équivalentes s'il existe une matrice A *régulière d'ordre* n *et une matrice* B *régulière d'ordre* m *telles que :*

$$M' = B^{-1}MA$$

ou bien par :

Deux (n, m)-matrices sont dites équivalentes si elles représentent une même application linéaire d'un espace de dimension n dans un espace de dimension m.

Ceci entraîne, en vertu de ce que nous avons vu au paragraphe précédent, que deux matrices équivalentes ont même rang. Mais, réciproquement, si elles ont même rang, elles peuvent représenter la même application puisqu'une matrice peut représenter toute application linéaire de son rang.

Deux (n, m)-matrices sont équivalentes si, et seulement si, elles sont de même rang.

Il est évident que l'équivalence de deux matrices est une relation d'équivalence.

3. Matrices semblables.

Cette notion, plus restrictive, ne s'applique qu'à des matrices carrées. E étant rapporté à la même base, qu'il soit considéré comme espace de départ ou comme espace d'arrivée, deux matrices carrées d'ordre n sont dites semblables si elles sont susceptibles de représenter un même endomorphisme d'un espace à n dimensions, ou, ce qui est équivalent, s'il existe une matrice régulière P d'ordre n, telle que :

$$M' = P^{-1}MP$$

P est la matrice de passage de la base (a'), à laquelle M' est relative à la

(*) Il est recommandé, pour éviter dans la pratique toute confusion sur ce que sont les matrices A et B, de refaire, si on hésite, le schéma de la figure 3.

base (a), à laquelle M est relative. Deux matrices semblables sont évidemment équivalentes.

La similitude des matrices est encore une relation d'équivalence.

Exercice 20. — On considère l'espace $\mathcal{P}_n \subset \mathbf{R}[x]$ des polynômes à coefficients réels, à une indéterminée, de degré au plus égal à n et l'endomorphisme φ de cet espace qui, à un polynôme, fait correspondre sa dérivée. Ecrire la matrice de φ en prenant comme base B_0 de $\mathcal{P}_n$ l'ensemble des monômes $\{x^k ; 0 \leqslant k \leqslant n\}$.

Montrer que l'ensemble des polynômes $p_0 = 1$, $p_1 = 1 + x, \ldots, p_n = 1 + x + \ldots + x^k + \ldots + x^n$ est une autre base B_1 de $\mathcal{P}_n$.

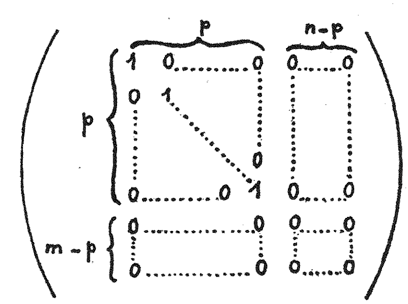
Ecrire les matrices de passage de B_0 à B_1 et de B_1 à B_0.

En déduire la matrice de φ dans la base B_1 ; vérifier en cherchant directement cette matrice.

§ 5. MATRICES REMARQUABLES EQUIVALENTES A UNE MATRICE DONNEE DETERMINATIONS PRATIQUES DU RANG D'UNE MATRICE

1. Matrice canonique d'une application de rang p.

Le rang d'une matrice est, d'après la définition, la dimension de $f(E)$ si f est une des applications linéaires que représente la matrice. Cherchons la forme la plus simple d'une (n, m)-matrice équivalente à une (n, m)-matrice donnée, de rang p.



Si la matrice est de rang p, c'est qu'il existe, dans $f(E)$, p éléments qui en constituent une base. Constituons une base de F en complétant cette base de $f(E)$ par $m\text{-}p$ vecteurs. D'autre part, pour constituer une base de E, considérons que $E = N \oplus H$, N représentant le noyau de f et H un supplémentaire, et prenons les images des p éléments de $f(E)$ par l'application réciproque de la restriction de f à H. Ces p images constituent une base de H ; complétons-la par $n\text{-}p$ éléments de N ; et cherchons la forme de la matrice de f dans la base ainsi définie.

On trouve la forme ci-dessus en remarquant que chacun des p vecteurs de $f(E)$ est l'image d'un vecteur de la base de H et que les vecteurs de N ont une image nulle.

Pour qu'une matrice soit de rang p, il est donc nécessaire et suffisant qu'elle soit équivalente à une matrice de cette forme, c'est-à-dire constituée en bordant la matrice I_p de lignes et de colonnes de zéros.

Cette forme particulièrement simple (matrice canonique d'une application de rang p) ne peut être obtenue que par un choix convenable des bases dans E et dans F. Nous allons voir qu'on peut obtenir des matrices de forme assez simple (matrices triangulaires), pour lesquelles le rang sera mis en évidence, en ne changeant de base que dans un seul des espaces E ou F, et nous allons indiquer des opérations à effectuer pratiquement pour obtenir ces formes.

2. Changement de la base de l'espace de départ.

Soit (α_i^j) la matrice d'une application f relative à des bases :

$$\{ a_i \,;\, 1 \leqslant i \leqslant n \} \text{ de E et } \{ b_j \,;\, 1 \leqslant j \leqslant m \} \text{ de F.}$$

Supposons que α_1^1 soit différent de zéro (s'il ne l'était pas, on pourrait toujours modifier l'ordre des éléments des deux bases pour qu'il le soit, dès l'instant que tous les coefficients de la matrice ne sont pas nuls). Choisissons alors pour E une nouvelle base $\{ a'_i \}$ définie par :

$$a'_1 = a_1 \text{ et pour } i \neq 1 \; a'_i = \alpha_1^1 \, a_i - \alpha_i^1 \, a_1$$

(Les a'_i forment bien une base, car ils sont linéairement indépendants puisque $\alpha_1^1 \neq 0$. Dans la matrice de f par rapport à cette nouvelle base, la première colonne est inchangée, la $i^{\text{ème}}$ colonne $(i \neq 1)$ est constituée par les coordonnées de :

$$f(a'_i) = \alpha_1^1 \, f(a_i) - \alpha_i^1 \, f(a_1) \qquad \text{pour } i \neq 1,$$

c'est-à-dire que chaque colonne autre que la première est remplacée par une colonne formée d'une combinaison linéaire des éléments de même indice j, d'elle-même et de la première colonne, combinaison qu'on pourrait écrire schématiquement :

$$\text{nouvelle } i^{\text{ème}} \text{ colonne} = \alpha_1^1 \, (i^{\text{ème}} \text{ colonne}) - \alpha_i^1 \, (1^{\text{ère}} \text{ colonne}).$$

La première ligne est alors formée de zéros, à l'exception du premier coefficient, car :

$$\alpha'^1_i = \alpha_1^1 \alpha_i^1 - \alpha_i^1 \alpha_1^1 = 0.$$

Si tous les coefficients situés dans l'angle en bas et à droite de α_1^1 ne sont pas nuls, on pourra toujours supposer $\alpha'^2_2 \neq 0$ (en modifiant, au besoin, l'ordre des éléments autres que a_1 et b_1 dans les bases). On recommence l'opération précédente en posant pour définir une nouvelle base :

$$a''_1 = a'_1 \qquad a''_2 = a'_2 \qquad \text{et pour } i \neq 1 \quad i \neq 2 \qquad a''_i = \alpha'^2_2 a'_i - \alpha'^2_i a'_2$$

Dans la matrice de f, la deuxième ligne à son tour sera formée de zéros à droite de la diagonale (*), la première ligne restant inchangée.

(*) Par abus de langage, nous désignerons ainsi l'ensemble des coefficients α_i^j avec $i = j$, même si la matrice n'est pas carrée.

On répétera cette opération jusqu'au moment où on butera sur un des bords de la matrice ou bien jusqu'au moment, où, dans l'angle en bas et à droite du dernier coefficient considéré de la diagonale, il ne se trouvera plus que des zéros. On aura alors obtenu une matrice équivalente à la matrice initiale et qui aura un des trois aspects suivants :

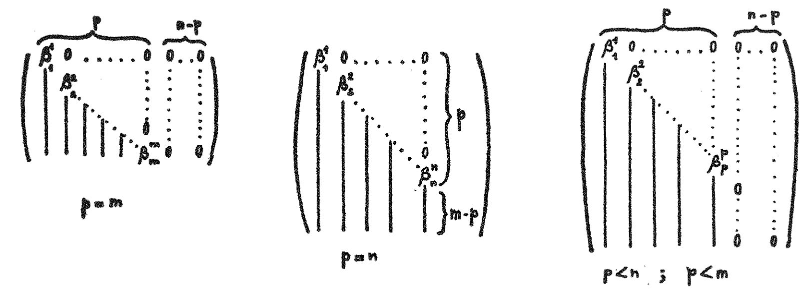


FIG. 3

(avec $\beta_1^1 = \alpha_1^1 \quad \beta_2^2 = \alpha'^2_2 \quad \beta_3^3 = \alpha''^3_3$, et ainsi de suite).

Il est inutile d'aller plus loin pour voir que ces matrices sont de rang p. Les n-p derniers vecteurs de la base de E sont en effet envoyés sur 0 ; quant aux p premiers, leurs images qui figurent dans les colonnes sont linéairement indépendantes. En effet, si $\{ e_1, \ldots, e_n \}$ est la base finalement choisie dans E,

$$\sum_{i=1}^{p} \lambda_i f(e_i) = 0$$

signifierait : $\quad \forall j \qquad \displaystyle\sum_{i=1}^{p} \lambda_i \beta_i^j = 0.$

Pour $j = 1$, on trouverait alors : $\lambda_1 \beta_1^1 = 0$ donc $\lambda_1 = 0$

.... $j = 2$, ... $\lambda_2 \beta_2^2 = 0$ .... $\lambda_2 = 0$

et ainsi de suite...

Les p vecteurs $f(e_i)$ pour $1 \leqslant i \leqslant p$ constituent donc une base (*) de $f(E)$; f et ses matrices sont de rang p.

(*) Nous avons ici un exemple d'une circonstance générale que l'on rencontre fréquemment. Si $\{ a_i \,;\, 1 \leqslant i \leqslant n \}$ est une base d'un espace vectoriel, les n vecteurs

$$\begin{array}{ll} b_1 = \lambda_1^1 a_1 & \lambda_1^1 \neq 0 \\ b_2 = \lambda_2^1 a_1 + \lambda_2^2 a_2 & \lambda_2^2 \neq 0 \\ \dots\dots\dots\dots & \\ b_i = \lambda_i^1 a_1 + \dots\dots + \lambda_i^i a_i & \lambda_i^i \neq 0 \\ \dots\dots\dots\dots\dots & \\ b_n = \lambda_n^1 a_1 + \dots\dots\dots\dots + \lambda_n^n a_n & \lambda_n^n \neq 0 \end{array}$$

sont linéairement indépendants et constituent donc une base de l'espace.

Si la base $\{ a_i \}$ est non pas finie, mais dénombrable, i décrivant $\mathbb{N}$, il est facile de voir que la famille $\{ b_i \}$ est encore une base, car elle est libre et tout vecteur a^i s'exprime linéairement en fonction des b_i (récurrence sur i). Exemple : dans l'espace $K[x]$, des polynômes à coefficients dans le corps K, toute famille comprenant un polynôme et un seul de chaque degré est une base.

Si on voulait maintenant obtenir la forme canonique décrite au début, il suffirait de prendre pour base de F l'ensemble $\{ f(e_1), ..., f(e_p) \}$, complété si $n > p$.

3. Changement de base dans l'espace d'arrivée.

Nous allons procéder par étapes, comme ci-dessus, chaque nouveau changement de base ayant, cette fois, pour effet de faire apparaître le plus grand nombre possible de zéros dans une colonne de la nouvelle matrice.

Le premier changement de base (dans F) consiste à prendre, si $\alpha_1^1 \neq 0$ (cas auquel on peut toujours se ramener si $f \neq 0$, en changeant les indices des éléments des bases de E et de F), pour nouvelle base :

$$b_1' = f(a_1) \qquad b_2' = b_2 \,.....\, b_n' = b_n.$$

Cet ensemble constitue bien une base dès l'instant que α_1^1 étant différent de zéro, $f(a_1)$ est indépendant de l'ensemble $\{ b_2 ... b_n \}$.

Pour chercher comment se transforme la matrice, exprimons d'abord les éléments de l'ancienne base en fonction de ceux de la nouvelle. On a :

$$\text{pour } i \neq 1 \qquad b_i = b'_i$$

et de :

$$b'_1 = \alpha_1^1 b_1 + \sum_{j=2}^{n} \alpha_1^j b_j$$

on tire :

$$b_1 = \frac{b_1'}{\alpha_1^1} - \sum_{j=2}^{n} \frac{\alpha_1^j}{\alpha_1^1} \; b_j'$$

ou, en choisissant une base (b''_i) déduite de (b'_i) par une homothétie de rapport $\dfrac{1}{\alpha_1^1}$

$$b_1 = b_1'' - \sum_{j=2}^{n} \alpha_1^j b''_j$$

ceci posé, $f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j$ s'écrit dans la nouvelle base :

$$f(a_i) = \alpha_i^1 \Big(b''_1 - \sum_{j=2}^{n} \alpha_1^j b''_j \Big) + \sum_{j=2}^{n} \alpha_i^j \alpha_1^1 \, b''_j$$

$$= \alpha_i^1 b''_1 + \sum_{j=2}^{n} \left(\alpha_1^1 \alpha_i^j - \alpha_1^j \alpha_i^1 \right) b''_j$$

Sur cette expression, on voit que la première ligne est inchangée ($\forall i \quad \alpha'^{1}_i = \alpha_i^1$) et que la $j^{\text{ème}}$ ligne est remplacée par une ligne formée par une combinaison linéaire des éléments de même indice i, d'elle-même et de la première ligne, schématiquement indiquée par :

$$\text{nouvelle } j^{\text{ème}} \text{ ligne} = \alpha_1^1 \; (j^{\text{ème}} \text{ ligne}) - \alpha_1^j \; (1^{\text{ère}} \text{ ligne}).$$

Et tous les coefficients de la première colonne, sauf α_1^1, sont nuls, comme il était prévu.

En supposant que le nouveau coefficient α'^{2}_2 soit différent de zéro

nous pourrons recommencer une opération analogue en prenant pour nouvelle base :

$$b_1''' = b_1'' \qquad b_2''' = f(a_2) \qquad b'''_3 = b''_3 ... \, b'''_n = b''_n$$

ou une base homothétique. Ceci reviendra à remplacer chaque fois, dans la matrice, un certain nombre de lignes par une combinaison linéaire de la ligne remplacée et d'une autre, le multiplicateur de la ligne remplacée devant être différent de zéro, et les multiplicateurs étant choisis de façon à annuler tous les coefficients d'une colonne au-dessous de la diagonale principale. On répétera l'opération jusqu'à ce qu'on obtienne une matrice de forme analogue à celles données ci-dessus, mais où on aura des zéros en-dessous et à gauche de la diagonale principale :

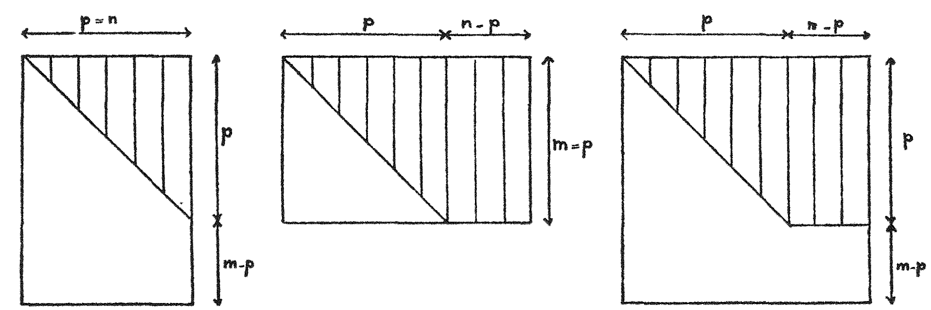


FIG. 4

(Sur ces dessins schématiques, les régions laissées en blanc sont celles où il n'y a que des zéros). Les p premiers termes de la diagonale principale sont différents de 0.

Sur ces formes, on peut encore reconnaître que les matrices sont de rang p ; directement pour la première (même raisonnement que dans le cas précédent) ; pour les deux autres, si on opérait maintenant un changement de base de E comme dans la section précédente, on obtiendrait une matrice équivalente dont on vérifie facilement qu'elle a les mêmes éléments dans sa diagonale principale, et que ces éléments sont les seuls à être différents de zéro, donc une matrice de rang p. (On n'aura donc pas besoin de faire effectivement cette deuxième transformation pour reconnaître le rang de la matrice).

Exercice 21. — Toutes les notions de ce chapitre ont été définies par rapport à un corps K supposé fixé. Il en est ainsi en particulier du rang p d'une (n, m)-matrice M dont les coefficients appartiennent à K. Mais, si K' est un surcorps de K, M peut aussi être considérée comme une matrice à coefficients dans K'. Elle représente alors des applications linéaires d'un espace de dimension n sur K' (par ex. K'^n) dans un espace de dimension m sur K' (par. ex. K'^m). Quel est alors son rang ?

CHAPITRE IV

DUALITÉ

§ 1. DUAL ET BIDUAL D'UN ESPACE VECTORIEL

1. Formes linéaires.

Nous avons vu que $\mathcal{L}$ (E, F) avait une structure d'espace vectoriel sur K. Pour F, prenons K lui-même. Une application linéaire de E dans K est appelée une *forme linéaire sur* E. L'espace des formes linéaires sur E est appelé espace dual de E ; il est désigné par E*, son élément générique par x^* ; $x^*(y)$ est la valeur prise par x^* pour y ; c'est un élément de K. On convient de la noter $< y, x^* >$.

Exemple de formes linéaires : formes coordonnées. E étant rapporté à une base $\{ a_i \}$, on peut considérer la forme linéaire définie de la façon suivante : à tout vecteur y, elle fait correspondre sa $i^{\text{ème}}$ coordonnée dans cette base :

$$y = \sum y^i a_i \longrightarrow y^i.$$

On appelle cette forme linéaire la *forme coordonnée d'indice i relative à la base* $\{ a_i \}$. Nous la noterons a^i ; et on a :

$$< y, a^i > = y^i.$$

2. Forme bilinéaire canonique sur E × E*.

Supposons maintenant que nous laissions y fixe et que x^* décrive E*, et considérons l'application de E* dans K, qui à x^* fait correspondre $< y, x^* >$. Cette application est linéaire puisque d'après les définitions des opérations sur les applications :

$$(x^* + z^*)(y) = x^*(y) + z^*(y)$$

et :

$$(\lambda x)^*(y) = \lambda x^*(y)$$

soit :

$$< y, x^* + z^* > = < y, x^* > + < y, z^* >$$
$$< y, \lambda x^* > = \lambda < y, x^* >.$$

L'application de E × E* dans K qui au couple (y, x^*) fait correspondre l'élément $< y, x^* >$ de K est donc bilinéaire : on l'appelle la *forme bilinéaire canonique sur* E × E*.

3. Bidual.

Une application linéaire de E* dans K est un élément du dual de E*, soit E**. Ce dernier espace est appelé *bidual* de E.

Nous venons de voir qu'à un élément $y \in \mathrm{E}$, on pouvait faire correspondre un élément de E^{**}, que nous noterons $\tilde{y}$, défini par :

$$\tilde{y}(x^*) = \langle y, x^* \rangle = x^*(y)$$

Ceci définit une application de E dans E^{**} :

$$y \longrightarrow \tilde{y}$$

dont nous allons montrer qu'elle est un homomorphisme.

$\forall x^*$ $\qquad \widetilde{(y+z)}(x^*) = x^*(y+z) = x^*(y) + x^*(z)$

$$= \tilde{y}(x^*) + \tilde{z}(x^*) = (\tilde{y} + \tilde{z})(x^*),$$

D'où, ceci étant vrai pour tout x^*,

$$\widetilde{y+z} = \tilde{y} + \tilde{z}.$$

De même $\forall x^*$ $\quad \widetilde{\lambda y}(x^*) = x^*(\lambda y) = \lambda x^*(y) = \lambda\tilde{y}(x^*)$.

D'où :

$$\widetilde{\lambda y} = \lambda\tilde{y}.$$

Montrons aussi que cette application est injective : $\tilde{y} = \tilde{z}$ signifie :

$\forall x^*$ $\qquad \tilde{y}(x^*) = \tilde{z}(x^*) \iff x^*(y) = x^*(z) \iff x^*(y-z) = 0.$

Toutes les formes linéaires sur E s'annulent pour le vecteur $y - z$. Ceci entraîne que $y - z = 0$ car si $y - z = a \neq 0$, on peut toujours choisir une base comprenant ce vecteur et la forme coordonnée relative à cette base et à l'indice de ce vecteur vaudrait 1 pour ce vecteur et non pas 0.

Donc, $\qquad \tilde{y} = \tilde{z} \implies y = z.$

Il est faux, si E est de dimension infinie, que l'homomorphisme injectif $y \to \tilde{y}$ soit un isomorphisme (cf. Exercice 23). Si on désigne par $\tilde{\mathrm{E}}$ l'image de E par cette application ($\tilde{\mathrm{E}} \approx \mathrm{E}$), $\tilde{\mathrm{E}}$ est alors un vrai sous-espace de E^{**}.

4. Cas où E est de dimension finie *n*.

On va montrer que $\dim \mathrm{E}^* = n$ en montrant que les formes coordonnées relatives à une base de E constituent une base de E^*.

Soit $\{a_i\}$ une base de E et soit a^i la forme coordonnée d'indice i définie comme nous l'avons dit plus haut par :

$$a^i(a_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \qquad (*)$$

Une forme linéaire quelconque va s'exprimer en fonction des a^i. En effet, une forme x^* est déterminée par les n valeurs x_i qu'elle prend pour les éléments de la base : $x^*(a_i) = x_i$ et on pourra écrire :

x^* $$x^* = \sum x_i a^i \quad a^i \in \mathrm{E}^*, \qquad a^i(y) = y_i$$

en effet $x^*(a_j) = \sum x_i a^i(a_j) = x_j$. $\qquad x^*(y) = \sum y_i x^*(a_i) = \sum y_i x_i = \sum x_i a^i(y)$

Les formes coordonnées constituent donc un système de générateurs de E^*. D'autre part, elles sont linéairement indépendantes. En effet, $\sum \lambda_i a^i = 0$ voudrait dire que la forme $\sum \lambda_i a^i$ prend la valeur 0 pour tout vecteur. Or,

$$\left(\sum \lambda_i a^i\right)(a_j) = \lambda_j.$$

(*) Ceci est noté parfois $a^i(a_j) = \delta^i_j$; δ^i_j, appelé symbole de Kronecker, étant défini par $\delta^i_j = 1$ si $i = j$, $\delta^i_j = 0$ si $i \neq j$.

Donc, ceci exigerait $\lambda_j = 0$ et ceci pour toutes les valeurs de j. La base $\{a^i\}$ est dite *base duale* de la base $\{a_i\}$. Il résulte de ce théorème que $\dim \mathrm{E}^{**} = n$; ceci entraîne puisque $\tilde{\mathrm{E}}$, qui est isomorphe à E, est de dimension n que $\tilde{\mathrm{E}} = \mathrm{E}^{**}$. L'application $y \longrightarrow \tilde{y}$ est alors un isomorphisme. Cet isomorphisme, dont la donnée résulte de la seule donnée de E, peut être qualifié de canonique.

Exercice 22. — Une base B étant donnée dans E de dimension finie, montrer que la base duale, dans E^{**}, de la base duale de B, dans E^*, est l'image de B dans l'isomorphisme canonique.

Exercice 23. — 1) Si E est de dimension infinie, montrer que l'ensemble des formes coordonnées relatives à une base est une partie libre de E^*, mais n'en est pas une base. (K [X] est un espace vectoriel de dimension $\aleph_0$ sur K admettant pour base l'ensemble des monômes X^n ($n = 0, 1, 2 \ldots$) ; on pourra donner un exemple de forme linéaire sur K[x] qui soit linéairement indépendante des formes coordonnées relatives à cette base).

2) En déduire que si E est de dimension infinie, $\tilde{\mathrm{E}}$ est un vrai sous-espace de E^{**}.

Identification de E *et* E^{**} (la dimension de E étant *finie*). L'isomorphisme $\mathrm{E} \approx \mathrm{E}^{**}$ étant ainsi établi quand E a un nombre fini de dimensions, on convient d'identifier y et $\tilde{y}$. On pose alors :

$$\langle y, x^* \rangle = x^*(y) = y(x^*),$$

ce dernier symbole ayant la signification primitivement attribuée à $\tilde{y}(x^*)$.

5. Expression de la forme bilinéaire canonique à l'aide des bases duales.

Soit : $y = \sum_i y^i a_i \in \mathrm{E}$ et $x^* = \sum_j x_j a^j \in \mathrm{E}^*$,

$a^j \in \mathrm{E}^*$ étant la forme coordonnée d'indice j. Nous aurons :

$$x^*(y) = \sum_i y^i x^*(a_i),$$

or, nous avons vu ci-dessus que $x^*(a_i) = x_i$; donc,

$$\langle y, x^* \rangle = \sum_i y^i x_i.$$

Remarque : E et E^*, ayant même dimension, sont isomorphes, d'une infinité de façons.

On peut se demander s'il existe un isomorphisme φ de E sur E^* qui jouirait de la propriété suivante : faire correspondre à tout élément y de E et à son image par un automorphisme arbitraire u deux formes linéaires qui prendraient la même valeur, respectivement pour un x quelconque de E et pour son image par u ; autrement dit, pour lequel :

(1) $\quad \langle x, \varphi(y) \rangle = \langle u(x), (\varphi \circ u)(y) \rangle \qquad \varphi(y) \in \mathrm{E}^* \qquad (\varphi \circ u)(y) \in \mathrm{E}^*$

pour tout x et tout y appartenant à E et pour tout $u \in \mathrm{GL(E)}$; ce qui revient encore à dire qu'un tel isomorphisme φ ferait correspondre au couple (x, y) de deux éléments de E un scalaire de K qui serait invariant dans tous les automorphismes. (Un tel isomorphisme serait qualifié de canonique parce qu'il ne dépendrait que de la structure d'espace vectoriel de E).

Exercice 24. — Montrer que l'existence d'un tel isomorphisme φ est impossible, en trouvant des automorphismes u pour lesquels l'égalité (1) ne peut avoir lieu pour tout x et tout y.

Exercice 25. — (a_i) étant une base de E il existe un isomorphisme φ de E sur E* pour lequel $\varphi(a_i) = a^i$.
Déterminer les automorphismes $u \in GL(E)$ pour lesquels

$$< x, \varphi(y) > \; = \; < u(x), (\varphi \circ u)(y) >$$

§ 2. ORTHOGONALITE

On dit que $x \in E$ et $x^* \in E^*$ sont orthogonaux si

$$< x, x^* > = 0.$$

On peut considérer l'ensemble de tous les éléments de E orthogonaux à $x^* \in E^*$, soit :

(1) $$\{ x \; ; \; < x, x^* > = 0 \} .$$

dim K = 1
$x^*(E)$ [illegible] donc dim 1 ou 0

C'est le noyau de la forme linéaire x^*, c'est-à-dire $\overset{-1}{x^*}(0)$. Mais x^* étant une application de E dans K, $x^*(E) = K$ (si x^* n'est pas identiquement nulle) ; or, K est un espace vectoriel sur lui-même, de dimension 1. L'ensemble (1) est donc un sous-espace de codimension 1. On le nomme *hyperplan de* E *orthogonal à* x^*. Si E est de dimension finie n, il est de dimension $n - 1$. On peut de même considérer

$$\{ x^* \; ; \; < x, x^* > = 0 \}$$

et définir ainsi l'hyperplan de E* orthogonal à x qui est le noyau $\tilde{x}^{-1}(0)$ de la forme $\tilde{x} \in E^{**}$. Si E *est de dimension finie*, c'est-à-dire si E peut être identifié à E**, cet hyperplan peut se noter $\overset{-1}{x}(0)$.

Continuons à nous placer dans le cas où E *est de dimension finie*, et soit maintenant $A \subset E$. Nous définirons un sous-ensemble de E*, appelé orthogonal de A, et noté $A^\perp$:

$$A^\perp = \{ x^* \; ; \; \forall x \in A \quad < x, x^* > = 0 \} .$$

Ce sous-ensemble peut être défini comme l'intersection des hyperplans relatifs à tous les x de A :

$$A^\perp = \bigcap \; \{ \overset{-1}{x}(0) \; ; \; x \in A \} .$$

Il en résulte que $A^\perp$ est un sous-espace vectoriel de E*. Considérons $A^{\perp\perp}$ qui est un sous-espace vectoriel de E. Il est évident que

$A^{\perp\perp} = \{x ; \forall x^* \in A^\perp \; <x, x^*> = 0\}$

$$A^{\perp\perp} \supset A \tag{1}$$

(Puisque $A^{\perp\perp}$ est l'ensemble des éléments orthogonaux à ceux de $A^\perp$, il contient les éléments de A qui étaient, par définition de $A^\perp$, orthogonaux aux éléments de $A^\perp$). La relation (1) entraîne

$$A^{\perp\perp\perp} \subset A^\perp \quad \text{et} \quad A^{\perp\perp\perp} \supset A^\perp \quad \Rightarrow$$

(car de façon générale $A \subset B \implies A^\perp \supset B^\perp$, un élément orthogonal à tous ceux du plus grand ensemble faisant partie de l'ensemble de ceux qui sont orthogonaux aux éléments du plus petit).

$A^\perp$ est l'ens des él. orth à ceux de A | $A^{\perp\perp}$ contient ts les él. orth à $A^\perp$
$A^{\perp\perp}$ ——————————— de $A^\perp$ | en particulier A

D'autre part, $A^{\perp\perp\perp}$ peut être considéré comme $(A^\perp)^{\perp\perp}$. Et la relation (1) appliquée à $A^\perp$ permet alors d'écrire :

$$A^{\perp\perp\perp} \supset A^\perp.$$

On en conclut donc (*) : $\quad A^{\perp\perp\perp} = A^\perp.$

Dans le cas où A *est un sous-espace vectoriel de* E, *on peut montrer que* $A^{\perp\perp} = A$. Ceci revient à montrer qu'on ne peut pas trouver de forme linéaire sur E* définie par un élément $x \notin A$ et qui appartienne à $A^{\perp\perp}$, donc qui s'annule pour tout élément de $A^\perp$; ou encore qu'étant donnée cette forme définie par $x \notin A$, il existe au moins un élément de $A^\perp$ qui ne l'annule pas. Constituons une base de E en prenant le vecteur x, en lui adjoignant une base de A et en complétant, si c'est nécessaire ; puis, considérons la forme coordonnée relative à x ; elle appartient à $A^\perp$ car elle s'annule pour tout élément de la base de A, donc pour tout élément de A. Elle ne s'annule pas pour x ; donc, la forme définie par x ne s'annule pas pour elle. Donc $\forall x \notin A$ on peut trouver une forme ds $A^\perp$ qui ne lui soit pas orthogonale

Le même raisonnement prouve que, pour A quelconque, $A^{\perp\perp}$ est le plus petit sous-espace vectoriel contenant A, car la démonstration précédente s'applique à $A^{\perp\perp}$ et au sous-espace vectoriel H engendré par A. (En effet, la forme coordonnée s'annule sur tout H si elle s'annule pour les éléments d'une base de H prise dans A) (**).

Il résulte de ceci qu'il existe une correspondance bijective entre les sous-espaces vectoriels de E et E* :

$$H^{\perp\perp} = H \subset E \quad \longleftrightarrow \quad H^\perp \subset E^*.$$

Exercice 26. — Montrer que : dimension $H^\perp$ = codimension H. Retrouver, en partant de là, le fait que $A^{\perp\perp\perp} = A^\perp$ et que $H^{\perp\perp} = H$.

(*) Observons que nous avons fait un raisonnement identique quand nous avons montré que $A^{+-+} = A^+$ (voir exercice 65, *Cours A.P.M. I*). Il est facile de voir que ce raisonnement est applicable dans une situation très générale. Soient deux ensembles X et Y, et supposons qu'une relation $\mathcal{R}$ soit définie entre les éléments $x \in X$ et $y \in Y$. Si on considère $A \subset X$, $\mathring{A}$, ensemble des éléments de Y en relation avec tous les éléments de A, $\overset{\circ\circ}{A}$, ensemble de tous les éléments de X en relation avec tous les éléments de $\mathring{A}$..., etc... On montrera de la même façon en considérant les sous-ensembles du produit $X \times Y$ que $A \subset B \implies \mathring{A} \supset \mathring{B}$, et que $\overset{\circ\circ\circ}{A} = \mathring{A}$.

Dans le raisonnement ci-dessus, on avait $X = E$, $Y = E^*$ et la relation $\mathcal{R}$ était celle d'orthogonalité. Dans le raisonnement de l'exercice cité, X et Y étaient le même ensemble ordonné et la relation $x \mathcal{R} y$ était $x \prec y$, si bien que si A était inclus dans X, $\overset{\circ\circ\circ}{A}$ était A^{+-+} et si $A \subset Y$, $\overset{\circ\circ}{A}$ se notait A^{-+}.

(**) On peut encore dire : $A \subset H \subset A^{\perp\perp}$ puisque $A^{\perp\perp}$ est un sous-espace vectoriel incluant A et H le plus petit sous-espace vectoriel incluant A. Mais ceci entraîne $A^\perp \supset H^\perp$ et par conséquent $A^{\perp\perp} \subset H^{\perp\perp}$. Or, d'après la démonstration précédente, $H^{\perp\perp} = H$. Donc, $H \subset A^{\perp\perp}$ et $A^{\perp\perp} \subset H$, d'où $H = A^{\perp\perp}$.

Exercice 27. — E étant un espace de dimension *finie* ou *infinie*, on pose pour tout sous-espace H de E :

$$H^{\perp} = \{ x^* ;\ \forall x \in H \quad < x, x^* > = 0 \}$$

Montrer que si H et K sont deux sous-espaces vectoriels de E

$$(H + K)^{\perp} = H^{\perp} \cap K^{\perp}$$
$$(H \cap K)^{\perp} = H^{\perp} + K^{\perp}$$

Conséquence si $E = H \oplus K$?

Exercice 28. — $\mathcal{P}_n$ étant l'espace des polynômes à une indéterminée à coefficients dans un corps K et de degré au plus égal à n, on se propose de déterminer tous les supplémentaires par rapport à $\mathcal{P}_n$ du sous-espace $\mathcal{C}$ constitué par les polynômes constants. A cet effet, 1° on cherchera de façon générale tous les supplémentaires d'un sous-espace de codimension 1 dans un espace vectoriel de dimension finie ; 2° on utilisera le résultat de l'exercice 27 pour remplacer la recherche des supplémentaire de $\mathcal{C}$ par la recherche des supplémentaires de $\mathcal{C}^{\perp}$ dans le dual et on reviendra à $\mathcal{C}$ pour conclure.

Exercice 29. — φ, $\varphi_1, \ldots, \varphi_p$ étant des formes linéaires sur un espace vectoriel de dimension finie, la condition nécessaire et suffisante pour que φ soit nulle pour tout vecteur annulant $\varphi_1, \varphi_2, \ldots, \varphi_p$ est qu'il existe p scalaires λ^i (appelés multiplicateurs de Lagrange) tels que $\varphi = \Sigma \lambda^i \varphi_i$.

§ 3. APPLICATION TRANSPOSEE D'UNE APPLICATION LINEAIRE

1. Définition.

Soient E et F, deux espaces vectoriels sur un même corps K, E* et F* leurs duals et u une application de E dans F. Soit $y^* \in F^*$ une application de F dans K. Nous pouvons considérer l'application composée

$$y^* \circ u$$

qui, à un élément de E, fait correspondre un élément de K et qui, étant linéaire puisque composée de deux applications linéaires, est donc un élément de E*. u étant choisi, à tout $y^* \in F^*$ correspond $y^* \circ u \in E^*$. Nous avons donc défini une application de F* dans E*, appelée transposée de l'application u et notée ${}^tu$.

$$\begin{array}{ccc} E & \xrightarrow{\quad u \quad} & F \\ x & & u(x) \\ E^* & \xleftarrow{\quad {}^tu \quad} & F^* \\ {}^tu(y^*) = y^* \circ u & & y^* \end{array}$$

Cette application est linéaire. En effet :

$$(y^* + z^*) \circ u = y^* \circ u + z^* \circ u$$
$$(\lambda y^*) \circ u = \lambda(y^* \circ u).$$

La définition de ${}^tu$:

$$\forall y^* \in E^* \qquad {}^tu(y^*) = y^* \circ u$$

signifie $\qquad \forall x \in E \qquad {}^tu(y^*)(x) = y^*(u(x)),$

soit, en utilisant la notation du paragraphe précédent :

$$< x, {}^tu(y^*) > = < u(x), y^* >.$$

Cette égalité (égalité entre deux scalaires dont le premier représente une valeur de la forme bilinéaire canonique sur $E \times E^*$ et le deuxième une valeur de la forme bilinéaire canonique sur $F \times F^*$) peut servir de définition à ${}^tu$ si elle a lieu *pour tout* $x \in E$ *et pour tout* $y^* \in F^*$.

2. Transposée d'une application composée.

Plaçons-nous dans la situation suivante : on donne trois espaces vectoriels E, F, G sur un même corps K, une application u de E dans F, une application v de F dans G et leur composée $v \circ u$, les transposées ${}^tu$ et ${}^tv$ de u et v et leur composée. Celle-ci est la transposée de $v \circ u$:

$$\begin{array}{c} E \xrightarrow{\ u\ } F \xrightarrow{\ v\ } G \quad (v \circ u : E \to G) \\ E^* \xleftarrow{\ {}^tu\ } F^* \xleftarrow{\ {}^tv\ } G^* \\ {}^tu \circ {}^tv = {}^t(v \circ u). \end{array}$$

Pour le montrer, appliquons à ${}^t(v \circ u)$ l'égalité de définition :

$$\forall x \in E \quad \forall z^* \in G^* \qquad < x, {}^t(v \circ u)(z^*) > = < (v \circ u)(x), z^* >.$$

Appliquons à v l'égalité de définition de sa transposée :

$$< v(u(x)), z^* > = < u(x), {}^tv(z^*) >.$$

Appliquons la maintenant à u :

$$< u(x), {}^tv(z^*) > = < x, ({}^tu \circ {}^tv)(z^*) >.$$

D'où $\quad \forall x \in E \quad \forall z^* \in G^* \qquad < x, {}^t(v \circ u)(z^*) > = < x, ({}^tu \circ {}^tv)(z^*) >,$

soit $\qquad {}^t(v \circ u) = {}^tu \circ {}^tv.$

Conséquence pour les automorphismes. Prenons $E = F$. Il suffit d'appliquer la définition pour voir que la transposée de l'automorphisme identique sur E est l'automorphisme identique sur E*. Ceci dit, si u est un automorphisme, il existe u^{-1} tel que :

$$u \circ u^{-1} = u^{-1} \circ u = I_E.$$

Il résulte de la propriété précédente que :

$${}^tu^{-1} \circ {}^tu = {}^tu \circ {}^tu^{-1} = {}^tI_E = I_{E^*},$$

${}^tu^{-1}$ constitue donc un inverse pour ${}^tu$ et *la transposée d'un automorphisme est un automorphisme.*

Son inverse ${}^tu^{-1} = ({}^tu)^{-1} = \check{u}$ est appelé automorphisme contragrédient de u et désigné par un symbole nouveau $\check{u}$ (qui se lit « tchèche »). Avec ce symbole on peut écrire :

$$x^* = {}^tu(y^*) \iff y^* = \check{u}(x^*),$$

et l'égalité de définition de ${}^tu$ se transcrit alors :

$$< x, x^* > = < u(x), \check{u}(x^*) >.$$

3. u et ${}^tu$ ont même rang.

Supposons E *et* F *de dimensions finies* m *et* n *respectivement.* Nous cherchons le rang :

$$\rho({}^tu) = \dim {}^tu(F^*).$$

Celui-ci est égal à la codimension du noyau H de l'application, qui est l'ensemble des formes linéaires sur F que ${}^tu$ envoie sur le zéro de E*, c'est-à-dire qui s'annulent pour tout $x \in \mathrm{E}$:

$$\begin{aligned} \mathrm{H} &= \{ y^* \,;\ \forall x \in \mathrm{E} \quad < x, {}^tu(y^*) > = 0 \} \\ &= \{ y^* \,;\ \forall x \in \mathrm{E} \quad < u(x), y^* > = 0 \} \\ &= \{ y^* \,;\ \forall y \in u(\mathrm{E}) \quad < y, y^* > = 0 \}\,. \end{aligned}$$

Mais $u(\mathrm{E})$ étant un sous-espace de F, on reconnaît là la définition de son orthogonal : $\mathrm{H} = u(\mathrm{E})^{\perp}$.

Or, on a vu (exercice 26), que $\dim u(\mathrm{E})^{\perp} = \operatorname{codim} u(\mathrm{E})$,
donc $\dim \mathrm{H} = \operatorname{codim} {}^tu(\mathrm{F}^*) = \operatorname{codim} u(\mathrm{E})$, d'où $\rho({}^tu) = \rho(u)$.

Exercice 30. — E et F étant deux espaces vectoriels de dimensions finies ou infinies, sur le même corps commutatif K, H étant un sous-espace de E on appelle $\mathcal{L}_{\mathrm{H}}(\mathrm{E}, \mathrm{F})$ le sous-espace de $\mathcal{L}(\mathrm{E}, \mathrm{F})$ constitué des applications f qui s'annulent sur H.

1) Montrer que

$$\mathcal{L}_{\mathrm{H}}(\mathrm{E}, \mathrm{F}) \approx \mathcal{L}(\mathrm{E}/\mathrm{H}, \mathrm{F}) \approx \mathcal{L}(\mathrm{G}, \mathrm{F})$$

G désignant un supplémentaire quelconque de H.

2) Montrer que

$$\mathrm{E}^* = \mathrm{G}^{\perp} \oplus \mathrm{H}^{\perp} \qquad \mathrm{H}^* \approx \mathrm{G}^{\perp} \qquad \mathrm{G}^* \approx \mathrm{H}^{\perp}$$

3) En déduire que, u étant une application linéaire de E dans F :

$${}^tu(\mathrm{F}^*) \approx [u(\mathrm{E})]^*$$

4. Transposée de la transposée.

Supposons à nouveau E *et* F *de dimensions finies*. Dans ce cas, on sait qu'on peut identifier E** et E auquel il est canoniquement isomorphe. Nous allons considérer la transposée de la transposée qui est une application de E** dans F** et montrer, les identifications de E et E**, de F et F** étant supposées faites, que ${}^{tt}u = u$:

$$\mathrm{E} \xrightarrow{\ u\ } \mathrm{F}$$

$$\mathrm{E}^* \xleftarrow{\ {}^tu\ } \mathrm{F}^*$$

$$\mathrm{E}^{**} \xrightarrow{\ {}^{tt}u\ } \mathrm{F}^{**}$$

${}^{tt}u$ peut être définie par :

$$\forall y^* \in \mathrm{F}^*,\ \forall x^{**} \in \mathrm{E}^{**} \qquad < y^*, {}^{tt}u(x^{**}) > = < {}^tu(y^*), x^{**} >,$$

soit, après l'identification :

$$< y^*, {}^{tt}u(x) > = < {}^tu(y^*), x > = < u(x), y^* >.$$

Or, cette égalité est réalisée pour tout x et tout y^* ; ceci entraîne ${}^{tt}u = u$. Car s'il existait un x pour lequel ${}^{tt}u(x)$ soit différent de $u(x)$, on pourrait toujours choisir une forme linéaire y^* qui prenne des valeurs différentes pour ces deux éléments.

5. Matrice d'une application transposée.

Soient deux espaces vectoriels E et F sur un corps K,
de bases (a_i) et (b_j),
leurs duals étant E* et F*
et les bases duales (a^i) et (b^j)

Soit $u \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ définie par $u(a_i) = \sum \alpha_i^j b_j$, c'est-à-dire par la donnée de la matrice (α_i^j).

Cherchons la matrice de ${}^tu$. Pour cela, nous devons chercher les images par ${}^tu$ des b^j et les exprimer en fonctions des a^i. Soient :

$${}^tu(b^j) = \sum_i \beta_i^j a^i$$

ces images. Pour respecter la convention « l'indice de sommation, c'est-à-dire celui qui ne figure pas dans le résultat, figure une fois en bas et une fois en haut dans chaque terme de la somme », nous sommes amenés à placer en haut l'indice de colonne et en bas l'indice de ligne dans chaque coefficient β_i^j de la matrice. Pour déterminer ${}^tu(b^j)$, cherchons l'image par cette application (${}^tu(b^j)$ est un élément de E*) des éléments de la base a_i de E. Il vient :

$$< {}^tu(b^j), a_i > = < u(a_i), b^j >$$

d'après la définition de ${}^tu$. Or,

$$< u(a_i), b^j > = < \sum_h \alpha_i^h b_h, b^j > = \alpha_i^j$$

puisque $b^j(b_h) = 0$ si $h \neq j$
$= 1$ si $h = j$.

D'autre part, $< {}^tu(b^j), a_i > = < \sum_k \beta_k^j\, a^k, a_i >$ qui, pour la même raison, vaut β_i^j.

On arrive donc à $\alpha_i^j = \beta_i^j$, α_i^j représentant le coefficient de la $i^{\text{ème}}$ colonne et $j^{\text{ème}}$ ligne de la matrice de u ; β_i^j représentant le coefficient de la $j^{\text{ème}}$ colonne et $i^{\text{ème}}$ ligne de la matrice de ${}^tu$.

La matrice de u*, relative à deux bases de* E *et* F*, et la matrice de* ${}^tu$*, relative aux bases duales des précédentes, se déduisent l'une de l'autre en intervertissant les lignes et les colonnes.*

Il en résulte, puisque u et ${}^tu$ ont même rang, que le *nombre de vecteurs colonnes linéairement indépendants de la matrice de* u (qui est le rang de u) *est égal au nombre de vecteurs lignes linéairement indépendants de la même matrice* (qui est le rang de ${}^tu$).

§ 4. REMARQUES SUR LES NOTATIONS

Lorsqu'en algèbre linéaire, on fait intervenir les coordonnées des vecteurs (et en algèbre multilinéaire, celles des tenseurs), on manipule constamment des expressions qui se présentent comme des sommes de produits de quantités indexées. Pour aérer l'écriture et fournir un moyen de contrôle des calculs, une convention classique, que nous avons adoptée ici, consiste à placer certains indices en haut et d'autres en bas de la lettre indexée, en imposant qu'un indice apparaissant dans chaque terme d'une somme figure à la même place (haut ou bas) dans le symbole désignant la somme, et que chaque indice de sommation, qui ne figure évidemment pas dans le résultat (et qui pour cette raison est appelé « indice muet »), apparaisse dans chaque terme de la somme une fois en haut et une fois en bas. Exemples :

$$\mathrm{A}_\alpha^\beta = \sum x_\alpha^i y_i^\beta$$

$$\mathrm{A}_{\alpha\beta}^\gamma = \sum_{i,j,k} x_{\alpha k}^{ij}\, y_{ij\beta}^{\gamma k}.$$

Raisonnant alors sur un certain nombre d'espaces vectoriels E, F, ... sur le même corps K, et si l'on fait la convention initiale (et forcément arbitraire) d'indexer en bas les éléments des bases utilisées dans chaque espace, la règle fixée ci-dessus impose la place des autres indices :

— 1° les coordonnées des vecteurs seront indexées en haut, en vertu de

$$x = \sum_i x^i a_i ;$$

— 2° les vecteurs des bases duales seront indexés en haut, en vertu de

$$a^i(x) = x^i ;$$

— 3° les coordonnées des vecteurs des espaces duals E*, F* seront indexées en bas, en vertu de

$$x^* = \sum_i x_i a^i,$$

et on aura :

$$< x, x^* > = \sum_i x^i x_i ;$$

— 4° l'image par une application linéaire f de E dans F du vecteur a_i de E aura des coordonnées α_i^j par rapport à la base $\{ b_j \}$ de F, données par

$$f(a_i) = \sum_j \alpha_i^j b_j.$$

Ceci conduit à :

$$y^j = \sum_j \alpha_i^j x^i.$$

— 5° l'image par une application linéaire g de E* dans F* du vecteur a^i de E* aura des coordonnées β_j^i par rapport à la base $\{ b^j \}$ de F*, données par

$$g(a^i) = \sum_j \beta_j^i b^j.$$

Ceci conduit à :

$$y_j = \sum_i \beta_j^i x_i ;$$

— 6° pour des applications u de E dans F*, ou v de E* dans F, on aura de manière analogue des formules du type :

$$u(a_i) = \sum_j \gamma_{ij} b^j$$

$$v(a^i) = \sum_j \delta^{ij} b_j.$$

Ces notations, qui sont celles du calcul tensoriel classique, ont fait leurs preuves. Même si l'on renonce à la suppression, commode et en fait très peu dangereuse, des Σ, elles ont l'avantage de permettre de distinguer les vecteurs de E, F, ... appelés vecteurs contravariants (de E, F, ... respectivement), et les vecteurs de E*, F*, ... appelés vecteurs covariants (respectivement des mêmes espaces E, F, ..., par abus de langage). Elles avertissent le calculateur de certaines inadvertances : une formule telle que $\sum \alpha_i^j x_i$ ne peut, dans ce système de notations, provenir que d'une erreur. Enfin, elles s'étendent à des tenseurs d'ordre quelconque.

En ce qui concerne les matrices, notre convention a été de ranger, dans tous les cas, en colonne, les coordonnées par rapport à la base choisie dans l'espace d'arrivée, de l'image des vecteurs de la base choisie dans l'espace de départ. Mais il faut alors remarquer que pour une application de E dans F, par exemple, l'élément α_i^j de la matrice appartient à la $i^{\text{ème}}$ colonne et à la $j^{\text{ème}}$ ligne, tandis que pour une application de F* dans E*, l'élément β_i^j appartient à la $j^{\text{ème}}$ colonne et à la $i^{\text{ème}}$ ligne. La seule donnée de l'écriture α_i^j, ou β_i^j, ne permet donc pas d'écrire la matrice (et donc de savoir si elle représente une application d'un espace E dans un espace F, ou une application d'un des espaces duals de ceux-ci dans un autre). On peut remédier à cet inconvénient en précisant les notations et en écrivant :

$${}_i\alpha^j \text{ au lieu de } \alpha_i^j$$

$${}^j\beta_i \text{ au lieu de } \beta_i^j.$$

Dans ces conditions, l'indice de gauche est toujours relatif à l'espace de départ, c'est-à-dire est, pour l'élément de la matrice, un indice de colonne, tandis que l'indice de droite est relatif à l'espace d'arrivée et est un indice de ligne. (La convention s'applique aussi bien aux cas évoqués en 6°).

Avec cette dernière précision, ce système de notations semble irréprochable : il est cohérent et donne le maximum de renseignements avec le minimum de signes.

Une autre convention peut être utilisée pour les matrices, dans le cas où n'interviendraient pas d'applications du type envisagé en 6°. Elle consiste, pour les applications d'un des espaces E*, F*, ..., dans un de ces mêmes espaces à ranger *en ligne* les coordonnées de l'image d'un élément de la base de l'espace de départ. Dans ce cas, l'indice de la colonne est toujours celui du bas. Mais cette convention réagit sur la règle du produit des matrices. On peut toutefois garder la même règle (associer les lignes de la matrice de gauche aux colonnes de la matrice de droite) à condition, si A est la matrice de f et B celle de g, de désigner par AB celle de $g \circ f$. Cela amène aussi à représenter par une matrice ligne X*(ou Y*) un vecteur x^*(ou y^*) de E*(ou F*) (vecteurs covariants) et à représenter l'équation $y^* = f(x^*)$ par l'égalité matricielle $Y^* = X^*A$.

La convention initiale étant conservée pour les matrices d'applications d'un des espaces E, F, ... dans un de ces mêmes espaces, une application de E dans F et sa transposée auront alors la même matrice, et la valeur de la forme bilinéaire canonique sur $E \times E^*$ s'écrit :

$$< x, x^* > = X^*X$$

où X* est une matrice ligne et X une matrice colonne.

CHAPITRE V

ÉTUDE DES ENDOMORPHISMES DES ESPACES VECTORIELS DE DIMENSION FINIE. FORME RÉDUITE DES MATRICES CARRÉES

§ 1. SOUS-ALGEBRE DE $\mathcal{L}$ (E, E) ENGENDREE PAR UN ENDOMORPHISME

1. Idéal annulateur.

Soit E un espace vectoriel de dimension n sur un corps K commutatif et f un endomorphisme de E, c'est-à-dire un élément de l'algèbre $\mathcal{L}$ (E, E).

Dans le but d'étudier l'endomorphisme f, nous allons commencer par étudier *la sous-algèbre engendrée par f et e* (e étant l'endomorphisme identique). Cette sous-algèbre comprend les homothéties λe (produit de l'identité par un scalaire), les puissances (ou itérées) de f et les composées et sommes de ces applications. Remarquons alors que :

1) le fait que f soit linéaire entraîne que

$$\lambda e \circ f = f \circ \lambda e = \lambda f$$

puisque ceci signifie $\forall x \in \mathrm{E} \quad \lambda f(x) = f(\lambda x)$;

2) le fait que la multiplication soit associative entraîne que

$$f^m \circ f^n = f^{m+n} = f^n \circ f^m.$$

Nous en déduisons que la sous-algèbre étudiée contient tous les endomorphismes représentés comme polynômes en f à coefficients dans K et que le composé de deux tels endomorphismes s'exprime par le produit des polynômes effectué comme s'il s'agissait de polynômes de $\mathrm{K}[x]$. C'est dire que l'ensemble $\mathrm{K}[f]$ des polynômes en f à coefficients dans K est l'image de $\mathrm{K}[x]$ par l'homomorphisme φ qui au polynôme formel $a_0 + a_1 x + \ldots + a_k x^k$ fait correspondre l'endomorphisme de E :

$$a_0 + a_1 f + \ldots + a_k f^k.$$

$\mathrm{K}[f]$ est donc une algèbre, c'est la sous-algèbre de $\mathcal{L}$ (E, E) engendrée par f et e. Notons qu'elle est commutative, alors que $\mathcal{L}$ (E, E) ne l'est pas.

Nous allons voir que l'homomorphisme φ n'est jamais un isomorphisme, en montrant que le noyau $\overset{-1}{\varphi}(0)$ de cet homomorphisme est différent de zéro. Ce noyau est un idéal puisqu'une algèbre est d'abord un anneau et que le noyau d'un homomorphisme d'anneau est un idéal. Il est constitué par les polynômes de $\mathrm{K}[x]$ dont l'image par φ est l'application linéaire nulle.

Soit $\{ u_0, u_1 \dots u_n \}$ une base de E. Considérons un des vecteurs u_i de cette base et ses images par les puissances de f ; les $n+1$ vecteurs u_i, $f(u_i)$, $f^{(2)}(u_i) \dots f^{(n)}(u_i)$, que nous convenons de noter dans ce paragraphe $fu_i \dots f^n u_i$, ne peuvent être linéairement indépendants ($\dim E = n$), donc

$$\exists\, \alpha_0, \alpha_1 \dots \alpha_n \in K, \text{ non tous nuls et tels que}$$

$$\alpha_0 u_i + \alpha_1 f u_i + \dots + \alpha_n f^n u_i = 0.$$

Autrement dit, il existe un polynôme non nul $\Pi_i[x] \in K[x]$ tel que :

$$\Pi_i(f) u_i = 0.$$

Déterminons ce polynôme Π_i pour chaque élément u_i et considérons le produit :

$$\Pi = \Pi_1 \Pi_2 \dots \Pi_n$$

de tous ces polynômes. L'endomorphisme $\Pi(f)$ envoie sur zéro tout élément de la base ; en effet, puisque la multiplication dans la sous-algèbre est commutative on peut, pour tout u_i, amener dans Π, Π_i en dernier ; on aura $\Pi_i(f)u_i = 0$, donc $\Pi_j \Pi_i(f) u_i = 0 \dots$ et $\Pi(f)u_i = 0$.

L'application linéaire $\Pi(f)$, qui annule tout élément de la base de E, annule tout E :

$$\Pi(f)E = 0.$$

Or, $\Pi(f)$ est l'image par φ d'un polynôme $\Pi \in K[x]$; ce polynôme $\Pi \neq 0$ appartient à l'idéal $\overset{-1}{\varphi}(0)$; φ n'est pas un isomorphisme.

Mais $K[x]$ étant un anneau principal, un idéal y est formé des multiples d'un polynôme unique ϖ qui est nécessairement un diviseur de Π. Les polynômes de l'idéal (ϖ) ont pour image par φ l'application linéaire nulle (qui envoie tout vecteur de E sur 0). On dit que l'idéal (ϖ) annule E.

Théorème : Etant donné un espace vectoriel de dimension finie et un endomorphisme f, il existe un polynôme non nul $\varpi \in K[x]$ qui engendre un idéal qui annule E.

2. Décomposition de E en somme directe de sous-espaces invariants.

Pour pousser plus loin cette étude, admettons que ce polynôme ϖ soit le produit de deux polynômes p et q premiers entre eux. Il existe donc deux polynômes λ et μ tels que :

$$\lambda p + \mu q = 1.$$

L'application $\lambda(f)p(f) + \mu(f)q(f)$ est alors l'application identique. Il en résulte que tout $x \in E$ peut s'écrire :

$$x = \lambda(f)p(f)x + \mu(f)q(f)x \qquad (1)$$

donc peut être mis sous forme d'une somme de deux éléments dont le premier appartient à l'image de E par $\lambda(f)p(f)$ et le deuxième à l'image de E par $\mu(f)q(f)$. Le premier de ces sous-espaces est annulé par l'application $q(f)$, car :

$$\forall x \in E \quad q(f)\lambda(f)p(f)x = \lambda(f)p(f)q(f)x = \lambda(f)\varpi(f)x = 0.$$

Nous l'appellerons E_q en posant par conséquent :

$$E_q = \lambda(f)p(f)E \text{ avec } q(f)E_q = 0.$$

Le deuxième est, de même, annulé par $p(f)$. On pose :

$$E_p = \mu(f)q(f)E \text{ et on a } p(f)E_p = 0.$$

On a donc : $E = E_p + E_q$.
D'autre part, $E_p \cap E_q = \{0\}$ car si $x \in E_p \cap E_q$,

$$p(f)x = 0 \qquad q(f)x = 0,$$

donc, d'après (1), $x = 0$. On peut donc écrire :

$$E = E_p \oplus E_q.$$

Ceci entraîne que les vecteurs de E_p sont les seuls à être annulés par p, car si un vecteur est annulé par p, sa composante dans E_q, qui est $\lambda(f)p(f)x$, est nulle. Le sous-espace E_p peut donc être défini comme l'ensemble des vecteurs de E annulés par p.

Il en résulte que $f(E_p) \subset E_p$.
En effet, $x \in E_p \Longrightarrow p(f)x = 0$; or, $p(f)fx = fp(f)x$ donc $p(f)fx = 0$, donc $fx \in E_p$.
De la même façon, $f(E_q) \subset E_q$.

On peut dire que l'endomorphisme f « ne mélange » pas les composantes, l'image de chacun des sous-espaces supplémentaires E_p et E_q étant incluse dans lui-même. Il en résulte évidemment que l'image de E_p (resp E_q), par n'importe quel endomorphisme de la sous-algèbre étudiée, est incluse dans E_p (resp E_q). La restriction à E_p (resp E_q) de f et de tout endomorphisme de $K[f]$ est un endomorphisme de E_p (resp E_q).

Enfin, remarquons que p engendre l'idéal annulant E_p. En effet, si ce n'était pas p, ce serait un diviseur p_1 de p. Le polynôme $p_1 q$ n'appartiendrait pas à l'idéal engendré par ϖ et pourtant $p_1(f)q(f)E$ serait égal à zéro puisque :

$$\forall x \in E \quad x = x_p + x_q \quad x_p \in E_p \quad x_q \in E_q$$

$$q(f)x = q(f)x_p \text{ donc } q(f)E = q(f)E_p,$$

mais, d'après la remarque ci-dessus :

$$q(f)E_p \subset E_p,$$

or $p_1(f)$ annule tous les éléments de E_p. Il y aurait donc contradiction.

En résumé, *si (ϖ) est l'idéal qui annule* E *et si $\varpi = pq$, p et q étant premiers entre eux,* E *est la somme directe de deux sous-espaces* E_p et E_q *dont chacun contient son image par f ; (p) et (q) sont respectivement les idéaux annulateurs des restrictions de f à E_p et E_q.*

Si p (ou q) est encore décomposable en deux polynômes premiers entre eux, on peut recommencer l'opération et E_p (ou E_q) est décomposé en somme directe. En définitive, si

$$\varpi = \varpi_1 \varpi_2 \dots \varpi_n,$$

$\varpi_1 \varpi_2 \dots \varpi_n$ étant des polynômes premiers entre eux deux à deux, on pourra dire que :

$$E = E_1 \oplus E_2 \oplus \dots \oplus E_n,$$

$E_1, E_2 \dots E_n$ étant des sous-espaces vectoriels ayant respectivement $(\varpi_1) \dots (\varpi_n)$ pour idéaux annulateurs.

3. Conséquence pour la représentation matricielle de *f*.

On choisira une base de E constituée par la réunion d'une base de E_1, d'une base de E_2, ... d'une base de E_n. D'après les propriétés précé-

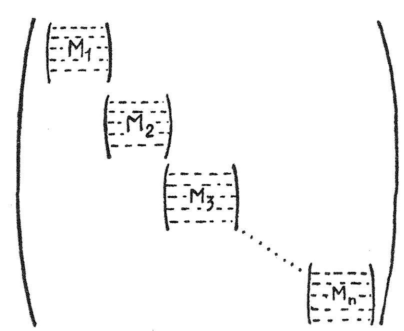


FIG. 5

dentes, les images d'un élément de la base de E_i sont dans E_i ; il en résulte que la matrice correspondante aura la forme ci-dessus, où $M_1, M_2 \dots M_n$ représentent les matrices des endomorphismes restrictions de f à $E_1, E_2 \dots E_n$ et où tout le reste de la matrice ne comprend que des zéros.

Si on ne sait rien du corps K, là s'arrête ce qu'on peut dire de f.

§ 2. CAS OU LE CORPS EST ALGEBRIQUEMENT CLOS MATRICES TRIANGULAIRES

1. *Si* K *est algébriquement clos,* le polynôme ϖ peut se factoriser de la façon suivante :

$$\varpi = \prod_i (x - \lambda_i)^{\alpha_i}.$$

Le polynôme annulateur de la restriction de f à un sous-espace E_i est la puissance $\alpha_i^{\text{ième}}$ d'un polynôme du premier degré. Nous sommes donc ramenés à étudier un endomorphisme dont l'idéal annulateur est de la forme $\left((x-\lambda)^{\alpha}\right)$.

Si l'idéal annulateur d'un endomorphisme f est engendré par le polynôme $(x-\lambda_i)^{\alpha_i}$, celui de l'endomorphisme $g = f - \lambda_i e$ est engendré par le polynôme x^{α_i}. Supprimons provisoirement l'indice i de α_i et de E_i.

2. Endomorphisme dont une puissance est nulle.

L'étude à faire est donc ramenée à celle d'un endomorphisme g d'un espace vectoriel E de dimension finie, tel que g^{α} soit l'endomorphisme nul et que g^j ne le soit pas pour $j < \alpha$. De $g(E) \subset E$, on déduit $g^2(E) \subset g(E) \dots$, $g^j(E) \subset g^{j-1}(E)$. Si l'on pose $E_j = g^j(E)$, la suite de sous-espaces :

$$E = E_0 \supset E_1 \supset E_2 \dots \supset E_j \supset E_{j+1} \quad \supset E_{\alpha} = \{0\}$$

est donc décroissante et, seul, E_{α} est réduit à $\{0\}$. Cette suite est strictement décroissante, car de $E_{j+1} = E_j$ on déduirait :

$$E_{j+2} = g(E_{j+1}) = g(E_j) = E_{j+1} = E_j$$

et par suite, de proche en proche, $E_j = E_{\alpha} = \{0\}$, ce qui est impossible pour $j \neq \alpha$.

On peut, à partir de cette remarque, décomposer E en une somme directe de sous-espaces F_j en prenant $F_{\alpha-1} = E_{\alpha-1}$ et, pour $F_j (j < \alpha - 1)$, un supplémentaire de E_{j+1} par rapport à E_j. On aura :

$$E = \bigoplus_{i=0}^{\alpha-1} F_j.$$

Aucun des sous-espaces F_j n'est réduit à $\{0\}$, leur nombre, qui est égal à α, est au plus égal à la dimension β de E. Si l'on choisit une base de chaque F_j, leur réunion sera une base de E, et si on ordonne cette base en prenant successivement les éléments de la base choisie en F_0, puis en $F_1, \dots$, puis en $F_{\alpha-1}$, on déduit de $g(F_j) \subset E_{j+1}$ que la matrice de g par rapport à cette base n'aura que des 0 sur la diagonale principale et au-dessus de cette diagonale. $g(E_{\alpha-1})$ étant nul, les dernières colonnes de la matrice, qui correspondent aux éléments de la base appartenant à $E_{\alpha-1} = F_{\alpha-1}$, ne seront formées que de zéros.

Si nous revenons à la matrice de $g + \lambda_i e$, par rapport à la même base, elle aura une diagonale principale formée de termes tous égaux à λ_i et

tous les termes situés au-dessus de la diagonale principale seront nuls. En outre, dans les dernières colonnes (correspondant à $F_{\alpha-1}$), il n'y aura que des zéros en-dessous de la diagonale.

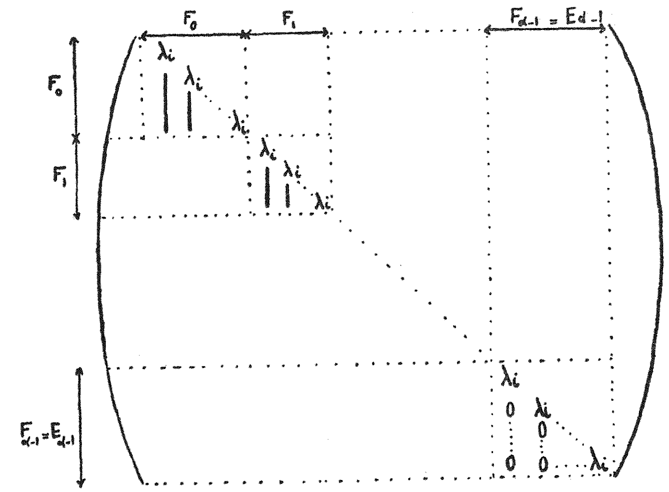


Fig. 6

Sous cette forme, on voit que les vecteurs de la base de $E_{\alpha-1}$ sont transformés par l'endomorphisme en leur produit par λ. Il en résulte qu'il en est de même pour tout vecteur de $E_{\alpha-1}$:

$$\forall x \in E_{\alpha-1} \qquad f(x) = (g + \lambda_i e)(x) = \lambda_i x.$$

La restriction de f à ce sous-espace est donc une homothétie de rapport λ_i. $E_{\alpha-1}$ étant différent de $\{0\}$, il existe au moins un sous-espace vectoriel de dimension 1 pour lequel il en est ainsi.

3. Retour au cas général.

Revenons maintenant à la matrice de l'application f dont le polynôme annulateur est :

$$\prod (x - \lambda_i)^{\alpha_i}.$$

Nous voyons, à la lumière de ce qui précède, que nous pouvons préciser la forme trouvée en (V, 1, 3) pour la matrice de l'application f, chacune des matrices M_i de la figure prenant la forme de la figure 6. Sur la diagonale principale ne figurent que les valeurs λ_i.

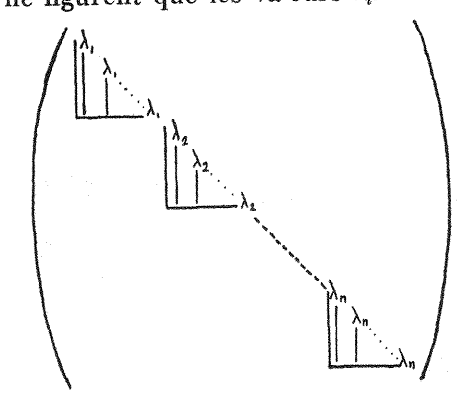


Fig. 7

Rappelons que l'ordre β_i de la $i^{\text{ème}}$ sous-matrice, dimension du sous-espace vectoriel annulé par $(x - \lambda_i)^{\alpha_i}$, est au moins égal à α_i.

Une matrice ayant comme celle de la figure ci-dessus tout le triangle situé au-dessus de la diagonale principale formé de zéros est dite *triangulaire* (*).

Exercice 31. — Montrer que les matrices (n, n) triangulaires forment une sous-algèbre de l'algèbre des matrices (n, n).

4. Vecteurs propres et valeurs propres.

Revenons sur certains des résultats précédents. Les nombres λ_i qui sont les racines, dans K, du polynôme ϖ sont appelés les *valeurs propres de f*. Les vecteurs *x non nuls* tels que $f(x) = \lambda_i x$ sont dits *vecteurs propres* relatifs à la valeur λ_i et nous venons de voir que pour tout λ_i il existe un sous-espace vectoriel de dimension 1 au moins formé de vecteurs propres.

Cas où toutes les racines du polynôme annulateur sont simples. Examinons le cas particulier où $\alpha_i = 1$. L'application $g - \lambda_i e$ est l'application nulle ; la restriction de f au sous-espace E_i est donc l'homothétie de rapport λ_i ; en d'autres termes, tous les vecteurs de E_i sont des vecteurs propres relatifs à λ_i (c'est le cas où E_i se confond avec son sous-espace $E_{\alpha-1}$).

Si maintenant pour tout i, $\alpha_i = 1$, on voit que tous les vecteurs de la base choisie sont des vecteurs propres et que, par rapport à cette base, la matrice de f est une matrice diagonale, c'est-à-dire une matrice qui n'a de coefficients différents de zéro que sur sa diagonale principale.

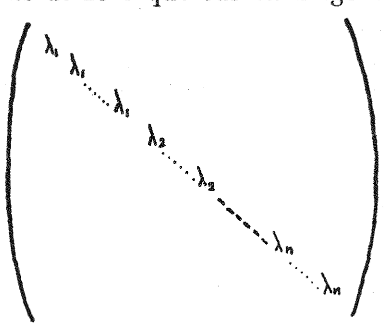


FIG. 8

On dit couramment, mais par abus de langage, qu'on a *diagonalisé* la matrice M qui définissait f dans une base initialement donnée, voulant dire par là qu'on a trouvé une matrice semblable de forme diagonale.

Réciproquement, si la matrice d'un endomorphisme f peut être diagonalisée sous la forme ci-dessus, le sous-espace E_i correspondant aux colonnes de la matrice où l'on trouve λ_i est annulé par $\lambda_i e - f$ et, en vertu du raisonnement fait au début du chapitre E, admet $\varpi = \Pi(\lambda_i - x)$ comme idéal annulateur. Les α_i sont donc tous égaux à 1.

On peut donc conclure que *la condition nécessaire et suffisante pour qu'une matrice d'un endomorphisme soit diagonalisable est que les racines du polynôme qui engendre l'idéal annulateur soient toutes simples.*

Nous allons plus loin (chap. VI, § 4), en utilisant la théorie des déterminants, montrer qu'il n'y a pas d'autres vecteurs propres que ceux que

(*) On peut aussi considérer des matrices triangulaires, pour lesquelles le triangle formé de zéros est situé au-dessous de la diagonale principale. Il faut évidemment dans l'exercice entendre qu'on considère les matrices triangulaires d'une seule des espèces indiquées.

nous avons trouvés et former une équation dont les racines seront encore les valeurs propres de f, mais dans laquelle chaque valeur propre λ_i aura pour multiplicité β_i, la dimension du sous-espace de E annulé par $(f - \lambda_i e)^{\alpha_i}$. Mais, auparavant, nous allons montrer qu'on peut obtenir une décomposition plus fine de E à laquelle sera associée une matrice plus simple.

§ 3. FORME REDUITE DE JORDAN

Nous introduisons d'abord les noyaux des applications g^i. Nous désignerons par N_i le noyau de l'endomorphisme g^i. De $g^{i+1} = g^i \circ g$, on déduit $N_{i+1} = \overset{-1}{g}(N_i)$ et de $\{0\} \subset N_1$, on déduit $N_1 \subset N_2, \ldots, N_i \subset N_{i+1}$. La suite de sous-espaces $N_0 = \{0\} \subset N_1 \subset N_2 \subset N_i \subset N_{i+1} \ldots \subset N = E_\alpha$ est donc croissante. Elle l'est strictement, car $N_1 = \{0\}$ signifierait que g est un automorphisme, et il en serait de même de g^α, alors que $g^\alpha = 0$. D'autre part, de $N_i = N_{i+1}$, on déduirait :

$$N_{i+2} = \overset{-1}{g}(N_{i+1}) = \overset{-1}{g}(N_i) = N_{i+1} = N_i$$

et par suite, de proche en proche, $N_i = N_\alpha = E$, d'où $g^i = 0$, ce qui n'a lieu que pour $i = \alpha$.

La connaissance de la suite des E_i, images de E par les puissances de g, et de la suite des N_i, noyaux des mêmes puissances, donne une idée du comportement de l'application g. On peut préciser cette idée en faisant intervenir les deux suites à la fois et en considérant tout d'abord les sous-espaces :

$$N_i \cap E_j \qquad \begin{matrix} 0 \leqslant j \leqslant \alpha - 1 \\ 1 \leqslant i \leqslant \alpha \end{matrix}$$

on a $N_i \cap E_j = E_j$ si $j \geqslant \alpha - i$, car $g^{\alpha-j}(E_j) = 0$ donc $E_j \subset N_{\alpha-j}$ et, *à fortiori*, E_j est contenu dans les noyaux d'indices supérieurs.

D'autre part, on peut énoncer :

Lemme : $\qquad g(N_i \cap E_j) = N_{i-1} \cap E_{j+1}.$

En effet, de $N_i = \overset{-1}{g}(N_{i-1})$ on déduit :

$$g(N_i) = N_{i-1} \cap g(E) = N_{i-1} \cap E_1.$$

D'autre part, $g(E_j) = E_{j+1}$, et l'image d'une intersection étant incluse dans l'intersection des images on a :

$$g(N_i \cap E_j) \subset N_{i-1} \cap E_1 \cap E_{j+1} = N_{i-1} \cap E_{j+1}.$$

Mais, réciproquement, si $x \in N_{i-1} \cap E_{j+1}$, l'ensemble $\overset{-1}{g}(x)$ est inclus dans N_i et a des éléments dans E_j. Il existe donc $y \in N_i \cap E_j$ tel que $g(y) = x$. Donc, $x \in g(N_i \cap E_j)$, c'est-à-dire $N_{i-1} \cap E_{j+1} \subset g(N_i \cap E_j)$, et on a bien l'égalité annoncée.

Nous allons maintenant successivement décomposer les noyaux N_i en somme directe de sous-espaces N_k^l qui nous permettront par groupement convenable de reconstituer aussi par somme directe les sous-espaces E_j, et tels que la restriction de g à N_k^l en sera un isomorphisme sur N_{k-1}^{l+1}.

Nous aurons à utiliser plusieurs fois les remarques suivantes dont la démonstration est évidente :

1re remarque : Si V et W sont deux sous-espaces vectoriels d'un espace E, si V' (resp W') est un supplémentaire de $V \cap W$ par rapport à V (resp W), la somme $V + W$ admet la décomposition en somme directe :

$$V' \oplus (V \cap W) \oplus W'.$$

2e remarque : Si V et W sont deux sous-espaces vectoriels de E, tels que la somme $V + W$ soit directe, il existe un supplémentaire de V par rapport à E qui contient W.

Nous commençons par attaquer N_1 et remarquons que la suite $N_1 \cap E_j$, $j = 1, 2, \ldots, \alpha - 1$, est croissante (mais pas forcément strictement) :

$$N_1 \cap E_{\alpha-1} \subset N_1 \cap E_{\alpha-2} \subset \ldots\ldots \subset N_1 \cap E_1 \subset N_1$$

et que $N_1 \cap E_{\alpha-1}$ est $E_{\alpha-1}$ et est certainement distinct de $\{0\}$. On peut alors décomposer N_1 en somme directe de sous-espaces N_1^i tels que :

$$N_1 \cap E_i = N_1^i \oplus (N_1 \cap E_{i+1})$$

(N_1^i est un supplémentaire de $N_1 \cap E_{i+1}$ par rapport à $N_1 \cap E_i$),

$$N_1 = N_1^{\alpha-1} \oplus N_1^{\alpha-2} \oplus \ldots\ldots \oplus N_1^0$$

Certains de ces sous-espaces N_1^i peuvent être réduits à $\{0\}$, mais ce n'est certainement pas le cas de $N_1^{\alpha-1}$, car :

$$E_{\alpha-1} = N_1 \cap E^{\alpha-1} = N_1^{\alpha-1} \oplus (N_1 \cap E\alpha) = N_1^{\alpha-1}$$

Nous représenterons dans un schéma la décomposition ci-dessus :

$N_1^{\alpha-1}$	$N_1^{\alpha-2}$		N_1^1	N_1^0

Rappelons que ce schéma ne doit pas induire en erreur sur le fait que N_1 est la somme directe des N_1^i et non pas leur réunion.

Nous attaquons alors N_2, et considérons $N_2 \cap E_{\alpha-2} = E_{\alpha-2}$ et la restriction de g à ce sous-espace. Le noyau de la restriction de g à un sous-espace U de E est $U \cap N_1$. Ici, c'est donc $N_1 \cap E_{\alpha-2} = N_1^{\alpha-1} \oplus N_1^{\alpha-2}$, et l'image de $N_2 \cap E_{\alpha-2}$ est $N_1 \cap E_{\alpha-1}$, en vertu du lemme. Si par suite $N_2^{\alpha-2}$ est un supplémentaire de $N_1^{\alpha-1} \oplus N_1^{\alpha-2}$ par rapport à $N_2 \cap E_{\alpha-2}$, la restriction de g à $N_2^{\alpha-2}$ en est un isomorphisme sur $N_1^{\alpha-1}$. On a donc :

$$E_{\alpha-2} = N_2 \cap E_{\alpha-2} = N_1^{\alpha-1} \oplus N_1^{\alpha-2} \oplus N_2^{\alpha-2}$$

D'autre part, $N_1^{\alpha-3}$ est un supplémentaire de $N_1 \cap E_{\alpha-2}$ par rapport à $N_1 \cap E_{\alpha-3}$; nous sommes dans la situation de la première remarque faite ci-dessus : $N_2 \cap E_{\alpha-2}$ et $N_1 \cap E_{\alpha-3}$, dont l'intersection est :

$$N_1 \cap E_{\alpha-2} = N_1^{\alpha-1} \oplus N_1^{\alpha-2},$$

jouant le rôle des sous-espaces V et W, la somme des quatre sous-espaces : $N_1^{\alpha-1}$, $N_1^{\alpha-2}$, $N_1^{\alpha-3}$, $N_2^{\alpha-2}$ est directe.

Nous complétons le schéma précédent :

$N_2^{\alpha-2}$					
$N_1^{\alpha-1}$	$N_1^{\alpha-2}$	$N_1^{\alpha-3}$	$N_1^{\alpha-4}$		N_1^0

Nous pouvons faire un raisonnement analogue pour $N_2 \cap E_{\alpha-3}$. La restriction de g à ce sous-espace a pour image $N_1 \cap E_{\alpha-2} = N_1^{\alpha-1} \oplus N_1^{\alpha-2}$ et pour noyau $N_1 \cap E_{\alpha-3} = N_1^{\alpha-1} \oplus N_1^{\alpha-2} \oplus N_1^{\alpha-3}$. Considérons alors un supplémentaire U du noyau $N_1 \cap E_{\alpha-3}$ par rapport à $N_2 \cap E_{\alpha-3}$, qui contienne $N_2^{\alpha-2}$, ce qui est possible, en vertu de la deuxième remarque du début, puisque la somme $(N_1 \cap E_{\alpha-3}) \oplus N_2^{\alpha-2}$ est directe. La restriction de g à U en est un isomorphisme sur $N_1^{\alpha-1} \oplus N_1^{\alpha-2}$. L'isomorphisme réciproque envoie $N_1^{\alpha-1}$ sur $N_2^{\alpha-2}$ et $N_1^{\alpha-2}$ sur un sous-espace que nous désignerons par $N_2^{\alpha-3}$. On a donc :

$$N_2 \cap E_{\alpha-3} = N_1^{\alpha-1} \oplus N_1^{\alpha-2} \oplus N_1^{\alpha-3} \oplus N_2^{\alpha-2} \oplus N_2^{\alpha-3}.$$

On peut prolonger le schéma :

2	$N_2^{\alpha-2}$	$N_2^{\alpha-3}$				
1	$N_1^{\alpha-1}$	$N_1^{\alpha-2}$	$N_1^{\alpha-3}$	$N_1^{\alpha-4}$		N_1^0

La somme des 6 espaces figurés à gauche est directe aussi, car $N_1^{\alpha-4}$ est un supplémentaire de $N_1 \cap E_{\alpha-3}$ par rapport à $N_1 \cap E_{\alpha-4}$, ce qui nous met encore dans la situation de la première remarque, $N_2 \cap E_{\alpha-3}$ et $N_1 \cap E_{\alpha-4}$, dont l'intersection est $N_1 \cap E_{\alpha-3}$, jouant cette fois le rôle de V et W. La restriction de g à chaque sous-espace de la ligne 2 en est un isomorphisme sur le sous-espace de la ligne 1 figuré dans la case immédiatement au-dessous de la sienne.

Le même raisonnement s'appliquera successivement aux espaces $N_2 \cap E_j$, puis aux espaces $N_3, N_4, \ldots$ Nous exposerons donc le raisonnement général.

On suppose que, pour $i < i_0$ et $j = 0, 1, \ldots, \alpha - 1$, les sous-espaces $N_i \cap E_j$ ont été décomposés en sommes directes :

$$N_i \cap E_j = \oplus \{ N_k^l ;\ 1 \leqslant k \leqslant i,\ j \leqslant l \leqslant \alpha - 1 \},$$

que les sous-espaces $N_{i_0} \cap E_j$ ont été décomposés de la même manière pour $j \geqslant j_0 + 1$, et que la restriction de g à N_k^l en est un isomorphisme sur N_{k-1}^{l+1}.

Considérons alors $N_{i_0} \cap E_{j_0}$ et la restriction de g à ce sous-espace. Le noyau en est $N_1 \cap E_{j_0}$ et l'image $N_{i_0-1} \cap E_{j_0+1}$. On applique la première remarque à $V = N_{i_0} \cap E_{j_0+1}$ et à $W = N_1 \cap E_{j_0}$ dont l'intersection est $N_1 \cap E_{j_0+1}$. Comme $N_1^{j_0}$ est un supplémentaire W' de $N_1 \cap E_{j_0+1}$ par rapport à $N_1 \cap E_{j_0}$, et que :

$$V' = \oplus \{ N_k^l ;\ 1 < k \leqslant i_0,\ j_0 + 1 \leqslant l \leqslant \alpha - 1 \}$$

est un supplémentaire de $N_1 \cap E_{j_0+1}$ par rapport à $N_{i_0} \cap E_{j_0+1}$, la somme $V' + (V \cap W) + W'$ est directe. Elle est, en outre, incluse dans $N_{i_0} \cap E_{j_0}$; et, en vertu de la 2e remarque, il existe un supplémentaire U de $N_1 \cap E_{j_0} = W$ par rapport à $N_{i_0} \cap E_{j_0}$ qui contient V'. La restriction de g à U en est un isomorphisme sur l'image $N_{i_0-1} \cap E_{j_0+1}$ qui peut s'écrire :

$$\oplus \{ N_k^l ;\ 1 \leqslant k \leqslant i_0 - 1,\ j_0 + 1 \leqslant l \leqslant \alpha - 1 \}.$$

Dans l'isomorphisme inverse, $N_{i_0-1}^{j_0+1}$ a pour image un sous-espace de $N_{i_0} \cap E_{j_0}$ que l'on notera $N_{i_0}^{j_0}$, qui sera tel que la restriction de g à $N_{i_0}^{j_0}$ en soit un isomorphisme sur $N_{i_0-1}^{j_0+1}$ et d'autre part tel que :

$$N_{i_0} \cap E_{j_0} = \oplus \{ N_k^l ;\ 1 \leqslant k \leqslant i_0,\ j_0 \leqslant j \leqslant \alpha - 1 \} .$$

Le résultat définitif peut être schématisé sur la figure :

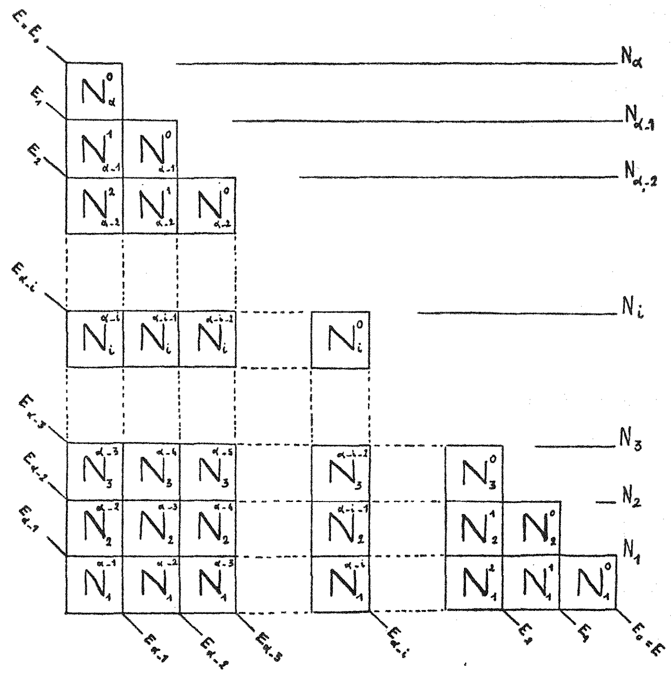


FIG. 9

— La somme directe des sous-espaces figurés dans les i premières lignes (en comptant à partir du bas) donne N_i, la somme directe des sous-espaces figurés dans les i premières diagonales (en comptant à partir du bas) donne $E_{\alpha-i}$.

— Les vecteurs de N_k^l sont les images de vecteurs de E par g^l et ont zéro pour image par g^k. La restriction de g à N_k^l en est un isomorphisme sur N_{k-1}^{l+1} qui est figuré dans la case immédiatement au-dessous.

— La décomposition indiquée n'est pas unique, mais une autre décomposition M_k^l jouissant des mêmes propriétés est telle que M_k^l est isomorphe à N_k^l.

— Enfin, le sous-espace $N_1^{\alpha-1} = E_{\alpha-1}$ et tous ceux de sa colonne qui lui sont isomorphes ne sont pas réduits à $\{0\}$. Les autres colonnes, par contre, peuvent ne représenter que des sous-espaces réduits à $\{0\}$.

A cette décomposition de E en somme directe, on peut associer une base par rapport à laquelle la matrice de g sera particulièrement simple.

A cet effet, on prendra une base dans chacun des sous-espaces N_i^0. Les images successives d'une base de N_i^0 sont des bases des sous-espaces de la même colonne. E étant la somme directe des N_k, la réunion des bases de N_k^l ainsi obtenues est une base de E. On en rangera les éléments de la manière suivante : on prend un élément de la base de N_α^0 suivi de toutes ses images par $g, g^2, ..., g^{\alpha-1}$, puis, s'il existe un deuxième élément de la base de N_α^0 suivi des images successives, par g et ceci jusqu'à épuisement de la base de N_α^0. On procédera alors de la même manière pour la base choisie dans $N_{\alpha-1}^0$, $N_{\alpha-2}^0$, etc...

Si u_1 est le premier élément de la base choisie en N_α^0, on aura donc, pour l'ordre dont on a muni la base de E :

$$g(u_1) = u_2 \quad g(u_2) = u_3 \dots g(u_{\alpha-1}) = u_\alpha \quad g(u_\alpha) = 0.$$

Si p est l'indice dans la base de E d'un élément de la base choisie dans N_i^0, on aura de même :

$$g(u_p) = u_{p+1} \dots g(u_{p+i-1}) = u_{p+i} \quad g(u_{p+i}) = 0.$$

la matrice de g sera donc formée en bordant de zéros des matrices disposées le long de la diagonale principale et qui seront du type :

$$\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

c'est-à-dire comportant des 1 sur la sous-diagonale principale et des zéros partout ailleurs. La première de ces matrices (éventuellement la seule qui existe) étant d'ordre α. La somme des ordres de ces matrices est la dimension de E. Si l'on revient à l'application f, sa matrice se déduira de la précédente en substituant λ aux zéros de la diagonale principale.

Exercice 32. — E est un espace vectoriel sur un corps commutatif K, dont l'unité sera notée 1.

1) On suppose dans cette question que la dimension de E est deux et on considère l'endomorphisme u de E dont la matrice relativement à une base donnée (e_1, e_2) est

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

Déterminer directement les vecteurs propres et les valeurs propres de u.

2) E étant de dimension quelconque (finie ou non), et H étant un sous-espace vectoriel de E de codimension un, on considère les endomorphismes u de E qui laissent invariant chaque vecteur de H.

Montrer que l'image $u(\dot{x})$ d'une classe mod H est une classe mod H. Montrer que l'application $\dot{x} \longrightarrow u(\dot{x})$ est un endomorphisme de E/H et est par suite de la forme $u(\dot{x}) = \lambda\dot{x}$, où λ est un élément de K ne dépendant que de u.

On dira que u est une transvection si $\lambda = 1$, et une dilatation si $\lambda \neq 1$.

Si e_0 est un vecteur n'appartenant pas à H, montrer que l'on a

$$u(e_0) = \lambda e_0 + e_1, \text{ où } e_1 \in \mathrm{H}.$$

3) Désignant par L un sous-espace supplémentaire de Ke_1 par rapport à H, tout vecteur $x \in E$ peut s'écrire de manière unique sous la forme

$$x = \xi_0 e_0 + \xi_1 e_1 + x_L$$
$$\text{où } \xi_0 \in K,\ \xi_1 \in K,\ x_L \in L.$$

Montrer qu'il y a exactement une droite (sous-espace vectoriel de dimension un) supplémentaire de H invariante par u, si u est une dilatation, et qu'il n'y en a pas, si u est une transvection différente de l'identité.

Dans le cas où la dimension de E est finie, en déduire que la matrice d'une dilatation peut prendre la forme diagonale (on donnera la valeur des éléments diagonaux), tandis que ceci est impossible pour une transvection différente de l'identité.

4) Qu'est $u(x) - x$, si u est une transvection ?

Montrer que si φ est une forme linéaire sur E telle que $\mathrm{H} = \overset{-1}{\varphi}(0)$, à toute transvection u correspond un vecteur $a \in \mathrm{H}$ tel que

$$u(x) = x + \varphi(x)a$$

Montrer que si E est de dimension finie, il existe pour toute transvection une base par rapport à laquelle sa matrice a tous ses éléments diagonaux égaux à 1, et au plus un élément non diagonal différent de zéro.

5) Montrer : *a*) que l'ensemble des automorphismes de E laissant invariant chaque vecteur de H est un sous-groupe G(H) de GL(E) (la loi est la composition des applications) ;

b) que l'ensemble $\Gamma(\mathrm{H})$ des transvections est un sous-groupe commutatif invariant de G(H), et que $\Gamma(\mathrm{H})$ est isomorphe à H ;

c) que le groupe quotient $G(H)/\Gamma(H)$ est isomorphe au groupe multiplicatif de K.

Exercice 33. — Soit E un espace vectoriel sur un corps K commutatif. On appellera involution un automorphisme f de E qui vérifie $f^2 = e$ (e : automorphisme identique).

1) On suppose que la caractéristique de K est différente de 2. Refaire sur cet exemple la théorie du cours (V. 1). En déduire que si on pose $g = e + f$, $h = e - f$, E est somme directe de $V = g(E)$ et $W = h(E)$, que la restriction de f à V est l'identité e_V et celle de f à W est $-e_W$.

En déduire que toute matrice carrée M sur K telle que $M^2 = I$ est semblable à une matrice

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$$

où I_p et I_q représentent les matrices unités d'ordre p et d'ordre q.

2) Si K est de caractéristique 2, montrer que f est de la forme $e + g$ avec $g^2 = 0$. n étant la dimension de E et p celle de $g(E)$, montrer que l'on a

$$2p \leqslant n.$$

Quel type d'application retrouve-t-on dans le cas $p = 1$ (cf. Exercice 32).

3) A quelle condition le produit de deux involutions est-il une involution ?

Exercice 34. — On considère en tant qu'anneau l'algèbre $K(f)$ engendrée par l'endomorphisme f. Caractériser les diviseurs de zéro de cet anneau. Montrer que tous les éléments qui ne sont pas diviseurs de zéro sont des automorphismes.

Exemple : f est l'endomorphisme de $\mathbf{R}^3$ de matrice

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

Déterminer les diviseurs de zéro de l'anneau $K(f)$. Préciser les rangs des applications linéaires correspondantes. (On remarquera d'abord que $f^3 = e$).

CHAPITRE VI

DÉTERMINANTS

§ 1. GROUPE SYMETRIQUE

1. Permutations.

Une permutation d'un ensemble E est une bijection de cet ensemble sur lui-même.

L'ensemble des bijections d'un ensemble E sur lui-même est un groupe pour la composition. On le désigne par $\mathcal{S}$ (E) et on l'appelle groupe symétrique de E.

> **Exercice 35.** — Les groupes symétriques de 2 ensembles de même cardinal sont isomorphes.

En particulier, on désigne par $\mathcal{S}_n$ le groupe symétrique des ensembles finis à n éléments et on pourra toujours supposer que cet ensemble est le segment $[1, n]$ de $\mathbf{N}$.

Une permutation $\sigma \in \mathcal{S}_n$ pourra être donnée sous la forme :

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} \tag{1}$$

$\mathcal{S}_n$ a $n\,!$ éléments. Il n'est pas commutatif pour $n > 2$ (*). On appelle *transposition* une permutation qui échange 2 éléments et laisse les autres invariants :

$$\tau_{ij}(i) = j \qquad \tau_{ij}(j) = i \qquad \forall k \neq i, j \qquad \tau_{ij}(k) = k$$

définit la transposition τ_{ij}.

On voit sans peine qu'une permutation peut être décomposée en un produit (non commutatif) de transpositions.

2. Signature d'une permutation.

Etant donnée la permutation (1), on peut considérer le produit :

$$\varepsilon_\sigma = \frac{\prod_{i>j} (i - j)}{\prod_{i>j} (\sigma(i) - \sigma(j))}$$

(*) On désigne aussi sous le nom de permutation l'image de 1, 2.... n par la bijection, c'est-à-dire l'ensemble ordonné $\sigma(1)\ \sigma(2) \dots \sigma(n)$. Il y a là un abus de langage qui n'est qu'un cas particulier de celui qui consiste à désigner la fonction f par sa valeur $f(x)$.

chacun des produits étant relatif à tous les couples d'indices i, j avec $i > j$. Puisque σ est une bijection, tous les facteurs du numérateur se retrouvent au signe près au dénominateur une fois et une seule. Ce produit vaut donc ± 1. On le nomme signature de la permutation. Il vaut $(-1)^{I}$, I étant le nombre de couples dont l'ordre est changé par la permutation, ou nombre d'inversions de la permutation.

La signature d'un produit de permutations est le produit des signatures. En effet, soient σ et σ' deux permutations :

$$\frac{\prod_{i>j}(i-j)}{\prod_{i>j}\big((\sigma\circ\sigma')(i)-(\sigma\circ\sigma')(j)\big)}$$

peut s'écrire :

$$\frac{\prod_{i>j}(i-j)}{\prod_{i>j}(\sigma'(i)-\sigma'(j))} \times \frac{\prod_{i>j}(\sigma'(i)-\sigma'(j))}{\prod_{i>j}\big(\sigma(\sigma'(i))-\sigma(\sigma'(j))\big)}$$

Le premier facteur est la signature $\varepsilon_{\sigma'}$; quant au deuxième, il suffit de remarquer que dans son numérateur toute différence $i-j$ ou son opposée figure une fois et une seule pour voir qu'il vaut $(-1)^{I}$, I étant le nombre de couples dont l'ordre est modifié par σ, donc qu'il est égal à ε_{σ}

Signature d'une transposition τ_{ij}. — Supposons $i < j$. Quand i est venu à la place de j, il est après $j-i$ nombres plus grands que lui, ce qui introduira $j-i$ facteurs négatifs au dénominateur de la fraction qui représente la signature. Mais, en outre, les $j-i-1$ nombres compris entre i et j se trouvent après j et introduisent autant de facteurs négatifs. La signature de τ_{ij} est donc :

$$(-1)^{j-i+j-i-1} = -1$$

La signature d'une transposition est toujours négative.

Il résulte de ces deux dernières propositions qu'une permutation est paire (c'est-à-dire de signature positive) ou impaire (c'est-à-dire de signature négative) suivant que le nombre des transpositions dont elle est le produit est pair ou impair.

Exercice 36. — Montrer que le groupe des permutations paires forme un sous-groupe de $\mathcal{S}_n$ qui est appelé groupe alterné $\mathcal{A}_n$.

Exercice 37. — Soient E et F deux ensembles, $\mathcal{S}$(E) le groupe symétrique de E et $\mathcal{F} = \mathcal{F}$(E, F) l'ensemble des applications de E dans F. A toute permutation σ de E ($\sigma \in \mathcal{S}$(E)) on fait correspondre l'application φ_σ de $\mathcal{F}$ dans $\mathcal{F}$ défini par

$$\varphi_\sigma(f) = f\circ\sigma^{-1}$$

1) Montrer que l'application φ_σ est une bijection, donc appartient au groupe symétrique $\mathcal{S}(\mathcal{F})$ de $\mathcal{F}$ et que l'application de $\mathcal{S}$(E) dans $\mathcal{S}(\mathcal{F})$ qui à σ fait correspondre φ_σ est un homomorphisme injectif, si F a au moins deux éléments. En vertu de ce résultat on désigne souvent φ_σ par la même lettre σ que la permutation de E à laquelle elle correspond et on écrit :

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

2) Si $\mathcal{G}$ est l'ensemble des applications de $\mathcal{F}$ dans un ensemble G, il

existe un homomorphisme injectif de $\mathcal{S}$(E) dans $\mathcal{S}(\mathcal{G})$. On désignera par σ l'élément de $\mathcal{S}(\mathcal{G})$ image de la permutation $\sigma \in \mathcal{S}$(E).

Exemple : I étant l'ensemble des n premiers entiers $\mathcal{S}(I) = \mathcal{S}_n$. Si E est un ensemble quelconque $\mathcal{F}$(I, E) n'est autre que E^n.

Si $x = (x_1, x_2 \dots x_n) \in E^n$ indiquer ce que vaut σx avec les conventions indiquées au 1).

Si F est un troisième ensemble et si $f \in \mathcal{F}(E^n, F)$, indiquer ce que vaut $\sigma f(x)$.

Si G est un quatrième ensemble et si $g \in \mathcal{F}(\mathcal{F}(E^n, F), G)$, indiquer ce que vaut $\sigma g(f(x))$.

§ 2. APPLICATIONS MULTILINEAIRES ET MULTILINEAIRES ALTERNEES

1. Applications multilinéaires.

Soient $E_1, E_2 \dots E_n$ des espaces vectoriels sur un corps K et F un autre espace vectoriel sur le même corps K. On dit qu'une application du produit cartésien $E_1 \times E_2 \times \dots \times E_n$ dans F :

$$\varphi : E_1 \times E_2 \times \dots \times E_n \longrightarrow F$$

est multilinéaire si elle est linéaire par rapport à chacune des composantes de l'élément du produit cartésien, les autres étant fixées, c'est-à-dire si on a :

avec $x_1 \in E_1$, $x_2 \in E_2$, ... $x_i \in E_i$, $y_i \in E_i$, ..., $x_n \in E_n$ $\lambda \in K$,

$$\varphi(x_1, x_2, \dots, x_i + y_i, \dots x_n) = \varphi(x_1, \dots, x_i, \dots x_n) + \varphi(x_1, \dots, y_i, \dots x_n),$$
$$\varphi(x_1, x_2, \dots, \lambda x_i, \dots, x_n) = \lambda\,\varphi(x_1, x_2, \dots, x_i, \dots, x_n).$$

Remarque : $E_1 \times E_2 \dots \times E_n$ peut recevoir une structure d'espace vectoriel sur K. Il importe de remarquer que φ, multilinéaire par rapport à $E_1 \dots E_n$, n'est pas linéaire par rapport au produit cartésien considéré comme espace vectoriel. En effet, ceci voudrait dire :

$$\varphi(\lambda x_1, \lambda x_2 \dots \lambda x_n) = \lambda\,\varphi(x_1, x_2 \dots x_n),$$

tandis que si φ est multilinéaire :

$$\varphi(\lambda x_1, \lambda x_2 \dots \lambda x_n) = \lambda^n \varphi(x_1 \dots x_n),$$

et de même si φ était linéaire on aurait :

$$\varphi(x_1 + y_1, x_2 + y_2 \dots x_n + y_n) = \varphi(x_1, x_2 \dots x_n) + \varphi(y_1, y_2 \dots y_n),$$

alors que φ multilinéaire donne :

$$\varphi(x_1 + y_1, x_2 + y_2 \dots x_n + y_n) = \Sigma\,\varphi(z_1, z_2 \dots z_n),$$

$z_1, z_2 \dots z_n$ pouvant représenter de toutes les façons possibles x_1 ou y_1, x_2 ou y_2 ..., ce Σ étant donc étendu à 2^n termes.

2. Applications multilinéaires alternées.

Plaçons-nous maintenant dans le cas où $E_1 = E_2 = \dots = E_n$, c'est-à-dire où l'on considère une application multilinéaire de E^n dans F, E^n étant le produit cartésien considéré comme il vient d'être dit (il n'est pas muni de la structure d'espace vectoriel produit). Les applications ainsi définies sont dites applications multilinéaires d'ordre n sur E.

Une application multilinéaire est dite alternée si elle s'annule pour tout élément dont deux composantes sont égales, ce qui entraîne qu'elle s'annule pour tout élément dont une composante est égale au produit de l'autre par un scalaire (le produit vectoriel nous fournit un exemple d'application bilinéaire alternée de $\mathbf{R}^3$ dans $\mathbf{R}^3$ puisqu'il s'annule quand les deux vecteurs sont colinéaires). En particulier, nous pourrons considérer les applications multilinéaires alternées de E^n dans K que l'on appellera *formes multilinéaires alternées.*

Une forme multilinéaire alternée change de signe quand on y échange les valeurs de deux composantes. Soit, en effet, $\varphi(x_1, ..., x_i, ..., x_j, ..., x_n)$ et $\varphi(x_1, ..., x_j, ..., x_i, ..., x_n)$, obtenue en permutant les valeurs des $i^{\text{ème}}$ et $j^{\text{ème}}$ composantes et en laissant les autres invariantes. Considérons :

$$\varphi(x_1, ..., x_i + x_j, ..., x_j + x_i, ..., x_n)$$

qui est nulle puisque c'est l'image d'un élément qui a deux composantes égales ; elle vaut :

$$\varphi(x_1,..,x_i,..,x_j,..,x_n)+\varphi(x_1,..,x_i,..,x_i,..,x_n)$$
$$+\ \varphi(x_1,..,x_j,..,x_j,..,x_n)+\varphi(x_1,..,x_j,..,x_i,..,x_n)\ ;$$

le deuxième et le troisième termes de cette somme étant nuls, le premier et le dernier sont opposés.

Plus généralement, si $\sigma \in \mathcal{S}_n$ est une permutation qui, aux indices $1, 2 ... n$ fait correspondre $\sigma(1), \sigma(2) ... \sigma(n)$, on aura :

$$\varphi(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = (-1)^k\ \varphi(x_1, x_2, ..., x_n) \qquad (1)$$

k étant le nombre de transpositions en lesquelles on peut décomposer σ, donc $(-1)^k$ représentant la signature de σ. Si on convient de noter σx l'image par σ du « n-uplet » qui constitue x, une forme multilinéaire alternée vérifiera donc :

$$\varphi(\sigma x) = \varepsilon_\sigma \varphi(x) \qquad (2)$$

Une forme qui vérifie cette égalité pour toutes les permutations σ est dite *antisymétrique.*

Exercice 38. — Montrer que réciproquement une forme antisymétrique est alternée si le corps n'est pas de caractéristique 2.

3. Formes *n*-linéaires alternées sur un espace vectoriel de dimension *n*.

Supposons que E soit de dimension n et rapporté à une base $(a_1, a_2, ..., a_n)$. Un élément de E^n est un « n-uplet » de vecteurs de E : $(x_1, x_2, ..., x_n)$ avec :

$$x_i = \sum \lambda_i^j a_j.$$

Etudions les formes multilinéaires alternées sur E^n. Soit une telle forme φ. La multilinéarité de φ entraîne que $\varphi(x_1, ..., x_n)$ devra être égale à la somme des quantités

$$\varphi(\lambda_1^{i_1} a_{i_1}, \lambda_2^{i_2} a_{i_2}, \ldots, \lambda_n^{i_n} a_{i_n}),$$

une telle quantité étant elle-même égale à :

$$\lambda_1^{i_1}\lambda_2^{i_2}, ..., \lambda_n^{i_n}\,\varphi(a_{i_1}, a_{i_2}, ..., a_{i_n}),$$

$i_1, i_2 ... i_n$ représentent chacun un des indices j, donc $\lambda_1^{i_1}$ représente un des λ_1^j, $\lambda_2^{i_2}$ un des λ_2^j, ..., etc..., et la somme envisagée s'étendant à tous les choix possibles de la suite $a_{i_1} ... a_{i_n}$ (et ayant par conséquent n^n termes). Mais dans cette somme, tous les termes qui portent sur des n-uplets à deux composantes égales sont nulles. Donc, il ne reste que les $n!$ termes où $a_{i_1} ... a_{i_n}$ sont tous différents, c'est-à-dire où ils sont l'image de $(a_1 ... a_n)$ par une permutation σ. On peut donc écrire que φ doit vérifier :

$$\varphi(x_1, ..., x_n) = \sum_{\sigma \in \mathcal{S}_n} \lambda_1^{\sigma(1)} \lambda_2^{\sigma(2)} \ldots \lambda_n^{\sigma(n)} \varphi(a_{\sigma(1)}, a_{\sigma(2)}, \ldots, a_{\sigma(n)}),$$

le signe Σ s'étendant à toutes les permutations $\sigma \in \mathcal{S}_n$. Mais, d'après l'égalité (2) précédente, ceci peut s'écrire :

$$\varphi(x_1, \ldots, x_n) = \varphi(a_1, a_2, ..., a_n) \sum_{\sigma \in \mathcal{S}_n} \varepsilon_\sigma \lambda_1^{\sigma(1)} \lambda_2^{\sigma(2)} \ldots \lambda_n^{\sigma(n)} \qquad (3)$$

Donc, au facteur $\varphi(a_1, a_2, ..., a_n)$ près, la valeur de la forme multilinéaire alternée est déterminée pour un n-uplet de vecteurs donnés.

Reste à vérifier que les formes ainsi construites répondent bien à la question. Leur multilinéarité est évidente. Quant au fait qu'elles sont alternées, il peut être vérifié de la façon suivante : supposons que les deux vecteurs x_i et $x_{i'}$ soient égaux (c'est-à-dire $\forall j \quad \lambda_i^j = \lambda_{i'}^j$) ; à tout terme de la somme précédente

$$\varepsilon_\sigma\, \lambda_1^{\sigma(1)} \ldots \lambda_i^{\sigma(i)} \ldots \lambda_{i'}^{\sigma(i')} \quad \lambda_n^{\sigma(n)}$$

correspond le terme de même valeur absolue :

$$\varepsilon_{\sigma'}\, \lambda_1^{\sigma(1)} \ldots \lambda_i^{\sigma(i')} \ldots \lambda_{i'}^{\sigma(i)} \ldots \lambda_n^{\sigma(n)}$$

Or, ce terme est celui que fournit une permutation σ' définie par:

$$\sigma' = \tau_{ii'} \circ \sigma.$$

Mais alors $\varepsilon_{\sigma'} = -\varepsilon_\sigma$ puisque $\varepsilon_{\tau_{ii'}} = -1$. Les deux termes considérés se détruisent donc dans la somme.

La condition (3) est donc nécessaire et suffisante et on peut conclure que, *sur un espace de dimension n, il existe une forme n-linéaire alternée déterminée à une constante multiplicative près.*

Exercice 39. — Montrer que les formes p-linéaires alternées sur un espace de dimension n forment un espace vectoriel dont on cherchera la dimension.

§ 3. DETERMINANT

1. Définition.

Le déterminant, par rapport à une base donnée, ordonnée, d'un espace à n dimensions, est celle des formes n-linéaires alternées qui prend la valeur 1 pour les n vecteurs de la base.

Autrement dit, on considère celle des formes pour laquelle $\varphi(a_1,a_2,...,a_n)=1$ et qui, par suite, pour les vecteurs $(x_1, x_2, ..., x_n)$ avec $x_i = \sum_j \lambda_i^j a_j$, prend la valeur :

$$\det\,[x_1, x_2, ..., x_n] = \sum_{\sigma \in \mathcal{S}_n} \varepsilon_\sigma \lambda_1^{\sigma(1)} \lambda_2^{\sigma(2)} \ldots \lambda_n^{\sigma(n)}.$$

Nous commettrons l'abus de langage traditionnel, mais peut-être pédagogiquement dangereux, qui consiste à appeler ce nombre déterminant des n vecteurs, et nous le noterons :

$$\det\,[x_1, ..., x_n].$$

Si les n vecteurs ont leurs coordonnées suivant la base écrites dans un tableau carré, λ_i^j $j^{\text{ème}}$ coordonnée du $i^{\text{ème}}$ vecteur étant dans la $j^{\text{ème}}$ ligne et la $i^{\text{ème}}$ colonne, on voit que le déterminant est la somme des produits obtenus en prenant un élément et un seul dans chaque colonne et dans chaque ligne, ce produit étant affecté du signe égal à la signature de la permutation du rang des lignes, les éléments étant rangés dans l'ordre naturel des colonnes. Mais, le produit des scalaires λ étant commutatif, on peut les ranger dans l'ordre naturel des lignes ; les indices inférieurs représenteront alors $\sigma^{-1}(1)$, $\sigma^{-1}(2) \dots \sigma^{-1}(n)$. Mais σ^{-1} a même signature que σ (puisque leur produit, l'identité, a 1 pour signature). Donc un terme du déterminant peut aussi bien être considéré comme ayant pour signe la signature de la permutation des indices des colonnes, les lignes étant prises dans l'ordre naturel. Il résulte de cette symétrie sur le rôle des lignes et des colonnes qu'un déterminant peut aussi bien être considéré comme une forme n-linéaire alternée sur les vecteurs lignes (c'est-à-dire vecteurs dont les éléments d'une ligne représentent les coordonnées par rapport à la base).

2. Développement du déterminant de n vecteurs suivant les éléments d'une colonne ou d'une ligne.

Dans l'expression du déterminant, nous pouvons grouper tous les termes qui contiennent chaque élément d'une ligne ou chaque élément d'une colonne. Le déterminant pourra alors s'exprimer sous la forme :

$$\lambda_i^1 A_i^1 + \lambda_i^2 A_i^2 + \dots + \lambda_i^j A_i^j + \dots + \lambda_i^n A_i^n$$

si on développe par rapport aux éléments de la $i^{\text{ème}}$ colonne.

Cherchons à déterminer le coefficient A_i^j de λ_i^j. Pour cela, commençons par déterminer A_1^1 (coefficient de λ_1^1 dans le développement par rapport aux éléments de la première ligne ou de la première colonne). Si nous groupons tous les termes où figure λ_1^1, nous voyons que ce nombre est facteur d'une somme qui n'est autre que le déterminant des $(n-1)$ vecteurs dont les coordonnées forment le tableau obtenu en barrant dans le tableau précédent la première ligne et la première colonne. [Il est clair que l'on trouve les mêmes produits partiels avec le même signe, car les permutations $\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$ et $\begin{pmatrix} 2 & \dots & n \\ \sigma(2) & \dots & \sigma(n) \end{pmatrix}$ ont même signature].

Pour trouver maintenant A_i^j, amenons la $i^{\text{ème}}$ colonne à la première place sans changer l'ordre des autres, ce qui signifie $i-1$ échanges de vecteurs, donc change $i-1$ fois le signe du déterminant puisque c'est la valeur d'une forme alternée. Amenons ensuite, de la même façon, la $j^{\text{ème}}$ ligne à la première place. Le déterminant étant alterné par rapport aux vecteurs lignes, nous en changerons le signe $j-1$ fois. Le déterminant finalement obtenu est égal à l'ancien multiplié par $(-1)^{i-1+j-1}$ soit $(-1)^{i+j}$. Le coefficient de λ_i^j y est multiplié par ce même nombre. C'est donc que *le coefficient de λ_i^j dans le développement du déterminant initial est égal au produit par $(-1)^{i+j}$ du déterminant relatif au tableau obtenu en barrant la $i^{\text{ème}}$ colonne et la $j^{\text{ème}}$ ligne.*

3. Déterminant d'une matrice ou d'un endomorphisme.

Se donner n vecteurs d'un espace de dimension n, c'est se donner l'endomorphisme où ces n vecteurs sont les images des vecteurs de la base ou encore c'est se donner la matrice de cet endomorphisme. Nous pourrons donc parler ou du déterminant d'une matrice, ou du déterminant d'un endomorphisme, car nous allons montrer que le déterminant d'un endomorphisme ne dépend pas de la base choisie.

Soient d'abord deux endomorphismes f et g d'un même espace vectoriel de matrices A et B par rapport à une base $(a_1, a_2, \dots, a_n)$. Soient

$x_1, x_2, \dots, x_n$ les images par f des éléments de la base et soient $g(x_1)$, $g(x_2) \dots, g(x_n)$ les images par g de ces éléments. L'application

$$(x_1, x_2, \dots, x_n) \longrightarrow \text{dét}\,[g(x_1), g(x_2), \dots, g(x_n)]$$

est une forme multilinéaire alternée en vertu :

1° de la multilinéarité du déterminant et de la linéarité de g ;

2° du fait que $x_i = x_j \implies g(x_i) = g(x_j) \implies \text{dét}\,[g(x_1) \dots g(x_n)] = 0$.

Elle est donc proportionnelle à toute autre forme n-linéaire alternée, par exemple au déterminant de $x_1\, x_2 \dots x_n$:

$$\det\,[g(x_1) \dots g(x_n)] = k \det\,[x_1 \dots x_n] \qquad k \in \mathrm{K}.$$

Pour déterminer la valeur de k, faisons $x_1 = a_1, \dots, x_n = a_n$, on trouve :

$$\det\,[g(a_1) \dots g(a_n)] = k,$$

d'où :

$$\det\,[g(x_1) \dots g(x_n)] = \det\,[g(a_1) \dots g(a_n)] \times \det\,[x_1 \dots x_n]$$

ou :

$$\det\,[g\circ f(a_1), \dots, g\circ f(a_n)] = \det\,[g(a_1), \dots, g(a_n)] \times \det\,[f(a_1), \dots, f(a_n)],$$

soit, en introduisant les matrices :

$$\det\,[\mathrm{BA}] = \det \mathrm{B} \times \det \mathrm{A}.$$

Le déterminant du produit des matrices est égal au produit des déterminants.

Cette propriété entraîne les conséquences suivantes :

— 1) Soit un automorphisme f, de matrice A, $\overset{-1}{f}$ son inverse de matrice A^{-1}. Ce qui précède entraîne :

$$\det \mathrm{A} \times \det \mathrm{A}^{-1} = 1.$$

Donc, le déterminant d'un automorphisme est différent de zéro ou encore le déterminant d'une matrice inversible est différent de zéro.

— 2) Soit un endomorphisme de matrice A par rapport à une base (a) et de matrice A′ par rapport à une base (a'). On sait (cf. III, 4, 3) qu'on a :

$$\mathrm{A}' = \mathrm{P}^{-1}\mathrm{AP},$$

P étant la matrice de passage d'une base dans l'autre. Les déterminants de P et P^{-1} étant inverses et la multiplication dans K étant commutative, ceci entraîne :

$$\det \mathrm{A}' = \det \mathrm{A}.$$

Le déterminant d'un endomorphisme d'un espace vectoriel est bien indépendant de la base à laquelle est rapporté cet espace.

Revenons à la propriété précédente et étudions-en la réciproque ou, ce qui revient au même, étudions le déterminant d'un endomorphisme qui ne soit pas un automorphisme, c'est-à-dire d'un endomorphisme f tel que $x_1 = f(a_1) \dots x_n = f(a_n)$ soient linéairement dépendants. Le déterminant de n vecteurs linéairement dépendants est nul. En effet, on peut alors mettre un des x_i, x_1 par exemple, sous la forme :

$$x_1 = \sum_{i=2}^{n} \mu_i x_i.$$

Mais alors le déterminant sera égal à une combinaison linéaire de $n-1$ déterminants de n-uplets comportant deux vecteurs égaux qui sont donc tous nuls.

On peut donc conclure :

$$f \text{ automorphisme} \implies \det f \neq 0$$
$$f \text{ non injectif} \implies \det f = 0$$

ou ce qui revient au même :

$$f \text{ automorphisme} \iff \det f \neq 0.$$

La condition nécessaire et suffisante pour qu'un endomorphisme soit un automorphisme est que son déterminant soit différent de zéro.

§ 4. APPLICATION DES DETERMINANTS A LA RECHERCHE D'UNE MATRICE DE FORME REMARQUABLE SEMBLABLE A UNE MATRICE DONNEE

Nous pouvons reprendre maintenant le problème du précédent chapitre avec l'aide des déterminants.

Cherchons *a priori* vecteurs propres et valeurs propres associés. Un vecteur propre de l'endomorphisme f est un vecteur $\xi \neq 0$ tel que $f(\xi) = \lambda\xi$. L'endomorphisme $f - \lambda e$ (où e désigne l'automorphisme identique) est tel qu'il existe un vecteur non nul ξ dont il donne une image nulle. Pour qu'il en soit ainsi, il est nécessaire et suffisant que cet endomorphisme $f - \lambda e$ ne soit pas un automorphisme, donc que son déterminant soit nul. Ce déterminant est celui de la matrice obtenue en retranchant λ à tous les termes de la diagonale principale. Il se présente comme un polynôme P de degré n par rapport à λ.

L'équation $P(\lambda) = 0$ *est appelée équation caractéristique de l'endomorphisme f.*

K étant algébriquement clos, ce polynôme peut s'écrire :

$$P(\lambda) = \Pi(\mu_j - \lambda)^{k_j} \qquad \text{avec} \quad \Sigma k_j = n.$$

Les μ_j sont les valeurs propres de l'endomorphisme f.

Nous allons montrer maintenant que ce polynôme caractéristique appartient à l'idéal annulateur de l'endomorphisme f, c'est-à-dire qu'il est divisible par ϖ, et qu'en outre il ne possède aucune racine μ_j qui ne soit une racine λ_i de ϖ, seul l'ordre de multiplicité des racines λ_i dans ϖ et P pouvant être différent.

Le polynôme P étant indépendant de la base choisie pour E, prenons celle qui a donné à la matrice de f la forme triangulaire trouvée en V, 2, 3 (fig. 7) et cherchons son polynôme caractéristique. Le déterminant d'une matrice triangulaire étant évidemment égal au produit des éléments de la diagonale principale, le polynôme caractéristique est :

$$P = \Pi(\lambda_i - \lambda)^{\beta_i}$$

où β_i est l'ordre de la matrice M_i, c'est-à-dire la dimension de l'espace vectoriel annulé par $(\lambda_i e - f)^{\alpha_i}$; et on voit qu'il est divisible par :

$$\varpi = \Pi(\lambda_i - \lambda)^{\alpha_i}$$

puisque : $\quad \forall i \quad \alpha_i \leqslant \beta_i.$

En d'autres termes, f qui annule ϖ annule aussi P. *f vérifie l'équation* $P(f) = 0$, ou, si l'on préfère, toute matrice M qui représente f vérifie $P(M) = 0$.

Signalons *le cas particulier très important où toutes les racines de* P *sont simples ;* c'est un cas particulier du cas déjà particulier envisagé en (V, 2, 3) où toutes les racines de ϖ étaient simples. Cette fois-ci :

$$\forall i \quad \alpha_i = \beta_i = 1.$$

La matrice de l'endomorphisme est diagonalisable, mais, en outre, tous ses coefficients $a_i^i = \lambda_i$ sont deux à deux distincts. Dans ce cas, la seule connaissance des racines du polynôme caractéristique fournit immédiatement la matrice diagonale semblable à la matrice donnée. Quant au vecteur propre correspondant à la valeur propre λ_i, ce sera la solution, unique à une homothétie près, de l'équation linéaire :

$$f(e_i) = \lambda_i e_i.$$

Remarque : Dans ce dernier cas, il est aisé de montrer directement qu'il existe une base de vecteurs propres. Il existe en effet par hypothèse n valeurs λ_i deux à deux distinctes et pour chacune d'elles un vecteur $e_i \neq 0$ tel que $f(e_i) = \lambda_i e_i$. Montrons par récurrence qu'ils forment une partie libre. Supposons que ce soit vrai pour $e_1, \ldots, e_{i-1}$, mais que l'on ait $e_i = \sum_{k=1}^{i-1} \alpha^k e_k$ avec des coefficients α^k non tous nuls (puisque $e_i \neq 0$). On a :

$$f(e_i) = \lambda_i e_i = \lambda_i \sum_{k=1}^{i-1} \alpha^k e_k,$$

Mais, d'autre part :

$$f\left(\sum_{k=1}^{i-1} \alpha^k e_k\right) = \sum_{k=1}^{i-1} \lambda_k \alpha^k e_k,$$

d'où, en rapprochant les deux valeurs de $f(e_i)$:

$$\sum_{k=1}^{i-1} (\lambda_i - \lambda_k)\alpha^k e_k = 0$$

ce qui, puisque $\{ e_1, \ldots, e_{i-1} \}$ est une partie libre et qu'au moins un α^k est différent de zéro, exigerait $\lambda_i = \lambda_k$, or cela est contraire à l'hypothèse.

Exercice 40. — f étant un endomorphisme et λ une de ses valeurs propres on appelle sous-espace propre correspondant à λ l'ensemble des vecteurs x tels que :

$$f(x) = \lambda x$$

1° Montrer que la somme des sous-espaces propres correspondant à des valeurs propres différentes est une somme directe.

2° En déduire que la condition nécessaire et suffisante pour que l'une des matrices de f soit diagonale est que E soit la somme directe des sous-espaces propres de f ; ou encore, que la dimension de chaque sous-espace propre soit égale à l'ordre de multiplicité, comme racine du polynôme caractéristique, de la valeur propre associée.

Une application linéaire qui vérifie cette propriété est dite « à spectre simple ».

3° Montrer que toute application linéaire à spectre simple et à valeurs propres différentes de zéro est inversible. Donner une condition nécessaire et suffisante portant sur les valeurs propres de l'application pour que celle-ci soit identique à son inverse (rapprocher du résultat de l'exercice 33).

4° *a)* f et φ étant deux endomorphismes d'un espace vectoriel de dimension finie, montrer que si f est à spectre simple, la condition nécessaire et suffisante pour que f et φ commutent est que les sous-espaces propres de f soient invariants par φ.

Quelle forme particulière cette condition prend-elle quand f a toutes ses valeurs propres distinctes ?

b) Montrer que ce résultat permet de retrouver celui de l'exercice 24.

c) Montrer comment ce résultat permet aussi de préciser celui de l'exercice 33 relatif au produit de deux involutions.

d) De façon plus générale, à quelle condition doit satisfaire un endomorphisme pour commuter avec une involution ?

Appliquer cette condition à chacun des types d'involution de $\mathbf{R}^3$.

Exercice 41. — Soit E un espace vectoriel de dimension n sur un corps K, $e_1, e_2, \ldots, e_n$ une base de E, f une application linéaire de E dans lui-même. Soit x_1 un vecteur de E ; on pose

$$x_2 = f(x_1) \ldots.. x_k = f(x_{k-1}) = f^{k-1}(x_1) \ldots \text{ etc...}$$

On considère le sous-espace vectoriel F engendré par les vecteurs x_k.

1° Montrer que si p est le plus grand entier tel que $x_1, x_2, \ldots, x_p$ soient linéairement indépendants, ces p vecteurs forment une base de F. Si F coïncide avec E tout entier on dira, dans la suite, que x_1 engendre E. Quelle est alors la valeur de p ?

2° On suppose que la matrice A de f par rapport à la base $e_1, e_2, \ldots, e_n$ est diagonale. Montrer que pour qu'il existe un vecteur x_1 engendrant E, il faut et il suffit que les valeurs propres de f soient deux à deux distinctes. (On posera $x_1 = \Sigma\, \xi^i e_i$ et l'on cherchera la condition pour que les vecteurs $x_1 \ldots x_n$ soient linéairement indépendants).

3° On suppose que le vecteur x_1 engendre E. Quelle est la forme de la matrice A' de f par rapport à $x_1, \ldots, x_n$? On désignera par a_i l'élément situé dans la $i^{\text{ème}}$ ligne et la dernière colonne de cette matrice. Montrer que le polynôme caractéristique de A' est

$$P(\lambda) = (-1)^n[\lambda^n - a_n\lambda^{n-1} - a_{n-1}\lambda^{n-2} - \ldots \ldots - a_1]$$

Vérifier directement que, conformément à la théorie générale, l'application linéaire $P(f)$ est l'application nulle.

CHAPITRE VII

STRUCTURE AFFINE

§ 1. DEFINITION

1. Groupe simplement transitif.

Soit un ensemble E, et soit G un groupe de permutations de E (ou bijections de E sur E), c'est-à-dire un sous-groupe du groupe symétrique de E. On dit de ce groupe G qu'il opère sur E. (Cf. *A.P.M. I*, II, 5). Nous rappelons les résultats suivants : un couple (x, y) d'éléments de E étant donné, on peut se demander s'il existe $u \in G$ tel que $y = u(x)$. L'existence d'un tel u définit sur E une relation d'équivalence dont les classes sont nommées « classes d'intransitivité » (les éléments d'une même classe sont ceux qui sont susceptibles d'être envoyés les uns sur les autres par les bijections du groupe G). Le groupe G est dit transitif sur l'ensemble E s'il existe une seule classe d'équivalence dans E. Il est dit *simplement transitif*, si, pour tout couple (x, y), il existe une seule bijection de G qui envoie x sur y. On peut la noter u_{xy}, et on écrit :

$$y = u_{xy}(x).$$

L'élément inverse dans G, c'est-à-dire la bijection réciproque, envoie y sur x. Si l'opération de G est notée additivement, on a donc :

$$u_{yx} = -u_{xy}.$$

La bijection composée de u_{xy} et u_{yz} envoie x sur z, c'est donc l'unique bijection u_{xz}, et on a (toujours en notation additive) :

$$u_{xz} = u_{xy} + u_{yz}.$$

2. Définition de la structure affine.

Ceci posé, soit un ensemble E, dont les éléments seront appelés points, et un espace vectoriel T sur un corps commutatif K.

On dira que E *est un espace affine* ou *a une structure affine par rapport à* T, *si le groupe additif de* T *opère sur* E *de façon simplement transitive.*

Il existe des espaces affines. En effet :

Un espace vectoriel est un espace affine par rapport à lui-même. Soit T un espace vectoriel sur K. Soit $x \in$ T, et $t \in$ T fixe. L'application, que nous nommerons translation,

$$(t) : \begin{array}{ccc} T & \longrightarrow & T \\ x & & x + t = y = (t)(x) \end{array}$$

est une bijection. L'ensemble (T) de ces bijections forme, quand t décrit T, un groupe isomorphe au groupe additif de T. D'autre part, si l'on se donne x et y, il existe un t et un seul tel que $x + t = y$. Donc, le groupe (T) opère sur T de façon simplement transitive.

En identifiant t et (t), on dira que le groupe additif de T opère sur lui-même de façon simplement transitive. T reçoit alors une structure d'espace affine par rapport à lui-même, un même élément de T pouvant être considéré comme point de l'espace affine ou comme vecteur de l'espace vectoriel. Les deux relations :

$$y = x + t \qquad\qquad y = t(x)$$

ont le même sens et peuvent donc être utilisées l'une pour l'autre. Cependant, originellement, la première formulation concerne trois vecteurs x, y, t de T, la deuxième deux points x et y de l'espace affine et un vecteur t opérant sur cet espace.

Inversement, si E est un espace affine par rapport à un espace vectoriel T, on peut doter E d'une structure d'espace vectoriel isomorphe à T. Soit en effet $a \in$ E fixe, par définition de la structure affine :

$$\forall x \in E \qquad \exists\, !\, t \in T \qquad x = t(a) \qquad (*)$$

Réciproquement, puisque T opère sur E :

$$\forall t \in T \qquad \exists\, !\, x \in E \qquad x = t(a).$$

(Quand on voudra rappeler que t est le vecteur de T tel que $x = t(a)$, on l'écrira t_{ax}).

Il y a donc bijection entre E et T : un point x de E, correspondant au vecteur t_{ax} de T. En posant que cette bijection est un isomorphisme, on dote E d'une structure d'espace vectoriel. *L'espace vectoriel ainsi défini est appelé espace vectoriel d'origine* a *et est noté* T_a. *On peut associer un tel espace vectoriel à tout point* a *de* E.

3. Structure affine et structure vectorielle.

Dans ce qui précède, structure affine et structure vectorielle apparaissent comme intimement liées, mais elles doivent cependant être soigneusement distinguées. On remarquera :

1° Que si E est un espace affine par rapport à T, les structures d'espace vectoriel isomorphes à T qu'on en a déduit sont deux à deux distinctes et qu'il n'y a aucune raison de faire jouer à l'une d'elles un rôle privilégié : doter un ensemble d'une structure affine, c'est, en effet, considérer, au contraire, que tous ses éléments sont appelés à jouer le même rôle dans la question étudiée.

2° Que les translations non nulles d'un espace vectoriel n'en sont pas des automorphismes (le translaté d'une somme n'est pas la somme des translatés, le zéro n'est pas invariant...). Au contraire, dans un espace affine, la translation t_{ab} réalise un isomorphisme de T_a sur T_b : c'est un automorphisme de la structure affine, c'est une bijection qui est affine (voir plus loin la définition de ce mot), ainsi que sa réciproque.

3° Le corps K opère sur l'espace vectoriel T, et T opère sur l'espace affine E, mais K n'opère pas sur E. La donnée d'une structure affine sur E, par rapport à T, ne fournit aucun moyen canonique de faire correspondre un élément de E à un couple (λ, m) de $K \times E$.

(*) $\exists$! se lit : « il existe un et un seul ».

La notion d'espace vectoriel précède logiquement la notion d'espace affine, si, comme nous l'avons fait, on fait intervenir la première notion dans la définition de la seconde. Mais il est possible de procéder en sens inverse : c'est ce que fait l'exposé traditionnel de la géométrie élémentaire qui introduit d'abord un ensemble de points (le plan, ou l'espace), puis un groupe particulier de bijections (les translations au sens élémentaire du mot), qui est simplement transitif et dont on montre ensuite que chaque translation est caractérisée par un « vecteur », la composée de deux translations étant caractérisée par la « somme géométrique » des deux vecteurs correspondants.

§ 2. VARIETES LINEAIRES AFFINES

1. Définition.

Soit a un point de E, T_a l'espace vectoriel d'origine a, H un sous-espace vectoriel de T et $H_a \approx H$ le sous-espace correspondant de T_a, que nous considérons aussi comme un ensemble de points de E. Soit un vecteur fixe $t \in T$ et soit m un point qui décrit H_a. On appelle *variété affine* parallèle à H l'ensemble des points $t(m)$. Si $m = u(a)$, on peut écrire :

$$t(m) = (t + u)(a)$$

u est un vecteur de H et décrit H considéré comme espace vectoriel quand m décrit H_a considéré comme espace affine.

L'ensemble $\{ t + u \; ; \; u \in H \}$ est une classe d'équivalence de T mod H.

Par tout point de l'espace affine, il passe une variété affine parallèle à H et une seule. Ce postulat d'Euclide généralisé résulte du fait que la relation d'équivalence $x R y \iff x - y \in H$ détermine une partition de T_a, donc de E.

Il faut observer que cette partition et la famille des variétés affines parallèles à H sont indépendantes de a. En effet, si on avait procédé de même à partir d'un point b avec le même vecteur $t \in T$, on aurait cherché l'ensemble des points :

$$\{ t(m) \; ; m \in H_b \} = \{ (t + u)(b) \; ; u \in H \} .$$

Mais $(t + u)(b) = (t + u + t_{ab})(a)$. Or, $(t + u + t_{ab})(a)$ est la variété affine parallèle à H définie par le vecteur $t + t_{ab}$.

2. Barycentre.

Nous nous proposons maintenant d'exprimer que des points de l'espace affine appartiennent à la même variété.

Soient des points x_0, x_i d'une variété affine V, x_0 étant un des points et $\{ x_i \}$ désignant un ensemble de points où i décrit un ensemble I fini ou infini d'indices (si I est l'ensemble 1, ..., n, nous considéronc donc $n + 1$ points $x_0, x_1, ..., x_n$, parmi lesquels nous particularisons le point x_0).

Les éléments de V sont déduits de a par $t + u$, u décrivant H et t étant fixé ; pour t nous pouvons prendre t_{ax_0}. Et pour les points x_i, on a :

$$t_{ax_i} = t_{ax_0} + t_{x_0 x_i},$$

les vecteurs $t_{x_0 x_i}$ pour tout $i \in I$ étant des vecteurs de H.

Toute combinaison linéaire finie

$$\sum_{i \in I} \lambda_i t_{x_0 x_i} \qquad (\lambda_i \in K, \lambda_i = 0 \text{ sauf pour un nombre fini d'indices})$$

appartient aussi à H et le vecteur

$$t_{ax} = t_{ax_0} + \sum_{i \in I} \lambda_i \, t_{x_0x_i}$$

définit un point x de V. On a :

$$t_{ax} = t_{ax_0} + \sum_{i \in I} \lambda_i \, (t_{ax_i} - t_{ax_0})$$

$$= t_{ax_0}(1 - \sum_{i \in I} \lambda_i) + \sum_{i \in I} \lambda_i \, t_{ax_i}$$

que l'on peut écrire si on désigne par $\widehat{I}$ la réunion de I et de { 0 } :

$$t_{ax} = \sum_{i \in \widehat{I}} \alpha_i \, t_{ax_i},$$

où les α_i sont des scalaires nuls sauf un nombre fini d'entre eux qui vérifient :

$$\sum_{i \in \widehat{I}} \alpha_i = 1.$$

Si les λ_i non nuls sont au nombre de n, les α_i sont au nombre de $n + 1$, en général, (de n, si $\sum_{i \in I} \lambda_i = 1$).

Si on avait pris b comme origine, on serait arrivé à une relation identique ; en effet, la précédente peut s'écrire :

$$t_{ab} + t_{bx} = \sum_{i \in \widehat{I}} (t_{ab} + t_{bx_i}) \, \alpha_i$$

$$= t_{ab} \sum_{i \in \widehat{I}} \alpha_i + \sum_{i \in \widehat{I}} \alpha_i \, t_{bx_i}$$

$$t_{bx} = \sum_{i \in \widehat{I}} \alpha_i \, t_{bx_i} \qquad (1).$$

Il en résulte qu'on peut donner un sens à l'égalité :

$$x = \sum_{i \in \widehat{I}} \alpha_i x_i \qquad (2)$$

sous réserve que $\Sigma\alpha_i = 1$, ce sens étant l'égalité entre vecteurs de H écrite en (1), égalité vraie quel que soit b.

Cette égalité (2) entre points de V définit un point x à partir d'un nombre fini de points de V. On l'appelle le *barycentre* des points x_i affectés de coefficients $\alpha_i (\Sigma\alpha_i = 1)$.

3. Indépendance affine. Dimension d'une variété.

Si les vecteurs $t_{x_0x_i}$ constituent un système de générateurs de H, tout vecteur de H peut être mis sous la forme $\Sigma\lambda_i t_{x\,x_i}$ et tout point x de V peut être obtenu sous la forme $\Sigma\alpha_i x_i$ (avec $\Sigma\alpha_i = 1$), c'est-à-dire comme barycentre d'un nombre fini de points x_i affectés de coefficients convenables.

Si, plus particulièrement, les vecteurs $t_{x_0x_i}$ constituent une base de H, tout x de V pourra être mis sous forme de barycentre des points x_i d'une façon et d'une seule, c'est-à-dire avec des coefficients déterminés.

Si la dimension de H est un nombre fini p, la *dimension de la variété* V est par définition égale à la dimension de H. Tout point de V sera alors obtenu d'une façon et d'une seule comme barycentre de $p + 1$ points de

V correctement choisis. Pour que ces $p + 1$ points $x_0, x_1, \dots x_p$ le soient, il est nécessaire et suffisant que les p vecteurs correspondants $\{ t_{x_0x_i} \}$ soient linéairement indépendants, c'est-à-dire que :

$$\sum_{i=1}^{p} \lambda_i \, t_{x_0x_i} = 0 \implies \forall i \quad \lambda_i = 0$$

ou

$$\sum_{i=1}^{p} \lambda_i \, (t_{ax_i} - t_{ax_0}) = 0 \implies \forall i \quad \lambda_i = 0$$

ou

$$-\sum_{i=1}^{p} \lambda_i \, t_{ax_0} + \sum_{i=1}^{p} \lambda_i \, t_{ax_i} = 0 \implies \forall i \quad \lambda_i = 0$$

ce qui peut s'écrire :

$$\sum_{i=0}^{p} \mu_i \, t_{ax_i} = 0 \implies \forall i \quad \mu_i = 0$$

les μ_i étant $p + 1$ scalaires tels que :

$$\Sigma\mu_i = 0.$$

Réciproquement, si

$$\sum_{i=0}^{p} \mu_i \, t_{ax_i} = 0 \quad \text{avec} \quad \Sigma\mu_i = 0$$

entraîne que $\forall i \quad \mu_i = 0$, les vecteurs $t_{x_0x_i}$ sont indépendants. La condition nécessaire et suffisante trouvée :

$$\left.\begin{aligned} \sum_{i=0}^{p} \mu_i \, t_{ax_i} &= 0 \\ \sum_{i=0}^{p} \mu_i &= 0 \end{aligned}\right\} \implies \forall i \quad \mu_i = 0$$

reste la même si on remplace a par un autre point b, car

$$\sum_{i=0}^{p} \mu_i t_{ax_i} = t_{ab} \sum_{i=0}^{p} \mu_i + \sum_{i=0}^{p} \mu_i \, t_{bx_i} = \sum_{i=0}^{p} \mu_i \, t_{bx_i}$$

ce qui permet d'écrire, sous réserve que $\Sigma\mu_i = 0$, la condition $\Sigma\mu_i t_{ax_i} = 0$, vraie si on remplace a par n'importe quel point de E sous la forme :

$$\sum_{i=0}^{p} \mu_i x_i = 0.$$

Quand les p vecteurs $t_{x_0x_i}$ sont linéairement indépendants, on dit que les $p + 1$ points x_0, x_i sont *affinement indépendants* et, avec la notation que l'on vient d'introduire, *la condition nécessaire et suffisante pour qu'ils le soient* s'écrit :

$$\left.\begin{aligned} \sum_{i=0}^{p} \mu_i x_i &= 0 \\ \sum_{i=0}^{p} \mu_i &= 0 \end{aligned}\right\} \implies \forall i \quad \mu_i = 0$$

Revenons maintenant au fait que nous avons trouvé, en passant, que $\sum_{i=0}^{n} \mu_i t_{ax_i}$ avec $\Sigma\mu_i = 0$ était un vecteur indépendant de a et ceci quel que soit n. Si les x_i appartiennent à une même variété affine V, nous pouvons placer a en un point de V. Cette somme devient la somme de vecteurs de H (sous-espace vectoriel auquel V est parallèle). Elle est donc un vecteur de H. $\{x_i\}$ étant un ensemble fini de points, nous avons donc donné un sens aux expressions :

$\Sigma\alpha_i x_i$ avec $\Sigma\alpha_i = 1$ qui représente un point de la variété à laquelle appartiennent les x_i,

et $\Sigma\mu_i x_i$ avec $\Sigma\mu_i = 0$ qui représente un vecteur du sous-espace parallèle à cette variété.

Prenons en particulier deux points x_1 et x_2 affectés des coefficients 1 et -1 ; $x_2 - x_1$ représente le vecteur $t_{ax_2} - t_{ax_1}$ qui est égal à $t_{x_1x_2}$:

$$x_2 - x_1 = t_{x_1x_2}.$$

Soit alors un point quelconque c ; la notation :

$$c + \sum_{i=0}^{n} \mu_i x_i \qquad \text{avec} \quad \Sigma\mu_i = 0$$

a un sens et représente le barycentre des points $c, x_0, x_1, \ldots, x_n$ affectés des coefficients $1, \mu_0, \mu_1, \ldots, \mu_n$,

$$c + \sum_{i=0}^{n} \mu_i x_i = d \tag{1}$$

Revenons à l'égalité vectorielle que traduit cette égalité. C'est :

$$t_{ac} + \sum_{i=0}^{n} \mu_i t_{ax_i} = t_{ad}$$

qui peut s'écrire :

$$\sum_{i=0}^{n} \mu_i t_{ax_i} = t_{ad} - t_{ac} = t_{dc} = d - c.$$

L'égalité (1) devient :

$$c + (d - c) = d.$$

La somme d'un point de l'espace affine et d'un vecteur de l'espace vectoriel associé est une notation qui prend désormais un sens et représente un autre point de l'espace affine.

Exercice 42. — Montrer que l'intersection d'une famille de variétés linéaires affines est vide ou est une variété linéaire affine. En déduire que si A est une partie de E, il existe une plus petite variété affine contenant A et qu'elle est l'ensemble des barycentres des points de A.

4. Variétés parallèles au sens large.

Il s'agit de généraliser la notion de droites parallèles à un plan.

Soient H et K deux sous-espaces vectoriels de T tels que $K \subset H \subset T$. Toutes les variétés parallèles à K sont dites parallèles à H.

5. Variétés linéaires affines supplémentaires.

Soient deux sous-espaces vectoriels de T tels que :

$$H \oplus K = T.$$

On peut montrer qu'une variété parallèle à H et une variété parallèle à K ont toujours un point commun et un seul. Si on prend une origine a et $T_a \approx T$ d'origine a, une variété parallèle à H est l'ensemble des points déduits de a par l'ensemble des translations :

$$H + t \qquad t \text{ fixe appartenant à } T_a$$

et une variété parallèle à K est l'ensemble des points déduits de a par l'ensemble des translations :

$$K + s \qquad s \text{ fixe appartenant à } T_a.$$

Les vecteurs t et s s'écrivent d'une seule façon :

$$t = t_H + t_K \qquad t_H \text{ et } s_H \in H$$
$$s = s_H + s_K \qquad t_K \text{ et } s_K \in K$$

et on peut dire :

$$t + H = t_K + H$$
$$s + K = s_H + K.$$

Considérons le point i déduit de a par la translation $t_H + s_K$; ce point appartient aux deux variétés. Il est unique ; en effet, un autre point commun j devrait être déduit de i par une translation t_{ij} appartenant à H et à K, qui ne peut donc être que nulle.

Réciproquement, si deux familles de variétés affines parallèles respectivement à deux sous-espaces vectoriels H et K sont telles que deux variétés quelconques appartenant respectivement aux deux familles aient toujours un point commun et un seul, les deux sous-espaces H et K sont supplémentaires. Il est d'abord évident que $H \cap K = \{0\}$, car si H et K avaient un élément commun $t \neq 0$, les variétés passant par a auraient en commun le point $a + t$. Pour prouver que H et K sont supplémentaires, il reste donc à prouver $H \oplus K = T$. Soit L un supplémentaire de $H \oplus K$ par rapport à T :

$$T = H \oplus K \oplus L.$$

Supposons que L contienne un vecteur u_L différent de zéro et soit α l'image de l'origine a par la translation u_L et soit la variété parallèle à H passant par α ; considérons son intersection avec K_a qui, en vertu de l'hypothèse, est un point b ; le vecteur t_{ab} appartient à K. Ceci signifie donc qu'il existe un vecteur $u_H \in H$ tel que :

$$u_L + u_H = t_{ab} \in K$$
$$u_L + u_H - t_{ab} = 0$$

zéro est décomposé en trois vecteurs appartenant aux trois sous-espaces supplémentaires. De l'unicité de la décomposition, on déduit $u_L = 0$, ce qui est contraire à l'hypothèse.

§ 3. APPLICATIONS AFFINES

1. Définition.

Considérons un espace affine E par rapport à un espace vectoriel S sur un corps K et un espace affine F par rapport à un espace vectoriel T sur le même corps K, et soit u une application de E dans F. On dira que c'est une application affine si elle respecte la structure d'espace affine. Cette définition peut être précisée de deux manières dont on établira l'équivalence.

1^re^ définition : Une application affine est une application dans laquelle l'image du barycentre d'un ensemble de points est le barycentre des images, affectées des mêmes coefficients :

$$\forall \{x_i\} \quad \forall \{\alpha_i\} \quad \text{vérifiant } \Sigma\alpha_i = 1$$
$$u(\Sigma\alpha_i x_i) = \Sigma\alpha_i u(x_i).$$

2^e^ définition : Soit a quelconque, mais fixe, et son image $u(a)$; un x quelconque déterminé par le vecteur $x - a \in \mathrm{S}$ et $u(x)$ déterminé par $u(x) - u(a) \in \mathrm{T}$. L'application u est dite affine si l'application :

$$x - a \longrightarrow u(x) - u(a)$$

est une application linéaire de S dans T :

$$u(x) = u(a) + v(t) \qquad t \in \mathrm{S} \qquad v \in \mathcal{L}(\mathrm{S}, \mathrm{T}).$$

Il est immédiat que la deuxième définition entraîne la première. Soit en effet x le barycentre des points $\{x_i\}$ affectés des coefficients α_i tels que $\Sigma\alpha_i = 1$.

$x = \Sigma\alpha_i x_i$ signifie $t_{ax} = \Sigma\alpha_i t_{ax_i}$.

On a donc $u(x) = u(a) + v(t_{ax}) = u(a) + \Sigma\alpha_i v(t_{ax_i})$.

Mais comme $\Sigma\alpha_i = 1$, on peut multiplier $u(a)$ par $\Sigma\alpha_i$ et ceci donne :

$$u(x) = \Sigma\alpha_i[u(a) + v(t_{ax_i})] = \Sigma\alpha_i u(x_i).$$

Pour montrer que la première définition entraîne la deuxième, il faut vérifier que si $u(a + t) = u(a) + v(t)$, v vérifie :

$$v(\lambda t) = \lambda v(t) \; ; \qquad (1)$$
$$v(t_1 + t_2) = v(t_1) + v(t_2). \qquad (2)$$

La première de ces égalités équivaut à :

$$u(a + \lambda t) - u(a) = \lambda v(t) \qquad (1').$$

Or, $a + \lambda t$ peut être considéré comme le barycentre de a et $a + t$ affectés des coefficients $1 - \lambda$ et λ, ce qui donne :

$$\begin{aligned} u(a + \lambda t) &= (1 - \lambda)u(a) + \lambda u(a + t) \\ &= (1 - \lambda)u(a) + \lambda u(a) + \lambda v(t) \\ &= u(a) + \lambda v(t), \end{aligned}$$

c'est-à-dire l'égalité (1').

L'égalité (2) est, de même, équivalente à :

$$u(a + t_1 + t_2) - u(a) = u(a + t_1) - u(a) + u(a + t_2) - u(a),$$

ou encore : $u(a + t_1 + t_2) = u(a + t_1) + u(a + t_2) - u(a)$ (2').

Mais $a + t_1 + t_2$ peut être considéré comme le barycentre de $a + t_1$, $a + t_2$, a affectés de 1, 1 et -1 respectivement, et la conservation du barycentre par l'application u donne l'égalité (2').

Si on convient, généralisant la terminologie de l'espace euclidien considéré comme espace affine sur $\mathbf{R}^3$, de désigner $a + t_1 + t_2$ comme le quatrième sommet du parallélogramme de sommet a construit sur les vecteurs t_1 et t_2, cette deuxième propriété peut aussi s'énoncer. « L'image par une transformation affine d'un parallélogramme est un parallélogramme. »

Exercice 43. — *a)* Définissant un parallélogramme dans un espace affine sur un corps K comme un ensemble ordonné de quatre points *(a, b, c, d)* tel que $b - a = c - d$, montrer que l'on a $d - a = c - b$, c'est-à-dire que *(a, d, c, b)* est un parallélogramme.

b) Montrer l'accord de cette défintion avec celle utilisée ci-dessus (parallélogramme construit à partir d'un point sur deux vecteurs t_1 et t_2).

c) Montrer que si la caractéristique de K est différente de 2, la propriété " *(a, b, c, d)* est un parallélogramme " est équivalente à " les couples *(a, c)* et *(b, d)* ont même milieu ".

d) Montrer que, si K est de caractéristique 2, si *(a, b, c, d)* est un parallélogramme, il en est de même de $(\sigma(a), \sigma(b), \sigma(c), \sigma(d))$ quelle que soit la permutation $\sigma \in \mathcal{S}_4$.

Exercice 44. — La composée de 2 applications affines est une application affine.

Exercice 45. — L'image par une application affine d'une variété linéaire affine est une variété linéaire affine.

L'image réciproque d'une variété linéaire affine est vide ou est une variété linéaire affine.

Exercice 46. — Les notations étant celles du présent paragraphe, le graphe d'une application affine de E dans F est un sous-ensemble de E $\times$ F ; ce dernier a une structure d'espace affine par rapport au produit S $\times$ T.

La condition nécessaire et suffisante pour qu'une application de E dans F soit affine est que son graphe soit une variété affine dont la projection sur E soit tout E.

2. Groupe affine de E.

Dans le cas où $\mathrm{F} = \mathrm{E}$, on peut considérer les applications affines de E dans E. Si on ne considère que celles qui sont bijectives, on voit qu'elles forment un groupe. En effet, la composée de deux bijections est une bijection et la composée de deux applications affines est une application affine (exercice 44).

D'autre part, la bijection identique est une application affine. Enfin, la bijection réciproque d'une application affine donnée par

$$y = u(x) = u(a) + v(x - a)$$

est donnée par $x = u^{-1}(y) = a + v^{-1}(y - u(a))$ qui est bien une application affine puisque v^{-1}, application réciproque d'une application linéaire, est linéaire.

Ce groupe des bijections affines est appelé groupe affine de E.

§ 4. ENSEMBLES CONVEXES

Supposons maintenant que $\mathrm{K} \supset \mathbf{R}$.

On appelle *segment* $[a, b]$ l'ensemble des éléments

$$\lambda a + \mu b \text{ avec } \lambda + \mu = 1 \quad \lambda \in \mathbf{R}^+ \quad \mu \in \mathbf{R}^+.$$

Ceci posé, nous dirons qu'*une partie d'un espace affine est convexe si elle contient tout segment* $[a, b]$ *dont elle contient les extrémités* a *et* b.

On voit immédiatement que l'intersection d'une famille de parties convexes est convexe. On peut donc définir la plus petite partie convexe qui contient une partie A d'un espace affine. *On l'appelle enveloppe convexe de* A.

Exercice 47. — Montrer que l'enveloppe convexe de A est l'ensemble des barycentres positifs des points de A, c'est-à-dire des barycentres de tout sous-ensemble fini x_1, x_2, ..., x_n de points de A affectés de coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathrm{R}^+$. (On montrera que ces barycentres doivent appartenir à l'enveloppe convexe et que leur ensemble est convexe).

CHAPITRE VIII

ÉQUATIONS LINÉAIRES

§ 1

1. Généralités.

Toute équation a pour origine une relation R entre éléments de deux ensembles E et F, et la résolution de l'équation désignée par $x_0 \mathrm{R} y$ (resp. $x \mathrm{R} y_0$) est la détermination de l'ensemble des éléments y de F (resp. x de E) qui sont en relation avec l'élément donné $x_0 \in$ E (resp. $y_0 \in$ E).

La situation plus favorable à laquelle on s'efforce de se ramener (théorie des fonctions implicites, uniformisation des applications réciproques de **C** dans **C**, etc...) est celle où la relation est une application f d'un ensemble E dans un ensemble F et où l'élément donné appartient à F. Une équation associée à l'application f est écrite :

$$f(x) = y_0.$$

La résoudre n'est donc pas autre chose que de déterminer $\overset{-1}{f}(y_0)$.

La remarque suivante, bien qu'évidente, doit être faite, car elle est le principe de toute discussion d'équation :

Pour qu'il existe des solutions, il est nécessaire et suffisant que $y_0 \in f(\mathrm{E})$.

S'il en est ainsi, l'ensemble des solutions de l'équation est le sous-ensemble $\overset{-1}{f}(y_0)$ (qui n'est pas vide).

2. Equations linéaires.

L'équation est dite linéaire si f est une application linéaire d'un espace vectoriel E sur un corps K dans un espace vectoriel F sur le même corps K (que nous supposerons ici, comme dans tout ce qui précède, commutatif).

Ici, nous savons que si $y_0 \in f(\mathrm{E})$, $\overset{-1}{f}(y_0)$ est constituée par une classe d'équivalence modulo le noyau de l'application, ou encore, en considérant E comme espace affine sur lui-même, $\overset{-1}{f}(y_0)$ est la variété linéaire affine parallèle à $\overset{-1}{f}(0)$ et passant par x_0, c'est-à-dire que si x_0 est une solution quelconque fixe de l'équation, toutes les autres sont données par

$$x = x_0 + u \quad \text{où} \quad u \in \overset{-1}{f}(0),$$

ce que l'on peut énoncer :

On obtient toutes les solutions d'une équation linéaire en ajoutant à une solution particulière de l'équation toutes les solutions de l'équation sans second membre.

(L'équation $f(x) = 0$ est dite équation sans second membre ou encore équation homogène associée à l'équation $f(x) = y_0$).

3. Equations et systèmes d'équations linéaires.

On peut se trouver en présence d'un système d'équations linéaires si $\{F_i\}$ ($i \in I$, ensemble arbitraire d'indices) étant une famille d'espaces vectoriels sur le même corps K, et si, pour chacun d'eux, une application

$$f_i : E \longrightarrow F_i$$

ayant été définie, on doit trouver $x \in E$ tel que :

$$\forall i \in I \qquad f_i(x) = y_i \qquad y_i \in F_i \qquad (1),$$

les y_i étant donnés.

Si nous considérons le produit cartésien $F = \Pi F_i$ muni de sa structure naturelle d'espace vectoriel et si nous définissons une application f,

$$f : E \longrightarrow \Pi F_i$$

qui à $x \in E$ fasse correspondre l'élément $(f_i(x))$ dont chaque composante est l'image de x par l'application de même indice, le système d'équations (1) aura mêmes solutions que l'équation

$$f(x) = (y_i) \qquad (2).$$

Réciproquement, une équation du type (2) a mêmes solutions que le système d'équations

$$\forall i \in I \quad pr_i \circ f(x) = y_i,$$

pr_i désignant la projection sur le sous-espace F_i (notation introduite dans l'exercice 18).

Il n'y a donc aucune différence essentielle entre une équation et un système d'équations. On préférera généralement l'écriture sous forme d'équation unique s'il s'agit d'une étude théorique, celle sous forme de système pour une résolution effective.

4. Exemples d'équations linéaires.

1° E est l'espace des fonctions réelles dérivables définies sur **R**, F est l'espace de toutes les fonctions réelles définies sur **R**. L'application f est l'application qui à $\varphi \in E$ fait correspondre $u\varphi + v\varphi'$, u et v étant deux éléments fixes de F, et φ' désignant la dérivée de φ, application dont on vérifie immédiatement qu'elle est linéaire. L'équation est de la forme :

$$u\varphi + v\varphi' = w.$$

C'est une équation différentielle linéaire du premier ordre. (On définirait de façon analogue les équations du second ordre, les systèmes d'équations différentielles, etc...).

2° E est un espace vectoriel quelconque. $F = \Pi F_i$ ($i \in I$), tous les F_i étant égaux à K. Chaque application f_i est alors une forme linéaire. Dans ce cas, on dit qu'on a affaire à des *équations scalaires.*

On peut se ramener à ce cas dès l'instant qu'on connaît effectivement une base (e_i ; $i \in I$) de F. On peut, en effet, comme nous l'avons vu ci-dessus, considérer, au lieu de l'application $f : E \longrightarrow F$, les applications

$$pr_i \circ f : E \longrightarrow Ke_i$$

telles que

$$(pr_i \circ f)(x) = \alpha^i e_i \qquad \alpha^i \in K.$$

Mais, si e^i désigne la forme coordonnée d'indice i sur F, on peut aussi considérer les formes linéaires φ^i sur E définies par :

$$\varphi^i = e^i \circ f.$$

L'équation $f(x) = y_0 \quad x \in E \quad y_0 \in F$ aura mêmes solutions que le système d'équations scalaires

$$\forall i \in I \qquad \varphi^i(x) = \alpha^i$$

si

$$y_0 = \Sigma \alpha^i e_i$$

(les α^i doivent être nuls sauf un nombre fini d'entre eux).

Pratiquement, la détermination d'une solution x se fera par celle de ses coordonnées ξ^λ par rapport à une base (a_λ) de E. ($x = \Sigma \xi^\lambda a_\lambda$, les ξ^λ sont nuls sauf un nombre fini d'entre eux).

L'équation scalaire d'indice i :

$$\varphi^i(x) = \alpha^i \quad \text{ou} \quad < x, \varphi^i > = \alpha^i$$

s'écrit encore :

$$\Sigma \; \xi^\lambda < a_\lambda, \varphi^i > = \alpha^i.$$

Mais les scalaires $< a_\lambda, \varphi^i >$ sont connues, dès que l'application f est connue (les donner est équivalent à donner f). Si on les désigne par μ^i_λ, l'équation d'indice i s'écrit :

$$\Sigma \mu^i_\lambda \xi^\lambda = \alpha^i.$$

§ 2. CAS DE p EQUATIONS SCALAIRES SUR UN ESPACE DE DIMENSION FINIE n

1. Rang d'un système.

Considérons le système

$$< x, \varphi^i > = \alpha^i \qquad i = 1, 2, \dots p$$

qui s'écrit, avec les notations du paragraphe précédent,

$$\sum_{\lambda=1}^{n} \mu^i_\lambda \xi^\lambda = \alpha^i \qquad i = 1, 2, \dots, p.$$

Ce système de p équations scalaires à n inconnues scalaires équivaut à :

$$u(x) = (\alpha^i),$$

u étant une application linéaire de E dans K^p.

On donne du rang d'un tel système les deux définitions suivantes :
1) *le rang du système est le rang de l'application* u, c'est-à-dire la dimension de $u(E)$, (ou encore le rang de la matrice (μ^i_λ)).

2) *Le rang du système est le rang du sous-espace* V *du dual de* E *engendré par les* p *formes linéaires,* c'est-à-dire le nombre maximum de formes linéairement indépendantes qu'on peut extraire de ces p formes.

Ces deux définitions sont équivalentes car, d'après la première, le rang $r = \text{codim}\, \overset{-1}{u}(0)$. Or, $\overset{-1}{u}(0)$ est constitué par le sous-espace annulé par toutes les formes de u, c'est-à-dire que $\overset{-1}{u}(0) = V^\perp$. Or, on a vu (exercice 26) que $\operatorname{codim} V^{\perp} = \dim V$. Donc, $r = \dim V$ (*). On peut encore remarquer que $n - r$ est la dimension de la variété affine décrite par les solutions.

Nous allons maintenant indiquer deux *schémas de résolution d'un tel système* (**).

2. Discussion.

Le premier est surtout théorique, mais met clairement en évidence les résultats de la discussion (théorème classique de Rouché-Fontené).

Supposons que le système soit de rang r, c'est-à-dire que r des formes linéaires φ^i, soit par exemple $\varphi^1, \varphi^2, \ldots, \varphi^r$ soient linéairement indépendantes, les autres $\varphi^{r+1}, \ldots, \varphi^p$, s'exprimant comme combinaison linéaire des premières. On peut construire une base du dual en prenant $\varphi^1, \varphi^2, \ldots, \varphi^r$ et en complétant. On peut alors choisir comme base de E la base duale, c'est-à-dire que $\varphi^1, \varphi^2, \ldots, \varphi^r$ seront des formes coordonnées. Soient $\eta^1, \eta^2, \ldots, \eta^n$ les coordonnées de x dans cette base :

$$< x, \varphi^i > = \eta^i.$$

Par conséquent, les r premières équations du système s'écrivent :

$$\begin{gathered} \eta^1 = \alpha^1 \\ \eta^2 = \alpha^2 \\ \vdots \\ \eta^r = \alpha^r \end{gathered}$$

Quant aux $p - r$ suivantes, on obtient leur forme en se rappelant que pour tout $j > r$, φ^j est une combinaison linéaire des $\varphi^i (1 \leqslant i \leqslant r)$ et est donc de la forme

$$\varphi^j = \lambda_1^j \varphi^1 + \lambda_2^j \varphi^2 + \ldots + \lambda_r^j \varphi^r$$

que l'on a en conséquence :

$$\forall j > r \qquad < x, \varphi^j > = \Sigma \lambda_i^j < x, \varphi^i > = \Sigma \lambda_i^j \eta^i.$$

Les $p - r$ dernières équations s'écrivent :

$$\forall j > r \qquad \Sigma \lambda_i^j \eta^i = \alpha^j.$$

Or, les valeurs de η^i sont fixées par les p premières équations. Pour que le système ait des solutions, il est donc nécessaire que les seconds membres vérifient les $p - r$ relations.

$$j = r + 1, r + 2, \ldots, p \qquad \alpha^j = \sum_{i=1}^{r} \lambda_i^j \alpha^i.$$

Or, dire que (α^i) vérifie $p - r$ relations, c'est dire qu'il annule $p - r$ formes du dual de K^p ; ces $p - r$ formes sont visiblement indépendantes

(*) Chaque ligne de la matrice u donnant les composantes de la forme linéaire correspondante par rapport aux formes coordonnées et le rang du sous-espace V étant donc celui du système des vecteurs lignes, nous retrouvons d'ailleurs simplement ici le fait signalé en IV, 3, 5, que le système des vecteurs lignes et le système des vecteurs colonnes d'une matrice ont même rang (qui est le rang de la matrice).

(**) Nous ne rappellerons pas ici la méthode bien connue de résolution et de discussion qui utilise les déterminants. Les déterminants permettent de donner explicitement les valeurs des coordonnées ξ_λ^i des solutions en fonction des coefficients μ_λ^i et des α^i. L'intérêt en demeure cependant plus théorique que pratique, du moins si l'on entend pratique au sens du calcul numérique.

puisqu'une d'entre elles et une seule contient chacun des α^j. La condition de possibilité du système est donc que (α^i) appartienne à l'orthogonal d'un sous-espace de dimension $p - r$ de K^{p*}. Cet orthogonal est de dimension $p - (p - r) = r$. Nous retrouvons le fait que $u(E)$ est de dimension r.

D'autre part, si (α^i) appartient bien à ce sous-espace, les coordonnées $\eta_{r+1} \ldots \eta_n$ ne sont soumises à aucune condition.

On retrouve le fait que la solution x, dont r coordonnées sont fixées et les autres arbitraires, décrit une classe modulo un sous-espace de dimension $n - r$, c'est-à-dire, encore, une variété linéaire affine de dimension $n - r$.

Le résultat de la discussion peut être résumé sous la forme suivante : La condition nécessaire et suffisante pour qu'un système de p équations scalaires à n inconnues scalaires

$$\varphi^i(x) = \alpha^i \qquad i = 1, 2, \ldots, p$$

de rang r ait des solutions est que les « seconds membres » α^i satisfassent aux relations linéaires qui lient les formes φ^i. L'ensemble des solutions est alors une variété linéaire affine de dimension $n - r$.

3. Un procédé pratique de résolution.

Au lieu de transformer la matrice de l'application u par un changement de la base de E, simplifions-la par un changement de la base de $F = K^p$, comme nous avons appris à le faire (cf. III, 5, 3). La matrice prendra une des formes indiquées alors. Sur cette forme apparaîtra la dimension de $f(E)$, donc le rang r du système (alors appelé p). Et en se plaçant dans le cas général (celui de la troisième forme, c'est-à-dire $n > r$, $p > r$), les équations s'écriront :

$$\begin{aligned} \lambda_1^1 \xi^1 + \ldots + \ldots + \lambda_r^1 \xi^r + \lambda_{r+1}^1 \xi^{r+1} + \ldots + \lambda_n^1 \xi^n &= \beta^1 \\ \lambda_2^2 \xi^2 + \ldots + \lambda_r^2 \xi^r + \lambda_{r+1}^2 \xi^{r+1} + \ldots + \lambda_n^2 \xi^n &= \beta^2 \\ \ddots \\ \lambda_r^r \xi^r + \lambda_{r+1}^r \xi^{r+1} + \ldots + \lambda_n^r \xi^n &= \beta^r \\ 0 &= \beta^{r+1} \\ \vdots \\ 0 &= \beta^p \end{aligned}$$

(Les $\beta^1, \beta^2, \ldots, \beta^p$ sont des combinaisons linéaires des α^i, (β^i) représentant le même vecteur de K^p que (α^i) mais décomposé dans la nouvelle base).

Sur cette forme on voit apparaître les $p - r$ conditions de possibilité

$$\beta^{r+1} = \beta^{r+2} = \ldots = \beta^p = 0.$$

(On retrouve encore le fait que $u(E)$ est de dimension r). D'autre part, ces conditions étant supposées réalisées, on voit qu'on peut déterminer $\xi^1 \ldots \xi^r$ de manière unique en fonction de $\xi^{r+1} \ldots \xi^n$, la $r^{\text{ième}}$ ligne donnant ξ^r, la $r - 1^{\text{ième}}$ ligne ξ^{r-1} et ainsi de suite, de proche en proche. $\xi^{r+1}, \ldots, \xi^n$ sont donc arbitraires et on retrouve le fait que la solution est une classe modulo un sous-espace de dimension $n - r$.

Il s'agit bien ici d'une méthode pratique de résolution qui a l'avantage de fournir à la fois le rang du système, les conditions de possibilité et les solutions.

4. Autre interprétation du système.

Récrivons le système sous forme d'une équation vectorielle unique.

$$\xi^1 u(a_1) + \dots + \xi^n u(a_n) = y_0.$$

Elle peut s'interpréter en disant qu'on cherche à décomposer le vecteur y_0 de F en vecteurs $\xi^i u(a_i)$ colinéaires aux n vecteurs $u(a_i) \dots u(a_n)$, les n composantes définissant l'inconnue x, cette décomposition étant possible si et seulement si y_0 appartient au sous-espace engendré par les vecteurs $u(a_i)$, et l'étant alors de façon unique ou non suivant que ces vecteurs forment une partie libre ou non.

Cette manière de présenter le problème peut être pédagogiquement intéressante pour n et p au plus égaux à 3.

CHAPITRE IX

FORMES BILINÉAIRES ET FORMES QUADRATIQUES

§ 1. PROPRIETES GENERALES

1. Formes bilinéaires. Espace $\mathcal{L}$ (E, F ; K).

Soient E et F deux espaces vectoriels sur le même corps K. Nous avons déjà défini (VI, 2, 1) les applications multilinéaires et en particulier les formes multilénaires. Nous allons maintenant étudier plus particulièrement les formes bilinéaires définies sur $E \times F$, c'est-à-dire les applications :

$$f : E \times F \longrightarrow K$$

satisfaisant à :

$$\begin{array}{ll} \forall\, x_1, x_2 \in E \quad \forall\, y \in F & f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y) \\ \forall\, y_1, y_2 \in F \quad \forall\, x \in E & f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2) \\ \left.\begin{array}{l} \forall\, x \in E \quad \forall\, y \in F \\ \forall\, \lambda \in K \end{array}\right\} & \left\{\begin{array}{l} f(x, \lambda y) = \lambda f(x, y) \\ f(\lambda x, y) = \lambda f(x, y). \end{array}\right. \end{array}$$

Observons que l'ensemble des formes bilinéaires sur $E \times F$:

$$\mathcal{L}\,(E, F\,;\, K)$$

a une structure d'espace vectoriel sur K, avec les définitions suivantes :

$$\begin{array}{ll} \forall\, x \in E,\ \forall\, y \in F & (f + g)\,(x, y) = f(x, y) + g(x, y) \\ \forall\, x \in E,\ \forall\, y \in F,\ \forall\, \lambda \in K & (\lambda f)\,(x, y) = \lambda f(x, y). \end{array}$$

Supposons les espaces E et F rapportés respectivement à des bases (a_i) et (b_j) et soient $x = \Sigma x^i a_i$, $y = \Sigma\, y^j b_j$. La bilinéarité de f entraîne que :

$$f(x, y) = \sum_{i,j} x^i y^j f(a_i, b_j) \tag{1}$$

la sommation étant étendue à tous les indices i et j si E et F sont de dimension finie, à tous les couples (en nombre fini) d'indices i et j pour lesquels $x^i y^j \neq 0$ si E ou F, ou E et F sont de dimension infinie.

Considérons les applications φ_{ij} définies par :

$$\varphi_{ij}(a_i, b_j) = 1 \qquad \varphi_{ij}(a_k, b_l) = 0 \text{ si non } \left\{\begin{array}{l} i = k \\ j = l \end{array}\right.$$

c'est-à-dire qui s'annulent pour tout couple d'indices différent du couple (i, j). On peut écrire :

$$\varphi_{ij}(x, y) = x^i y^j$$

et par conséquent :

$$\forall\, x \in E,\ \forall\, y \in F \qquad f(x, y) = \sum_{i,j} \varphi_{ij}(x, y) f(a_i, b_j)$$

et en posant : $\alpha_{ij} = f(a_i, b_j)$

$$\forall\, x \in E,\ \forall\, y \in F \qquad f(x, y) = \sum_{i,j} \alpha_{ij}\varphi_{ij}(x, y).$$

Cette sommation est étendue à tous les couples d'indices (i, j) si E et F et par conséquent $E \times F$ sont de dimensions *finies*.

On voit sous cette forme que les φ_{ij} constituent alors un système de générateurs de l'espace $\mathcal{L}$ (E, F ; K). Comme elles sont linéairement indépendantes [même démonstration que pour les formes coordonnées en (IV, 3, 4)], elles constituent une base de $\mathcal{L}$. La forme bilinéaire f peut alors s'écrire :

$$f = \sum_{i,j} \alpha_{ij}\varphi_{ij}$$

et est déterminée par la donnée des α_{ij} qui sont ses coordonnées dans la base $\{ \varphi_{ij} \}$.

Par contre, si $E \times F$ est de dimension infinie, les formes φ_{ij} constituent encore une partie libre, mais ne constituent plus un système de générateurs de $\mathcal{L}$. En effet, de la même façon que les formes coordonnées n'engendrent pas l'espace dual tout entier (voir exercice 23), les formes φ_{ij} engendrent seulement l'espace des formes qui ne prennent une valeur différente de zéro que pour un nombre fini de couples (a^i, b^j).

2. Cas où E et F sont de dimensions finies.

Soit : $\dim E = m$; $\dim F = n$. $\mathcal{L}$ (E, F ; K) sera de dimension mn et la donnée d'une forme bilinéaire est, dans ce cas, la donnée des mn coefficients α_{ij}. Ces coefficients peuvent être rangés en une (m, n)-matrice A (m colonnes ; n lignes ; i indice de colonne ; j indice de ligne). La quantité :

$$f(x, y) = \sum_{i,j} \alpha_{ij} x^i y^j$$

peut alors être obtenue comme le produit de la matrice ligne $(y^1 y^2 \ldots y^j \ldots y^n)$ de la matrice A et de la matrice colonne $\begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}$

$$f(x, y) = (y^1 y^2 \ldots y^n) \begin{pmatrix} \alpha_{1n} & \cdots & \alpha_{mn} \\ \vdots & & \vdots \\ \vdots & \cdots & \vdots \\ \vdots & & \vdots \\ \alpha_{11} & \cdots & \alpha_{m1} \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ \vdots \\ x^m \end{pmatrix}$$

La matrice colonne $X = \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}$ représente le vecteur x.

Quant à la matrice ligne, c'est la transposée de la matrice colonne $Y = \begin{pmatrix} y^1 \\ \vdots \\ y^n \end{pmatrix}$ qui représente le vecteur y. On pourra donc écrire :

$$f(x, y) = {}^tYAX$$

si A est la matrice qui définit la forme f.

3. Cas E = F. Définition de la forme quadratique associée.

Plaçons-nous dans le cas où $E = F$. La restriction de l'application $E \times F \longrightarrow K$ à la diagonale définit une application Q de E dans K (qui prend pour valeur la valeur de la forme bilinéaire quand les éléments x et y sont égaux).

$$Q : E \longrightarrow K \qquad Q(x) = f(x, x).$$

Cette application est appelée forme quadratique associée à la forme bilinéaire f.

Elle jouit de la propriété d'*homogénéité* suivante, conséquence de la bilinéarité de f :

$$Q(\lambda x) = f(\lambda x, \lambda x) = \lambda^2 f(x, x) = \lambda^2 Q(x)$$
$$Q(\lambda x) = \lambda^2 Q(x).$$

Si E est de dimension finie, la matrice A est carrée et la formule $f(x, y) = \Sigma\, \alpha_{ij} x^i x^j$ donne :

$$Q(x) = \sum_{i=1}^{n} \alpha_{ii}(x^i)^2 + \sum_{(i,j)} (\alpha_{ij} + \alpha_{ji}) x^i x^j.$$

la sommation $\sum_{(i,j)}$ portant sur l'ensemble des C_n^2 *combinaisons* de deux entiers compris entre 1 et n.

Nous remarquons, sous cette forme, que seule la somme des coefficients $\alpha_{ij} + \alpha_{ji}$ pour toute combinaison (i, j) intervient dans l'expression de Q. La même forme quadratique provient donc d'une infinité de formes bilinéaires, seuls les coefficients de la diagonale principale et la somme des coefficients symétriques par rapport à cette diagonale étant déterminés.

Parmi ces formes, une seule est symétrique ; c'est celle pour laquelle on choisit des coefficients β_{ij} tels que :

$$\beta_{ij} = \beta_{ji} = \frac{\alpha_{ij} + \alpha_{ji}}{2}$$

(cette forme symétrique n'existe donc que si K n'est pas de caractéristique 2). La forme quadratique associée à cette forme bilinéaire symétrique s'écrira :

$$Q(x) = \sum_{i=1}^{n} \alpha_i (x^i)^2 + 2 \sum_{(i,j)} \beta_{ij} x^i x^j$$

(en remplaçant α_{ii} par α_i). Si nous excluons le cas des corps de caractéristique 2 on pourra donc toujours supposer qu'une forme quadratique, dont on sait qu'elle provient d'une forme bilinéaire, provient d'une autre forme, celle-ci symétrique.

Sachant qu'une application de E dans K est une forme quadratique (donc qu'elle provient d'une forme bilinéaire et par conséquent d'une infinité de telles formes), et, sans aborder pour le moment la question des conditions pour qu'il en soit ainsi, nous pouvons déterminer ces formes bilinéaires. Soit Q la forme quadratique et f une des formes bilinéaires cherchées telles que :

$$f(x, x) = \mathrm{Q}(x).$$

Considérons $\mathrm{Q}(x+y) = f(x+y, x+y)$ qui en vertu de la bilinéarité s'écrit :

$$\begin{aligned}\mathrm{Q}(x+y) &= f(x, x) + f(x, y) + f(y, x) + f(y, y)\\ &= \mathrm{Q}(x) + f(x, y) + f(y, x) + \mathrm{Q}(y).\end{aligned}$$

Cette égalité donne la valeur de $f(x, y) + f(y, x)$ pour tout couple (x, y).

$$f(x, y) + f(y, x) = \mathrm{Q}(x+y) - \mathrm{Q}(x) - \mathrm{Q}(y).$$

Si on choisit pour f la forme symétrique, f vérifie :

$$f(x, y) = \frac{\mathrm{Q}(x+y) - \mathrm{Q}(x) - \mathrm{Q}(y)}{2}$$

et est parfaitement définie par cette égalité (*).

f est la forme bilinéaire symétrique associée à Q, appelée aussi forme polaire de Q.

Quant aux autres formes bilinéaires auxquelles est associée la forme Q, remarquons que nous pouvons les obtenir toutes en ajoutant à la forme polaire une forme bilinéaire alternée quelconque. En effet, pour que g soit une telle forme, il est nécessaire et suffisant que $g - f$ soit une forme bilinéaire et vérifie :

$$(g - f)\,(x, x) = g(x, x) - f(x, x) = 0,$$

donc que $g - f$ soit une forme bilinéaire alternée.

Remarque 1. Une forme quadratique peut d'ailleurs être définie comme une application de E dans K satisfaisant à :

$$\mathrm{Q}(\lambda x) = \lambda^2\, \mathrm{Q}(x)$$

et telle que : $\dfrac{\mathrm{Q}(x+y) - \mathrm{Q}(x) - \mathrm{Q}(y)}{2} = f(x, y)$

soit une forme bilinéaire. On vérifie immédiatement que $f(x, x)$ est bien égale à $\mathrm{Q}(x)$ car :

$$f(x, x) = \frac{\mathrm{Q}(2\,x) - 2\,\mathrm{Q}(x)}{2} = \frac{4\,\mathrm{Q}(x) - 2\,\mathrm{Q}(x)}{2} = \mathrm{Q}(x).$$

Par conséquent, cette définition est équivalente à celle que nous avons donnée.

Remarque 2. On peut écrire :

$$\begin{aligned}\mathrm{Q}\,(x+y) &= \mathrm{Q}(x) + 2\,f(x, y) + \mathrm{Q}(y)\\ \mathrm{Q}\,(x-y) &= \mathrm{Q}(x) - 2\,f(x, y) + \mathrm{Q}(y)\end{aligned}$$

(*) Un calcul analogue aboutit à l'expression :

$$f(x, y) = \frac{\mathrm{Q}(x) + \mathrm{Q}(y) - \mathrm{Q}(x-y)}{2}$$

car $f(x, -y) = -f(x, y)$ et $\mathrm{Q}(-y) = (-1)^2\,\mathrm{Q}(y)$. D'où en ajoutant :

$$\mathrm{Q}(x+y) + \mathrm{Q}(x-y) = 2\,[\mathrm{Q}(x) + \mathrm{Q}(y)]\ (*).$$

Exercice 48. — Soit E un espace vectoriel sur **R**, et Q une application de E dans **R**, vérifiant les conditions suivantes :

1° Pour tout couple (x, y) d'éléments de E, on a

$$(\mathrm{Q}) \qquad \mathrm{Q}\,(x+y) + \mathrm{Q}\,(x-y) = 2\,\mathrm{Q}(x) + 2\,\mathrm{Q}(y)$$

2° Pour tout couple d'éléments fixes (x, y) de E, l'application de **R** dans **R** définie par $\lambda \longrightarrow \mathrm{Q}\,(x + \lambda y)$ est continue.

Montrer que si l'on pose :

$$f\,(x, y) = \frac{\mathrm{Q}\,(x+y) - \mathrm{Q}\,(x-y)}{4}$$

f est une forme bilinéaire symétrique sur E, dont Q est la forme quadratique associée.

Remarque 3. Toute application linéaire de E dans lui-même qui conserve une forme quadratique conserve sa forme polaire et réciproquement :

$$f\,[u(x), u(y)] = f(x, y) \iff \mathrm{Q}\,(u(x)) = \mathrm{Q}(x).$$

4. Influence des changements de base sur l'expression d'une forme bilinéaire (espace de dimension finie).

Nous avons vu que :

$$f(x, y) = {}^t\mathrm{Y\,A\,X}$$

ce qui donne, dans le cas E = F, pour la forme quadratique associée :

$$\mathrm{Q}(x) = {}^t\mathrm{X\,A\,X}.$$

Soit à effectuer un changement de base dans E, donné par la matrice de passage M, tel que :

$$\mathrm{X} = \mathrm{M\,X'} \qquad \mathrm{Y} = \mathrm{M\,Y'}.$$

L'expression de $f(x, y)$ dans la nouvelle base est :

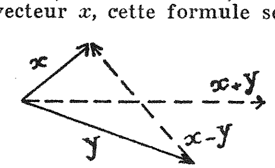
$${}^t(\mathrm{M\,Y'})\,\mathrm{A\,M\,X'} = {}^t\mathrm{Y'}\,{}^t\mathrm{M\,A\,M\,X'}.$$

On voit que la forme bilinéaire est exprimée dans la nouvelle base par une nouvelle matrice :

$$\mathrm{A'} = {}^t\mathrm{M\,A\,M}.$$

Cette formule est à rapprocher de la formule du changement de base pour un endomorphisme : $\mathrm{A'} = \mathrm{M}^{-1}\,\mathrm{A\,M}$. Les considérations qui suivent vont expliquer analogie et différence entre ces deux formules.

(*) Quand nous aurons donné, ultérieurement, dans certains cas, à $\mathrm{Q}(x)$ le sens du carré de la longueur du vecteur x, cette formule se traduira par le théorème dit



« de la médiane » (la somme des carrés de deux côtés d'un triangle, etc...) ou sa variante : la somme des carrés des quatre côtés d'un parallélogramme est égale à la somme des carrés des diagonales.

§ 2. ISOMORPHISME $\mathcal{L}$ (E, F ; G) $\approx$ $\mathcal{L}$ (F, $\mathcal{L}$ (E, G)) ET SES CONSEQUENCES

1.

Considérons les applications bilinéaires de $E \times F$ dans G ; E, F et G étant trois espaces vectoriels sur le même corps K, de dimensions finies ou infinies.

Théorème : $\mathcal{L}(E, F ; G) \approx \mathcal{L}(F, \mathcal{L}(E, G))$.

Soit f une application bilinéaire de $E \times F$ dans G. Si on fixe $y \in F$ égal à une valeur y_0, la correspondance :

$$g_{y_0} : x \longrightarrow f(x, y_0)$$

est une application linéaire de E dans G qu'on pourra noter g_{y_0}. Nous avons donc défini une application :

$$y_0 \in F \longrightarrow g_{y_0} \in \mathcal{L}(E, G) \qquad (1).$$

Vérifions que cette application est linéaire :

$$\forall x \in E \quad g_{y_0+z_0}(x) = f(x ; y_0 + z_0) = f(x, y_0) + f(x, z_0)$$
$$= g_{y_0}(x) + g_{z_0}(x),$$

donc, $g_{y_0+z_0} = g_{y_0} + g_{z_0}$.

$$\forall x \in E \quad g_{\lambda y_0}(x) = f(x, \lambda y_0) = \lambda f(x, y_0) = \lambda g_{y_0}(x),$$

donc, $g_{\lambda y_0} = \lambda g_{y_0}$.

L'application (1) appartient donc à $\mathcal{L}(F ; \mathcal{L}(E, G))$.

Ainsi, nous avons défini une application :

$$\mathcal{L}(E, F ; G) \longrightarrow \mathcal{L}\Big(F ; \mathcal{L}(E, G)\Big),$$

faisant correspondre à f l'application (1) et dont nous allons établir qu'elle est un isomorphisme.

a) *D'abord, elle est linéaire.* Soient en effet :

$$\begin{array}{l} f \in \mathcal{L}(E \times F ; G) \\ \varphi \in \mathcal{L}(E \times F ; G) \end{array} \quad \text{et leurs images} \quad \begin{array}{l} y \to g_y \\ y \to \psi_y \end{array}$$

L'image de $f + \varphi$ est l'application qui, à y, fait correspondre l'application u_y définie par :

$$\forall x \in E \qquad u_y(x) = (f + \varphi)(x, y)$$
$$= f(x, y) + \varphi(x, y)$$
$$= g_y(x) + \psi_y(x),$$

donc, $\qquad u_y = g_y + \psi_y$.

Ceci étant vrai pour tout $y \in F$, l'application $y \longrightarrow u_y$ est bien la somme des applications $y \longrightarrow g_y$ et $y \longrightarrow \psi_y$.

De même, l'image de λf est une application qui à y_0 fait correspondre l'application χ_y définie par :

$$\forall x \in E \quad \chi_y(x) = (\lambda f)(x, y) = \lambda f(x, y) = \lambda g_y(x),$$

donc, $\chi_y = \lambda g_y$.

et ceci étant vrai pour tout $y \in F$, l'application $y \longrightarrow \chi_y$ est bien le produit par λ de l'application $y \longrightarrow g_y$.

b) *Ensuite, elle est injective :* gardons les mêmes notations et supposons $f \neq \varphi$. Il faut comparer $y \longrightarrow g_y$ et $y \longrightarrow \psi_y$.

$$f \neq \varphi \quad \Longrightarrow \quad \exists (x_0, y_0) \qquad f(x_0, y_0) \neq \varphi(x_0, y_0)$$
$$\text{donc,} \quad \exists y_0 \qquad g_{y_0}(x_0) \neq \psi_{y_0}(x_0),$$

ce qui implique : $g_{y_0} \neq \psi_{y_0}$, ce qui signifie que les applications $y \longrightarrow g_y$ et $y \longrightarrow \psi_y$ sont distinctes.

c) *Enfin, elle est surjective.* En effet, si on a un élément de :

$$\mathcal{L}\Big(F, \mathcal{L}(E, G)\Big),$$

c'est-à-dire une application linéaire :

$$y \longrightarrow g_y \qquad (1)$$

où g_y représente une application linéaire en x, on peut définir une application :

$$(x, y) \longrightarrow f(x, y) = g_y(x),$$

et cette application est bilinéaire (linéaire en x du fait de la linéarité de g, linéaire en y du fait de la linéarité de l'application (1)). C'est donc un élément de $\mathcal{L}(E, F ; G)$.

L'isomorphisme est donc établi.

Remarque : Si, plus généralement, on considère l'ensemble des applications du produit cartésien de deux ensembles dans un troisième ensemble et si, dans le raisonnement précédent, on néglige ce qui est relatif à la linéarité des applications, tout en conservant le reste, on arrive au fait qu'il existe une bijection entre :

$$\mathcal{F}(E \times F, G) \qquad \text{et} \qquad \mathcal{F}\Big(F, \mathcal{F}(E, G)\Big).$$

Si on désigne par a, b, c les cardinaux des trois ensembles, il en résulte que c^{ab} cardinal du premier ensemble est égal à $(c^a)^b$ cardinal du deuxième ; d'où l'égalité entre nombres cardinaux :

$$c^{ab} = (c^a)^b.$$

2. Application aux formes bilinéaires. (Cas E = F ; G = K).

$\mathcal{L}(E, K)$ est le dual E^*. Le théorème précédent s'écrit donc :

$$\mathcal{L}(E, E ; K) \approx \mathcal{L}(E, E^*).$$

A toute forme bilinéaire f sur E^2 correspond une application linéaire g de E dans E^* qui, à $y \in E$, fait correspondre la forme linéaire sur E, $g(y)$ précédemment notée g_y et définie par $g_y(x) = f(x, y)$.

$$g : y \in E \longrightarrow g_y \in E^*.$$

Or, $g_y(x)$, valeur en x de la forme linéaire g, est la valeur de la forme bilinéaire canonique $< x, g_y >$ ou, avec la nouvelle notation, $< x, g(y) >$.

On peut donc écrire :

$$f(x, y) = < x, g(y) >.$$

On peut naturellement définir de façon analogue l'application :

$$\gamma : x \in E \longrightarrow \gamma_x \in E^*$$

avec $\gamma_x(y) = f(x, y)$, et γ_x sera notée dorénavant $\gamma(x)$. On aura encore :

$$f(x, y) = < y, \gamma(x) >.$$

Supposons E *de dimension finie* et rapporté à une base (e_i), et E^* rapporté à la base duale (e^j). L'application γ est alors définie par la donnée des éléments de E^* :

$$\gamma(e_i) = \sum_j \alpha_{ij} e^j$$

(notation conforme à ce qui a été prévu en IV, 4, 6°).

Si $$x = \Sigma\, x^i\, e_i \qquad \gamma(x) = \Sigma\, x^i\, \alpha_{ij}\, e^j$$
et $$< y, \gamma(x) > = < \Sigma\, y^k\, e_k,\ \Sigma\, \alpha_{ij}\, x^i\, e^j > = \Sigma\, \alpha_{ij}\, x^i\, y^j.$$

On retrouve l'expression de la forme bilinéaire f écrite en (IX, 1, 1) ; et on constate l'identité des coefficients α_{ij} introduits *à priori* ou comme nous venons de le faire, donc l'identité de la matrice de l'application bilinéaire avec la matrice de l'application γ.

D'autre part, remarquons que l'application γ de E dans E* admet pour transposée une application de E** dans E ; mais nos espaces étant à un nombre fini de dimensions, E** peut être identifié à E et la transposée de γ est une application de E dans E* définie par :

$$< \gamma(x), y > = < x, {}^t\gamma(y) >.$$

Mais comme on avait $f(x, y) = < \gamma(x), y > = < x, g(y) >$, on voit que l'application g n'est autre que la transposée de l'application γ.

Remarquons enfin que le fait que f soit symétrique est équivalent à $\gamma = g$, c'est-à-dire $g = {}^t g$.

Changement de base : Supposons maintenant qu'on prenne une nouvelle base a_i de E, et soit (a^j) la nouvelle base duale de E*. On sait que la matrice de γ avec les nouvelles bases se trouve en écrivant :

$$\begin{tikzcd}
\underset{(e_i)}{E} \arrow[r, "\gamma"] & \underset{(e^j)}{E^*} \\
\underset{(a_i)}{E} \arrow[u] & \underset{(a^j)}{E^*} \arrow[u]
\end{tikzcd}$$

$$\gamma = I_{E^*} \circ \gamma \circ I_E,$$

où I_E représente l'identité dans E rapporté à deux bases différentes, et I_{E^*} l'identité dans le dual rapporté aux deux bases duales, le changement se faisant en sens inverse. Mais l'identité dans le dual est la transposée de l'identité dans E (l'égalité

$$< u(x), y^* > = < x, {}^t u(y^*) >$$

de la théorie générale, devenant

$$< x, y^* > = < x, {}^t u(y^*) >,$$

donc, ${}^t u(y^*) = (y^*)$. Or, on a vu (IV, 3, 5) que les espaces E et F étant rapportés à des bases et leurs duals aux bases duales, la matrice de l'application u de E dans F est la transposée de la matrice de l'application ${}^t u$ de F* dans E*. Appliquée ici, où F représente E* et F* représente E (par identification avec E**), cette règle montre que la matrice de I_{E^*} est la transposée de la matrice de I_E. On retrouve donc :

$$A' = {}^t\, M\, A\, M,$$

cette formule prenant ici le sens d'une formule de changement de base pour l'application linéaire γ.

3. Orthogonalité.

Deux vecteurs x et y d'un espace vectoriel E *sont dits orthogonaux ou conjugués par rapport à la forme quadratique* Q *s'ils annulent la forme polaire de* Q. Cette relation est symétrique.

Nous avons donné précédemment (IV, 2) une autre définition de l'orthogonalité concernant deux éléments x et x^* appartenant respectivement à E et E* :

$$x \perp x^* \iff < x, x^* > = 0.$$

Or, $f(x, y) = 0 \iff < x, \gamma(y) > = 0 \iff < y, g(x) > = 0.$

Par conséquent, dire : deux éléments d'un même espace vectoriel sont orthogonaux par rapport à φ, équivaut à dire : un de ces éléments est orthogonal à l'image de l'autre dans le dual par la transformation γ (resp. g).

Si nous cherchons alors l'ensemble des éléments orthogonaux par rapport à φ à un élément x, c'est l'ensemble des éléments orthogonaux (à l'ancien sens du mot) à l'élément $g(x)$ du dual ; c'est donc un hyperplan de l'espace, sous-espace de codimension 1, si $g(x) \neq 0$.

§ 3. FORME CANONIQUE DES FORMES QUADRATIQUES SUR UN ESPACE DE DIMENSION FINIE

1. Décomposition en carrés.

Soit n la dimension de l'espace. Il peut arriver que la forme quadratique $Q(x) = f(x, x)$ s'annule sur tout E. Laissons ce cas à part et supposons :

$$\exists\, a_1 \qquad f(a_1, a_1) \neq 0\ ; \quad \text{donc } g(a_1) \neq 0.$$

Considérons l'hyperplan :

$$H_1 = \{ x\ ;\ f(a_1, x) = 0 \} \qquad \dim H_1 = n - 1,$$

a_1 n'appartient pas à H_1 car $f(a_1, a_1) \neq 0$.

Il peut arriver que $\forall x \in H_1, f(x, x) = 0$. Supposons qu'il n'en soit pas ainsi et que :

$$\exists\, a_2 \in H_1 \qquad f(a_2, a_2) \neq 0, \text{ donc } g(a_2) \neq 0.$$

Les vecteurs de H_1 orthogonaux à a_2 forment un hyperplan de H_1, donc un sous-espace vectoriel de E de dimension $n - 2$, soit H_2, qui est l'ensemble des vecteurs orthogonaux à a_1 et a_2. Il se peut que $\forall x \in H_2 \quad f(x, x) = 0$. S'il n'en est pas ainsi :

$$\exists\, a_3 \qquad \begin{matrix} a_3 \perp a_1 \\ a_3 \perp a_2 \end{matrix} \qquad f(a_3, a_3) \neq 0 \qquad \begin{matrix} a_3 \neq a_1 \\ a_3 \neq a_2 \end{matrix}$$

Nous considérerons l'hyperplan H_3 de H_2 de dimension $n - 3$ formés des vecteurs orthogonaux à a_1, a_2, a_3 et ainsi de suite. Nous poursuivrons ce processus jusqu'à ce que nous tombions sur un sous-espace H_r de dimension $n - r$ tel que :

$$\forall x \in H_r \qquad Q(x) = f(x, x) = 0.$$

Il en résulte que si x et y sont deux vecteurs de H_r, on a $f(x, y) = 0$

$$\left(\text{en vertu de } f(x, y) = \frac{Q(x + y) - Q(x) - Q(y)}{2} \right).$$

r peut être égal à la dimension n de l'espace ou inférieur à n, et on a mis en évidence r vecteurs $a_1\, a_2 \ldots a_r$ qui sont deux à deux orthogonaux Ces vecteurs ont été choisis de façon telle qu'ils forment une base d'un sous-espace supplémentaire de H_r. En effet :

$$E = K a_1 \oplus H_1,\ H_1 = K a_2 \oplus H_2, \ldots H_{r-1} = K a_r \oplus H_r.$$

Si donc $r = n$, $a_1, a_2 \dots a_n$ constituent une base de E. Si $r < n$, en prenant une base de $n - r$ vecteurs dans H_r, nous compléterons $(a_1, a_2 \dots a_r)$ en une base. Nous pouvons donc énoncer :

Une forme quadratique étant donnée sur un espace vectoriel de dimension finie, celui-ci possède une base formée d'éléments deux à deux orthogonaux.

Les vecteurs de cette base vérifient :

$$\begin{array}{ll} \forall i \neq j & f(a_i, a_j) = 0 \\ \forall i \leqslant r & f(a_i, a_i) \neq 0 \\ \forall i > r & f(a_i, a_i) = 0 \end{array}$$

Par rapport à cette base, la matrice de la forme f prend donc la forme diagonale (*), les $n - r$ derniers coefficients de cette diagonale étant nuls.

$$f(x, y) = \sum_{i=1}^{r} \alpha_i x_i^2.$$

L'intérêt de cette décomposition est que la valeur de r est une propriété intrinsèque de f indépendante des choix successifs de $a_1, a_2 \dots$, etc..., c'est-à-dire *indépendante du choix de la base*. En effet, r est la codimension de H_r qui est l'ensemble des vecteurs orthogonaux à tous les vecteurs de l'espace (puisqu'ils sont tous orthogonaux entre eux et orthogonaux aux vecteurs $a_1, a_2 \dots a_r$). En utilisant la première définition de l'orthogonalité, nous voyons que H_r est donc l'orthogonal de l'image dans le dual de tout l'espace E.

$$H_r = g(E)^{\perp}.$$

D'où (voir exercice 26) $r = \operatorname{codim} H_r = \dim g(E)$, *$r$ est un invariant de f*. On l'appelle *le rang de la forme quadratique* (c'est aussi le rang des applications g et γ et le rang de la matrice A).

2. Signe des coefficients des carrés. Signature d'une forme quadratique.

Nous allons maintenant supposer que $K \subset \mathbf{R}$ ou du moins que K est un corps totalement ordonné, et nous allons montrer que, dans ce cas, le nombre des coefficients α_i positifs est aussi une propriété intrinsèque de la forme, invariante par changement de base.

Supposons en effet que nous ayons deux décompositions en carrés :

— l'une par rapport à une base (a_i) avec p coefficients positifs et r-p coefficients négatifs ;

— l'autre par rapport à une base (b_j) avec q coefficients positifs et r-q coefficients négatifs.

On peut toujours ranger les éléments de ces bases de façon que les p coefficients positifs de la première forme et les q coefficients positifs de

---

(*) La diagonalisation obtenue ici de la matrice de la forme bilinéaire ne doit pas être confondue avec la diagonalisation recherchée précédemment des matrices d'endomorphismes. Nous avons vu que la matrice de f était celle de l'application g, application de E dans E*. La diagonalisation que nous venons d'obtenir diagonalise cette matrice au prix d'un changement de base dans E et d'un changement corrélatif dans E* (puisqu'on suppose toujours E* rapporté à la base duale). Le problème est donc tout différent de celui de diagonaliser une matrice d'endomorphisme grâce à un seul changement de base, celui de l'espace E sur lequel est défini l'endomorphisme.

la deuxième forme soient les premiers. Considérons les sous-espaces engendrés, le premier par :

$$a_{p+1}, a_{p+2} \dots a_r, a_{r+1} \dots a_n \quad \text{de dimension } n - p$$

le deuxième engendré par :

$$b_1, b_2 \dots b_q \quad \text{de dimension } q$$

Pour tout vecteur x du premier : $Q(x) \leqslant 0$.

Pour tout vecteur $x \neq 0$ du deuxième : $Q(x) > 0$ (strictement).

Il y aurait donc contradiction pour les vecteurs non nuls de l'intersection et celle-ci doit être réduite à zéro. Or, la somme des dimensions des deux sous-espaces est $n - p + q$. Si q était supérieur à p, cette somme serait supérieure à n et l'intersection ne pourrait être réduite à zéro. Donc $q \leqslant p$. Mais on montrerait de la même façon que p ne peut être supérieur à q et on conclut $p = q$. *p est un invariant.*

On appelle *signature* d'une forme quadratique la donnée des entiers qui caractérisent la nature de la décomposition en carrés. Ce pourra être les trois entiers :

$$p,\ r - p,\ n - r$$

qui indiquent que la forme décomposée en carrés possède :

p carrés à coefficients positifs,
$r - p$ carrés à coefficients négatifs,
$n - r$ carrés à coefficients nuls.

Cas $p = r$. Dans ce cas, la forme diagonalisée ne comprend que des carrés à coefficients positifs. Elle s'écrit par rapport à toute base formée de vecteurs deux à deux orthogonaux :

$$Q(x) = \sum_{i=1}^{r} \alpha_i (x^i)^2 \qquad \alpha_i > 0.$$

Si, plus particulièrement, $p = r = n$:

$$Q(x) = \sum_{i=1}^{n} \alpha_i (x^i)^2 \qquad \alpha_i > 0.$$

Cette forme ne prend que des valeurs positives, et la valeur zéro seulement si toutes les coordonnées du vecteur sont nulles.

$$Q(x) = 0 \implies x = 0.$$

Une telle forme est dite *définie positive.*

Si $p = r < n$, la forme ne prend aussi que des valeurs positives ou nulles, mais elle peut prendre la valeur zéro pour des vecteurs non nuls (vecteurs appartenant à H_r). On dit qu'elle est *semi-définie positive.*

Notons enfin que des vecteurs colinéaires à des vecteurs orthogonaux sont orthogonaux. Si le corps K contient les racines carrées de ses éléments positifs (ce qui est vérifié pour $K = \mathbf{R}$), on pourra remplacer les vecteurs $a_1 a_2 \dots a_r$ par des vecteurs $\lambda_1 a_1, \lambda_2 a_2 \dots \lambda_r a_r$ tels que :

$$Q(\lambda_1 a_1) = (\lambda_1)^2 Q(a_1) = \pm 1.$$

Par rapport à la nouvelle base ainsi choisie, la forme quadratique s'exprimera par :

$$Q(x) = \sum_{i=1}^{n} \varepsilon_i (x^i)^2 \qquad \text{avec } \varepsilon_i = +1, -1 \text{ ou } 0.$$

C'est le cas, par exemple, de l'espace de la relativité restreinte où on considère la forme :

$$Q(x) = (x^1)^2 + (x^2)^2 + (x^3)^2 - (x^4)^2$$

avec x^1, x^2, x^3 coordonnées d'un point dans $\mathbf{R}^3$ et $x^4 = ct$, c représentant la vitesse de la lumière et t le temps.

3. Cas $r = n$. Identification de E et E*.

Revenons au cas où $r = n$. L'application g est un isomorphisme. On sait (voir exercice 24) que E et E* peuvent être isomorphes d'une infinité de façons, mais qu'aucune ne réalise un isomorphisme canonique φ pour lequel on aurait, quel que soit l'automorphisme u de E :

$$< u(x), \varphi \circ u(y) > = < x, \varphi(y) > \qquad (1).$$

Mais si f joue un rôle particulier, il est naturel de considérer plus particulièrement l'isomorphisme g et *on peut convenir d'identifier* E* *et* E *au moyen de* g.

L'égalité (1) n'est alors pas réalisée pour tous les automorphismes u, mais pour ceux qui vérifient :

$$f(u(x), u(y)) = f(x, y)$$

égalité équivalente à l'égalité (1), où φ est fait égal à g. L'isomorphisme établi par g est donc respecté par tous les automorphismes qui conservent la forme bilinéaire f à laquelle g est associée, et par ceux-là seulement.

Cette identification une fois faite, on voit se confondre les deux définitions de l'orthogonalité. En effet :

2° définition : $x \perp y \iff f(x, y) = 0 \quad$ si $x \in \mathrm{E} \quad y \in \mathrm{E}$,

1re définition : $x \perp g(y) \iff < x, g(y) > = f(x, y) = 0$
si $x \in \mathrm{E} \quad g(y) \in \mathrm{E}^*$.

Si on identifie $g(y)$ et y, toute différence disparaît, mais il importe de noter que la définition de l'orthogonalité de deux vecteurs de E n'est pas intrinsèque pour la structure d'espace vectoriel de E ; elle est relative à une structure plus riche correspondant à la double donnée d'une structure d'espace vectoriel et d'une forme quadratique, choisie une fois pour toutes, qu'on appelle forme quadratique fondamentale.

Par exemple, dans $\mathbf{R}^3$ rapporté à un trièdre orthonormé $Oxyz$, cette identification signifie qu'on assimile le vecteur de coordonnées (u, v, w) à la forme linéaire :

$$(x, y, z) \longrightarrow ux + vy + wz$$

forme qui a pour coordonnées (u, v, w) dans la base duale. Que les deux vecteurs $\mathrm{U}(u, v, w)$ et $\mathrm{U}'(u', v', w')$ soient orthogonaux s'écrit alors avec nos notations $< \mathrm{U}, \mathrm{U}' > = 0$ ou $uu' + vv' + ww' = 0$, c'est-à-dire que U' appartient au plan $ux + vy + wz = 0$.

Sous-espaces orthogonaux. Nous supposons toujours $r = n$. Soit H un sous-espace de E. Il a, suivant la première définition de l'orthogonalité, un orthogonal H' dans E* et dim H' = codim H. Il a, suivant la deuxième définition, un orthogonal K dans E qui est l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de H, et qui a pour image H' par l'isomorphisme g de E sur E*. K est donc un sous-espace de E et

$$\dim \mathrm{K} = \operatorname{codim} \mathrm{H}.$$

Si, de plus, *la forme quadratique est définie positive*, seul 0 peut être orthogonal à lui-même, par suite $\mathrm{H} \cap \mathrm{K} = \{0\}$ *et* K *est un supplémentaire de* H.

§ 4. — NORME ET DISTANCE ASSOCIEES A UNE FORME QUADRATIQUE DEFINIE POSITIVE. ESPACE EUCLIDIEN. ESPACE HILBERTIEN REEL.

Etant donné un espace vectoriel de dimension finie ou infinie sur $\mathbf{R}$, une forme quadratique sur cet espace est dite définie positive, si elle ne prend que des valeurs positives ou nulles et ne s'annule que pour le vecteur nul.

La forme bilinéaire symétrique associée à une forme quadratique définie positive est appelée produit scalaire. Le produit scalaire de deux vecteurs est la valeur de la forme pour ce couple de vecteurs et sa nullité exprime l'orthogonalité de ces deux vecteurs. On rencontre pour le produit scalaire de x et y les diverses notations suivantes :

$$x . y, \quad < x, y >, \quad (x|y), \quad (x, y), \quad [x, y], \quad \text{etc...}$$

On peut associer à toute forme quadratique définie positive une norme et une distance.

Définition générale d'une norme :

Une norme sur un espace vectoriel E sur $\mathbf{R}$ (ou sur $\mathbf{C}$) est une application de E dans $\mathbf{R}+$ qui au vecteur $x \in \mathrm{E}$ fait correspondre le réel positif $\|x\|$, appelé norme de x, et qui satisfait aux trois conditions suivantes :

1) $\quad \|x\| = 0 \iff x = 0$,
2) $\quad \forall \lambda \in \mathbf{R}$ (ou $\mathbf{C}$) $\quad \|\lambda x\| = |\lambda| \, \|x\|$,
3) $\quad \|x + y\| \leqslant \|x\| + \|y\|$.

Cette dernière condition est souvent appelée inégalité triangulaire.

Lorsqu'une norme est définie sur un espace vectoriel, celui-ci est qualifié de *normé*.

Exercice 49. — Montrer que dans un espace vectoriel normé la boule unité $\{ x ; \|x\| \leqslant 1 \}$ est convexe.

Norme relative à une forme quadratique définie positive sur un espace vectoriel sur $\mathbf{R}$.

Soit Q la forme quadratique, et f sa forme polaire. On pose :

$$\|x\| = \sqrt{f(x, x)} = \sqrt{\mathrm{Q}(x)}.$$

Montrons qu'elle satisfait aux axiomes des normes.

$\mathrm{Q}(x) = 0$ entraîne $x = 0$ puisque la forme est définie positive.

$$\|\lambda x\| = \sqrt{\mathrm{Q}(\lambda x)} = \sqrt{\lambda^2 \mathrm{Q}(x)} = |\lambda| \, \|x\|.$$

Enfin, l'inégalité triangulaire sera établie par *l'inégalité de Schwarz*. Soit $f(x + \lambda y, x + \lambda y)$; cette quantité doit être positive ou nulle pour tout λ. Elle peut s'écrire :

$$f(x, x) + 2\lambda f(x, y) + \lambda^2 f(y, y)$$

et, pour que ce trinôme en λ soit positif ou nul pour tout λ, il est nécessaire que :

$$[f(x, y)]^2 \leqslant f(x, x) \times f(y, y) \qquad (1)$$

(inégalité de Schwarz). Nous en déduisons :

$$|f(x, y)| \leqslant \|x\| . \|y\|.$$

Or
$$\begin{aligned} \|x + y\|^2 = f(x + y, x + y) &= \|x\|^2 + \|y\|^2 + 2f(x, y) \\ &\leqslant \|x\|^2 + \|y\|^2 + 2\|x\| \, \|y\| \\ \|x + y\|^2 &\leqslant (\|x\| + \|y\|)^2. \end{aligned}$$

Remarque : Observons que l'égalité entre les deux membres de (1) signifie qu'il existe λ tel que $f(x + \lambda y, x + \lambda y) = 0$, ce qui entraîne

$x + \lambda y = 0$. Autrement dit, l'égalité entre les deux membres exprime que les vecteurs x et y sont colinéaires.

Cosinus de l'angle de deux vecteurs.

Il est défini par :

$$\cos V = \frac{f(x, y)}{\|x\| \, \|y\|}$$

quantité qui est toujours comprise entre -1 et $+1$, en vertu de l'inégalité de Schwarz.

Exemple : Soit $E = \mathcal{C}([0, 1], \mathbf{R})$, espace vectoriel des fonctions réelles sur $[0, 1]$; on peut considérer la forme :

$$f(x, y) = \int_0^1 x(t)y(t)dt$$

qui est bilinéaire et associée à la forme quadratique :

$$\int_0^1 x^2(t)dt$$

définie positive puisqu'elle ne s'annule que pour la fonction nulle ; la norme d'une fonction sera :

$$\|x\| = \sqrt{\int_0^1 x^2(t)dt}$$

et le cosinus de deux fonctions sera :

$$\cos(x, y) = \frac{\int_0^1 x(t)y(t)dt}{\sqrt{\int_0^1 x^2(t)dt \int_0^1 y^2(t)dt}}$$

Distance. On appelle distance, de deux vecteurs d'un espace vectoriel normé, la norme de leur différence :

$$d(x, y) = \|x - y\|.$$

C'est une application $E^2 \longrightarrow \mathbf{R}+$ qui satisfait aux axiomes suivants :

$$d(x, y) = 0 \iff x = y$$
$$d(x, y) = d(y, x)$$
$$d(x, y) \leqslant d(x, z) + d(z, y).$$

Cette distance est invariante par translation :

$$d(a + x, a + y) = d(x, y).$$

Distance sur un espace affine par rapport à un espace vectoriel normé T.

On peut définir une distance sur E, en posant pour tout couple de points (p, q) de E :

$$d(p, q) = \|t_{pq}\|$$

où t_{pq} est le vecteur de T représentant la translation de E qui envoie p sur q.

Cette distance est invariante par toutes les translations de E.

Terminologie. On peut donner le nom d'*espace de Hilbert* à tout espace vectoriel sur **R**, muni d'une forme quadratique, définie, positive. Mais une terminologie plus précise distingue :

1) les *espaces préhilbertiens réels,* de dimension infinie ;

2) les *espaces hilbertiens réels* ou « de Hilbert, réels », de dimension infinie et complets pour la distance associée à la norme (cf. *A.P.M. I ;* IV, 1) ;

3) les *espaces euclidiens,* de dimension finie, qui sont toujours complets (ceci sera démontré dans *A.P.M. III ;* cf., dès maintenant, l'exercice 50 pour l'espace euclidien de dimension 2). Ce cas peut être considéré comme un cas particulier du précédent.

Le terme d'espace euclidien est aussi utilisé pour désigner l'espace affine et métrique associé à un espace vectoriel normé du type précédent.

Base orthonormée d'un espace euclidien.

Nous avons vu (IX, 3, 2) que l'on pouvait choisir, pour un espace de dimension finie, une base formée de vecteurs a_i, deux à deux orthogonaux et tels que la forme quadratique s'écrivait $Q(x) = \sum_{i=1}^{n} \varepsilon_i (x^i)^2$, avec $\varepsilon_i = +1, -1$ ou 0. Pour un espace euclidien où la forme Q est définie, positive, il existe donc des bases telles que :

$$Q(x) = \sum_{i=1}^{n} (x^i)^2.$$

Une telle base caractérisée par le fait que *les éléments de la base ont tous pour norme* 1, *et sont deux à deux orthogonaux,* est dite base *orthonormée.*

Par rapport à elle, le produit scalaire a pour expression :

$$(x|y) = \sum_{i=1}^{n} x^i y^i.$$

On a en particulier $(x|a_i) = x^i$.

La $i^{\text{ème}}$ *coordonnée d'un vecteur est donc le produit scalaire de ce vecteur par le* $i^{\text{ème}}$ *vecteur de la base.*

Remarque : Il faut bien noter que le fait, pour une base, d'être orthonormée n'est pas une propriété intrinsèque pour la structure d'espace vectoriel, mais une propriété relative à la forme quadratique fondamentale choisie. Inversement, choisir une base quelconque d'un espace vectoriel et la déclarer orthonormée revient à choisir une forme quadratique. En effet, $< a_i, g(a_j) >$ est connu (sa valeur est δ_{ij}) pour tout couple (i, j) ; donc, $f(a_i, a_j)$ l'est aussi et par conséquent $f(x, y)$ aussi pour tout couple (x, y).

Plus généralement, on peut considérer une *famille orthonormée,* formée de vecteurs a_i, de norme 1, deux à deux orthogonaux. Une telle famille est une partie libre, car :

$$V = \sum_{i=1}^{n} \lambda^i a_i = 0 \implies \forall k \quad (a_k|V) = 0 \implies \forall k \quad \lambda^k = 0.$$

Il suffit donc qu'une telle famille ait un nombre d'éléments égal à la dimension de l'espace pour qu'elle en constitue une base.

Exercice 50. — On considère l'espace vectoriel V des solutions de l'équation différentielle

$$y'' + 2\,ay' + (b^2 + a^2)\,y = 0$$

(a et b constantes données ; $a > 0$).
Déterminer une base B de V.
On définit un produit scalaire dans V par

$$\{ y_1, y_2 \} = \frac{b}{\pi} \int_0^{\frac{2\pi}{b}} e^{2ax} y_1 y_2 \, dx$$

Calculer son expression en fonction des coordonnées de y_1 et y_2 dans B.
Calculer $\|y\|$.
Déterminer une base orthonormée de V.
Définir une suite de Cauchy. Montrer qu'elle est convergente.

Reprendre les mêmes questions avec le produit scalaire suivant :

$$[y_1, y_2] = \int_0^{\infty} y_1 y_2 dx.$$

§ 5. ISOMETRIES DE L'ESPACE EUCLIDIEN

1. Toute isométrie de l'espace euclidien est une transformation affine.

Nous venons de voir que si T est un espace vectoriel de dimension n muni d'une forme quadratique définie positive Q, on peut lui donner une structure d'espace vectoriel normé, et, à l'espace affine associé E, une structure d'espace métrique.

Cherchons les isométries de cet espace métrique, c'est-à-dire les applications de E dans E qui conservent la distance.

Nous avons vu que les translations de l'espace affine E étaient des isométries ; il en résulte qu'on peut composer une isométrie avec une translation de manière à obtenir une isométrie laissant invariant un point O de l'espace affine, ou que toute isométrie est le produit d'une translation et d'une isométrie laissant un point invariant.

Considérons alors l'espace vectoriel T_0 ayant ce point O pour origine et une isométrie u de E laissant O invariant. u peut être considéré comme application de T_0 dans lui-même, donc aussi de T dans lui-même, et nous sommes ramenés à chercher les applications u de T dans lui-même qui conservent la distance donnée ici par $d(x, y) = \sqrt{Q(x - y)}$, c'est-à-dire les applications u qui vérifient :

$$Q(x - y) = Q(u(x) - u(y))$$

et, à cause de l'invariance de O :

$$Q(x) = Q(u(x)).$$

Mais si f est la forme polaire de Q, les produits scalaires relatifs à Q sont donnés par :

$$(x|y) = f(x, y) = \frac{Q(x) + Q(y) - Q(x - y)}{2}$$

et $(u(x)|u(y)) = f(u(x), u(y)) = \dfrac{Q(u(x) + Q(u(y)) - Q(u(x) - u(y))}{2}$

D'où : $\quad (x|y) = (u(x)|u(y)).$

Le produit scalaire est donc conservé par u.

Il en résulte que si (a_i) est une base orthonormée de T, les vecteurs $u(a_i)$ sont, d'une part, de norme 1, et d'autre part, deux à deux orthogonaux. Ils sont au nombre de n, ils constituent donc une base orthonormée de T.

Mais si x est un vecteur de T, on a :

$$x = \sum_{i=1}^{n} x^i a_i \qquad \text{avec} \quad x^i = f(x, a_i).$$

On a par suite : $\quad f(u(x), u(a_i)) = x^i$

et : $\quad u(x) = \sum_{i=1}^{n} x^i u(a_i).$

u est donc un automorphisme de l'espace vectoriel T, qui transforme toute base orthonormée en base orthonormée.

Autrement dit, toutes les isométries de E sont des bijections affines particulières de E sur lui-même.

2. Transformations orthogonales. Groupe orthogonal.

Cherchons à caractériser de manière générale les automorphismes de T qui laissent Q invariante. Nous avons vu qu'ils conservaient f. La réciproque est évidente. Il est donc équivalent de chercher les automorphismes u de T tels que :

$$f(u(x), u(y)) = f(x, y).$$

Notons que ce sont les automorphismes dont on a parlé en (IX, 3, 3).

Si g est l'application linéaire de T dans T* associée à f, on a, en faisant intervenir la transposée de u :

$$f(u(x), u(y)) = \langle u(x), g \circ u(y) \rangle = \langle x, {}^t u \circ g \circ u(y) \rangle.$$

La condition que u doit vérifier est donc :

$$\forall (x, y) \in E \times E \qquad \langle x, g(y) \rangle = x, {}^t u \circ g \circ u(y) \rangle$$

qui est équivalente à :

$$g = {}^t u \circ g \circ u.$$

Un automorphisme de T, jouissant de cette popriété, est dit *transformation orthogonale* de T (par rapport à Q, ou f, ou g).

Si on a choisi une base dans T et la base duale dans T* et si, par rapport à ces bases, les applications u, g, ${}^t u$ ont respectivement pour matrices O, G, ^{t}O, celles-ci vérifient :

$$G = {}^t O . G . O. \tag{1}$$

Mais, si E est rapportée à une base orthonormée, G est la matrice unité I, et l'égalité précédente donne :

$${}^t O . O = I$$

ou $\quad {}^t O = O^{-1} \qquad (2)$

Les transformations orthogonales peuvent donc être caractérisées par l'égalité (1) relative à leurs matrices ou par l'égalité (2) si l'espace est rapporté à une base orthonormée (*). Dans ce dernier cas, leurs coefficients satisfont à :

$$\forall i \quad \sum_{j=1}^{n} (\alpha_i^j)^2 = 1 \qquad \forall (i, h) \quad \sum_{j=1}^{n} \alpha_i^j \alpha_h^j = 0.$$

Ces relations, nécessaires et suffisantes pour que le produit de la matrice O par sa transposée soit la matrice unité, expriment d'ailleurs

(*) L'égalité (2) est encore vraie si la base est orthogonale sans être orthonormée car alors G est diagonale et la multiplication par une matrice diagonale étant commutative $G = {}^t O.G.O$. implique encore $G = {}^t O.O.G$ qui implique ${}^t O.O = I$ car une matrice diagonale ne peut être diviseur de zéro de l'anneau des matrices carrées de son ordre.

que les vecteurs $u(a_i)$ de coordonnées respectives α_i^j forment une base orthonormée.

Les transformations orthogonales forment évidemment un sous-groupe du groupe, linéaire, appelé *groupe orthogonal*. On le désigne par $O(\mathbf{R}^n)$ si E est un espace de dimension n sur $\mathbf{R}$, donc isomorphe à $\mathbf{R}^n$.

3. Transformations orthogonales positives et négatives.

La relation $g = {}^t u \circ g \circ u$ entraîne, dans tous les cas, que le déterminant de u ait pour carré 1. On peut distinguer deux classes parmi les transformations orthogonales suivant le signe de leur déterminant :

— celles, dont le déterminant vaut 1, sont dites positives ou directes ;
— celles, dont le déterminant vaut — 1, sont dites négatives ou inverses, ou indirectes.

Les transformations directes forment un sous-groupe invariant du groupe orthogonal.

Si nous considérons alors tous les sous-ensembles de $\mathbf{R}^n$ images d'un même sous-ensemble dans une transformation orthogonale (toutes les figures isométriques d'une même figure dans le langage de la géométrie élémentaire), leur ensemble est partagé en deux classes, et ils appartiennent à l'une ou à l'autre suivant que la transformation orthogonale qui les a fournis est directe ou inverse. On passe par une transformation orthogonale directe de l'un à l'autre de deux ensembles quelconques de la même classe. La propriété pour deux tels ensembles de pouvoir se déduire l'un de l'autre dans une transformation orthogonale ou isométrie directe est donc une relation d'équivalence (elle est indépendante de l'ensemble à partir duquel on a fait la classification). C'est cette propriété que l'on exprime en disant que les deux sous-ensembles ont *même orientation*.

Le fait, pour deux ensembles isométriques, d'appartenir respectivement à chacune des deux classes, donc d'être homologues dans une isométrie inverse (également indépendante du point de départ de la classification), se traduira en disant qu'ils ont des *orientations contraires* (*).

4. Forme canonique des transformations orthogonales.

Nous allons étudier maintenant les valeurs propres et vecteurs propres d'une transformation orthogonale et nous verrons si on peut en déduire une simplification des matrices orthogonales.

Valeurs propres. Ce sont les racines du polynôme caractéristique. Nous ne nous occuperons d'abord que des racines réelles, les seules qui ont directement une signification géométrique.

Ces racines réelles ne peuvent être égales qu'à ± 1. Supposons, en effet, que u admette un vecteur propre x de valeur propre λ :

$$f(u(x), u(x)) = f(x, x)$$

(*) Cette notion d'orientation peut d'ailleurs être étendue à des sous-ensembles qui ne sont pas isométriques mais qui sont affinement isomorphes. Si deux sous-ensembles se correspondent dans une transformation affine de rang égal à la dimension n de l'espace, on pourra dire qu'ils ont même orientation ou les orientations contraires suivant que le déterminant de la matrice de l'application linéaire associé à la transformation affine est positif ou négatif.

En particulier l'image d'une base de l'espace vectoriel par une application de rang n est une base. Et, par une transformation affine de rang n l'image de l'ensemble de n vecteurs formant base de l'espace vectoriel d'origine a est un ensemble de n vecteurs formant base de l'espace vectoriel d'origine l'image de a. Ces deux bases pourront avoir même orientation ou des orientations contraires. Dans le cas $n = 3$ on retrouve la notion de sens des trièdres.

mais : $\quad u(x) = \lambda x$, donc :

$$f(u(x), u(x)) = f(\lambda x, \lambda x) = \lambda^2 f(x, x)$$

d'où : $\quad \lambda^2 = 1.$

D'autre part, remarquons que dans $\mathbf{R}^n$, avec n impair, le polynôme caractéristique étant de degré impair, on trouvera au moins une valeur propre réelle.

Quand on aura trouvé une valeur propre réelle, on cherchera les vecteurs propres correspondants. Nous renvoyons à l'exercice 51 pour voir comment l'utilisation de vecteurs propres permet de simplifier les formes des matrices orthogonales.

Exemple : Transformations orthogonales dans $\mathbf{R}^2$.

Supposons $\mathbf{R}^2$ rapporté à une base orthonormée l_1, l_2. Les coefficients d'une matrice orthogonale satisfont à :

$$(\alpha_1^1)^2 + (\alpha_1^2)^2 = 1 \qquad (\alpha_2^1)^2 + (\alpha_2^2)^2 = 1$$

on peut donc toujours poser :

$$\alpha_1^1 = \cos\alpha \qquad \alpha_2^1 = \cos\beta$$
$$\alpha_1^2 = \sin\alpha \qquad \alpha_2^2 = \sin\beta$$

Mais ils doivent en outre vérifier :

$$\cos\alpha\ \cos\beta + \sin\alpha\ \sin\beta = 0,$$

soit

$$\alpha = \beta + \frac{\pi}{2} + k\pi \qquad (k \in \mathbf{Z})$$

D'où les deux matrices :

$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ de déterminant 1 (isométrie positive, rotation autour de l'origine) ;

$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$ de déterminant — 1 (isométrie négative).

Une valeur propre de la première satisferait à :

$$(\lambda - \cos\alpha)^2 + \sin^2\alpha = 0.$$

Il n'en existe donc pas de réelle sauf dans le cas $\sin\alpha = 0$, cas qui se subdivise en :

$\sin\alpha = 0 \quad \cos\alpha = 1$	$\sin\alpha = 0 \quad \cos\alpha = -1$
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
identité	demi-tour autour de l'origine

Dans ces deux transformations, tous les vecteurs sont propres.

Quant à la matrice de déterminant — 1, elle admet pour valeur propre les racines de :

$$(\cos\alpha - \lambda)(-\cos\alpha - \lambda) - \sin^2\alpha = 0,$$

soit

$$\lambda = \pm 1.$$

Les vecteurs propres correspondants, définis par leurs coordonnées dans la base $l_1\, l_2$ ($x = x^1 l_1 + x^2 l_2$, $x' = x'^1 l_1 + x'^2 l_2$) vérifient :

pour $\lambda = 1$ $\qquad x^1(\cos\alpha - 1) + x^2 \sin\alpha = 0,$

d'où $\qquad x^1 = \cos\dfrac{\alpha}{2} \qquad x^2 = \sin\dfrac{\alpha}{2}$;

pour $\lambda = -1$ $\qquad x'^1(\cos\alpha + 1) + x'^2 \sin\alpha = 0,$

d'où $\qquad x'^1 = -\sin\frac{\alpha}{2}$,

$$x'^2 = \cos\frac{\alpha}{2}.$$

Ces vecteurs x et x' constituent une base orthonormée par rapport à laquelle la matrice s'écrit :

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

C'est la matrice d'une symétrie par rapport au support du vecteur x'. Les rotations de centre O et les symétries par rapport à des droites passant par O sont les seules transformations orthogonales de $\mathbf{R}^2$, respectivement directes et inverses. Mais si on classe les transformations du point de vue de leurs valeurs propres et de leurs vecteurs propres, nous devons mettre à part les deux cas particuliers signalés et nous avons donc les quatre types de transformations dont les matrices admettent les quatre formes canoniques :

$$\underbrace{\underset{\alpha \neq k\pi}{\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{positives}} \quad \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{négative}}$$

Quant aux transformations orthogonales de l'espace affine, elles s'obtiennent comme produit d'une translation par une des transformations précédentes. La géométrie élémentaire apprend quelles formes canoniques on peut donner à ces produits.

Exercice 51.

a) Dans une transformation orthogonale, quand un vecteur est propre, le sous-espace orthogonal est invariant. Plus généralement, quand un sous-espace est invariant, son orthogonal est aussi invariant.

b) Deux vecteurs propres correspondant à des valeurs propres différentes sont orthogonaux.

c) Déduire de ces propriétés une forme simple que l'on peut donner à une matrice orthogonale. Montrer, en particulier, que si la transformation a toutes ses valeurs propres réelles on peut la diagonaliser par rapport à une base orthonormée ; que dans le cas contraire, on peut, toujours par rapport à une base orthonormée, la mettre sous la forme

$$\begin{pmatrix} (M) & O \\ O & (M') \end{pmatrix}$$

M représentant une matrice diagonale et M' une matrice orthogonale à valeurs propres imaginaires.

d) Caractériser l'ensemble des vecteurs propres de la transformation.

5. Valeurs propres et vecteurs propres imaginaires.

E étant un espace de dimension n sur $\mathbf{R}$ et $(a_1\, a_2 \dots a_n)$ en étant une base, rappelons que E est isomorphe à $\mathbf{R}^n$ par l'isomorphisme :

$$x \longleftrightarrow (x^1 \dots x^n)$$

si $x^1 \dots x^n$ sont les coordonnées de $x \in$ E dans la base $(a_1 \dots a_n)$. $\mathbf{R}^n$ est un sous-ensemble de $\mathbf{C}^n$; E peut donc être considéré comme un sous-ensemble d'un espace F de dimensions n sur $\mathbf{C}$ ayant $(a_1 \dots a_n)$ pour base.

Si f est une forme bilinéaire associée à une forme quadratique, définie positive, elle a pour valeur :

$$f(x, y) = \Sigma\, x^i y^j$$

si la base $(a_1 \dots a_n)$ a été choisie orthonormée par rapport à f. La même expression définit une forme bilinéaire sur F et on peut encore considérer la forme quadratique $f(x, x)$, mais celle-ci peut prendre des valeurs négatives, ou des valeurs complexes pour des x appartenant à F, et dont toutes les coordonnées ne sont pas réelles.

Soit alors une transformation orthogonale u de E. On peut considérer l'automorphisme $\tilde{u}$ de F défini par la même matrice par rapport à la base $(a_1 \dots a_n)$. Cet automorphisme vérifie encore dans F :

$$f(x, y) = f(\tilde{u}(x), \tilde{u}(y))$$

puisque la matrice de $\tilde{u}$ continue à vérifier ${}^t\mathrm{O} = \mathrm{O}^{-1}$ par rapport à la base qui est orthonormée.

Cet automorphisme admet n valeurs propres distinctes ou non. Soit λ une de ces valeurs propres supposée imaginaire. Il lui correspond au moins un vecteur propre x (certainement imaginaire car si x était réel la matrice de $\tilde{u}$ étant à coefficients réels $\tilde{u}(x)$ le serait aussi) et on a toujours :

$$f(x, x) = \lambda^2 f(x, x).$$

Mais, cette fois, $\lambda^2 \neq 1$ et cette égalité entraîne :

$$f(x, x) = 0.$$

Tout *vecteur propre imaginaire est orthogonal à lui-même*. Un tel vecteur orthogonal à lui-même est qualifié de *vecteur isotrope*.

A deux valeurs propres imaginaires conjuguées correspondront toujours des vecteurs imaginaires conjugués, c'est-à-dire de coordonnées imaginaires conjuguées par rapport à la base. Le sous-espace (sur $\mathbf{C}$) engendré par ces deux vecteurs est invariant par la transformation $\tilde{u}$.

Exercice 52. — Montrer que ce sous-espace contient un sous-ensemble qui est un sous-espace vectoriel sur $\mathbf{R}$.

En déduire, si u est une transformation orthogonale sans valeurs propres réelles, que $\mathbf{R}^{2p}$ est somme directe de p sous-espaces E_i de dimension 2, deux à deux orthogonaux et que u est le produit de p transformations u_i induisant une rotation dans E_i et laissant invariant $\bigoplus_{j \neq i} E_j$. En déduire la forme la plus simple que l'on puisse donner à une matrice orthogonale dans ce cas et dans le cas général.

Exercice 53. — Caractériser tous les types de transformations orthogonales de $\mathbf{R}^3$ et $\mathbf{R}^4$. (Forme des matrices ; vecteurs propres ; sous-espaces invariants).

Exercice 54. — E est un espace vectoriel sur le corps des complexes $\mathbf{C}$; il est donc aussi espace vectoriel sur le corps des réels $\mathbf{R}$. On précisera celle des deux structures précédentes qui est considérée pour E en la désignant par E($\mathbf{C}$) (espace vectoriel sur $\mathbf{C}$) ou E($\mathbf{R}$) (espace vectoriel sur $\mathbf{R}$).

1) Une base de E($\mathbf{C}$) étant donnée, en déduire une base de E($\mathbf{R}$).
Si dim E($\mathbf{C}$) $= n$, que vaut dim E($\mathbf{R}$) ?

2) Indiquer la forme des applications linéaires de $\mathbf{C}$ dans $\mathbf{C}$ et de $\mathbf{R}^2$ dans $\mathbf{R}^2$.

3) Montrer que si f est une forme linéaire sur E($\mathbf{C}$) c'est une application linéaire de E($\mathbf{R}$) dans $\mathbf{C}$.

Montrer que l'inclusion $\mathcal{L}(\mathrm{E}(\mathbf{C}), \mathbf{C}) \subset \mathcal{L}(\mathrm{E}(\mathbf{R}), \mathbf{C})$ est stricte.

4) Soit f une forme linéaire sur E($\mathbf{C}$). Pour tout x de E on désigne par $\varphi(x)$ et $\psi(x)$ respectivement les parties réelles et imaginaires de $f(x)$.

Montrer que les applications φ et ψ ainsi définies appartiennent à $\mathcal{L}(\mathrm{E}(\mathbf{R}), \mathbf{R})$ et que l'on a

$$\varphi(ix) = -\psi(x)$$

5) Montrer que, réciproquement, à tout élément φ de $\mathcal{L}$ (E(**R**), **R**) correspond un élément f unique de $\mathcal{L}$ (E(**C**), **C**) tel que $\varphi(x)$ soit la partie réelle de $f(x)$.

Quelle conséquence en déduit-on pour les deux espaces

$\mathcal{L}$(E(**C**), **C**) et $\mathcal{L}$ (E(**R**), **R**) ?

6) Soit E un espace vectoriel sur **R** ; montrer que la condition nécessaire et suffisante pour qu'il puisse recevoir une structure d'espace vectoriel sur **C** est qu'il existe un isomorphisme u de E sur E (pour sa structure E(**R**) d'espace vectoriel sur **R**) tel que pour tout x de E, $u \circ u(x) = -x$.

Exercice 55. — Etant donné un espace vectoriel de E sur **R**, normé à l'aide d'une forme définie positive f et étant donné sur cet espace une forme bilinéaire symétrique φ, on veut montrer qu'il est possible de trouver une base orthonormée de E par rapport à laquelle la matrice de φ soit diagonalisée. Dans ce but on montrera :

1) pour que l'hyperplan orthogonal à un vecteur soit confondu avec l'hyperplan conjugué de ce vecteur par rapport à φ il faut et il suffit que ce vecteur soit vecteur propre de l'application linéaire de E dans E, $g^{-1} \circ \psi$, g et ψ représentant les applications de E dans E* respectivement associées à f et à φ.
2) cette application $g^{-1} \circ \psi$ a toutes ses valeurs propres réelles.

§ 6. FORMES SESQUILINEAIRES ET HERMITIENNES

1. Définition.

Dans ce qui précède, nous avons fait jouer à $\mathbf{C}^n$ le rôle d'un intermédiaire. Considérons maintenant $\mathbf{C}^n$ en lui-même, ou tout autre espace vectoriel sur **C**, soit E. Quelle que soit la forme bilinéaire $f(x, y)$ définie sur lui, la forme $f(x, x)$ ne possède plus la propriété fondamentale qui a permis de définir une norme à partir d'elle.

Par exemple, dans $\mathbf{C}^n$, la forme quadratique :

$$\sum_{i=1}^{n} (x^i)^2$$

s'annule pour des vecteurs x non nuls.

Mais si x_i et $\bar{x}_i$ sont des nombres imaginaires conjugués,

$$\Sigma\, x_i \bar{x}_i = \Sigma |x_i|^2$$

est un nombre réel positif qui ne s'annule que si tous les x_i sont nuls. Ceci amène à considérer l'application :

$$f : \mathrm{E}^2 \longrightarrow \mathbf{C}$$

qui au couple (x, y) d'éléments de E fait correspondre :

$$f(x, y) = \sum_{i,j} \alpha_{ij}\, x^i \bar{y}^j,$$

les α_{ij} étant des coefficients qui satisfont à :

$$\forall (i, j) \qquad \alpha^{ij} = \bar{\alpha}_{ji}$$

(ce qui entraîne que les α_{ii} soient réels).

Une telle application est linéaire par rapport à x. Par rapport à y, elle jouit de propriétés voisines de la linéarité :

$$f(x, y_1) + f(x, y_2) = f(x, y_1 + y_2) \qquad\qquad (1).$$
$$f(x, \lambda y) = \bar{\lambda} f(x, y)$$

Elle est telle aussi que :

$$\forall (x, y) \in \mathrm{E}^2 \qquad f(x, y) = \overline{f(y, x)} \qquad\qquad (2).$$

Ceci nous amène à poser les définitions suivantes.

On appelle *forme semi-linéaire* sur E une application φ de E dans **C** telle que :

$$\varphi(y_1 + y_2) = \varphi(y_1) + \varphi(y_2)$$
$$\varphi(\lambda y) = \bar{\lambda}\varphi(y)$$

et *forme sesquilinéaire* une application f de E^2 dans **C**, linéaire par rapport à une des variables et semi-linéaire par rapport à l'autre.

On considère parmi ces formes celles qui jouissent de la propriété (2) (symétrie hermitienne).

Remarquons d'ailleurs que si une application de E^2 dans **C** est linéaire par rapport à x et jouit de la symétrie hermitienne, elle est semi-linéaire par rapport à y, donc sesquilinéaire. En effet :

$$f(x, y_1 + y_2) = \overline{f(y_1 + y_2, x)} = \overline{f(y_1, x)} + \overline{f(y_2, x)}$$
$$= f(x_1, y_1) + f(x_1, y_2)$$
$$f(x, \lambda y) = \overline{f(\lambda y, x)} = \bar{\lambda}\overline{f(y, x)} = \bar{\lambda} f(x, y).$$

Considérons donc une application f de E^2 dans **C** linéaire par rapport à x et jouissant de la symétrie hermitienne, et cherchons quelle forme elle doit avoir pour satisfaire à ces hypothèses.

Si E est rapporté à une base (a_i) et si $x = \Sigma\, x^i a_i$, $y = \Sigma\, y^i a_i$, la sesquilinéarité permet d'écrire :

$$f(x, y) = \sum_{i,j} x^i \bar{y}^j f(a_i, a_j) = \sum_{i,j} \alpha_{ij}\, x^i \bar{y}^j$$

en posant :

$$\alpha_{ij} = f(a_i, a_j).$$

Mais, d'après l'hypothèse de symétrie hermitienne,

$$\alpha_{ji} = \bar{\alpha}_{ij},$$

cherchons la valeur de la forme pour $x = y$.

$$f(x, x) = \sum_{i,j} (\alpha_{ij}\, x^i \bar{x}^j + \bar{\alpha}_{ij}\, \bar{x}^i x^j) + \sum_i \alpha_{ii}\, x^i \bar{x}^i$$

est un nombre réel. L'application :

$$\mathrm{H} : x \longrightarrow f(x, x)$$

est donc une application de E dans **R** qui satisfait à :

$$f(\lambda x, \lambda x) = \lambda \bar{\lambda} f(x, x) = |\lambda|^2 f(x, x).$$

Cette application est appelée une *forme hermitienne*. Une forme hermitienne est donc une application de E dans **R**, restriction à la diagonale d'une application sesquilinéaire possédant la symétrie hermitienne.

Dans le cas où E est de dimension finie, les coefficients α_{ij} peuvent être rangés dans une matrice dite *matrice hermitienne.*

Remarquons que, de la connaissance de H, on peut déduire celle de la forme sesquilinéaire dont elle est issue ; on a, en effet :

$$f(x, y) = \frac{\mathrm{H}(x + y) - \mathrm{H}(x - y)}{4} + i\, \frac{\mathrm{H}(x + iy) - \mathrm{H}(x - iy)}{4}.$$

2. Applications linéaires associées à une forme hermitienne.

Nous avons associé à toute forme bilinéaire et, par suite, à toute forme quadratique, deux applications de E dans E* qui se confondaient dans le cas où f était symétrique, et avaient alors pour matrice par rapport à une base de E et à la base duale de E* précisément la matrice de la forme f. Nous pouvons faire une étude analogue ici, un peu plus compliquée du fait que f est sesquilinéaire. Nous pourrons cependant nous ramener à des applications linéaires de la manière suivante :

Si E est un espace vectoriel sur **C** et si λx désigne le composé par l'opération externe de $\lambda \in \mathbf{C}$ et de $x \in$ E, nous pouvons définir sur E une autre opération externe dont le composé sera noté $\lambda * x$, en posant :

$$\lambda * x = \bar{\lambda} x.$$

Cette nouvelle opération, associée à la même loi de groupe commutatif sur E, lui confère, comme il est aisé de le vérifier, une structure d'espace vectoriel distincte de la précédente ; nous désignerons par $\bar{\mathrm{E}}$ l'ensemble E muni de cette structure. Il faut bien remarquer que E et $\bar{\mathrm{E}}$ désignent le même ensemble mais muni de deux structures différentes. Il est clair, en vertu de la relation très simple entre les deux structures, que toute base de E est une base de $\bar{\mathrm{E}}$.

Si u est une application de E dans un espace vectoriel F sur **C**, c'est aussi une application de $\bar{\mathrm{E}}$ dans F, mais si u est linéaire pour E, elle est semilinéaire pour F et réciproquement. En effet, l'additivité est conservée et si :

$$u(\lambda x) = \lambda\, u(x) \qquad \text{(linéarité sur E),}$$

on aura : $\quad u(\lambda * x) = u(\bar{\lambda} x) = \bar{\lambda}\, u(x) \quad$ (semilinéarité sur $\bar{\mathrm{E}}$).

Si, au contraire : $\quad u(\lambda * x) = \lambda\, u(x) \quad$ (linéarité sur $\bar{\mathrm{E}}$),

on aura : $\quad u(\lambda x) = u(\bar{\lambda} * x) = \bar{\lambda}\, u(x) \quad$ (semilinéarité sur E).

Le dual de E est l'espace des formes linéaires sur E ; le dual $\bar{\mathrm{E}}^*$ de $\bar{\mathrm{E}}$ est l'espace des formes linéaires sur $\bar{\mathrm{E}}$ et semilinéaires sur E, étant bien entendu que l'opération externe sur $\bar{\mathrm{E}}^*$ est définie pour :

$$\lambda \in \mathbf{C} \qquad u \in \bar{\mathrm{E}}^* \qquad x \in \bar{\mathrm{E}} \text{ par } (\lambda u)(x) = \lambda u(x),$$

où $\lambda u(x)$ est le produit dans **C** des nombres complexes λ et $u(x)$.

Ceci étant, à partir de la forme sesquilinéaire f, de façon analogue à ce que nous avons fait en (IX, 2, 1 et 2), nous pourrons définir pour tout y de E (ou $\bar{\mathrm{E}}$ puisque c'est le même ensemble), l'application g_y définie par :

$$\forall x \in \mathrm{E} \qquad g_y(x) = f(x, y) \ ;$$

f étant linéaire en x, g_y appartient à E*, et on peut écrire :

$$f(x, y) = \langle x, g_y \rangle,$$

où $\langle x, x^* \rangle$ désigne la valeur de la forme bilinéaire canonique sur E $\times$ E*.

On a ainsi défini une application de E (ou $\bar{\mathrm{E}}$) dans E* qui envoie y sur g_y. Désignons cette application par g, et remplaçons la notation g_y par la notation $g(y)$.

L'additivité de f en y entraîne l'additivité de g. D'autre part, $f(x, \lambda y) = \lambda f(x, y)$ s'exprime de manières équivalentes par :

$$g(\lambda y) = \bar{\lambda} g(y) \qquad \text{ou} \qquad g(\lambda * y) = \lambda g(y),$$

c'est-à-dire que g est une application semilinéaire de $\bar{\mathrm{E}}$ dans E* et une application linéaire de E dans E*.

C'est ce dernier point de vue que nous allons adopter. La transposée ${}^t g$ de g est une application de E** dans $\bar{\mathrm{E}}^*$. Si E est de dimension finie, on a E** = E par l'application canonique. Dans le cas où la dimension de E est infinie, on peut restreindre ${}^t g$ au sous-espace $\tilde{\mathrm{E}}$ de E** canoniquement isomorphe à E. Dans les deux cas, on aura donc :

$$\forall x \in \mathrm{E} \qquad \forall y \in \bar{\mathrm{E}} \qquad f(x, y) = \langle x, g(y) \rangle = \langle {}^t g(x), y \rangle.$$

La symétrie hermitienne de f exige que l'on ait :

$$f(x, y) = \overline{\langle g(x), y \rangle}.$$

Considérons alors l'application de $\bar{\mathrm{E}}$ dans **C** :

$$\gamma_x : y \longrightarrow \overline{\langle g(x), y \rangle}.$$

Elle est linéaire car elle est additive et :

$$\overline{\langle g(x), \lambda * y \rangle} = \overline{\langle g(x), \bar{\lambda} y \rangle} = \overline{\bar{\lambda} \langle g(x), y \rangle} = \lambda\, \overline{\langle g(x), y \rangle}.$$

C'est donc un élément de $\bar{\mathrm{E}}^*$.

Cette application linéaire γ_x est l'image de x par une application γ de E dans $\bar{\mathrm{E}}^*$. Nous noterons dorénavant γ_x par $\gamma(x)$ et on peut écrire :

$$\langle \gamma(x), y \rangle = \overline{\langle g(x), y \rangle} = \langle x, g(y) \rangle.$$

On vérifie immédiatement que γ est une application linéaire. De l'égalité $\langle \gamma(x), y \rangle = \langle {}^t g(x), y \rangle$ vraie pour tout couple (x, y) d'éléments de E, on déduit $\gamma = {}^t g$, égalité qui généralise l'égalité analogue $(g = {}^t g)$ que nous avons trouvée dans l'étude des formes bilinéaires symétriques. Partant de cette égalité, nous allons retrouver la condition que doit remplir la matrice d'une forme hermitienne, définie sur un espace de dimension finie. Supposons alors que l'on ait choisi une base (a_i) dans E ; nous rapporterons $\bar{\mathrm{E}}$ à la même base, E* et $\bar{\mathrm{E}}^*$ aux bases duales correspondantes, (a^i) dans E* et (b^i) dans $\bar{\mathrm{E}}^*$. Il y a lieu de remarquer que, bien que a^i et b^i soient caractérisées par les mêmes égalités :

$$\langle a_i, a^j \rangle = \delta_i^j \qquad \langle a_i, b^j \rangle = \delta_i^j,$$

elles sont différentes, car les a^j sont des formes linéaires sur E et les b^j des formes linéaires sur $\bar{\mathrm{E}}$ ou semi-linéaires sur E. Si x est un élément de E et si ses coordonnées dans E sont x^i, dans $\bar{\mathrm{E}}$ elles sont $\bar{x}^i$. On a, en effet :

$$x = \Sigma\, x^i a_i = \Sigma\, \bar{x}^i * a_i,$$

d'où $\qquad a^j(x) = x^j \qquad b^j(x) = \bar{x}^j.$

Comparons alors les matrices de g (application de $\bar{\mathrm{E}}$ dans E*) et de γ (application de E dans $\bar{\mathrm{E}}^*$) par rapport à ces bases. On a :

$$g(a_i) = \Sigma\, \alpha_{ij}\, a^j \qquad \gamma(a_i) = \Sigma\, \beta_{ij}\, b^j.$$

On doit avoir :

$$\overline{\langle g(a_i), a_k \rangle} = \langle \gamma(a_i), a_k \rangle,$$

d'où $\quad \bar{\alpha}_{ik} = \beta_{ik}$.

De $\gamma = {}^t g$, on déduit donc pour la matrice G de g la condition :

$$\forall (i, k) \qquad \bar{\alpha}_{ik} = \alpha_{ki},$$

ce que l'on écrit symboliquement : $\quad {}^t\mathrm{G} = \bar{\mathrm{G}}$.

La condition nécessaire et suffisante pour que G *soit la matrice d'une forme hermitienne est que sa transposée soit égale à son imaginaire conjuguée.*

La théorie de la diagonalisation de la matrice G se déroule parallèlement à celle des matrices des formes quadratiques sur **R** et conduit à des résultats analogues :

Les coefficients diagonaux sont réels, les nombres de coefficients positifs, négatifs et nuls sont liés à la forme hermitienne, qui est dite définie positive si le nombre des coefficients positifs est égal à la dimension de E. Dans ce cas, on a :

$$\forall x \in \mathrm{E} \qquad \mathrm{H}(x) \geqslant 0 \quad \text{et} \quad \mathrm{H}(x) = 0 \iff x = 0.$$

Cette dernière propriété caractérise par définition les formes hermitiennes définies positives lorsque la dimension de E est infinie.

Exercice 56. — H étant une forme hermitienne définie positive et f la forme sesquilinéaire associée, montrer que ($\mathcal{R}z$ désignant la partie réelle du nombre complexe z)

$$[\mathcal{R}f\,(x, y)]^2 \leqslant (f(x, x)\,f\,(y, y) \qquad (1)$$

En déduire que $\sqrt{\mathrm{H}\,(x)} = \sqrt{f(x,\, x)}$ est une norme sur E. Montrer qu'on peut aussi déduire de (1) que

$$|(f(x, y)|^2 < f(x, x)\; f(y, y)$$

3. Applications unitaires.

A la notion d'application orthogonale correspond ici celle d'application unitaire. On désigne ainsi un endomorphisme de E tel que :

$$\forall (x, y) \in \mathrm{E}^2 \qquad f\,(x, y) = f(u(x),\ u(y))$$

f étant la forme sesquilinéaire associée à une forme hermitienne définie positive.

Pour en démontrer l'existence et en faire l'étude, remarquons d'abord qu'une application de E dans E, qui est un endomorphisme pour une des deux structures d'espaces vectoriels sur C que nous y avons définies, l'est aussi pour l'autre. En effet, de $u(\lambda x) = \lambda\, u(x)$, on déduit :

$$u\,(\lambda * x) = u\,(\bar{\lambda}\, x) = \bar{\lambda}\, u(x) = \lambda * u(x)$$

et réciproquement.

Cependant, la matrice de u sera différente suivant que l'on considère une structure ou l'autre. Si U est la matrice de u pour E rapporté à la base (a_i), $\overline{\mathrm{U}}$ sera la matrice pour $\overline{\mathrm{E}}$ rapporté à la même base (en effet, $u(a_i) = \sum_j \alpha_i^j a_j = \sum_j \overline{\alpha_i^j} * a_j$). Les transposées seront également distinctes, l'une applique E^* dans E^*, l'autre $\overline{\mathrm{E}}^*$ dans $\overline{\mathrm{E}}^*$. Pour ces raisons, nous désignerons par u l'endomorphisme de E et par v l'endomorphisme de $\overline{\mathrm{E}}$, qui ne sont en fait qu'une seule et même application.

Dire que u est unitaire s'écrit donc :

$$f\,(x, y) = f(u(x)),\ u(y))$$

ou, en faisant intervenir l'application g et en utilisant la convention de distinguer u et v :

$$< x, g(y) > = < u(x), g\circ v(y) >.$$

En introduisant la transposée :

$$< x, g(y) > = < x, {}^t u\circ g\circ v(y) >.$$

Une condition nécessaire et suffisante pour que u soit unitaire est donc :

$$g = {}^t u\circ g\circ v$$

ce qui donne, dans le cas des espaces de dimension finie, pour les matrices correspondantes :

$$\mathrm{G} = {}^t\mathrm{U}\,\mathrm{G}\,\overline{\mathrm{U}}.$$

En particulier, si G est la matrice unité, on a : $\overline{\mathrm{U}}^{-1} = {}^t\mathrm{U}$ qui est équivalent à $\mathrm{U}^{-1} = {}^t\overline{\mathrm{U}}$.

L'étude de la diagonalisation des matrices unitaires est parallèle à celle des matrices orthogonales et conduit à des résultats analogues.

SOLUTION DES EXERCICES

Exercice 1.

Nous savons qu'on peut faire correspondre à tout $x \in [0, 1]$ un développement décimal illimité et un seul à condition d'éliminer tous les développements qui sont à partir d'un certain rang constitués de 9 et que nous appellerons développements impropres.

Ceci posé, considérons la correspondance $t \longrightarrow (x, y)$ définie comme il est indiqué au début de l'énoncé. Elle est surjective, car, étant donné un couple (x, y) quelconque dont les deux composantes ont un développement propre, il lui correspondra un t défini par un développement propre. Mais est-elle injective ? A deux t et t' distincts donnés par leurs développements propres correspondent deux couples (x, y), (x', y') à développements formellement différents. Toutefois, nous ne pouvons affirmer $(x, y) \neq (x', y')$ que si nous sommes sûrs que les quatre développements sont propres. Or, pour qu'un des quatre développements soit impropre, il suffit (et il faut) que les décimales de rang pair (ou impair) de t ou de t' soient toutes des 9 à partir d'un certain rang.

Si $t = 0{,}23459696999891 \ldots..$ nous aurons :

$$x = 0{,}24999 \ldots.. \qquad y = 0{,}356981 \ldots..$$

Ce point est le même que $x = 0{,}25$, $y = 0{,}356981 \ldots..$, qui serait obtenu à partir de $t' = 0{,}235506090801 \ldots.. \neq t$, et l'application ainsi définie n'est donc pas injective.

En prenant au contraire la deuxième application définie dans l'énoncé, toutes les tranches de décimales considérées commençant par un chiffre différent de 9, nous sommes sûrs que les deux développements de x et de y, provenant d'un développement propre de t, sont eux-mêmes propres. La nouvelle application est encore surjective et, cette fois, elle est injective puisque $t \neq t'$ entraîne que x et x' d'une part, y et y' d'autre part ont des développements propres formellement différents, donc sont différents.

L'existence de cette bijection établit donc que $c^2 = c$.

Cherchons maintenant si l'application $t \longrightarrow (x, y)$ est continue. Soit t_0 et (x_0, y_0) son image. Si t tend vers t_0, peut-on affirmer que x tend vers x_0 et y tend vers y_0 ? Si t_0 n'est pas décimal, c'est-à-dire si son développement propre comprend une infinité de chiffres significatifs, il est clair que la réponse est positive, car, quand t tend vers t_0, le rang n de la première décimale de t différente de la décimale correspondante de t_0 tend vers l'infini, et il en sera de même des rangs des premières décimales différentes de x et x_0, de y et y_0.

Mais si t_0 est une fraction décimale, c'est-à-dire si son développement propre ne comprend que des 0 à partir d'un certain rang, et si t tend vers t_0 par valeurs inférieures, il n'en sera pas ainsi. Soit, par exemple :

$$t_0 = 0{,}83500 \ldots..\ 00 \ldots.. \qquad x_0 = 0{,}85 \quad y_0 = 0{,}3.$$

Considérons les valeurs de t inférieures à t_0 et tendant vers t_0. Elles ont des développements de la forme :

$$t = 0,\overline{8}\ \overline{3}\ \overline{4\ 9\ 9 \ldots\ldots 9}\ \overline{\alpha_1}\ \overline{\alpha_2} \ldots$$

le nombre de 9 qui suit le 4 tendant vers l'infini quand t tend vers t_0. On aura :

$$x = 0,8499 \ldots\ldots 9 \quad \alpha_2 \ldots\ldots$$

qui tendra vers 0,85, mais :

$$y = 0,3\,\alpha_1 \quad \alpha_3 \ldots\ldots$$

qui ne tend pas vers 0,3 (il n'a pas de limite). Il est clair que cette circonstance est générale, la discontinuité portant sur x ou sur y suivant la parité du nombre de tranches de chiffres significatifs de t_0.

La bijection $t \longrightarrow (x, y)$ *est donc discontinue à gauche pour toutes les valeurs de* t *de la forme* $\dfrac{k}{10^n}$ (les valeurs correspondantes de x et y sont elles-mêmes de cette forme).

Inversement, pour que l'application $(x, y) \longrightarrow t$ soit discontinue pour un point (x_0, y_0), il faut que, quand x tend vers x_0 et y tend vers y_0, le rang minimum des décimales de x et y, qui restent différentes de celles de même rang de x_0 ou y_0, reste fini. Ceci ne peut se produire (et se produit effectivement) que si x *ou* y est une fraction décimale,

ex. : $\quad x_0 = 0,84 \quad y_0 = 0,6666 \ldots 6 \ldots \quad t_0 = 0,86460606 \ldots 06 \ldots$

quand, y tendant vers y_0, x tend vers x_0 par valeur inférieure, c'est-à-dire quand x est de la forme : $x = 0,8399 \ldots 9\,\alpha_1\,\alpha_2 \ldots$ avec un nombre de 9 qui tend vers l'infini, on trouve $t = 0,86399 \ldots 96\,\alpha_1\,6\,\alpha_2 \ldots$, qui tend vers $t = 0,864 \neq t_0$.

La bijection $(x, y) \longrightarrow t$ *est discontinue pour tous les points du carré dont une coordonnée est de la forme* $\dfrac{k}{10^n}$.

Remarquons d'ailleurs qu'aucune application continue :

$$f : \mathrm{M}(x, y) \longrightarrow t$$

où M représente un point du carré, ne peut être bijective. Soient en effet

$$\begin{array}{ll} \mathrm{M}_0(x_0, y_0) \text{ et son image } t_0 = f(\mathrm{M}_0) & \\ \mathrm{M}_1(x_1, y_1) \text{ et son image } t_1 = f(\mathrm{M}_1) & \quad x_0 \neq x_1 \end{array}$$

Si $t_0 = t_1$, f n'est pas bijective. Supposons $t_0 \neq t_1$. Considérons deux chemins continus, sans points communs, allant de M_0 à M_1, définis par exemple par deux fonctions continues :

$$\begin{array}{lcl} x \longrightarrow y = \varphi(x) & & t = f(x, \varphi(x)) = \Phi(x) \\ x \longrightarrow y = \psi(x) & \text{on a alors :} & t = f(x, \psi(x)) = \Psi(x) \end{array}$$

Φ et Ψ sont deux fonctions continues de x satisfaisant à :

$$\Phi(x_0) = \Psi(x_0) = t_0,$$
$$\Phi(x_1) = \Psi(x_1) = t_1.$$

Si $t_2 \in\ ]t_0, t_1[$, les propriétés des fonctions continues permettent d'affirmer :

$$\exists\, \mathrm{M}_2(x_2, y_2) \quad x_2 \in\ ]x_0, x_1[\quad \Phi(x_2) = t_2 \quad \text{c'est-à-dire } f(\mathrm{M}_2) = t_2$$
$$\exists\, \mathrm{M}'_2(x'_2, y'_2) \quad x'_2 \in\ ]x_0, x_1[\quad \Psi(x'_2) = t_2 \quad \text{c'est-à-dire } f(\mathrm{M}'_2) = t_2.$$

D'après les hypothèses, $\mathrm{M}_2 \neq \mathrm{M}'_2$ et f n'est pas bijective.

Exercice 2.

On sait qu'à tout irrationnel correspond une suite illimitée de quotients incomplets $q_0, q_1, q_2, \ldots, q_n \ldots$ entiers, tous strictement positifs, sauf q_0 qui peut être positif, négatif ou nul et est justement nul pour un

nombre de $]0, 1[$; qu'inversement une telle suite $q_1, q_2, \ldots, q_n \ldots$ d'entiers strictement positifs définit un irrationnel qui admet cette suite comme suite de ses quotients incomplets. Le procédé du développement en fraction continue définit donc une bijection entre l'ensemble des irrationnels de $[0, 1]$ d'une part et les suites $q_1, q_2, \ldots, q_n \ldots$ d'entiers positifs d'autre part. Or, retirer un ensemble dénombrable d'un ensemble de puissance c ne change pas la puissance de cet ensemble. L'ensemble des irrationnels de $[0, 1]$ a donc pour puissance c et l'ensemble des suites d'entiers a donc aussi pour puissance c. Or, une telle suite est une application :

$$\begin{array}{ccc} \mathbf{N}^* & \longrightarrow & \mathbf{N}^* \\ n & & u_n \end{array}.$$

La puissance de l'ensemble $\mathcal{F}(\mathrm{N}^*, \mathrm{N}^*)$ de ces applications est par définition de l'exponentiation $\aleph_0{}^{\aleph_0}$. D'où :

$$\aleph_0{}^{\aleph_0} = c.$$

Exercice 3.

1) La donnée d'une suite de réels compris entre 0 et 1 est celle d'une application :

$$\begin{array}{ccc} \mathbf{N}^* & \longrightarrow & [0, 1] \\ n & & u_n \end{array}$$

et l'ensemble $\mathcal{F}$ de ces suites a pour puissance $c^{\aleph_0}$. Il reste à montrer que l'application indiquée dans le texte :

$$\begin{array}{ccc} \mathcal{F} & \longrightarrow & [0, 1] \\ \{x_n\} & & x \end{array}$$

est bien une bijection pour établir que $c^{\aleph_0} = c$.

Or, cette application est injective car, si $\{x_n\}$ et $\{x'_n\}$ sont deux suites différentes, il existe au moins un rang n pour lequel $x_n \neq x'_n$; et il existe alors au moins un rang p pour lequel les tranches de décimales correspondantes sont différentes : $\alpha_n^p \neq \alpha'^p_n$, et ces tranches de décimales se retrouveront au même rang dans les développements de x et de x' qui, en conséquence, seront différents. Ces développements étant propres grâce au procédé utilisé (cf. ex. 1), x et x' sont différents :

$$\{x_n\} \neq \{x'_n\} \quad \Longrightarrow \quad x \neq x'.$$

D'autre part, l'application est surjective. En effet, soit :

$$x = 0,\ a_1\ a_2\ a_3 \ldots ;$$

les a représentant des tranches de décimales définies toujours par le même procédé. On posera :

$$\begin{array}{lllll} x_1 = 0, & a_1 & a_2 & a_4 & a_7 \ldots\ldots \\ & & / & / & / \\ x_2 = 0, & a_3 & a_5 & a_8 \ldots\ldots & \\ & & / & / & \\ x_3 = 0, & a_6 & a_9 \ldots\ldots & & \\ & & / & & \\ x_4 = 0, & a_{10} \ldots\ldots & & & \\ \ldots\ldots & & & & \end{array}$$

les décimales de x étant distribuées comme le schéma l'indique entre les différents termes de la suite. La suite ainsi définie a bien x comme image dans l'application et on peut conclure que cette application est bijective.

2) Remarquons qu'une application continue de $[0, 1]$ dans $\mathbf{R}$ est déterminée par sa restriction à l'ensemble des rationnels de $[0, 1]$. En effet, tout réel x est la limite de la suite $\{u_n\}$ de ses valeurs décimales par défaut à 10^{-n} près. Quand une application f est définie sur l'ensemble de

ces nombres et qu'elle est continue, elle est définie pour x puisque $f(x) = \lim f(u_n)$ quand u_n tend vers x, c'est-à-dire quand n tend vers l'infini. La puissance de $\mathcal{C}$ est donc celle d'une partie de l'ensemble des applications de l'ensemble des rationnels de $[0, 1]$, soit $[0, 1] \cap \mathbf{Q}$ dans $\mathbf{R}$. Or, $[0, 1] \cap \mathbf{Q}$ est dénombrable, $\mathbf{R}$ a pour puissance c, l'ensemble des applications continues a donc une puissance inférieure ou égale à $c^{\aleph_0} = c$. D'autre part, l'ensemble des applications constantes de $[0, 1]$ dans $\mathbf{R}$, dont la puissance est c, est contenu dans $\mathcal{C}$.

Donc, $c \leqslant \text{card } \mathcal{C} \leqslant c$; donc, $\text{card } \mathcal{C} = c$,

tandis que l'ensemble $\mathcal{F}([0, 1], \mathbf{R})$ de toutes les applications de $[0, 1]$ dans $\mathbf{R}$ a pour puissance c^c.

On peut montrer que $c^c > c$. En effet, nous avons montré (I, 3, 3) que 2^c est strictement supérieur à c. Or, $c^c \geqslant 2^c$, car l'ensemble des applications d'un ensemble E de puissance c dans un ensemble E' de puissance c contient évidemment l'ensemble des applications de E dans une partie à deux éléments de E'. Donc, $c^c \geqslant 2^c > c$.

Exercice 4.

1° Un sous-module est d'abord un sous-groupe ; la réponse à la deuxième question est donc évidente.

D'autre part, si g est un sous-groupe d'un groupe G :

$$\forall a \in g \qquad \forall n \in \mathbf{Z} \qquad na \in g$$

g est donc stable pour l'opération externe et les axiomes de la structure de module relativement à cette opération sont vérifiés. On peut conclure :

$$\text{sous-groupe} \iff \text{sous-module.}$$

2° Soit A un anneau. Observons d'abord que sous-anneaux, sous-modules et idéaux sont des sous-groupes de A pour sa première loi de composition.

Soit I un tel sous-groupe. Pour qu'il soit un sous-module sur A ou pour qu'il soit un idéal, il faut et il suffit qu'il soit stable pour l'opération externe :

$$\forall a \in \mathrm{A} \qquad \forall x \in \mathrm{I} \qquad ax \in \mathrm{I} \qquad (1)$$

Il y a donc équivalence, dans A, entre les notions d'idéal et de sous-module.

D'autre part, de (1) il résulte en particulier :

$$\forall y \in \mathrm{I} \qquad \forall x \in \mathrm{I} \qquad xy \in \mathrm{I} \qquad (2)$$

ce qui suffit, I étant déjà un sous-groupe pour la première opération, pour que I soit un anneau.

On peut conclure :

$$\begin{tikzcd} \text{idéal} \arrow[rr, Leftrightarrow] \arrow[dr, Rightarrow] & & \text{sous-module} \arrow[dl, Rightarrow] \\ & \text{sous-anneau} & \end{tikzcd}$$

Il est clair qu'inversement, (2) n'implique pas (1). On doit pouvoir trouver des exemples de sous-anneaux qui ne soient pas des idéaux.

Exemples : a) Prenons pour A l'anneau des entiers de Gauss (nombres de la forme $a + bi$; $a \in \mathbf{Z}$, $b \in \mathbf{Z}$; $\mathbf{Z}$ est un sous-anneau de cet anneau qui n'est pas un idéal (ou sous-module).

b) Prenons pour A l'anneau $\mathbf{R}[x]$ des polynômes réels à une indéterminée ; l'ensemble des polynômes dont tous les monômes sont de degré pair en constitue un sous-anneau qui n'est pas un idéal.

3° Un sous-corps ou un sous-espace vectoriel est d'abord un sous-groupe pour la première opération de K. Mais ceci ne suffit pas pour être un sous-corps qui doit contenir les produits et les inverses de ses éléments. Ceci ne suffit pas non plus pour être un sous-espace vectoriel qui doit contenir les produits de ses éléments par ceux de k.

Aucune des deux propriétés n'implique donc l'autre.

Exemple : Considérons $\mathbf{C}$ comme espace vectoriel sur $\mathbf{R}$. Il admet comme sous-corps le corps des nombres $a + bi$ ($a \in \mathbf{Q}$; $b \in \mathbf{Q}$), qui n'est pas un sous-espace vectoriel sur $\mathbf{R}$. Il admet comme sous-espace vectoriel l'ensemble des complexes d'arguments $\theta + k\pi$ (avec $\theta \neq k\pi$), qui n'est pas un sous-corps.

Exercice 5.

Tout x appartenant à E peut s'écrire :

$$\Sigma \lambda^i a_i + \lambda^j b_j = \Sigma \lambda^i a_i + \Sigma \lambda^j (b'_j - a)$$
$$= \underbrace{\Sigma \lambda^i a_i - a \Sigma \lambda^j}_{\mathrm{V}_1} + \underbrace{\Sigma \lambda^j b'_j}_{\mathrm{V}'_2}$$

a étant combinaison linéaire finie des a_i, on a :

$$x = \underbrace{\Sigma \mu^i a_i}_{\mathrm{V}_1} + \underbrace{\Sigma \lambda^j b'_j}_{\mathrm{V}'_2}$$

Cette décomposition est unique ; en effet :

$$\Sigma \mu^i a_i + \Sigma \lambda^j b'_j = \Sigma \mu'^i a_i + \Sigma \lambda'^j b'_j$$

entraînerait :

$$\Sigma (\mu^i - \mu'^i) a_i = \Sigma (\lambda'^j - \lambda^j) b'_j.$$

Un vecteur du sous-espace V'_2 engendré par les b'_j serait égal à un vecteur de V_1. Or, une telle égalité :

$$\Sigma \nu^i a_i = \Sigma \rho^j b'_j$$

s'écrirait, en revenant aux b_j :

$$\Sigma \nu^i a_i = \Sigma \rho^j (b_j + a)$$
$$\Sigma \nu^i a_i - a \Sigma \rho^j = \Sigma \rho^j b_j$$

et signifierait l'égalité d'un vecteur de V_1 et d'un vecteur de V_2. Elle n'est possible que si ces vecteurs sont nuls, c'est-à-dire si :

$\forall j \quad \rho^j = 0$, ce qui entraîne $\quad \lambda'^j = \lambda^j$

et, par suite : $\quad \forall i \quad \nu^i = 0$, » $\quad \mu'^i = \mu^i$

on en conclut que :

$$\{ a_i \} \cup \{ b'_j \}$$

constitue une base. Mais ceci exprime aussi que E est somme directe des sous-espaces V_1 et V'_2.

V'_2 est donc un supplémentaire de V_1. Or, il est distinct de V_2, car :

$$a \notin \mathrm{V}_2 \implies b'_j \notin \mathrm{V}_2.$$

Soit maintenant un autre élément $\bar{a} \in \mathrm{V}_1$ et considérons, de même que précédemment, le sous-espace $\overline{\mathrm{V}}'_2$ engendré par $\{ \overline{b}'_j = b_j + \bar{a} \}$. Montrons que V'_2 et $\overline{\mathrm{V}}'_2$ sont distincts si $a \neq \bar{a}$.

Soit, en effet, un vecteur :

$$\Sigma \lambda^j (b_j + a) \in \mathrm{V}'_2.$$

Pour qu'il appartienne à $\overline{V}'_2$, il faut qu'on ait une égalité de la forme :

$$\Sigma \lambda_j(b_j + a) = \Sigma \overline{\lambda}_j(b_j + \overline{a})$$

qui peut s'écrire :

$$\Sigma (\lambda_j - \overline{\lambda}_j) b_j = \overline{a}\, \Sigma \overline{\lambda}_j - a\, \Sigma \lambda_j$$

cette égalité exprime l'égalité d'un vecteur de V_2 et d'un vecteur de V_1 ; elle n'est possible que s'ils sont nuls, c'est-à-dire si l'on a :

$$\forall j \quad \lambda_j = \overline{\lambda}_j \; ; \qquad \overline{a}\, \Sigma \overline{\lambda}_j - a\, \Sigma \lambda_j = 0.$$

La première de ces deux conditions permet de mettre la seconde sous la forme :

$$(\overline{a} - a)\, \Sigma \lambda_j = 0.$$

Pour tout vecteur pour lequel $\Sigma \lambda_j \neq 0$, l'égalité est impossible. Une infinité (si K est infini) de vecteurs de $\overline{V}'_2$ n'appartiennent pas à V'_2.

Exercice 6.

La condition est nécessaire. Supposons en effet qu'il existe $i < n$ tel que $(H_1 + \ldots\ldots H_i) \cap H_{i+1} \neq \{ 0 \}$ et soit $a \neq 0$ appartenant à cette intersection.

$$a \in H_{i+1} \qquad \exists\, a_1 \in H_1,\ a_2 \in H_2 \ldots a_i \in H_i \quad a = a_1 + \ldots + a_i.$$

Soit alors $x \in E$ et sa décomposition :

$$x = x_1 + x_2 + \ldots + x_i + x_{i+1} + x_{i+2} + \ldots + x_n$$

on peut écrire aussi :

$$x = (x_1 + a_1) + (x_2 + a_2) + \ldots + (x_i + a_i) + (x_{i+1} - a) + x_{i+2} + \ldots + x_n.$$

Cette décomposition est distincte de la première. H n'est pas somme directe des H_i.

La condition est suffisante.

$$\forall x \in E \quad \exists\, x_n \in H_n, \quad y_n \in H_1 + \ldots + H_{n-1}$$
$$x = x_n + y_n.$$

De $(H_1 + \ldots H_{n-1}) \cap H_n = \{ 0 \}$, on déduit que la décomposition précédente est unique, donc que x_n est déterminée, et y_n également.

Mais alors :

$$\exists\, x_{n-1} \in H_{n-1} \quad y_{n-1} \in H_1 + H_2 + \ldots + H_{n-2}$$
$$y_n = x_{n-1} + y_{n-1},$$

et cette décomposition est encore unique puisque :

$$(H_1 + \ldots + H_{n-2}) \cap H_{n-1} = \{ 0 \}$$

et ainsi de suite de proche en proche. Nous arriverons au fait que :

$$y_3 = x_2 + y_2 \quad x_2 \in H_2 \quad y_2 \in H_1,$$

x_2 et y_2 (que l'on peut appeler x_1) étant déterminés de manière unique.

La décomposition de $x = x_n + x_{n-1} + \ldots + x_2 + x_1$ est unique. H est somme directe des H_i.

La condition $H_i \cap H_j = \{ 0 \}$ pour $i \neq j$ n'est pas suffisante. Etant donnés deux espaces H_1 et H_2 tels que $H_1 \cap H_2 = \{ 0 \}$, on peut, en effet, toujours trouver un espace H_3 tel que :

$$H_1 \cap H_3 = \{ 0 \} \qquad H_2 \cap H_3 = \{ 0 \} \qquad H_3 \subset H_1 + H_2.$$

(Il suffira pour cela de définir, par exemple, H_3 par :

$$H_3 = \{ \lambda(x_1 + x_2) \; ; \; \lambda \in K \},$$

$x_1 \in H_1$ et $x_2 \in H_2$ étant deux vecteurs fixes).

On a alors $H_1 + H_2 + H_3 = H_1 + H_2$ et $H_1 + H_2 + H_3$ n'est pas

une somme directe. Nous pourrons en effet écrire un élément $x = x_1 + x_2 + x_3$ avec $x_1 \in H_1$, $x_2 \in H_2$, $x_3 \in H_3$ sous la forme $x = x_1 + x_2 + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2)$ avec $x_1 + y_1 \in H_1$ $x_2 + y_2 \in H_2$, ce qui donne une décomposition de x distincte de la première.

Exercice 7.

1° La somme de deux applications $\varphi \in \mathcal{H}$ et $\varphi' \in \mathcal{H}$ appartient à $\mathcal{H}$ car $(\varphi + \varphi')(a)$ n'est différent de zéro que pour un ensemble fini d'éléments (inclus dans la réunion de ceux pour lesquels $\varphi(a) \neq 0$ et de ceux pour lesquels $\varphi'(a) \neq 0$). Le produit d'une application $\varphi \in \mathcal{H}$ par un scalaire de K appartient encore à $\mathcal{H}$ puisqu'elle ne prend de valeurs différentes de zéro que pour les mêmes a que l'application φ. $\mathcal{H}$ est donc un sous-espace vectoriel de $\mathcal{F}$ (A, K).

L'application $\varphi \in \mathcal{H}$ telle que :

$$\varphi(a_1) = \mu_1 \quad \varphi(a_2) = \mu_2 \ldots \varphi(a_n) = \mu_n$$

et

$$\forall a \in A - \{ a_1, a_2, \ldots, a_n \} \qquad \varphi(a) = 0$$

peut s'écrire : $\varphi = \mu_1 \varphi_{a_1} + \mu_2 \varphi_{a_2} + \ldots + \mu_n \varphi_{a_n}$.

Toute application $\varphi \in \mathcal{H}$ peut donc s'exprimer comme combinaison linéaire finie des applications du type φ_a. L'ensemble de ces applications constitue donc un système de générateurs de $\mathcal{H}$. D'autre part, elles forment une partie libre, une d'entre elles φ_a ne pouvant être combinaison linéaire de plusieurs autres $\varphi_{a'}$, $\varphi_{a''}$, ..., etc. puisque $\varphi_a(a) = 1$, tandis que toute combinaison linéaire de $\varphi_{a'}$ avec $a' \neq a$ prend pour a la valeur zéro.

L'ensemble des applications φ_a constitue donc une base de $\mathcal{H}$. Il y a une bijection évidente entre A et l'ensemble des φ_a, ce qui permet, si l'on veut, d'identifier les deux ensembles et d'écrire, si cela n'a pas d'inconvénient pour la question à traiter, les combinaisons linéaires formelles d'éléments de A sous la forme $\Sigma \{ \mu_a a ; a \in A \}$ au lieu de $\Sigma \{ \mu_a \varphi_a ; a \in A \}$.

2° Tout $\varphi \in \mathcal{H}$ s'écrit de manière unique :

$$\varphi = \Sigma \{ \mu_a \varphi_a \; ; \; a \in A \}.$$

Il est alors naturel de lui faire correspondre l'élément x de H :

$$x = \Sigma \{ \mu_a a \; ; \; a \in H \}$$

et l'application de $\mathcal{H}$ dans H qui à φ fait correspondre x est linéaire d'après sa définition même. Elle est surjective, puisque H étant engendré par A, les combinaisons linéaires finies d'éléments de A dans E décrivent tout H. Pour qu'elle soit injective, il faut que tout élément de H soit obtenu d'une seule façon comme combinaison linéaire finie d'éléments de A. Il en sera ainsi, si et seulement si A est une base de H, c'est-à-dire une partie libre de E.

L'application de $\mathcal{H}$ sur H fait correspondre à une combinaison linéaire formelle d'éléments de A, une combinaison linéaire dans E. Si on utilisait ici la notation indiquée à la fin de la première question, toutes deux seraient notées de la même façon, bien que, lorsque l'application n'est pas injective, plusieurs (et, à vrai dire, une infinité) combinaisons linéaires formelles ont pour image la même combinaison linéaire dans E.

Exercice 8.

1° Un élément $x \in E$ s'écrit $\{ x_i \; ; \; i \in I \}$. E a une structure d'espace vectoriel sur K pour les deux lois

$$\{ x_i \} + \{ y_i \} = \{ x_i + y_i \}$$
$$\lambda \{ x_i \} = \{ \lambda x_i \}$$

Considérons les éléments dont toutes les composantes sont nulles sauf la $i^{\text{ème}}$. Appelons H_i leur ensemble. Il est évident qu'il forme un sous-espace vectoriel de E. La correspondance entre x_i (élément de E_i) et l'élément de H_i qui a x_i pour $i^{\text{ème}}$ composante est une bijection dont il est évident qu'elle est un isomorphisme.

2° E_0, plus petit sous-espace vectoriel contenant $\cup H_i$, contient toutes les sommes finies d'éléments appartenant à $\cup H_i$, donc tous les éléments de E qui n'ont qu'un nombre fini de composantes différentes de zéro. Considérons donc *l'ensemble des éléments de* E *qui n'ont qu'un nombre fini de composantes non nulles.*

Cet ensemble constitue un sous-espace vectoriel (car la somme de deux de ses éléments est un de ses éléments, ainsi que le produit de l'un de ses éléments par un scalaire). C'est donc le sous-espace vectoriel E_0 cherché.

D'après sa construction même, E_0 est la somme des H_i. D'autre part, la décomposition d'un élément de E_0 en ses composantes dans les H_i est unique. Donc :

$$E_0 = \oplus H_i.$$

3° Les nombres cardinaux c_x et $c_{x'}$ des ensembles des composantes non nulles de deux éléments x et x' de E vérifient :

$$c_x < a \qquad c_{x'} < a,$$

ce qui entraîne $c_{x+x'} \leqslant c_x + c_{x'} < a + a$; a étant un cardinal infini $a + a = a$. D'où $x + x' \in E_a$.

Pour deux éléments de F_a on a de même :

$$c_{x+x'} \leqslant c_x + c_{x'} \leqslant a + a = a$$
$$x + x' \in F_a.$$

Il est évident, d'autre part, que

$$x \in E_a \implies \lambda x \in E_a \quad \text{et} \quad x \in F_a \implies \lambda x \in F_a$$

Si $a = \aleph_0$, $c_x < a$ veut dire « c_x fini ». On retrouve E_0 :

$$E_{\aleph_0} = E_0.$$

Exercice 9.

P_n est bien un espace vectoriel sur **R** car il vérifie tous les axiomes de cette structure. On en obtient immédiatement une base B formée de tous les monômes de degré inférieur ou égal à n, soit $n + 1$ monômes de degré n, n monômes de degré $n - 1$, ..., etc., soit au total :

$$1 + 2 + 3 + \dots + n + 1 = \frac{(n+1)(n+2)}{2} = \mathrm{N}$$

monômes. Ce nombre représente la dimension de P_n.

Les deux premiers et le dernier sous-ensembles envisagés ne sont pas des sous-espaces vectoriels, la somme de deux éléments du sous-ensemble n'appartenant pas forcément au sous-ensemble (ni, pour le dernier, le produit par un scalaire).

Tous les autres sous-ensembles sont au contraire des sous-espaces vectoriels, à condition d'ajouter 0 à l'ensemble des polynômes homogènes de degré fixe.

Polynômes homogènes de degré fixe $n_0 < n$. — Leur sous-espace est de dimension $n_0 + 1$. Un sous-espace supplémentaire est fourni par celui des polynômes qui n'ont pas de termes de degré n_0. Sa dimension est $\mathrm{N} - (n_0 + 1)$. Des bases des 2 supplémentaires s'obtiennent immédiatement par partition de la base B.

Polynômes de degré inférieur ou égal à n à une indéterminée. (Cet espace vectoriel est bien inclus dans P_n car c'est celui qu'on obtient en

considérant les polynômes de P_n où seuls des monômes en x seul ont des coefficients non nuls). Sa dimension est $n + 1$. Un supplémentaire est fourni par l'ensemble des polynômes n'ayant pas de monômes en x seul. Sa dimension est $\mathrm{N} - (n + 1)$. Des bases s'obtiennent encore par partition de B.

Polynômes pairs. — Les polynômes impairs forment un sous-espace supplémentaire. En effet, d'une part, tout polynôme $P(x, y)$ peut s'écrire :

$$P(x, y) = \frac{P(x, y) + P(-x, -y)}{2} + \frac{P(x, y) - P(-x, -y)}{2}$$

c'est-à-dire sous forme de somme d'un polynôme pair et d'un polynôme impair ; d'autre part, l'intersection des deux sous-espaces se réduit au polynôme nul. Des bases s'obtiennent encore par partition de B en monômes pairs et impairs.

Si $n = 2\lambda$, la première contient $1 + 3 + 5 + \dots + 2\lambda + 1 = (\lambda + 1)^2$ monômes, la deuxième en contient $2 + 4 + 6 + \dots + 2\lambda = \lambda(\lambda + 1)$ et on vérifie bien que $(\lambda + 1)^2 + \lambda(\lambda + 1) = \dfrac{(n+1)(n+2)}{2}$.

Si $n = 2\lambda + 1$, la première contient

$$1 + 3 + 5 + \dots + 2\lambda + 1 = (\lambda + 1)^2$$

monômes, la deuxième en contient

$$2 + 4 + 6 + \dots + 2\lambda + 2 = (\lambda + 1)(\lambda + 2)$$

et on vérifie encore que $(\lambda + 1)^2 + (\lambda + 1)(\lambda + 2) = \dfrac{(n+1)(n+2)}{2}$.

Polynômes symétriques. — Les polynômes antisymétriques (tels que $P(x, y) = -P(y, x)$) forment un sous-espace supplémentaire car, d'une part, tout polynôme $P(x, y)$ peut s'écrire :

$$P(x, y) = \frac{P(x, y) + P(y, x)}{2} + \frac{P(x, y) - P(y, x)}{2}$$

c'est-à-dire sous forme de somme d'un polynôme symétrique et d'un polynôme antisymétrique ; d'autre part, l'intersection des deux sous-espaces est réduite au polynôme nul puisqu'un polynôme à la fois symétrique et antisymétrique doit vérifier $P(y, x) = -P(y, x)$, c'est-à-dire $P = 0$. Cette fois, les éléments de B n'appartiennent ni à l'un, ni à l'autre des sous-espaces. Mais tout polynôme symétrique peut s'écrire, en groupant les termes qui ont même coefficient, sous la forme :

$$\Sigma\, a_{p,q}\, (x^p y^q + x^q y^p) \qquad \text{avec } p + q \leqslant n \quad p \leqslant q.$$

Sous cette forme on voit que l'ensemble :

$$\{ x^p y^q + x^q y^p \,;\, p + q \leqslant n \quad p \leqslant q \}$$

constitue un système de générateurs du sous-espace des polynômes symétriques. Il en est, d'autre part, une partie libre, aucun d'eux ne pouvant être obtenu comme combinaison linéaire des autres. Il en constitue donc une base.

De même, l'espace des polynômes antisymétriques de P_n admet pour base l'ensemble :

$$\{ x^p y^q - x^q y^p \,;\, p + q \leqslant n \quad p < q \}.$$

Si $n = 2\lambda$, ces bases ont respectivement :

pour les polynômes symétriques,

$$1 + 1 + 2 + 2 + \dots + \lambda + \lambda + \lambda + 1 = (\lambda + 1)^2 \text{ éléments ;}$$

pour les polynômes antisymétriques,

$$1 + 1 + 2 + 2 + \dots + \lambda + \lambda = \lambda(\lambda + 1) \text{ éléments.}$$

Si $n = 2\lambda + 1$, elles en ont :
pour les symétriques,
$$1 + 1 + 2 + 2 + \ldots + \lambda + \lambda + \lambda + 1 + \lambda + 1 = (\lambda + 1)(\lambda + 2) ;$$
pour les antisymétriques,
$$1 + 1 + 2 + 2 + \ldots + \lambda + \lambda + \lambda + 1 = (\lambda + 1)^2.$$
Dans les deux cas, la somme des dimensions est bien la dimension de P_n. La réunion des deux bases précédentes constitue une base de P_n distincte de B.

Polynômes de degré inférieur ou égal à $p < n$. — Leur sous-espace admet comme supplémentaire celui des polynômes dont les $p + 1$ premiers coefficients sont nuls. La base B se partage en donnant les deux bases qui ont, l'une $\dfrac{(p+1)(p+2)}{2}$ éléments, l'autre $N - \dfrac{(p+1)(p+2)}{2}$ éléments.

Polynômes s'annulant pour $x = a$, $y = b$. — Tout polynôme P tel que $P(a\ b) = c$ pouvant s'écrire :
$$P(x, y) = \underbrace{P(x, y) - c} + c,$$
et $P(x, y) - c$ appartenant au sous-espace considéré Π, on voit que P peut se mettre sous forme de la somme d'un polynôme de Π et d'un réel :
$$P_n = \Pi + \mathbf{R}.$$
Mais $\Pi \cap \mathbf{R} = \{0\}$ car le seul polynôme constant qui s'annule pour $x = a$, $y = b$ est le polynôme nul, donc :
$$P_n = \Pi \oplus \mathbf{R}.$$
La dimension de $\mathbf{R}$ est 1 ; celle de Π est celle de l'ensemble des polynômes sans terme constant en $X = x - a$ et $Y = y - b$; elle est donc $N - 1$ et la vérification relative aux dimensions est encore faite.

Pour base de Π on peut prendre l'ensemble des polynômes :
$$\{(x - a)^p (y - b)^q \,;\, p + q \leqslant n \,;\, p + q \neq 0\}$$
qui forme bien une base de Π puisque l'ensemble des X^pY^q forme bien une base de l'ensemble des polynômes en X et Y nuls à l'origine.

La réunion de cette base et d'une constante arbitraire (base de $\mathbf{R}$) constitue encore une nouvelle base de P_n.

Exercice 10.

1° Cherchons le laplacien de $f = (x^2 + y^2)^\lambda P(x, y)$. Il vient :
$$\Delta f = (x^2 + y^2)^\lambda \Delta P + 4\lambda (x^2 + y^2)^{\lambda - 1}\left(x\frac{\partial P}{\partial x} + y\frac{\partial P}{\partial y}\right) + 4\lambda^2 (x^2 + y^2)^{\lambda - 1} P,$$
P étant homogène et de degré n :
$$x\frac{\partial P}{\partial x} + y\frac{\partial P}{\partial y} = n\,P,$$
et la formule précédente devient :
$$\Delta f = (x^2 + y^2)^\lambda \Delta P + 4(x^2 + y^2)^{\lambda - 1}\lambda(n + \lambda)\,P.$$
Si $\Delta f = 0$, il en résulte que :
$$(x^2 + y^2)\Delta P + 4\lambda(n + \lambda)\,P = 0,$$
donc que P est divisible par $x^2 + y^2$ s'il n'est pas le polynôme nul.

On a donc : $f = (x^2 + y^2)^{\lambda + 1} P_1$.

Mais le même raisonnement s'appliquant à P_1, celui-ci sera divisible par $x^2 + y^2$ et ainsi de suite... P étant de degré n, au bout de $\dfrac{n}{2}$ opérations

au plus, nous aurons un quotient qui ne pourra plus être divisible par $x^2 + y^2$ sans être identiquement nul. Donc, $P = 0$.

2° $\mathcal{P}_n$ est un espace vectoriel sur $\mathbf{R}$. Il est immédiat que $\mathcal{H}_n$ et $\mathcal{A}_n$ en sont des sous-espaces.

La première question montre que $\mathcal{A}_n \cap \mathcal{H}_n = \{0\}$ puisque si un polynôme harmonique est divisible par $x^2 + y^2$, il est nul. Si donc les dimensions de ces espaces ont pour somme la dimension de $\mathcal{P}_n$, la réunion de leurs bases constituera une base de $\mathcal{P}_n$ et $\mathcal{P}_n$ sera la somme directe de $\mathcal{A}_n$ et $\mathcal{H}_n$. Or, $\dim \mathcal{P}_n = n + 1$.

D'autre part, tout élément de $\mathcal{A}_n$ pouvant s'écrire $(x^2 + y^2)Q(x, y)$, $Q(x, y)$ représentant un polynôme homogène arbitraire de degré $n - 2$, $_n$ est de dimension $n - 1$.

Reste à trouver la dimension de $\mathcal{H}_n$; pour cela, cherchons les polynômes harmoniques homogènes de degré n. Il est commode de passer en coordonnées polaires. Pour une fonction :
$$f(x, y) = f(\rho\cos\theta, \rho\sin\theta) = \varphi(\rho, \theta),$$
le laplacien s'écrit :
$$\Delta f = \Delta\varphi = \frac{\partial^2\varphi}{\partial\rho^2} + \frac{1}{\rho^2}\frac{\partial^2\varphi}{\partial\theta^2} + \frac{1}{\rho}\frac{\partial\varphi}{\partial\rho}$$
Les polynômes homogènes de degré n vérifient $P(x, y) = \rho^n P(\cos\theta, \sin\theta)$ et leur laplacien vaut :
$$\Delta P = \rho^{n-2} n^2 P + \rho^{n-2}\frac{\partial^2\rho}{\partial\theta^2}$$
Ceux qui sont harmoniques sont donc les solutions de l'équation différentielle :
$$n^2 P + P'' = 0$$
soit :
$$P = a\cos n\theta + b\sin n\theta$$
(qui sont bien des polynômes en $\sin\theta$ et $\cos\theta$ homogènes et de degré n) et qui, dépendant linéairement de deux scalaires arbitraires, forment un espace vectoriel de dimension 2 sur $\mathbf{R}$.

On a donc : $\dim \mathcal{H} = 2$ et on a bien $(n - 1) + 2 = n + 1$
$$\mathcal{P}_n = \mathcal{H}_n \oplus \mathcal{A}_n.$$

3° Tout polynôme homogène de degré n peut donc se mettre de manière unique sous la forme :
$$Q_n = H_n + AQ_{n-2} \qquad H_n \in \mathcal{H}_n \qquad AQ_{n-2} \in \mathcal{A}_n$$
Mais Q_{n-2} étant homogène de degré $n - 2$ peut se mettre de manière unique sous la forme :
$$Q_{n-2} = H_{n-2} + AQ_{n-4} \qquad H_{n-2} \in \mathcal{H}_{n-2} \qquad AQ_{n-4} \in \mathcal{A}_{n-2},$$
d'où $Q_{n-4} \in \mathcal{P}_{n-4}$.
D'où $Q_n = H_n + AH_{n-2} + A^2Q_{n-4}$.

On décompose Q_{n-4} de la même façon... et ainsi de suite et on arrive pour Q_n à la forme indiquée où H_p est un polynôme homogène du premier degré ou une constante. Et chacun des H_i a été déterminé de manière unique.

Un polynôme Q non nécessairement homogène est somme de polynômes homogènes ; chacun des termes de cette somme étant mis sous la forme indiquée, il suffit de mettre A^k en facteur dans la somme des termes de la forme A^kH_k pour obtenir la forme indiquée et pour l'obtenir de manière unique.

4° La restriction de Q à E est la même que celle du polynôme harmonique $H_0 + H_1 + \dots + H_p$ déterminé de manière unique. Autrement dit, ce polynôme prolonge dans le plan la fonction polynômiale sur E.

5° Le théorème de Stone-Weierstrass garantit pour tout ε l'existence d'un polynôme $\psi(x, y)$ approximant à ε près la fonction f sur le compact $C(x^2 + y^2 = 1)$. La 4° question garantit l'existence d'un polynôme harmonique φ ayant la même restriction que ψ sur C. A la suite $\varepsilon_n = \frac{\varepsilon}{2^n}$ correspondra une suite de polynômes harmoniques φ_n qui convergera uniformément vers f sur C. Les φ_n étant harmoniques, la convergence uniforme de la suite $\{ \varphi_n \}$ sur C entraîne sa convergence uniforme à l'intérieur de C (principe du maximum) vers une fonction qui sera harmonique, en tant que limite uniforme de fonctions harmoniques.

Exercice 11.

$\mathcal{V}(E)$ ordonné par inclusion forme un treillis : deux sous-espaces vectoriels ont pour intersection un sous-espace vectoriel qui est le plus grand sous-espace inclus dans chacun d'eux et qui est donc leur plus grand minorant. Ils ont pour somme un sous-espace vectoriel (en effet,
$x \in V_1 + V_2,\ y \in V_1 + V_2 \implies x + y \in V_1 + V_2$ et $\lambda x \in V_1 + V_2$)
qui est le plus petit sous-espace vectoriel qui les contient tous deux, donc leur plus petit majorant.

Etudier si ce treillis est distributif, c'est comparer d'une part :

$$A = V_1 + (V_2 \cap V_3) \quad \text{et} \quad B = (V_1 + V_2) \cap (V_1 + V_3)$$

d'autre part :

$$C = V_1 \cap (V_2 + V_3) \quad \text{et} \quad D = (V_1 \cap V_2) + (V_1 \cap V_3).$$

On trouve immédiatement, en se reportant aux définitions, les inclusions suivantes :

$$D \subset C \subset V_1 \subset A \subset B$$

et nous devons montrer qu'on peut choisir les sous-espaces de telle façon que $D \neq C$ et $A \neq B$.

Soient V_1, V_2, V_3 trois sous-espaces à une dimension, distincts les uns des autres, ce qui implique :

$$V_1 \cap V_2 = V_2 \cap V_3 = V_3 \cap V_1 = \{0\} \qquad (1)$$

et tels que :

$$V_1 \subset V_2 + V_3 \qquad (2)$$

La condition (1) implique que les trois sommes $V_1 + V_2$, $V_2 + V_3$, $V_3 + V_1$ soient des sommes directes, donc que ces trois sous-espaces soient à deux dimensions. Or, la condition (2) implique immédiatement que :

$$V_1 + V_2 \subset V_2 + V_3 \qquad V_1 + V_2 \subset V_1 + V_3.$$

Deux sous-espaces de même dimension dont l'un est inclus dans l'autre sont nécessairement confondus. On a donc :

$$V_1 + V_2 = V_1 + V_3 = V_2 + V_3 = W$$

W étant un sous-espace de dimension 2. (Si $E = \mathbf{R}^3$, V_1, V_2, V_3 sont trois droites distinctes du même plan).
Pour ces sous-espaces on a :

$$A = V_1 + \{0\} = V_1$$
$$B = W \cap W = W \qquad A \neq B$$

$$C = V_1 \cap W = V_1$$
$$D = \{0\} + \{0\} = \{0\} \qquad C \neq D$$

Remarque 1. — Il était superflu de montrer à la fois $C \neq D$ et $A \neq B$, si on tenait compte de la propriété générale suivante des treillis : quand une des opérations « inf » et « sup » est distributive par rapport à l'autre, la deuxième est distributive par rapport à la première.

Supposons en effet que, pour tout a, b, c appartenant à un treillis, on ait :

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \qquad (1)$$

et supposons qu'on veuille évaluer :

$$A = (a \vee b) \wedge (a \vee c)$$

La distributivité exprimée par (1) permet d'écrire en considérant $a \vee b$ globalement [c'est-à-dire en lui faisant jouer le rôle du « a » de la formule (1)] :

$$A = [(a \vee b) \wedge a] \vee [(a \vee b) \wedge c].$$

Or, $(a \vee b) \wedge a = a$ d'après les définitions du inf et du sup. Evaluons $(a \vee b) \wedge c$ en utilisant à nouveau la distributivité de l'opération « inf ». Il vient :

$$(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c)$$

et

$$A = a \vee (a \wedge c) \vee (b \wedge c).$$

Dans la recherche du sup de trois éléments, deux d'entre eux peuvent être remplacés par leur sup ; $a \vee (a \wedge c) = a$. Donc :

$$A = a \vee (b \wedge c)$$

ce qui démontre la distributivité de l'opération « sup » par rapport à l'opération « inf ».

Remarque 2. — Observons que si le treillis des sous-espaces d'un espace vectoriel n'est pas distributif, il jouit au moins de la propriété d'être modulaire, qui est une propriété de distributivité faible (ou sous condition). Nous avons démontré cette propriété (cf. exercice 26, *Cours A.P.M. I*) de façon générale pour le treillis des sous-groupes d'un groupe. Il est facile de la vérifier directement dans le cas particulier présent, où elle s'écrit :

$$V_1 \subset V_3 \implies V_1 + (V_2 \cap V_3) = (V_1 + V_2) \cap V_3$$

(V_3 étant identique à $V_1 + V_3$).

2° Si $V_1 \cap V_2 = \{0\}$, $V_1 + V_2$ est une somme directe et :

$$\dim (V_1 + V_2) = \dim V_1 + \dim V_2.$$

Si maintenant $V_1 \cap V_2 \neq \{0\}$, considérons une base de $V_1 + V_2$ constituée de la façon suivante : une base de $v = V_1 \cap V_2$; une base de W_1, supplémentaire de V_1 par rapport à v ; une base de W_2, supplémentaire de V_2 par rapport à v. La réunion de ces trois bases engendre bien la partie $V_1 \cup V_2$, donc $V_1 + V_2$, plus petit sous-espace vectoriel contenant cette partie. De par sa construction, cette base est partagée en trois sous-ensembles disjoints et cette partition de la base conduit à écrire :

$$V_1 + V_2 = v \oplus W_1 \oplus W_2$$

ce qui entraîne : $\dim (V_1 + V_2) = \dim v + \dim W_1 + \dim W_2$
Mais $\dim V_1 = \dim v + \dim W_1 \qquad \dim V_2 = \dim v + \dim W_2$
d'où $\dim (V_1 + V_2) = \dim V_1 + \dim V_2 - \dim v$,
soit la formule demandée.

Exercice 12.

1° $f \longrightarrow \frac{df}{dx}$. — Le noyau de l'application est formé par l'ensemble des constantes. Pour supplémentaire du noyau, on peut prendre l'ensemble des fonctions qui s'annulent pour une valeur x_0 de $[0, 1]$, car, pour toute fonction f, on peut écrire :

$$f(x) = f(x_0) + \underbrace{f(x) - f(x_0)}.$$

La restriction de l'application à ce sous-espace est bien injective puisque :

$$\frac{df}{dx} = \frac{dg}{dx} \Longrightarrow f = g \qquad \text{si } f(x_0) = g(x_0).$$

2° $f \longrightarrow \frac{d^pf}{dx^p}$. — Le noyau est formé de l'ensemble des polynômes de degré inférieur à p. Toute fonction p fois dérivable peut s'écrire :

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \ldots + \frac{(x - x_0)^{p-1}}{(p-1)!} f^{(p-1)}(x_0) + \varphi(x)$$

et apparaître comme la somme d'un polynôme de degré inférieur à p et d'une fonction qui, pour la valeur x_0, s'annule, ainsi que ses $p - 1$ premières dérivées. L'ensemble de ces dernières fonctions constitue un sous-espace supplémentaire du noyau.

La restriction de l'application à ce sous-espace est encore injective puisque, pour les fonctions de ce sous-espace, l'égalité des dérivées $p^{\text{èmes}}$ entraîne celle des fonctions.

3° $f \longrightarrow g$ avec $g(x) = \int_0^x f(t)dt$. — Le noyau est réduit au zéro de l'espace de départ (fonction nulle sur tout $[0, 1]$). L'application est donc injective et est donc un isomorphisme sur l'image. (Pas un isomorphisme sur tout l'espace d'arrivée, comme nous le verrons dans l'exercice suivant).

4° $f \longrightarrow \int_0^1 f(x)dx$. — Le noyau est constitué par l'ensemble des fonctions de valeur moyenne nulle sur l'intervalle $[0, 1]$. Pour toute fonction f de valeur moyenne μ, on peut écrire :

$$f(x) = \varphi(x) + \mu$$

φ étant une fonction de valeur moyenne nulle sur $[0, 1]$. On peut donc prendre pour supplémentaire du noyau le sous-espace des fonctions constantes sur $[0, 1]$, sous-espace isomorphe à $\mathbf{R}$ (μ est l'image de la fonction constante et égale à μ).

5° $\{u_n\} \longrightarrow \{v_n = u_n - u_{n-1}\}$. — Le noyau est constitué par l'ensemble $\mathcal{C}$ des suites constantes ($\forall n \;\; u_n = u_{n-1}$). On peut prendre pour sous-espace supplémentaire celui, $\mathcal{H}$, des suites telles que $u_p = 0$ pour un p quelconque fixé, car on peut toujours écrire :

$$\{u_n\} = \{u_n - u_p\} + \{u_p\} \qquad \{u_n - u_p\} \in \mathcal{H} \quad \{u_p\} \in \mathcal{C}$$

Le fait que la restriction à $\mathcal{H}$ de l'application soit injective se vérifie encore, car :

$$\{v_n\} = \{v'_n\} \Longrightarrow \qquad u_n - u_{n-1} = u'_n - u'_{n-1}$$

donc $\forall n \;\; u_n - u'_n = u_p - u'_p = 0$, donc $\{u_n\} = \{u'_n\}$.

Exercice 13.

1° Dire que *f est un isomorphisme,* c'est dire qu'elle est surjective et injective, et alors $\dim f(E) = \dim F = n$.

Réciproquement, si *f est de rang n,* f est d'une part injective, car la dimension du noyau est zéro, d'autre part surjective ; en effet, si $f(E)$ n'était pas F, $f(E)$ serait un sous-espace vectoriel de F et aurait, dans F, un supplémentaire ; celui-ci aurait une dimension au moins égale à 1 et la dimension de F dépasserait celle de E d'au moins une unité, f est donc un isomorphisme.

Si maintenant *f est surjective,* ceci entraîne immédiatement $\dim f(E) = n$, et on est ramené à la précédente hypothèse.

Enfin, si *f est injective,* le noyau de l'application est réduit à zéro ; donc $\dim f(E) = n$, et on est encore ramené à l'hypothèse « f de rang n ».

2° *Contre-exemples :* I. Soient E et F deux espaces vectoriels à bases dénombrables :

$$\{a_n\,;\; n \in \mathbf{N}\} \qquad \text{et} \qquad \{b_n\,;\; n \in \mathbf{N}\}.$$

a) L'application f définie par $a_n \longrightarrow b_{2n}$ est injective, mais n'est pas surjective.

b) L'application g définie par $a_{2n+1} \longrightarrow b_n$, $a_{2n} \longrightarrow b_n$ est surjective, mais n'est pas injective.

(E et F pourront être, par exemple, l'espace vectoriel des polynômes $K[x]$ à coefficients dans un corps K et qui admet pour base $\{x^m\,;\; m \in \mathbf{N}\}$).

II. *a*) Prenons pour espace vectoriel E sur $\mathbf{R}$ celui des fonctions réelles continues sur un intervalle fixé et s'annulant pour une valeur fixe x_0 de cet intervalle. (La somme de deux fonctions appartenant à E et le produit par un scalaire d'une fonction appartenant à E appartiennent à E).

Prenons pour application de E sur lui-même celle qui à g fait correspondre la fonction G donnée par $G(x) = \int_{x_0}^x g(t)\,dt$ (G est celle des primitives de g qui appartient à E).

On vérifie immédiatement que cette application est bien linéaire, car la primitive de $h + g$ est $H + G$ et celle de λg est λG. Cette application est injective, car G_1 ne peut égaler G_2 sans que $g_1 = g_2$, mais elle n'est pas surjective car toutes les fonctions continues ne sont pas dérivables et ne peuvent donc être ainsi obtenues.

b) Les fonctions réelles indéfiniment dérivables sur un intervalle fixé forment encore un espace vectoriel sur $\mathbf{R}$. Comme application f de cet espace sur lui-même, prenons la dérivation. On vérifie immédiatement que c'est une application linéaire :

$$\frac{d}{dx}(g + h) = \frac{d}{dx}g + \frac{d}{dx}h\,; \qquad \frac{d}{dx}\lambda g = \lambda\frac{d}{dx}g.$$

Cette application est surjective, car toute fonction continue a une primitive ; elle n'est pas injective puisqu'une infinité de fonctions ont la même dérivée.

Exercice 14.

1° L'application $f \longrightarrow f'$ est injective. En effet :

$$f' = g' \Longrightarrow \psi \circ f \circ \overset{-1}{\varphi} = \psi \circ g \circ \overset{-1}{\varphi} \Longrightarrow f = g$$

car on sait que de telles égalités sont simplifiables à droite et à gauche quand les multiplicateurs sont des isomorphismes (voir exercice 7, *Cours A.P.M. I*).

De même, $f \longrightarrow f'$ est surjective, car :

$$\forall f' \in \mathcal{L}(E', F') \qquad \exists\, f \text{ telle que } f' = \psi \circ f \circ \overset{-1}{\varphi}$$

Il suffit en effet de prendre $f = \overset{-1}{\psi} \circ f' \circ \varphi$.

Enfin, $f \longrightarrow f'$ est un homomorphisme pour la structure d'espace vectoriel sur $\mathcal{L}(E, F)$. En effet, l'image de $h = f + g$ est :

$$h' = \psi \circ (f + g) \circ \overset{-1}{\varphi}$$

et puisque la composition des applications est distributive par rapport à la somme :

$$h' = \psi \circ f \circ \overset{-1}{\varphi} + \psi \circ g \circ \overset{-1}{\varphi} = f' + g'.$$

De même, en vertu des lois :

$$f \circ \lambda g = \lambda f \circ g = \lambda (f \circ g)$$

on a :

$$(\lambda f)' = \psi \circ (\lambda f) \circ \overset{-1}{\varphi} = \psi \circ \lambda (f \circ \overset{-1}{\varphi}) = \lambda (\psi \circ f \circ \overset{-1}{\varphi}) = \lambda f'.$$

2° Vérifier que l'isomorphisme précédent s'étend à la structure d'algèbre, c'est montrer que $(f \circ g)' = f' \circ g'$. Or, en tenant compte du fait qu'ici $\varphi = \psi$, on peut écrire :

$$(f \circ g)' = \varphi \circ (f \circ g) \circ \overset{-1}{\varphi} = \varphi \circ f \circ \overset{-1}{\varphi} \circ \varphi \circ g \circ \overset{-1}{\varphi} = f' \circ g'.$$

Exercice 15.

Quels que soient les ensembles E et F et quelle que soit la nature des applications, l'hypothèse permet déjà d'affirmer que : f est injective [car $f(x_1) = f(x_2)$ entraînerait que, dans $f \circ g$, x_1 et x_2 n'auraient qu'une seule image] ; g est surjective [sans quoi l'image par g de $f(E)$ ne pourrait être E tout entier].

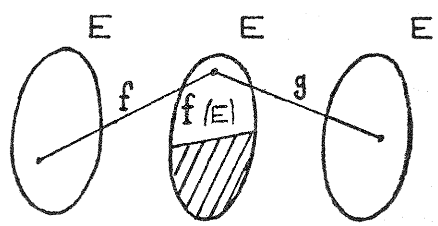


FIG. 10

Soit maintenant E un espace vectoriel de dimension n, de base $\{a_i\}$ et $\{f(a_i)\}$ son image par f.

$$\{g(f(a_i))\} = \{a_i\}$$

est une partie libre, donc $\{f(a_i)\}$ est une famille libre de n éléments ; elle constitue donc une base de E, ce qui entraîne $f(E) = E$.

Il suffit alors de se reporter à l'exercice 13 pour voir que f et g sont des isomorphismes, donc ont des inverses.

L'exercice 13 nous fournira aussi les contre-exemples demandés :

1) En considérant, sur un espace vectoriel E à base dénombrable $\{a_n ; n \in \mathbf{N}\}$, les applications f et g définies à l'exercice 13. $f(E)$ est le sous-espace engendré par l'ensemble des éléments d'indice pair de la base $\{a_n\}$.

2) En prenant pour espace E celui défini dans l'exemple II-a), pour f l'application qui y est définie et pour g la dérivation. $g \circ f$ est encore l'identité, mais $f(E)$ est l'ensemble des fonctions continûment dérivables sur l'intervalle, c'est-à-dire ayant une dérivée continue, qui n'est qu'une partie de l'ensemble des fonctions continues.

Exercice 16.

Soit x un élément différent de zéro de E et soit f un automorphisme qui laisse x invariant.

$$(u \circ f)(x) = u(x) \qquad (f \circ u)(x) = f(u(x))$$

$u(x)$ est donc aussi un élément invariant de f.

Considérons alors un automorphisme qui n'admet pas d'autres éléments invariants que ceux de la forme λx. [Pour construire un tel automorphisme f, il suffit de décomposer E en somme directe du sous-espace $F = \{\lambda x ; \lambda \in K\}$ et d'un sous-espace supplémentaire F', puis de prendre pour image de tout élément de E l'élément qui a même composante dans F et composante opposée dans F'].

Pour un tel automorphisme, l'hypothèse exige donc $u(x) = \lambda x$, λ étant un scalaire dépendant de x.

Montrons maintenant que λ est le même pour tous les x. Soient en effet x_1 et x_2 deux éléments quelconques de E tels que :

$$u(x_1) = \lambda_1 x_1 \qquad u(x_2) = \lambda_2 x_2$$

et choisissons un automorphisme f qui fasse correspondre x_2 à x_1 :

$$(u \circ f)(x_1) = u(x_2) = \lambda_2 x_2$$

$(f \circ u)(x_1) = f(\lambda_1 x_1) = \lambda_1 f(x_1)$ puisque f est une application linéaire, donc $(f \circ u)(x_1) = \lambda_1 x_2$.

L'hypothèse n'est donc vérifiée que si $\lambda_1 = \lambda_2$.

Exercice 17.

1° Le théorème fondamental (III, 1) permet d'écrire :

$$e' = c(\varphi) + \dim \varphi(E)$$
$$e = n(\varphi) + \dim \varphi(E).$$

On doit avoir :

$$e' - e = c(\varphi) - n(\varphi)$$
$$e' - e = b - a \qquad b \leqslant e' \qquad a \leqslant e$$
$$d(\varphi) = e' - e$$

2° La relation précédente appliquée aux trois applications donne :

$$d(\varphi'') = e'' - e = (e'' - e') + (e' - e) = d(\varphi') + d(\varphi).$$

3° Pour qu'une application φ réponde aux conditions posées, il suffit d'écrire si $\{\alpha_n ; n \in \mathbf{N}\}$ et $\{\beta_n ; n \in \mathbf{N}\}$ sont respectivement les bases de E et E' :

$$\varphi(\alpha_1) = \varphi(\alpha_2) = \ldots = \varphi(\alpha_a) = 0$$

et $\qquad \forall n \in \mathbf{N} \qquad \varphi(\alpha_{a+n}) = \beta_{b+n}$

car $\{\alpha_1, \alpha_2 \ldots \alpha_a\}$ engendreront $\overset{-1}{\varphi}(0)$ de dimension a et $\{\beta_1 \ldots \beta_b\}$ engendreront un supplémentaire de $\varphi(E)$ de dimension b.

4° $n'' = n(\varphi'')$ est la dimension de l'ensemble :

$$\{x \in E ; \varphi''(x) = 0\}.$$

Or, pour que $\varphi''(x) = 0$ il faut que $\varphi(x)$ appartienne à $\overset{-1}{\varphi'}(0)$ et il faut aussi qu'il appartienne à $\varphi(E)$. Il est donc nécessaire que :

$$\varphi(x) \in \overset{-1}{\varphi'}(0) \cap \varphi(E) = L'$$

et c'est aussi suffisant. La dimension l' de ce sous-espace est inférieure ou égale à celle de $\overset{-1}{\varphi'}(0)$ qui est $n(\varphi') = n'$ par définition.

D'autre part, $n(\varphi'')$ est la dimension de son image réciproque par φ, soit $l' + n$. On a donc :

$$n'' = n + l' \leqslant n + n'.$$

Si n et n' sont finis, n'' est fini.

D'autre part, $c(\varphi'') = c''$ est la dimension d'un supplémentaire C″ de $\varphi''(\mathrm{E})$ par rapport à E″.

$$\mathrm{E}'' = \varphi''(\mathrm{E}) \oplus \mathrm{C}''.$$

On a de même :

$$\mathrm{E}' = \varphi(\mathrm{E}) \oplus \mathrm{C} \qquad \dim \mathrm{C} = c = c(\varphi)$$
$$\mathrm{E}'' = \varphi'(\mathrm{E}') \oplus \mathrm{C}' \qquad \dim \mathrm{C}' = c' = c(\varphi').$$

E′ étant la somme directe de $\varphi(\mathrm{E})$ et de C, son image par φ' est la somme (pas nécessairement directe car φ' n'étant pas injective, des éléments de $\varphi(\mathrm{E})$ et de C peuvent avoir des images confondues) de $\varphi'(\varphi(\mathrm{E})) = \varphi''(\mathrm{E})$ et de $\varphi'(\mathrm{C})$. Un supplémentaire de $\varphi''(\mathrm{E})$ par rapport à $\varphi'(\mathrm{E}')$ est donc de dimension inférieure ou égale à celle de $\varphi'(\mathrm{C})$, elle-même inférieure ou égale à c. On a donc :

$$\varphi'(\mathrm{E}') = \varphi''(\mathrm{E}) \oplus \mathrm{C}_1 \qquad c_1 = \dim \mathrm{C}_1 \leqslant c$$

et
$$\mathrm{E}'' = \varphi''(\mathrm{E}) \oplus \mathrm{C}_1 \oplus \mathrm{C}',$$

d'où
$$\mathrm{C}'' \approx \mathrm{C}_1 \oplus \mathrm{C}' \qquad c'' = c_1 + c' \leqslant c + c',$$

c'' est donc fini quand c et c' le sont.

5° Dans le cas où n et n' sont aussi finis, précisons cette évaluation de c'', donc de c_1. Pour cela, décomposons E′ en somme directe de la façon suivante :

1) le sous-espace $\mathrm{L}' = \overset{-1}{\varphi'}(0) \cap \varphi(\mathrm{E})$ de dimension l' déjà considéré ;

2) un supplémentaire de ce sous-espace par rapport à $\overset{-1}{\varphi'}(0)$, soit G′ ; sa dimension est $g' = n' - l'$;

3) un supplémentaire de L′ par rapport à $\varphi(\mathrm{E})$, soit H′ ;

4) un supplémentaire du sous-espace engendré par les trois précédents, soit K′. On peut écrire indifféremment :

$$\mathrm{E}' = \mathrm{L}' \oplus \mathrm{G}' \oplus \mathrm{H}' \oplus \mathrm{K}' = \overset{-1}{\varphi'}(0) \oplus \mathrm{H}' \oplus \mathrm{K}' \qquad (1)$$
$$= \varphi(\mathrm{E}) \oplus \mathrm{G}' \oplus \mathrm{K}'$$

$\mathrm{G}' \oplus \mathrm{K}'$ représente le sous-espace C de la question 4). Le théorème fondamental nous permet d'écrire en vertu de (1) :

$$\varphi'(\mathrm{E}') \approx \mathrm{H}' \oplus \mathrm{K}' \qquad (2).$$

D'autre part,

$$\varphi(\mathrm{E}) = \mathrm{L}' \oplus \mathrm{H}'.$$

Mais il résulte des définitions des différents sous-espaces que L′ est le noyau de la restriction de φ' à $\varphi(\mathrm{E})$. Donc, le théorème fondamental permet encore d'écrire :

$$\varphi''(\mathrm{E}) = \varphi'(\varphi(\mathrm{E})) \approx \mathrm{H}' \qquad (3).$$

Nous avions posé au 4° :

$$\varphi'(\mathrm{E}') \approx \varphi''(\mathrm{E}) \oplus \mathrm{C}_1.$$

Il suffit de rapprocher cette dernière égalité de (2) et (3) pour voir que $\mathrm{K}' \approx \mathrm{C}_1$. On a donc :

$$c_1 = k' \text{ et puisque } \mathrm{G}' \oplus \mathrm{K}' = \mathrm{C}$$
$$c_1 = k' = c - g' = c - (n' - l')$$

Finalement :
$$c'' = c' + c + l' - n'.$$

Formons :

$$d'' = c'' - n'' = c' + c + l' - n' - (n + l')$$
$$= c' - n' + c - n,$$
$$\text{soit } d(\varphi'') = d(\varphi') + d(\varphi).$$

Exercice 18.

1) La composante du vecteur somme dans E_i étant la somme des composantes :

$$\mathrm{pr}_i(x + y) = \mathrm{pr}_i(x) + \mathrm{pr}_i(y),$$

et de même :

$$\mathrm{pr}_i(\lambda x) = \lambda \mathrm{pr}_i(x),$$

pr_i est linéaire.

$$\sum_{i=1}^{m} \mathrm{pr}_i(x) = x_1 + x_2 + \ldots + x_m = x,$$

donc :

$$\sum_{i=1}^{m} \mathrm{pr}_i = e_{\mathrm{E}} \qquad (1),$$

$\mathrm{pr}_{i'}(x) = x_{i'}$; $\mathrm{pr}_i(x_{i'})$ est nulle puisque la $i^{\text{ème}}$ composante de $x_{i'}$ est nulle :

$$\text{Si} \quad i' \neq i \qquad \mathrm{pr}_i \circ \mathrm{pr}_{i'} = 0.$$

Au contraire, $\mathrm{pr}_i \circ \mathrm{pr}_i = \mathrm{pr}_i$.

2) L'égalité (1) permet d'écrire :

$$f = e_{\mathrm{F}} \circ f \circ e_{\mathrm{E}} = \sum_j \mathrm{pr}_j^{\mathrm{F}} \circ f \circ \sum_i \mathrm{pr}_i^{\mathrm{E}}$$

Mais la composition des applications étant distributive par rapport à la somme (à droite et à gauche), ceci donne :

$$f = \sum_{i,j} \mathrm{pr}_j^{\mathrm{F}} \circ f \circ \mathrm{pr}_i^{\mathrm{E}}$$

Posons : $f_i^j = \mathrm{pr}_j^{\mathrm{E}} \circ f \circ \mathrm{pr}_i^{\mathrm{E}}$.

On a donc :

$$f = \sum_{i,j} f_i^j$$

f_i^j est une application linéaire qui est nulle sur E_i' et qui peut être définie par sa restriction à E_i ; l'image de E par cette application étant dans F_j, f_i^j est définie par la donnée d'un élément φ_i^j de $\mathcal{L}(\mathrm{E}_i, \mathrm{F}_j)$.

Considérons alors les espaces $\mathcal{L}(\mathrm{E}_i, \mathrm{F}_j)$ et leur somme directe ; soit

$$\varphi \in \bigoplus_{i,j} \mathcal{L}(\mathrm{E}_i, \mathrm{F}_j),$$

φ ayant pour tout couple (i, j), pour composante dans $\mathcal{L}(\mathrm{E}_i, \mathrm{F}_j)$ l'élément φ_i^j qui définit f_i^j.

Il est clair que l'application :

$$f = \sum_{i,j} f_i^j \in \mathcal{L}(\mathrm{E}, \mathrm{F}) \longrightarrow \varphi = \sum_{i,j} \varphi_i^j \in \bigoplus \mathcal{L}(\mathrm{E}_i, \mathrm{F}_j)$$

est bijective. Il est clair aussi que c'est un isomorphisme (car dans $f + g$ les composantes de mêmes indices s'ajoutent comme dans $\varphi + \gamma$; et dans λf elles sont multipliées par λ comme dans $\lambda\varphi$). Donc,

$$\mathcal{L}(\mathrm{E}, \mathrm{F}) \approx \bigoplus_{i,j} \mathcal{L}(\mathrm{E}_i, \mathrm{F}_j).$$

3) La démonstration précédente est, dans le cas des sommes directes infinies, fausse dès l'égalité (1), car on ne peut considérer $\Sigma\, \mathrm{pr}_i$ qui porterait sur un nombre infini d'indices. On peut voir, par ailleurs, que le résultat est également faux. En effet, soit encore :

$$\bigoplus_{i,j} \mathcal{L}(\mathrm{E}_i, \mathrm{F}_j)$$

et soit φ un élément de cette somme directe. Il lui correspond comme précédemment une application f de $\mathcal{L}$ (E, F), somme des applications f_i^j dont chacune a pour restriction à E_i la composante φ_i^j de φ. L'ensemble de ces applications f constitue un sous-espace de $\mathcal{L}$ (E, F) isomorphe à $\oplus \mathcal{L}(E_i, F_j)$. Mais φ n'a qu'un nombre fini de composantes non nulles. Donc, seul un nombre fini de sous-espaces E_i ne seront pas envoyés sur zéro par l'application f. Si donc on considère une application g de E dans F qu'on définit en envoyant sur des vecteurs de F différents de zéro tous les éléments des bases de tous les E_i (dont la réunion constitue une base de E), cette application g ne pourra appartenir au sous-espace des applications f qui est donc un vrai sous-espace de $\mathcal{L}$ (E, F).

4) D'après la deuxième question :

$$g \circ f = \Big(\sum_{k,j} \mathrm{pr}_k^G \circ g \circ \mathrm{pr}_j^F\Big) \circ \Big(\sum_{i,j'} \mathrm{pr}_{j'}^F \circ f \circ \mathrm{pr}_i^E\Big)$$

Mais les propriétés de distributivité permettent d'écrire ceci :

$$g \circ f = \sum_{k,j,j',i} \mathrm{pr}_k^G \circ g \circ \mathrm{pr}_j^F \circ \mathrm{pr}_{j'}^F \circ f \circ \mathrm{pr}_i^E$$

Mais $\mathrm{pr}_j^F \circ \mathrm{pr}_{j'}^F = 0$ si $j \neq j'$. Seuls dans la somme précédente ne sont pas nuls les termes où $j = j'$. Cette somme se réduit donc à :

$$g \circ f = \sum_{k,j,i} \mathrm{pr}_k^G \circ g \circ \mathrm{pr}_j^F \circ f \circ \mathrm{pr}_i^E,$$

ou encore :

$$\sum_{k,i} \sum_j (\mathrm{pr}_k^G \circ g \circ \mathrm{pr}_j^F) \circ (\mathrm{pr}_j^F \circ f \circ \mathrm{pr}_i^E)$$

la deuxième parenthèse représente le f_i^j de la première question. On peut de même poser :

$$\mathrm{pr}_k^G \circ g \circ \mathrm{pr}_j^F = g_j^k$$

et on aura :

$$(g \circ f) = \sum_{k,i} \sum_j g_j^k \circ f_i^j.$$

Si on introduit à son tour la composante de $g \circ f$ relative à E_i et G_k (restriction de $g \circ f$ à E_i, à valeurs dans G_k), on aura donc :

$$(g \circ f)_i^k = \sum_j g_j^k \circ f_i^j \qquad (2)$$

Supposons alors que E, F, G soient rapportés à des bases, réunions des bases des E_i (resp : F_j, G_k) ; la matrice de l'application f_i^j est formée par la matrice de φ_i^j (notation de la deuxième question) entourée de zéros. Les matrices de g_j^k et de $(g \circ f)_i^k$ seront de la même forme. La matrice de f, somme de telles matrices, sera constituée par la juxtaposition des blocs identiques aux matrices des φ_i^j. Les matrices de g et $g \circ f$ sont constituées de la même façon et la formule (2) est la formule qui donne la valeur d'un bloc de la matrice produit en fonction des blocs des matrices facteurs. Si on appelle M la matrice de f, N celle de g et P celle du produit et si on désigne les blocs par les mêmes indices que les sous-espaces correspondants, on a :

$$P_i^k = \sum_j N_j^k M_i^j.$$

On voit que cette formule est une généralisation de celle qui donne un élément de la matrice produit en fonction des éléments des matrices facteurs.

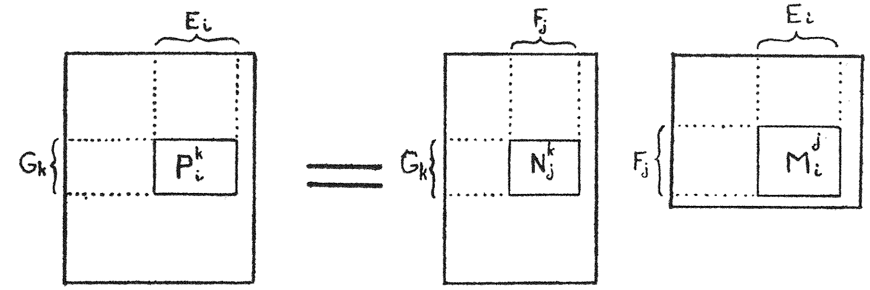


FIG. 11

Exercice 19.

Le plan P et la droite D sont des sous-espaces vectoriels supplémentaires puisque la somme de leur dimension est 3 et leur intersection est réduite au vecteur nul. Les deux projections envisagées sont donc des projections au sens plus général envisagé dans l'exercice précédent et sont, par conséquent, des applications linéaires. Les deux théorèmes : « La projection de la somme est la somme des projections » et « La projection du vecteur λV est égale au produit par λ de la projection du vecteur V », ne sont d'ailleurs que la traduction en langage géométrique de cette linéarité.

Soit un vecteur d'origine O et d'extrémité M de coordonnées x, y, z. Les points de la droite parallèle à D passant par M ont des coordonnées de la forme :

$$x + \alpha\rho,\ y + \beta\rho,\ z + \gamma\rho$$

et le point où cette droite rencontre P a pour coordonnées :

$$x' = \frac{(v\beta + w\gamma)x - v\alpha y - w\alpha z}{u\alpha + v\beta + w\gamma} = \frac{1}{k}[v\beta + w\gamma)x - v\alpha y - w\alpha z]$$

(en posant $k = u\alpha + v\beta + w\gamma$), y' et z' s'en déduisant par permutation circulaire de u, v, w et α, β, γ. On en déduit :

$$A = \frac{1}{k}\begin{pmatrix} v\beta + w\gamma & -v\alpha & -w\alpha \\ -u\beta & w\gamma + u\alpha & -w\beta \\ -u\gamma & -v\gamma & u\alpha + v\beta \end{pmatrix}$$

Le plan parallèle à P passant par M a pour équation :

$$u(X - x) + v(Y - y) + w(Z - z) = 0$$

et son intersection avec D a des coordonnées x'', y'', z'' telles que :

$$\frac{x''}{\alpha} = \frac{y''}{\beta} = \frac{z''}{\gamma} = \frac{ux + vy + wz}{k}$$

On en déduit la matrice :

$$B = \frac{1}{k}\begin{pmatrix} \alpha u & \alpha v & \alpha w \\ \beta u & \beta v & \beta w \\ \gamma u & \gamma v & \gamma w \end{pmatrix}$$

Géométriquement, il est évident que $AB = BA = 0$, puisque la composition des deux projections envoie sur O tout vecteur de l'espace. C'est bien ce que l'on retrouve en appliquant la règle du produit des matrices.

Exercice 20.

$\varphi(x^n) = nx^{n-1}$. Donc, la matrice M de φ relativement à B_0 s'écrit :

$$M = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & 0 \cdots & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & n \\ 0 & & & & 0 & 0 \end{pmatrix}$$

Les polynômes p_k étant de degré k avec $0 \leqslant k \leqslant n$, leur ensemble constitue bien une base (voir note III, 5, 2).

La matrice de passage de B_1 à B_0, dont la $(k+1)^{\text{ème}}$ colonne donne les coordonnées de p_k dans la base B_0, s'écrit :

$$P = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & \vdots & \vdots & 0 & 1 \end{pmatrix}$$

Pour trouver P^{-1}, matrice de passage de B_0 à B_1, observons que :

$$\forall k \neq 0 \qquad x^k = p_k - p_{k-1} \qquad x^0 = p_0$$

ce qui donne :

$$P^{-1} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & -1 \\ 0 & 0 & \vdots & & 1 \end{pmatrix}$$

La matrice M' de φ dans la base B_1 est $M' = P^{-1}MP$. On trouve :

$$MP = \begin{pmatrix} 0 & 1 & 1 & \cdots & \cdots & 1 \\ 0 & 0 & 2 & 2 & \cdots & 2 \\ 0 & 0 & 0 & 3 & 3 \cdots & 3 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & n \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix} \qquad M' = \begin{pmatrix} 0 & 1 & -1 & -1 & \cdots & -1 \\ 0 & 0 & 2 & -1 & -1 \cdots & -1 \\ 0 & 0 & 0 & 3 & -1 \cdots & -1 \\ \vdots & & & \ddots & \ddots & -1 \\ \vdots & & & & 0 & n \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

Vérification : cherchons directement M', c'est-à-dire cherchons à exprimer $\varphi(p_k)$ en fonction des polynômes $p_{k'}$. La formule :

$$\varphi(p_k) = kp_{k-1} - \sum_{1}^{k-2} p_{k'}$$

se démontre aisément par récurrence sur k. C'est bien elle que traduit la matrice M'.

Exercice 21.

Les opérations effectuées sur la matrice M pour obtenir une matrice triangulaire équivalente de rang p sont l'échange des lignes, ou des colonnes (changement dans l'ordre des éléments de la base de E, ou de la base de F), et le remplacement d'une ligne (ou d'une colonne) par une combinaison linéaire à coefficients dans K de lignes (ou de colonnes). Mais, puisque $K \subset K'$, ces combinaisons sont aussi des combinaisons linéaires à coefficients dans K' ; par suite, en tant que représentant une application de K'^n dans K'^m, la matrice est équivalente à une matrice triangulaire qui a exactement p éléments non nuls sur la diagonale principale. Elle est donc de rang p.

Ce résultat peut paraître *a priori* difficilement conciliable avec le fait que la dimension d'un espace vectoriel sur un corps K dépend du corps. Il importe de remarquer que si f et f' sont deux des applications représentées par M suivant qu'on la considère comme matrice à coefficients dans K ou comme matrice à coefficients dans K', f et f' sont définies dans deux espaces E et E' différents et prennent leurs valeurs dans des espaces différents F et F'. L'égalité des rangs exprime que $f(E)$ est un sous-espace de F de dimension p sur K, et que $f'(E')$ est un sous-espace de F' de dimension p sur K'.

Exercice 22.

Soit $\{a_i\}$ une base de E, $\{a^i\}$ la base duale de E^* définie par :

$$\begin{matrix} & a^i(a_i) = 1 \\ \text{pour} \quad j \neq i & a^i(a_j) = 0 \end{matrix}$$

Définissons la base duale $\{\alpha_i\}$ de E^{**}. Elle est constituée des n formes linéaires qui satisfont pour tout i à :

$$\begin{matrix} & \alpha_i(a^i) = 1 \\ \text{pour} \quad j \neq i & \alpha_i(a^j) = 0 \end{matrix} \qquad (1)$$

Or, l'élément $\tilde{a}_i$ de E^{**}, image de a_i dans l'isomorphisme canonique, est défini par :

$$\forall y^* \in E^* \qquad \tilde{a}_i(y^*) = y^*(a_i)$$

ce qui donne, en faisant décrire $\{a^i\}$ à y^* :

$$\begin{matrix} \text{pour} \quad j \neq i & \tilde{a}_i(a^j) = a^j(a_i) = 0 \\ & \tilde{a}_i(a^i) = a^i(a_i) = 1 \end{matrix} \qquad (2)$$

D'où, en rapprochant (1) et (2) qui prouvent que les deux formes linéaires $\tilde{a}_i$ et α_i prennent la même valeur pour tous les éléments de la base duale $\{a^i\}$:

$$\alpha_i = \tilde{a}_i$$

et, comme il en est ainsi pour tout i, $\{\alpha_i\}$ est bien l'image de $\{a_i\}$ dans l'isomorphisme canonique.

Exercice 23.

1° Montrer que l'ensemble des formes coordonnées a^i (i appartenant à un ensemble $\mathcal{I}$ infini d'indices) est partie libre, c'est montrer qu'une combinaison linéaire finie :

$$\sum_{i \in I} \lambda_i a^i \qquad \text{(I ensemble fini d'indices} \quad I \subset \mathcal{I})$$

ne peut être nulle que si tous les λ_i le sont.

Or
$$\sum_{i \in I} \lambda_i a^i = 0$$

veut dire
$$\forall x \in E \quad \sum_{i \in I} \lambda_i a^i(x) = 0.$$

Prenons $x = a_j$, j appartenant à I :

$$\sum_{i \in I} \lambda_i a^i(a_j) = \lambda_j,$$

d'où on déduit $\lambda_j = 0$; ce raisonnement pouvant être repris pour tout $j \in I$, l'ensemble des a^i est une partie libre.

Etudions le sous-espace de E* engendré par les formes coordonnées ; tous ses éléments sont des combinaisons linéaires finies des formes, du type envisagé ci-dessus. Soit x^* une de ces formes.

On aura
$$x^*(a_j) = \lambda_j \neq 0$$
pour tous les indices j d'une partie finie I de $\mathcal{I}$

et
$$x^*(a_j) = 0 \qquad \text{si } j \notin I.$$

Seul un nombre fini d'éléments de la base auront donc une image par x^* différente de zéro.

Si nous considérons alors une forme linéaire y^* définie par :

$$\forall i \in I \qquad y^*(a_i) = 1,$$

y^* ne peut être exprimée comme combinaison linéaire finie des formes coordonnées.

L'ensemble des a^i n'engendre donc pas E*.

Exemple : Pour E, prenons K$[x]$; la forme linéaire y^* envisagée ci-dessus est celle qui, à tout polynôme $P \in K[x]$, fait correspondre sa valeur pour $x = 1$ puisqu'à tout monôme elle fait correspondre son coefficient.

2° L'isomorphisme canonique de E sur $\tilde{E}$ donne pour image de la base (a_i) de E une base $(\tilde{a}_i)$ de $\tilde{E}$, les applications $\tilde{a}_i$ vérifiant :

$$\tilde{a}_i(a^j) = a^j(a_i) = \delta_i^j.$$

Par suite, tout élément de $\tilde{E}$ est une forme linéaire sur E* qui s'annule pour tous les a^j sauf un nombre fini d'entre eux. Or, les a^i constituant une partie libre de E*, il existe des formes linéaires sur E* prenant des valeurs non nulles sur une infinité de a^j. $\tilde{E}$ est donc strictement inclus dans E**.

Exercice 24.

Nous allons chercher des automorphismes u, pour lesquels l'égalité

$$< x, \varphi(y) > = < u(x), (\varphi \circ u)(y) > \qquad (1)$$

n'est pas vérifiée pour tout x et tout y de E.

Nous donnerons plusieurs contre-exemples différents parce qu'ils n'ont pas le même domaine d'application.

1° Prenons pour automorphisme u l'homothétie définie par :

$$u(x) = \lambda x \qquad \lambda \in K \quad \lambda \neq 0.$$

Le deuxième membre de l'égalité (1) s'écrit :

$$< u(x), (\varphi \circ u)(y) > = < \lambda x, \varphi(\lambda y) >.$$

Mais φ, devant être un isomorphisme pour une structure d'espace vectoriel, est linéaire donc :

$$\varphi(\lambda y) = \lambda \varphi(y).$$

D'autre part, $\varphi(y)$ est une forme linéaire, donc :

$$< \lambda x, \varphi(y) > = \lambda < x, \varphi(y) >$$

et
$$< \lambda x, \varphi(\lambda y) > = \lambda < \lambda x, \varphi(y) > = \lambda^2 < x, \varphi(y) >.$$

On devrait donc, pour que (1) soit réalisée, avoir pour tout x et tout y :

$$< x, \varphi(y) > = \lambda^2 < x, \varphi(y) >,$$

ce qui exige que $\lambda^2 = 1$.

Dans tout espace vectoriel sur un corps K, où tous les éléments n'ont pas 1 pour carré, il est donc possible de trouver une homothétie qui contredise l'égalité (1).

Les seuls corps K où tous les éléments ont pour carré 1 sont $\mathbf{Z}_2$ et $\mathbf{Z}_3$. (En effet, l'équation $\lambda^2 = 1$ ne peut avoir plus de deux solutions ; un corps où elle est toujours vérifiée ne peut avoir plus de deux éléments différents de zéro).

L'impossibilité de trouver φ vérifiant l'égalité (1) est donc établie pour tous *les espaces vectoriels sur des corps autres que* $\mathbf{Z}_2$ *et* $\mathbf{Z}_3$.

2° Soit un couple x, y. Choisissons un automorphisme u qui satisfasse à :

$$u(x) = \lambda x, \qquad u(y) = y \qquad \lambda \in K \quad \lambda \neq 0.$$

Un tel choix de u est possible si x et y ne sont pas colinéaires.

$$< u(x), (\varphi \circ u)(y) > = <\lambda x, \varphi(y) > = \lambda < x, \varphi(y) >.$$

L'égalité (1) ne peut donc être vérifiée s'il existe dans K un λ différent de 1, c'est-à-dire, cette fois, si K est différent de $\mathbf{Z}_2$.

Mais, comme cette démonstration a supposé qu'on pouvait choisir x et y non colinéaires, elle n'est donc valable que dans *les espaces vectoriels de dimension supérieure à 1 sur des corps autres que* $\mathbf{Z}_2$.

3° $y \neq 0$ et φ étant choisis, considérons l'hyperplan $[\varphi(y)]^\perp$. Choisissons x hors de cet hyperplan, ce qui est toujours possible puisque, φ étant un isomorphisme, $\varphi(y)$ n'est pas le zéro du dual, donc ne s'annule pas sur tout E. On a donc :

$$x \notin [\varphi(y)]^\perp, \quad \text{c'est-à-dire} \quad < x, \varphi(y) > \neq 0.$$

Ceci posé, choisissons l'automorphisme u tel que :

$$u(y) = y \qquad u(x) \in [\varphi(y)]^\perp.$$

Si un tel choix est possible, on aura :

$$< u(x), (\varphi \circ u)(y) > = < u(x), \varphi(y) > = 0.$$

L'égalité (1) est contredite par cet automorphisme.

Reste à vérifier qu'un tel choix est possible. Pour qu'il le soit, il faut qu'à x et y non colinéaires il fasse correspondre $u(x)$ et y non colinéaires. Ce sera toujours possible si l'hyperplan est à deux dimensions au moins, même si cet hyperplan contient y. L'hyperplan sera à deux dimensions si l'espace est à trois dimensions (on ne peut pas savoir s'il ne serait pas possible de trouver φ tel que $\forall y,\ y \in [\varphi(y)]^{\perp}$, ce qui rendrait impossible le choix de u utilisé, dans le cas où $[\varphi(y)]^{\perp}$ serait de dimension 1).

La démonstration est donc valable pour *les espaces de dimension supérieure à 2*. Les domaines de validité des trois démonstrations sont schématisés ci-dessous. Quant aux trois cases restées vierges, on peut montrer qu'elles correspondent effectivement à trois cas d'exception.

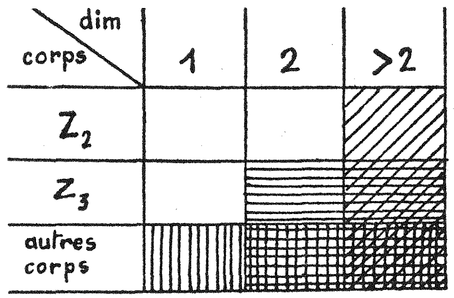


FIG. 12

Z_2, n'ayant qu'un élément différent de zéro, ne possède pas d'automorphisme autre que l'identité et le problème ne s'y pose pas.

Z_3 possède deux éléments, 1 et 2, différents de 0, et par conséquent un seul automorphisme u différent de l'identité :

$$u : \begin{array}{l} 1 \longrightarrow 2 \\ 2 \longrightarrow 1 \end{array}$$

Il est clair que cet automorphisme vérifie la condition trouvée lors de notre première démonstration et que l'égalité (1) est vérifiée quel que soit φ (il y a, en fait, deux isomorphismes φ possibles).

Soit enfin Z_2^2 qui possède une base à deux éléments e_1, e_2 et qui possède en tout les quatre éléments 0, e_1, e_2, $e_1 + e_2 = e_3$. Son dual admet la base duale e^1, e^2 et les quatre éléments 0, e^1, e^2, $e^1 + e^2 = e^3$. Considérons l'isomorphisme φ de E dans E* défini par $\varphi(e_1) = e^2$ et $\varphi(e_2) = e^1$, ce qui entraîne $\varphi(e_3) = e^3$.

Nous constatons que

$$\begin{aligned} < e_1, \varphi(e_1) > &= < e_1, e^2 > = 0 \\ < e_2, \varphi(e_2) > &= < e_2, e^1 > = 0 \\ < e_3, \varphi(e_3) > &= < e_1 + e_2, e^1 + e^2 > \\ &= 1 + 1 = 0 \end{aligned}$$

ce qui peut se résumer par :

$$< e_i, \varphi(e_i) > = 0 \quad \text{pour} \quad i = 1, 2, 3$$

ou en disant que l'hyperplan relatif à tout vecteur de Z_2^2 contient ce vecteur. Nous sommes dans le cas d'exception signalé où la 3e démonstration ne s'appliquait pas.

D'autre part, un tel hyperplan étant à une dimension ne contient qu'un seul vecteur différent de zéro (car tout sous-espace de Z_2^2 à une dimension ne contient que zéro et un élément différent de zéro). Donc,

$$\text{si} \quad j \neq i \quad e_j \notin \varphi(e_i)^{\perp}$$
$$< e_j, \varphi(e_i) > \neq 0 \quad \text{donc} \quad < e_j, \varphi(e_i) > = 1.$$

Remarquons alors que u étant un automorphisme, donc étant bijectif, il fera correspondre à x et y égaux deux éléments $u(x)$ et $u(y)$ égaux, à x et y différents deux éléments $u(x)$ et $u(y)$ différents. Dans le premier cas on aura :

$$\begin{aligned} < x, \varphi(y) > &= < e_i, \varphi(e_i) > = 0 \\ < u(x), (\varphi \circ u)(y) > &= < e_{i'}, \varphi(e_{i'}) > = 0. \end{aligned}$$

Dans le deuxième :

$$\begin{aligned} < x, \varphi(y) > &= < e_i, \varphi(e_j) > = 1 \\ < u(x), (\varphi \circ u)(y) > &= < e_{i'}, \varphi(e_{j'}) > = 1, \end{aligned}$$

donc toujours $< x, \varphi(y) > = < u(x), (\varphi \circ u)(y) >$.

Exercice 25.

Pour qu'un automorphisme u satisfasse à :

$$< x, \varphi(y) > = < u(x), \varphi \circ u(y) > \qquad (1)$$

pour tout x il est suffisant qu'il y satisfasse pour x décrivant la base $\{ a_i \}$ de E du fait de la linéarité de la forme bilinéaire par rapport à x. Mais φ étant linéaire, la forme bilinéaire $< x, \varphi(y) >$ est aussi linéaire par rapport à y. Si (1) est vérifiée pour y décrivant $\{ a_i \}$ elle le sera pour tout y. Il suffit donc de vérifier :

$$< a_i, \varphi(a_j) > = < u(a_i), (\varphi \circ u)(a_j) > \qquad (2).$$

Calculons les deux membres de cette égalité. Pour le premier, il vient :

$$< a_i, \varphi(a_j) > = < a_i, a^j > = \delta_i^j$$

(symbole de Kronecker).

Donnons-nous maintenant u par sa matrice, c'est-à-dire par les coefficients u_i^l, en posant :

$$u(a_i) = \sum_l u_i^l a_l.$$

On a de même :

$$u(a_j) = \sum_k u_j^k a_k$$

$$(\varphi \circ u)(a_j) = \sum_k u_j^k a^k.$$

D'où $< u(a_i), (\varphi \circ u)(a_j) > = < \sum_l u_i^l a_l, \sum_k u_j^k a^k > = \sum_l \sum_k u_i^l u_j^k a^k(a_l)$.

Mais $a^k(a_l) = \delta_l^k$; donc, dans cette somme de n^2 termes (si dim E $= n$), seuls ne sont pas nuls ceux où $l = k$. L'expression précédente vaut donc $\sum_l u_i^l u_j^l$. Pour que l'égalité (2) soit vérifiée, il faut et il suffit donc que $\sum_l u_i^l u_j^l = \delta_i^j$, c'est-à-dire que les automorphismes u répondant à la question sont ceux qui vérifient :

$$\forall i \qquad \sum_l (u_i^l)^2 = 1$$

$$\forall i \ \forall j \neq i \qquad \sum_l u_i^l u_j^l = 0 \ (*).$$

(*) Les notations utilisées dans le corrigé de l'exercice 25 ne respectent pas les principes énoncés en IV, 4, à cause de l'égalité $\varphi\ (a_i) = a^i$ qu'il faudrait écrire

Exercice 26.

Constituons une base de E en prenant une base $\{a_1 \dots a_h\}$ de H et en la complétant par une base $\{b_1 \dots b_{n-h}\}$ d'un supplémentaire de H.

$H^\perp$ est l'ensemble des éléments du dual qui s'annulent sur tout H. Pour qu'un élément du dual appartienne à $H^\perp$, il est donc nécessaire et suffisant qu'il s'annule pour $a_1 \dots a_h$. Considérons une forme linéaire $x^* \in E^*$. Elle peut s'écrire :

$$x^* = \sum_{i=1}^{h} x_i a^i + \sum_{j=1}^{n-h} y_j b^j.$$

Nous avons $x^*(a_i) = x_i$. Pour que x^* s'annule pour tous les a_i, il est donc nécessaire et suffisant que pour tout i, $x_i = 0$, donc que x^* appartienne au sous-espace engendré par $\{b^1, b^2 \dots b^{n-h}\}$. $H^\perp$ est donc ce sous-espace. Il est de dimension $n-h$, donc $\dim H^\perp = \operatorname{codim} H$.
Ceci entraîne aussi $\operatorname{codim} H^\perp = h = \dim H$.

Appliquons cette propriété à $A^{\perp\perp\perp}$; nous avons :

$$\dim A^{\perp\perp\perp} = \operatorname{codim} A^{\perp\perp}, \quad \text{mais} \quad \operatorname{codim} A^{\perp\perp} = \dim A^\perp,$$
$$\text{donc } \dim A^{\perp\perp\perp} = \dim A^\perp.$$

Or, la définition du biorthogonal d'un ensemble entraîne qu'il est contenu dans cet ensemble :

$$A^{\perp\perp\perp} \subset A^\perp.$$

Deux sous-espaces vectoriels dont l'un est inclus dans l'autre ne peuvent avoir même dimension que s'ils sont confondus. D'où :

$$A^{\perp\perp\perp} = A^\perp.$$

Si A n'est plus une partie quelconque mais un sous-espace vectoriel, le même raisonnement montrera que :

$$\dim H^{\perp\perp} = \dim H \quad \text{avec} \quad H^{\perp\perp} \subset H, \quad \text{donc} \quad H^{\perp\perp} = H.$$

Exercice 27.

Observons d'abord que si E *est un espace de dimension finie*, les égalités s'obtiennent immédiatement. En effet, dans ce cas, la relation d'orthogonalité est une bijection entre sous-espaces vectoriels de E et E* et

$$A \subset B \implies A^\perp \supset B^\perp \qquad (1).$$

$\varphi(a_i) = \sum_l \alpha_{il} a^l$ où les α_{il}, coefficients de la matrice de φ, vérifient $\alpha_{il} = 0$ si $i \neq l$ et $\alpha_{ii} = 1$.

On vérifie aisément que la formule

$$\sum_l u_i^l u_j^l = \delta_i^j$$

s'écrirait alors

$$\sum_{m,k,l} \alpha_{mk} u_i^l u_j^m \delta_l^k = \sum_l \alpha_{jl} \delta_i^l$$

Cette dernière formule aurait son intérêt dans une théorie générale qui voudrait faire constamment appel aux coordonnées ; elle est inutilement compliquée pour traiter le seul cas particulier étudié dans l'exercice 25.

Or, l'ensemble $\mathcal{E}$ des sous-espaces vectoriels de E et l'ensemble $\mathcal{E}^*$ des sous-espaces vectoriels de E*, ordonnés par inclusion, forment deux treillis avec :

$$\forall H, K \in \mathcal{E} \quad \begin{array}{l} \sup(H, K) = H + K \\ \inf(H, K) = H \cap K \end{array}$$

et les relations analogues dans $\mathcal{E}^*$.

La relation (1) entraîne qu'un majorant d'un élément quelconque A de E a pour orthogonal un minorant de $A^\perp$; le plus petit des majorants de deux éléments a donc pour orthogonal le plus grand des minorants de leurs orthogonaux, soit :

$$(H+K)^\perp = H^\perp \cap K^\perp$$

et le plus grand des minorants a pour orthogonal le plus petit des majorants des orthogonaux :

$$(H \cap K)^\perp = H^\perp + K^\perp.$$

Mais si E *est de dimension infinie*, on ne peut plus identifier $A^{\perp\perp}$ et $A^\perp$ et le raisonnement précédent ne s'applique plus.

Comparons $(H+K)^\perp$ et $H^\perp \cap K^\perp$.

$$(H+K)^\perp = \{x^* ;\ \forall x \in H \quad \forall y \in K \quad x^*(x+y) = 0\}.$$

Ceci entraîne évidemment, en faisant $y = 0$:

$$\forall x \in H \qquad x^*(x) = 0,$$

et de même :

$$\forall y \in K \qquad x^*(y) = 0,$$

d'où $\qquad x^* \in H^\perp \quad x^* \in K^\perp,$

et par conséquent :

$$x^* \in (H+K)^\perp \implies x^* \in H^\perp \cap K^\perp.$$

Soit maintenant $y^* \in H^\perp \cap K^\perp$.
Ceci implique $\quad \forall x \in H \quad y^*(x) = 0,$
$\qquad \forall y \in K \quad y^*(y) = 0,$
et par conséquent $\forall (x+y) \in H + K \quad y^*(x+y) = 0.$
D'où $\quad y^* \in (H+K)^\perp,$

$$y^* \in H^\perp \cap K^\perp \implies y^* \in (H+K)^\perp.$$

On peut donc conclure :

$$H^\perp \cap K^\perp = (H+K)^\perp.$$

Comparons maintenant $(H \cap K)^\perp$ et $H^\perp + K^\perp$.
Le premier peut s'écrire :

$$(H \cap K)^\perp = \{x^* \in E^* ;\ \forall x \in H \cap K \quad x^*(x) = 0\}$$

et le deuxième :

$$H^\perp + K^\perp = \{x^* \in E^* ;\ \exists y^*, \exists z^*, x^* = y^* + z^*, \forall x \in H\ y^*(x) = 0, \forall x \in K\ z^*(x) = 0\}.$$

Or, cette dernière propriété des x^* entraîne :

$$\forall x \in H \cap K \qquad x^*(x) = 0,$$

donc $\qquad H^\perp + K^\perp \subset (H \cap K)^\perp.$

Soit maintenant x^* une forme appartenant à $(H \cap K)^\perp$. Elle est définie par ses valeurs pour tous les éléments d'une base de E. Nous pouvons constituer une telle base en réunissant : une base $\{e_i\}$ de $H \cap K$, une base $\{f_j\}$ d'un supplémentaire de $H \cap K$ par rapport à H, une base $\{g_k\}$ d'un supplémentaire de $H \cap K$ par rapport à K, et l'espace vectoriel engendré par la réunion de ces trois bases étant $H + K$, on complétera par une base $\{h_l\}$ d'un supplémentaire de $H + K$ par rapport à E.

Si $x^* \in (H \cap K)^\perp \quad \forall i \quad x^*(e_i) = 0$.

Nous pourrons décomposer x^* de la façon suivante :

$$x^* = y^* + z^*,$$

y^* et z^* satisfaisant à :

$$\forall j \quad y^*(f_j) = x^*(f_j) \qquad \forall k \quad y^*(g_k) = 0$$
$$\forall j \quad z^*(f_j) = 0 \qquad \forall k \quad z^*(g_k) = x^*(g_k)$$
$$\forall i \quad y^*(e_i) = z^*(e_i) = 0.$$

Les valeurs $y^*(h_l)$ et $z^*(h_l)$ sont choisies arbitrairement de telle sorte qu'elles aient pour somme $x^*(h_l)$; on a : $y^* \in K^\perp \quad z^* \in H^\perp$, donc $x^* \in H^\perp + K^\perp$.

$$(H \cap K)^\perp \subset H^\perp + K^\perp$$

l'égalité $(H \cap K)^\perp = H^\perp + K^\perp$ est démontrée.

Cas où $E = H \oplus K$: Dans ce cas, $H \cap K = \{0\}$; $(H \cap K)^\perp$, ensemble de toutes les formes linéaires qui s'annulent pour zéro, est E^* tout entier. Donc :

$$H^\perp + K^\perp = E^*.$$

D'autre part, $(H + K)^\perp = E^\perp$ est réduit au zéro de E^*.

Donc, $H^\perp \cap K^\perp = \{0\}$. On peut conclure :

$$E^* = H^\perp \oplus K^\perp.$$

Exercice 28.

1° Etant donné un sous-espace V d'un espace vectoriel sur un corps K, nous avons vu dans l'exercice 5 un procédé qui permet, connaissant un supplémentaire de V, d'en obtenir une infinité d'autres (si K est infini). Nous allons montrer, dans le cas où V est de codimension 1, que ce procédé fournit tous les supplémentaires de V. Soit, en effet, Kb_1, le supplémentaire connu de V ; b_1 est un vecteur n'appartenant pas à V. Soit $\{b_i\}$, $i = 2, ..., n$ une base de V. Tout supplémentaire de V est de la forme Kb, où b est un vecteur qui n'appartient pas à V, donc s'écrit $b = \lambda^1 b_1 + a$, avec $a \in V$ et $\lambda^1 \neq 0$. Comme b n'est déterminé qu'à une homothétie près, on peut prendre $\lambda^1 = 1$. On obtiendra donc tous les sous-espaces supplémentaires de V sous la forme Kb avec $b = b_1 + \sum_{i=2}^{n} \lambda^i b_i$,

où les λ^i sont des scalaires arbitraires.

2° Soit H un supplémentaire de $\mathcal{C}$ par rapport à $\mathcal{P}^n$. Il résulte de l'exercice 27 que $\mathcal{C}^\perp \oplus H^\perp = \mathcal{P}_n^*$

et de l'exercice 26 que codim $\mathcal{C}^\perp = \dim \mathcal{C} = 1$.

Prenons dans $\mathcal{P}_n$ la base constituée des $n + 1$ monômes $1, x, ..., x^n$, et dans $\mathcal{P}_n^*$ la base duale constituée des formes coordonnées, a^i :

$$a^i(x^i) = 1 \text{ et si } j \neq i \quad a^i(x^j) = 0.$$

$\mathcal{C}^\perp$ est l'ensemble des formes qui s'annulent pour les constantes, c'est-à-

dire dont la coordonnée par rapport à a^0 est nulle. $\mathcal{C}^\perp$ est engendré par $a^1, ..., a^n$ et Ka^0 est un de ses supplémentaires. Tout supplémentaire $H^\perp$ de $\mathcal{C}^\perp$ est donc un sous-espace de $\mathcal{P}_n^*$ à une dimension engendré par une forme du type

$$a^0 + \sum_{1 \leq i \leq n} \lambda_i a^i,$$

où les λ_i sont des scalaires arbitraires.

L'orthogonal H de $H^\perp$ est l'ensemble des polynômes qui annule une telle forme, d'où le résultat :

Tout supplémentaire de $\mathcal{C}$ par rapport à $\mathcal{P}_n$ est constitué des polynômes dont les coefficients vérifient une relation du type

$$\alpha^0 + \sum_{1 \leq i \leq n} \lambda_i \alpha^i = 0,$$

les λ_i étant n scalaires fixes, arbitrairement choisis.

Remarque : Il résulte de l'exercice 10, que les polynômes s'annulant pour $x = b$ constituent un sous-espace supplémentaire de $\mathcal{C}$. Ceci correspond à $\lambda_i = b^i$, et rentre bien dans le cas général.

Exercice 29.

Il est évident que la condition est suffisante. Si, inversement, φ s'annule quand $\varphi_1, ..., \varphi_p$ sont nulles, cela signifie que l'orthogonal de φ contient l'orthogonal de l'ensemble $\{\varphi_1, ..., \varphi_p\}$, ce qui entraîne que le biorthogonal de φ, c'est-à-dire $K\varphi$, est inclus dans le biorthogonal de $\{\varphi_1, ..., \varphi_p\}$, c'est-à-dire le sous-espace engendré par $\{\varphi_1, ..., \varphi_p\}$, ce qui est équivalent au fait que φ s'écrive $\sum_{i=1}^{p} \lambda^i \varphi_i$.

On retrouve là les résultats relatifs aux faisceaux de droites, aux faisceaux et réseaux de plans de la géométrie analytique.

Si une base a été choisie dans E et si on a :

$$\varphi(x) = \Sigma\, \alpha_j x^j \qquad \varphi_i = \Sigma\, \alpha_{ij} x^j,$$

la condition nécessaire et suffisante pour que φ s'annule pour les vecteurs annulant les φ_i est qu'il existe des scalaires λ^i tels que pour $j = 1, 2, ..., n$:

$$\alpha_j - \sum_{i=1}^{p} \lambda^i \alpha_{ij} = 0.$$

C'est sous cette forme que ce résultat est souvent utilisé en Calcul des variations et en Mécanique (équations de Lagrange dans le cas de liaisons non holonomes).

Exercice 30.

1) Soit G un supplémentaire de H. Tout x de E étant somme d'un élément de G et d'un élément de H, il est clair que $f \in \mathcal{L}_H(E, F)$ est déterminée par sa restriction φ à G et que celle-ci peut être choisie arbitrairement. L'application

$$f \in \mathcal{L}_H(E, F) \longrightarrow \varphi \in \mathcal{L}(G, F)$$

est donc bijective, et étant linéaire est un isomorphisme. D'autre part, à tous les éléments d'une même classe de E (mod H), classe qui peut s'écrire :

$$a + H \qquad a \in G,$$

f fait correspondre l'élément $f(a) = \varphi(a)$.

On peut donc considérer l'application ψ de E/H dans F définie par :

$$\forall a \in G \qquad \psi(a + H) = \varphi(a).$$

L'application $\varphi \in \mathcal{L}(G, F) \longrightarrow \psi \in \mathcal{L}(E/H, F)$ apparaît alors comme bijective et, étant linéaire, est un isomorphisme.

2) Le fait que $E^* = G^{\perp} \oplus H^{\perp}$ a été établi dans l'exercice 27. On retrouve d'ailleurs directement ce résultat en remarquant que toute application f de E dans F est déterminée de manière unique par ses restrictions à H et à G, c'est-à-dire que f peut être décomposée de manière unique en somme de deux applications, l'une s'annulant sur H, l'autre s'annulant sur G, c'est-à-dire encore que :

$$\mathcal{L}(E, F) = \mathcal{L}_H(E, F) \oplus \mathcal{L}_G(E, F).$$

Si nous prenons $F = K$, les espaces vectoriels précédents deviennent respectivement E^*, $H^{\perp}$ et $G^{\perp}$.

D'autre part, le résultat de la première question devient :

$$H^{\perp} \approx \mathcal{L}(G, K) = G^*.$$

On montrerait de même que $G^{\perp} \approx H^*$.

3) ${}^tu(F^*)$ est le sous-espace de E^*, image par ${}^tu$ de F^*. Il est isomorphe à un supplémentaire du noyau de cette application, noyau dont on a montré (IV, 3, 3) qu'il était $[u(E)]^{\perp}$. Si donc nous posons :

$$F^* = [u(E)]^{\perp} \oplus L \qquad (1),$$

nous aurons

$${}^tu(F^*) \approx L \qquad (2).$$

Mais d'autre part on peut écrire : $F = u(E) \oplus N$, ce qui entraîne en vertu de la deuxième question :

$$F^* = u(E)^{\perp} \oplus N^{\perp} \qquad (3)$$

et

$$[u(E)]^* \approx N^{\perp} \qquad (4).$$

Mais d'autre part il suffit de rapprocher (1) et (3) pour avoir :

$$L \approx N^{\perp} \qquad (5).$$

Mais alors, (2), (5) et (4) permettent d'écrire :

$${}^tu(F^*) \approx [u(E)]^*.$$

Exercice 31.

Montrer que les matrices triangulaires forment une sous-algèbre de l'algèbre des matrices carrées d'ordre n, c'est montrer que leur ensemble est stable par rapport à l'addition, la multiplication interne, la multiplication par un scalaire. La première et la troisième de ces propriétés sont évidentes. Examinons ce qu'est l'élément a_i^j du produit de deux matrices triangulaires. Il est obtenu à partir de la $j^{\text{ème}}$ ligne de la matrice de gauche et la $i^{\text{ème}}$ colonne de celle de droite.

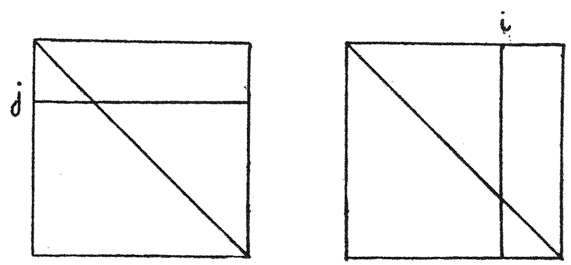


FIG. 13

La $j^{\text{ème}}$ ligne a $n - j$ éléments nuls, la $i^{\text{ème}}$ colonne a $i - 1$ éléments nuls. Les $n - j$ derniers et les $i - 1$ premiers produits partiels dont la somme constitue a_i^j sont nuls.

Si $n - j + i - 1 \geqslant n$, tous les produits partiels sont nuls et $a_i^j = 0$. Il en est ainsi si $i \geqslant j + 1$, c'est-à-dire sur tout le triangle supérieur. L'ensemble des matrices triangulaires est stable par rapport à la multiplication.

Interprétation par les endomorphismes : Considérons les matrices carrées d'ordre n comme des matrices d'endomorphismes d'un espace vectoriel à n dimensions. Dire que la matrice d'un endomorphisme est triangulaire, c'est dire que l'image de e_i par cet endomorphisme n'a de composantes non nulles que sur les e_j d'indice $j \geqslant i$. Si on compose cet endomorphisme avec un autre de la même famille, l'image de e_i continuera à n'avoir de composantes que sur les e_j d'indice $j \geqslant i$. La matrice de l'application composée sera donc aussi triangulaire.

Exercice 32.

1) $f(e_1) = e_1$, donc e_1 est un vecteur propre ; la valeur propre correspondante est 1. Cherchons si u admet un deuxième vecteur propre. Pour que $\mu e_1 + \nu e_2$ le soit, il faut qu'il soit colinéaire à :

$$\begin{aligned} f(\mu e_1 + \nu e_2) &= \mu e_1 + \nu(e_1 + \lambda e_2) \\ &= (\mu + \nu)\, e_1 + \nu\lambda e_2. \end{aligned}$$

Il faut donc que $\dfrac{\mu + \nu}{\mu} = \lambda$ ou $\mu(\lambda - 1) = \nu$.

Deux cas sont donc à distinguer :

a) $\lambda = 1$. L'égalité précédente exige $\nu = 0$. L'endomorphisme n'a pas d'autre vecteur propre que e_1.

b) $\lambda \neq 1$. On prend $\mu = 1$, $\nu = \lambda - 1$. Le vecteur $e_1 + (\lambda - 1)e_2$ est propre. La valeur propre correspondante est λ.

2) Une classe d'équivalence de E mod H (aussi appelée variété affine parallèle à H ; ce serait dans $\mathbf{R}^3$ un plan parallèle à xOy si H était le plan xOy) peut s'écrire $\dot{x} = x + H$ et son image est $u(\dot{x}) = u(x) + H$ qui est aussi une variété affine parallèle à H.

On sait que E/H a une structure d'espace vectoriel isomorphe à un supplémentaire de H, donc d'espace vectoriel à une dimension. L'application $\dot{x} \longrightarrow u(\dot{x})$ est un endomorphisme. En effet :

$$u(\widehat{x + y}) = u(x + y) + H = u(x) + u(y) + H = u(\dot{x}) + u(\dot{y})$$

$$u(\dot{\widehat{\lambda x}}) = u(\lambda x) + H = \lambda u(x) + H = \lambda \dot{u}(x).$$

Or, les seuls endomorphismes d'espaces vectoriels à une dimension sont les endomorphismes $u(\dot{x}) = \lambda \dot{x}$ où $\lambda \in K$. Etant donné $e_0 \notin H$, son image doit appartenir à :

$$u(\dot{e}_0) = \lambda \dot{e}_0 = \lambda e_0 + H.$$

On a donc :

$$u(e_0) = \lambda e_0 + e_1 \qquad e_1 \in H.$$

3) Le sous-espace Ke_0 de dimension 1 est un supplémentaire de H ; E est somme directe des trois sous-espaces Ke_0, Ke_1 et L et on peut écrire :

$$x = \xi_0 e_0 + \xi_1 e_1 + x_L.$$

Cherchons si un tel x peut être colinéaire à son image :

$$u(x) = \xi_0 \lambda e_0 + (\xi_0 + \xi_1)\, e_1 + x_L.$$

Si x_L n'était pas nul cela exigerait $u(x) = x$, donc $\xi_o = 0$, et x appartiendrait à H.

Il faut donc $x_L = 0$. On est alors ramené au problème du premièrement, où e_o joue le rôle de e_2 puisque $u(e_1) = e_1$, $u(e_o) = \lambda e_o + e_1$. On sait donc qu'il n'y a pas de vecteur propre autre que e_1 si $\lambda = 1$ et que, si $\lambda \neq 1$, le sous-espace $K((\lambda - 1)\, e_o + e_1)$ est invariant.

Si E est de dimension finie n et si on en constitue une base en prenant une base de H et le vecteur $(\lambda - 1)\, e_o + e_1$ comme $n^{\text{ième}}$ élément, la matrice aura la forme diagonale, les $(n - 1)$ premiers éléments de la diagonale valant 1 et le dernier λ.

Si u est une transvection différente de l'identité, aucun vecteur n'appartenant pas à H n'est propre. Il n'existe donc aucune base de vecteurs propres, c'est-à-dire aucune base par rapport à laquelle la matrice de u soit diagonale.

4) Dans le cas d'une transvection $u(e_o) = e_o + e_1$, $u(e_o) - e_o = e_1$. Un vecteur x quelconque peut être décomposé en un vecteur $\xi_o e_o$ et un vecteur $x_H \in$ H. Or, $u(x_H) = x_H$; donc, si on considère l'endomorphisme

$$x \longrightarrow u(x) - x,$$

l'image qu'il donne de x est celle de sa composante $\xi_o e_o$.
D'où $u(x) - x = \xi_o e_1$.

Or, si on considère une forme linéaire φ s'annulant sur H :

$$\varphi(x) = \varphi(\xi_o e_o) = \xi_o \varphi(e_o).$$

On voit que $u(x) - x$ lui est proportionnel. Il suffit de poser $a = \dfrac{e_1}{\varphi(e_o)}$ pour pouvoir écrire $u(x) = x + a\varphi(x)$. Le vecteur a (ou le vecteur e_1) définit donc complètement la transvection.

Si on prend pour base de E une base de H complétée par le vecteur e_o, la matrice de u prendra la forme ci-dessous, tous les éléments non indiqués étant égaux à zéro.

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

FIG. 14

5) *a*) Le produit de deux automorphismes qui laissent les vecteurs de H invariants est un automorphisme qui laisse les vecteurs de H invariants ; de même, l'inverse d'un automorphisme qui laisse les vecteurs de H invariants est un automorphisme qui laisse les vecteurs de H invariants.

b) Le produit de deux transvections est un automorphisme appartenant à G(H) qui laisse globalement invariante toute variété affine parallèle à H, donc est une transvection. De même, l'inverse d'une transvection sera une transvection. Les transvections forment un sous-groupe.

Montrons qu'il est invariant. Une dilatation d de rapport λ (qui

transforme la variété $e_o + \mathrm{H}$ en la variété $\lambda e_o + \mathrm{H}$) a pour inverse une dilatation de rapport $\dfrac{1}{\lambda}$. Le produit $d^{-1} \circ \tau \circ d$, où τ représente une transvection, transforme $e_o + \mathrm{H}$ en $e_o + \mathrm{H}$, car τ laisse globalement invariante la variété $\lambda e_o + \mathrm{H}$ que d^{-1} retransforme en $e_o + \mathrm{H}$. $d^{-1} \circ \tau \circ d$ est donc une transvection et $\Gamma(\mathrm{H})$ est invariant.

Soient τ et τ' deux transvections définies par la donnée du vecteur défini au 4). (a pour τ ; a' pour τ'). On a :

$$\forall x \in \mathrm{E} \qquad \tau(x) = x + a\, \varphi(x)$$
$$\tau'(\tau(x)) = \tau(x) + a'\, \varphi(\tau(x)).$$

Mais φ s'annulant sur H, $\varphi(\tau(x)) = \varphi(x)$ et

$$\tau'(\tau(x)) = x + (a + a')\varphi(x),$$

$\tau' \circ \tau$ est donc identique à la transvection définie par le vecteur $a + a'$. Si on considère l'application $\Gamma(\mathrm{H}) \longrightarrow \mathrm{H}$ faisant correspondre à une transvection son vecteur caractéristique, nous venons de montrer qu'à la composée de deux transvections correspond la somme des deux vecteurs. Le groupe $\Gamma(\mathrm{H})$ est donc isomorphe au groupe additif de H.

c) Dans G(H) la classe d'équivalence, mod $\Gamma(\mathrm{H})$, d'une dilatation d de rapport λ, est constituée par l'ensemble des dilatations $\tau \circ d$, où τ décrit $\Gamma(\mathrm{H})$. Ces dilatations font toutes correspondre à la variété $e_o + \mathrm{H}$, la variété $\lambda e_o + \mathrm{H}$.

Réciproquement, toute dilatation d' de rapport λ appartient à cet ensemble. En effet, soit d' une dilatation de rapport λ défini par :

$$d'(e_o) = \lambda e_o + e'_1.$$

On peut trouver une transvection τ telle que $d' = \tau \circ d$. En effet, si $d(e_o) = \lambda e_o + e_1$:

$$\tau(d(e_o)) = d(e_o) + a\, \varphi(d(e_o))$$
$$= \lambda e_o + e_1 + a\, \lambda\, \varphi(e_o).$$

Il suffit de choisir a tel que $e'_1 = e_1 + a\, \lambda\, \varphi(e_o)$ pour que $d'(e_o) = (\tau \circ d)(e_o)$ donc $d' = \tau \circ d$.

Par conséquent, l'ensemble $\mathcal{D}_\lambda = \{ \tau \circ d \; ; \; \tau \in \Gamma(\mathrm{H}) \}$ représente l'ensemble des dilatations de rapport λ.

A une telle classe $\mathcal{D}_\lambda$ correspond bijectivement le scalaire λ.

$$\psi : \; \mathcal{D}_\lambda \in \mathrm{G}(\mathrm{H})/\Gamma(\mathrm{H}) \longrightarrow \lambda \in \mathrm{K}.$$

Le produit $\mathcal{D}_\lambda \times \mathcal{D}_{\lambda'}$ est, par définition, la classe du produit $d \circ d'$ si $d \in \mathcal{D}_\lambda$, $d' \in \mathcal{D}_{\lambda'}$. Or, $d \circ d'$ fait correspondre à la variété affine $e_o + \mathrm{H}$ la variété $\lambda\lambda' e_o + \mathrm{H}$. D'où $\psi(\mathcal{D}_\lambda \times \mathcal{D}_{\lambda'}) = \lambda\lambda'$.

Le groupe $g(\mathrm{H})/\Gamma(\mathrm{H})$ est donc isomorphe au groupe multiplicatif de K.

Variante de c : On a vu au 2) que tout endomorphisme $u \in \mathrm{G}(\mathrm{H})$ définissait un endomorphisme de E/H de la forme : $\dot{x} \longrightarrow \lambda \dot{x}$. On peut donc définir ainsi une correspondance :

$$\psi : u \in \mathrm{G}(\mathrm{H}) \longrightarrow \lambda \in \mathrm{K}.$$

Cette correspondance est un homomorphisme ; la loi de composition considérée étant, à gauche, la composition des applications et à droite la multiplication du corps K. En effet, si :

$$\psi(u) = \lambda \qquad \psi(v) = \lambda',$$

c'est que $\quad u(\dot{e}_o) = \lambda \dot{e}_o \qquad v(\dot{e}_o) = \lambda' \dot{e}_o$:

$$(u \circ v)(\dot{e}_o) = u(\lambda' \dot{e}_o) = \lambda'\, u(\dot{e}_o) = \lambda' \lambda \dot{e}_o,$$

donc

$$\psi(u \circ v) = \lambda\lambda'.$$

Or, dans un homomorphisme de groupe, le groupe image est isomorphe au groupe quotient du groupe par le noyau de l'homomorphisme. Celui-ci est, ici, l'ensemble des éléments de G(H) qui sont envoyés sur l'élément neutre de l'image, c'est-à-dire sur 1. Ce noyau est donc constitué par Γ(H) et K, en tant que groupe multiplicatif, est isomorphe à G(H)/Γ(H).

Exercice 33.

1° F vérifie la relation $e - f^2 = 0$ qui peut s'écrire :

$$(e - f)(e + f) = 0 \text{ ou } h \circ g = g \circ h = 0.$$

(Le polynôme $1 - x^2$ est le polynôme ϖ de la théorie générale, $1 + x$ et $1 - x$ les polynômes p et q de cette théorie).

On peut écrire : $e - f + e + f = 2\,e$. Il en résulte que :

$$\forall x \in \mathrm{E} \qquad g(x) + h(x) = 2\,x \qquad (1).$$

Si le corps K n'est pas de caractéristique 2, ceci peut s'écrire :

$$x = \frac{g(x)}{2} + \frac{h(x)}{2},$$

$\dfrac{g(x)}{2} \in \mathrm{V}$, $\dfrac{h(x)}{2} \in \mathrm{W}$. Tout $x \in \mathrm{E}$ peut être mis sous forme de la somme d'un élément appartenant à chacun des sous-espaces V et W.

D'autre part, $\mathrm{V} \cap \mathrm{W} = \{0\}$. En effet, $h(\mathrm{V}) = (h \circ g)(\mathrm{E}) = 0$; $g(\mathrm{W}) = (g \circ h)(\mathrm{E}) = 0$. Donc, si $x \in \mathrm{V} \cap \mathrm{W}$, $h(x) = g(x) = 0$ ce qui, en vertu de (1), entraîne $x = 0$. On peut conclure $\mathrm{E} = \mathrm{V} \oplus \mathrm{W}$. Appliquons f à un élément x de V ; comme $\forall x \in \mathrm{V} \quad \exists y \in \mathrm{E} \quad g(y) = x$, on a :

$$f(x) = (f \circ g)(y) = f \circ (e + f)(y) = f(y) + f^2(y)$$
$$= f(y) + e(y) = g(y) = x.$$

Donc, $\forall x \in \mathrm{V} \qquad f(x) = x.$

La restriction de f à V est l'identité. On montrerait de même que la restriction de f à W est $-e_{\mathrm{W}}$.

Ceci posé, il suffit de construire une base de E en prenant la réunion d'une base de V et d'une base de W pour que la matrice de f prenne la forme indiquée.

2° Si K est de caractéristique 2, l'égalité (1) devient :

$$g + h = 0 \qquad g = h.$$

Et on a alors :

$$g \circ h = g^2 = 0$$

et $g = e + f$ donne $f = g + e$.

$g^2(\mathrm{E}) = 0$ signifie que $g(\mathrm{E})$ est inclus dans le noyau de l'application g, ce qui exige :

$$\dim g(\mathrm{E}) \leqslant \dim \overset{-1}{g}(0),$$

soit
$$p \leqslant n - p \qquad 2\,p \leqslant n.$$

Si $p = 1$, $\overset{-1}{g}(0)$ est un sous-espace de codimension 1 et si x appartient à ce sous-espace :

$$f(x) = (g + e)(x) = x,$$

f est donc une application qui laisse invariants tous les vecteurs d'un sous-espace de codimension 1 ; c'est une application du type considéré à l'exercice 32. En outre, pour un $x \in \mathrm{E}$ quelconque, on a :

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

Or, $g(x) \in g(\mathrm{E}) \subset \overset{-1}{g}(0)$. L'image par f de la variété affine $x + \overset{-1}{g}(0)$ est donc elle-même. L'application est une *transvection*.

3° Soient deux involutions f et φ : $f^2 = e$, $\varphi^2 = e$. Pour que leur produit soit une involution, il faut et il suffit que :

$$(f \circ \varphi)^2 = f \circ \varphi \circ f \circ \varphi = e,$$

condition qui équivaut à celle obtenue en multipliant les deux membres à gauche par f et à droite par φ, soit :

$$\varphi \circ f = f \circ e \circ \varphi = f \circ \varphi.$$

Pour que le produit de deux involutions soit une involution, il est nécessaire et suffisant que ces deux involutions commutent.

Exercice 34.

Nous avons vu que K(f), considéré comme anneau, était image homomorphe de K(x) dans un homomorphisme dont le noyau est le polynôme $\varpi(x)$, donc que K(f) était isomorphe à l'anneau des polynômes modulo $\varpi(x)$. Il résulte de cet isomorphisme que les diviseurs de zéro de l'anneau K(f) sont les images des diviseurs de zéro de $\mathrm{K}(x)/\varpi(x)$. Ceux-ci sont les classes de polynômes non premiers avec $\varpi(x)$ (car ϖ ne peut diviser $p_1 p_2$ sans diviser p_1 ou p_2 que si ceux-ci ne sont pas premiers avec lui). *Les diviseurs de zéro de* K(f) *sont les endomorphismes $p(f)$ où p est un polynôme de degré inférieur à ϖ et non premier avec lui.*

Supposons maintenant que p soit premier avec ϖ. Le théorème de Bezout permet d'affirmer qu'il existe deux polynômes q et q' tels que : $pq + \varpi q' = 1$, ce qui entraîne $p(f) \times q(f) = e$.

L'endomorphisme $p(f)$ admet $q(f)$ comme inverse ; c'est un automorphisme.

Exemple. — Soient a_1, a_2, a_3 les vecteurs de la base par rapport à laquelle est donnée la matrice. On voit que :

$$f(a_1) = a_2 \qquad f(a_2) = a_3 \qquad f(a_3) = a_1.$$

f réalise une permutation circulaire sur les éléments de la base, ce dont il résulte immédiatement que $f^3 = e$.

Le polynôme $x^3 - 1 = (x - 1)(x^2 + x + 1)$ appartient à l'idéal annulateur. Le polynôme ϖ de la théorie générale doit donc diviser $x^3 - 1$; or, ce n'est ni $x - 1$ (car f n'est pas l'identité), ni $x^2 + x + 1$ car $f^2 + f + e$ a pour matrice :

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

qui n'est pas la matrice de l'application nulle. On a donc : $\varpi = x^3 - 1$.

Les diviseurs de zéro de K(f) *sont, d'une part,* $f^2 + f + e$ dont nous venons d'écrire la matrice ; *d'autre part, les endomorphismes de la forme :*

$$(f - e)(\alpha f + \beta e) = \alpha f^2 + (\beta - \alpha) f - \beta$$

avec α et β réels non simultanément nuls, dont la matrice est :

$$\begin{pmatrix} -\beta & \alpha & \beta - \alpha \\ \beta - \alpha & -\beta & \alpha \\ \alpha & \beta - \alpha & -\beta \end{pmatrix}$$

$f^2 + f + e$ donne de tout vecteur une image colinéaire à $a_1 + a_2 + a_3$. *Il est donc de rang 1.*

Pour les autres, cherchons le rang de la matrice ci-dessus en faisant des combinaisons linéaires sur les lignes ou sur les colonnes. On trouve successivement les matrices équivalentes :

$$\begin{pmatrix} -\beta & \alpha & \beta - \alpha \\ \beta - \alpha & -\beta & \alpha \\ 0 & 0 & 0 \end{pmatrix} \text{(3}^{\text{e}}\text{ ligne remplacée par la somme des lignes).}$$

$$\begin{pmatrix} -\beta & \alpha & 0 \\ \beta-\alpha & -\beta & 0 \\ 0 & 0 & 0 \end{pmatrix}$$ (3ᵉ colonne remplacée par la somme des colonnes).

Si $\beta = 0$ (α est alors différent de 0), la matrice est $\begin{pmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et est de rang 2.

Si $\beta \neq 0$, une combinaison linéaire des colonnes donne la matrice équivalente $\begin{pmatrix} -\beta & 0 & 0 \\ \beta-\alpha & \alpha\beta-\alpha^2-\beta^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ qui est encore de rang 2, car $\alpha\beta-\alpha^2-\beta^2$ ne peut s'annuler pour α et β réels.

Les applications linéaires du deuxième type sont de rang 2.

Exercice 35.

Soient deux ensembles E et F de même cardinal. Il existe une bijection $\varphi : E \longrightarrow F$.

Soit alors $u \in \mathcal{S}(E)$ une bijection. Considérons l'application de F dans F :

$$v = \varphi^{-1} \circ u \circ \varphi,$$

c'est une bijection donc un élément de $\mathcal{S}(F)$. Nous avons donc défini une application :

$$\psi : \mathcal{S}(E) \longrightarrow \mathcal{S}(F)$$
$$u \qquad\qquad v = \varphi^{-1} \circ u \circ \varphi$$

Cette application est bijective. En effet :

$v = v' \Longrightarrow \varphi^{-1}\circ u \circ \varphi = \varphi^{-1} \circ u' \circ \varphi \Longrightarrow u = u'$; elle est injective.

$\forall v \in \mathcal{S}(F) \quad \exists\, u = \varphi \circ v \circ \varphi^{-1} \in \mathcal{S}(E) \quad \psi(u) = v$; elle est surjective.

D'autre part, c'est un homomorphisme, car :

$$\psi(u \circ u') = \varphi^{-1} \circ u \circ u' \circ \varphi = \varphi^{-1} \circ u \circ \varphi \circ \varphi^{-1} \circ u' \circ \varphi = \psi(u) \circ \psi(u').$$

ψ est donc un isomorphisme de $\mathcal{S}(E)$ sur $\mathcal{S}(F)$.

Exercice 36.

Le produit de deux permutations paires est une permutation paire puisque sa signature est positive. La permutation identique appartient au sous-ensemble des permutations paires. L'inverse d'une permutation paire est une permutation paire. Celles-ci forment donc un sous-groupe.

Celui-ci est invariant ; en effet, soit p une permutation paire et σ une permutation quelconque. La permutation $\sigma \circ p \circ \sigma^{-1}$ a pour signature :

$$\varepsilon_\sigma\, \varepsilon_p\, \varepsilon_\sigma^{-1} = \varepsilon_p,$$

puisque $\varepsilon_\sigma \varepsilon_\sigma^{-1} = 1$. Donc, $\sigma \circ p \circ \sigma^{-1}$ a une signature positive et appartient au groupe alterné qui est donc invariant.

Exercice 37.

1) φ_σ est injective. En effet, $\varphi_\sigma(f) = \varphi_\sigma(f')$ signifie $f \circ \sigma^{-1} = f' \circ \sigma^{-1}$; σ^{-1} étant une bijection, cette égalité est simplifiable, donc :

$$\varphi_\sigma(f) = \varphi_\sigma(f') \Longrightarrow f = f'.$$

φ_σ est surjective. En effet, si $g \in \mathcal{F}$:

$$f \circ \sigma^{-1} = g \Longrightarrow f = g \circ \sigma.$$

Donc, $\quad \forall g \in \mathcal{F} \qquad \exists\, f = g \circ \sigma \in \mathcal{F} \qquad \varphi_\sigma(f) = g,$

φ_σ est donc un élément de $\mathcal{S}(\mathcal{F})$.

Considérons maintenant l'application :

$$\sigma \in \mathcal{S}(E) \longrightarrow \varphi_\sigma \in \mathcal{S}(\mathcal{F}),$$

elle est injective. En effet, $\varphi_\sigma = \varphi_{\sigma'}$ signifie :

$$\forall f \in \mathcal{F} \qquad f \circ \sigma^{-1} = f \circ \sigma'^{-1}$$

c'est-à-dire $\quad \forall x \in E \qquad f(\sigma^{-1}(x)) = f(\sigma'^{-1}(x)).$

Si σ était différent de σ' (et par conséquent $\sigma^{-1} \neq \sigma'^{-1}$), il existerait au moins une valeur x pour laquelle $\sigma^{-1}(x)$ serait différent de $\sigma'^{-1}(x)$. On pourrait alors, pourvu que F ait au moins deux éléments, trouver une application f pour laquelle l'égalité ci-dessus ne serait pas satisfaite ; donc, celle-ci exige $\sigma = \sigma'$.

Montrons que cette application est un homomorphisme pour l'opération composition des applications en cherchant ce qu'est $\varphi_\sigma \circ \varphi_{\sigma'}$:

$$\forall f \in \mathcal{F} \quad \varphi_\sigma \circ \varphi_{\sigma'}(f) = \varphi_\sigma(f \circ \sigma'^{-1}) = f \circ \sigma'^{-1} \circ \sigma^{-1} = f \circ (\sigma \circ \sigma')^{-1} = \varphi_{\sigma\circ\sigma'}(f).$$

Cette égalité, vraie pour tout f, signifie :

$$\varphi_\sigma \circ \varphi_{\sigma'} = \varphi_{\sigma\circ\sigma'}.$$

L'application envisagée est bien un homomorphisme.

2) L'application qui, à $\sigma \in \mathcal{S}(E)$, fait correspondre $\sigma \in \mathcal{S}(\mathcal{F})$ est un homomorphisme injectif. On peut définir de la même façon une application de $\mathcal{S}(\mathcal{F})$ dans le groupe des bijections de l'ensemble des applications de $\mathcal{F}$ dans un ensemble G, c'est-à-dire de $\mathcal{S}(\mathcal{F})$ dans $\mathcal{S}(\mathcal{G})$; cette application à $\sigma \in \mathcal{S}(\mathcal{F})$ fait correspondre ce que l'on note encore $\sigma \in \mathcal{S}(\mathcal{G})$ et est encore un homomorphisme injectif.

L'application $\sigma \in \mathcal{S}(E) \longrightarrow \sigma \in \mathcal{S}(\mathcal{G})$, produit de deux homomorphismes injectifs, est un homomorphisme injectif.

Voyons de plus près sur l'exemple en quoi consiste cet homomorphisme. Nous noterons provisoirement $\bar{\sigma}$, $\bar{\bar{\sigma}}$, etc... les images successives de σ par les homomorphismes envisagés. D'abord, $\mathcal{F}(I, E)$ est bien E^n puisque la donnée d'une application de I dans E est la donnée d'un n-uplet ordonné d'éléments de E, donc un élément de E^n. D'après les conventions d'écriture, $\bar{\sigma}x = x \circ \sigma^{-1}$, x étant considéré comme l'application qui fait correspondre x_i à i ($i = 1, 2, \ldots, n$).

$x \circ \sigma^{-1}$ fait correspondre $x_{\sigma^{-1}(i)}$ à i suivant le schéma :

$$i \longrightarrow \sigma^{-1}(i) \longrightarrow x_{\sigma^{-1}(i)},$$

donc $\quad \bar{\sigma}x = (x_{\sigma^{-1}(1)}, x_{\sigma^{-1}(2)}, \ldots x_{\sigma^{-1}(n)}) \in E^n = \mathcal{F}(I, E)$ et $\bar{\sigma}$ est l'application qui à $x \in E^n$ fait correspondre $\bar{\sigma}(x) = \bar{\sigma}x \in E^n$. Donc, $\bar{\sigma} \in \mathcal{S}(E^n)$.

Soit maintenant $f \in \mathcal{F}(E^n, F)$. On lui fait correspondre :

$$\bar{\bar{\sigma}} f = f \circ \bar{\sigma}^{-1} \qquad \bar{\bar{\sigma}} \in \mathcal{S}\,[\mathcal{F}(E^n, F)],$$

où $\bar{\sigma}$ représente l'élément que nous venons de définir :

$$\bar{\bar{\sigma}} f(x) = (f \circ \bar{\sigma}^{-1})(x) = f(\bar{\sigma}^{-1}(x)),$$

or $\qquad \bar{\sigma}^{-1}(x) = (x_{\sigma(1)} \ldots x_{\sigma(n)}).$

D'où $\qquad \bar{\bar{\sigma}} f(x) = f(x_{\sigma(1)} \ldots x_{\sigma(n)}).$

Si on refait encore une fois la même opération en considérant : $g \in \mathcal{F}(\mathcal{F}(E^n, F), G)$, on aura :

$$\bar{\bar{\bar{\sigma}}}\, g = g \circ \bar{\bar{\sigma}}^{-1} \qquad \bar{\bar{\bar{\sigma}}} \in \mathcal{S}\,[\mathcal{F}(\mathcal{F}(E^n, F), G)],$$

ce qui veut dire :

$$\bar{\bar{\sigma}}\, g(f) = g \circ \bar{\bar{\sigma}}^{-1} f$$

$$\bar{\bar{\sigma}}\, g\,[f(x)] = g \circ \bar{\bar{\sigma}}^{-1} f(x) = g\,[f(x_{\sigma^{-1}(1)} \dots x_{\sigma^{-1}(n)})].$$

Exercice 38.

Soit $x \in E^n$ de composantes $x_1 \dots x_n$ (E, espace vectoriel sur K). Dire que f est une forme antisymétrique sur E^n, c'est dire que pour tout $\sigma \in \mathcal{S}_n$:

$$\forall x \in E \qquad f(\sigma(x)) = \varepsilon_\sigma f(x).$$

Prenons, pour permutation, la transposition τ_{ij} :

$$\forall x \in E \qquad f(\tau_{ij}(x)) = -f(x).$$

Soit alors x tel que $x_i = x_j$. Pour un tel x :

$$f(\tau_{ij}(x)) = f(x).$$

Donc, $f(x) = -f(x)$.

Si le corps K n'est pas de caractéristique 2, on en déduit $f(x) = 0$. La forme f s'annule pour tout $x \in E^n$ dont deux composantes sont égales. Elle est donc alternée.

Si K est de caractéristique 2, la symétrie et l'antisymétrie sont la même propriété et sont distinctes de la propriété : être alternée.

Exercice 39.

Une forme p-linéaire f sur un espace de dimension n rapporté à une base $(a_1\, a_2 \dots a_n)$ prend pour le p-uplet de vecteurs $x_1 \dots x_p$ donnés par :

$$x_i = \sum_{j=1}^{n} \xi_i^j a_j$$

la valeur

$$\Sigma\, \xi_1^{\lambda(1)}\, \xi_2^{\lambda(2)} \dots \xi_p^{\lambda(p)} f(a_{\lambda(1)}, a_{\lambda(2)}, \dots a_{\lambda(p)}),$$

où λ représente une application de l'ensemble $(1, \dots, p)$ dans l'ensemble $(1, \dots, n)$ et la sommation étant étendue à l'ensemble de ces applications (n^p termes).

Si la forme f doit être alternée, ceux des nombres $f(a_{\lambda(1)}, \dots, a_{\lambda(p)})$ relatifs aux applications λ non injectives portent sur un p-uplet à deux composantes égales et sont nuls. Si $p > n$, aucune application λ n'est injective et la seule forme p-linéaire alternée est la forme 0. Supposons donc maintenant $p \leqslant n$. Il ne reste donc dans la somme ci-dessus que les $n(n-1) \dots (n-p+1)$ termes relatifs à des injections de l'ensemble $\{1, \dots, p\}$ dans l'ensemble $\{1 \dots n\}$. Mais ceux de ces termes où l'ensemble $\{\lambda(1), \dots \lambda(p)\}$ est le même à l'ordre près sont égaux au signe près, une forme alternée étant antisymétrique. Ces $p\,!$ termes pourront donc être groupés en :

$$f(a_{\lambda(1)}, \dots, a_{\lambda(p)}) \sum_{\sigma \in \mathcal{S}_p} \varepsilon_\sigma\, \xi_1^{\sigma\lambda(1)}\, \xi_2^{\sigma\lambda(2)} \dots \xi_p^{\sigma\lambda(p)} \qquad (1)$$

σ étant une permutation de l'ensemble $\{\lambda(1), \dots \lambda(p)\}$ et la sommation étant étendue aux $p\,!$ permutations.

Les C_n^p combinaisons de p vecteurs, pris parmi les n vecteurs de la base $\{a_i\}$, étant arbitrairement indexés de 1 à C_n^p, nous désignons par c_k le p-uplet des vecteurs de la k^e combinaison rangés dans l'ordre naturel

de leurs indices. Soit λ l'injection correspondant au p-uplet c_k, nous posons :

$$\sum_{\sigma \in \mathcal{S}_p} \varepsilon_\sigma\, \xi_1^{\sigma\lambda(1)}\, \xi_2^{\sigma\lambda(2)} \dots \xi_p^{\sigma\lambda(p)} = \mu^k \qquad (*)$$

Le terme (1) s'écrira $\mu^k f(c_k)$ et nous aurons :

$$f(x_1, \dots, x_p) = \sum_{k=1}^{C_n^p} \mu^k f(c_k) = \sum_{k=1}^{C_n^p} \mu^k \eta_k \qquad (2)$$

On voit que la valeur de f est déterminée par la donnée des C_n^p valeurs $f(c_k) = \eta_k$.

Considérons alors les C_n^p formes p-linéaires alternées f^k définies par la condition qu'elles prennent, la valeur 1 pour la combinaison de même indice et la valeur 0 pour toute autre :

$$f^k(c_k) = 1, \quad \text{pour } k' \neq k \quad f^k(c_{k'}) = 0.$$

Cherchons la valeur de $f^k(x_1, \dots, x_p)$:

$$f^k(x_1, \dots, x_p) = \sum_{l=1}^{C_n^p} \mu^l f^k(c_l).$$

Mais $f^k(c_l) = 0$ si l n'est pas égal à k, d'où :

$$f^k(x_1, \dots, x_p) = \mu^k.$$

On peut alors écrire l'égalité (2) :

$$f(x_1, \dots, x_p) = \sum_{k=1}^{C_n^p} \eta_k f^k(x_1, \dots, x_p).$$

Ceci étant vrai pour tout p-uplet de vecteurs de E, on peut conclure :

$$f = \sum_{k=1}^{C_n^p} \eta_k f^k$$

égalité qui montre que les formes p-linéaires alternées sur E forment un espace vectoriel qui admet les C_n^p formes f^k comme système de générateurs. Nous aurons donc montré que *les formes p-linéaires alternées forment un espace vectoriel à C_n^p dimensions* si nous montrons que les formes f^k (qui sont une généralisation des formes coordonnées de l'espace dual) sont linéairement indépendantes.

Si elles étaient dépendantes, c'est qu'il existerait des scalaires λ_k non tous nuls tels que :

$$\sum_{k=1}^{C_n^p} = \lambda_k f^k$$

soit la forme identiquement nulle, donc s'annule pour tout p-uplet de vecteurs. Prenons pour p-uplet la combinaison c_k de vecteurs de la base. La valeur de la forme ci-dessus se réduit à λ_k. Il faudrait donc $\lambda_k = 0$, et ceci pour toute valeur de k. L'indépendance est donc établie.

(*) μ^k est un déterminant ; celui de la matrice obtenue en prenant dans la matrice qui donne les coordonnées de $x_1 \dots x_p$ par rapport à la base (matrice à p, colonnes et n lignes), les p lignes correspondant aux coordonnées relatives aux éléments du p-uplet c_k, pris dans l'ordre naturel de leurs indices.

Exercice 40.

1° Soit H_i le sous-espace propre relatif à la valeur λ_i. Nous montrerons qu'un ensemble de vecteurs non nuls $x_1, x_2, ..., x_n$, tels que pour tout i, x_i appartienne à H_i, est nécessairement une partie libre, en répétant mot pour mot la démonstration de ce fait donnée dans le cours pour le cas où toutes les valeurs propres sont distinctes.

Soit alors un vecteur de $H_1 + H_2 + ... + H_n$. S'il admettait deux décompositions distinctes :

$$y_1 + y_2 + ... y_i + ... + y_n = z_1 + z_2 + ... + z_i + ... + z_n$$

avec $\forall i \quad y_i \in H_i,\ z_i \in H_i$, on aurait $\Sigma\,(y_i - z_i) = 0$, et les différences $y_i - z_i$ non toutes nulles, d'une part, constitueraient une partie libre, et, d'autre part, auraient une somme nulle, ce qui est absurde.

2° Si E est la somme des H_i, on peut constituer une base de E par la réunion d'une base de chaque H_i ; par rapport à cette base, la matrice de f est diagonale.

Réciproquement, si la matrice de f est diagonalisée, c'est qu'on a pu trouver une base formée de vecteurs propres. E est donc la somme des H_i qui est directe d'après 1°.

Ceci peut s'exprimer autrement. La dimension de chaque H_i est au plus égale à la multiplicité β_i de la racine λ_i du polynôme caractéristique. (H_i n'est autre que le sous-espace $E_{\alpha_i - 1}$ de la théorie, dont la dimension était au plus égale à α_i, lui-même au plus égal à β_i). Pour que E soit somme directe des H_i, il est nécessaire et suffisant que :

$$\Sigma\, \beta_i = \dim E = \Sigma \dim H_i$$

ce qui équivaut, en vertu de $\dim H_i \leqslant \beta_i$, à : $\quad \forall i \qquad \beta_i = \dim H_i$.

3° Un sous-espace propre relatif à une valeur propre différente de zéro coïncide avec son image par f. Dire que f est somme directe de tels sous-espaces, c'est dire que $f(E)$ a même dimension que E, donc que f est un automorphisme.

On peut dire aussi : la matrice M d'une application f à spectre simple pouvant être diagonalisée, si les valeurs propres qui figurent dans sa diagonale principale sont toutes différentes de zéro, il est évident que cette matrice admet pour inverse la matrice diagonale dont les éléments sont les inverses des éléments de même indice de M.

Pour que f soit identique à son inverse, *il faut et il suffit que chaque valeur propre* soit sa propre inverse, donc *soit 1 ou — 1*. f identique à son inverse signifie $f^2 = e$; nous avons bien établi à l'exercice 33 que la matrice d'une involution pouvait être diagonalisée, tous les éléments de la diagonale valant 1 ou — 1.

4° *a*) Il existe, par hypothèse, une base $\{ e_i \}$ de E constituée de vecteurs propres de f. Soit λ_i la valeur propre correspondant à e_i (les λ_i peuvent ne pas être deux à deux distinctes). Une application linéaire étant définie par les images des éléments d'une base, la condition nécessaire et suffisante pour que f et φ commutent est que, pour $i = 1, 2, ..., n$, on ait :

$$f \circ \varphi(e_i) = \varphi \circ f(e_i)$$

c'est-à-dire : $\qquad f[\varphi(e_i)] = \varphi(\lambda_i e_i) = \lambda_i \varphi(e_i)$

c'est-à-dire que $\varphi(e_i)$ est un vecteur propre de f relativement à la valeur λ_i. La condition est donc que φ donne de tout vecteur propre de f relatif à λ_i un vecteur propre relatif à la même valeur, ou sous une forme équivalente, que les *sous-espaces propres de f soient invariants par φ*.

Si les λ_i sont deux à deux distincts, les sous-espaces propres de f sont à une dimension et la condition est que l'ensemble des vecteurs propres de f soit inclus dans l'ensemble des vecteurs propres de φ.

b) Pour qu'un endomorphisme φ commute avec tous les autres, il

faut déjà qu'il commute avec tous ceux dont les valeurs propres sont distinctes, ce qui exige qu'il admette pour vecteurs propres tous les vecteurs propres de ces endomorphismes. Soit alors un vecteur quelconque x ; on peut toujours construire un endomorphisme à valeurs propres distinctes admettant x pour vecteur propre ; il devra être propre pour φ aussi ; d'où :

$$\forall x \in E \qquad \varphi(x) = \lambda_x x.$$

Pour prouver que φ est une homothétie, il reste seulement à vérifier que λ_x est indépendant de x. Or, si on a :

$$\varphi(x_1) = \lambda_1 x_1 \qquad \varphi(x_2) = \lambda_2 x_2$$

on ne peut avoir $\varphi(x_1 + x_2) = \lambda(x_1 + x_2)$ que si $\lambda_1 = \lambda_2$.

c) Une involution est une opération à spectre simple puisque, d'après l'exercice 33, E est somme directe des deux sous-espaces propres relatifs aux valeurs 1 et — 1.

Pour que deux involutions commutent (et aient pour produit une involution), il est donc nécessaire et suffisant que les sous-espaces propres de l'une soient invariants par l'autre.

d) De façon plus générale, *une involution commute avec un endomorphisme si et seulement si celui-ci laisse invariants les sous-espaces propres de l'involution.*

Dans $\mathbf{R}^3$, si on laisse de côté la transformation identique, les involutions sont de trois espèces :

α) — Les trois racines du polynôme caractéristique sont égales à — 1. E tout entier est espace propre pour cette transformation. E étant invariant par tout endomorphisme, une transformation de ce type commute avec tout autre endomorphisme. (Il s'agit d'ailleurs d'une homothétie particulière).

β) — Deux racines du polynôme sont égales à — 1, la troisième est égale à 1. E est somme directe d'un sous-espace propre de dimension 2 et de valeur propre — 1 et d'un sous-espace propre de dimension 1 et de valeur propre 1. Une involution de ce type commute avec les endomorphismes qui laissent invariants ces deux sous-espaces et avec ceux-là seulement.

γ) — Une racine du polynôme est égale à — 1, les deux autres sont égales à 1. E est somme directe d'un sous-espace propre de dimension 1 et de valeur propre — 1 et d'un sous-espace propre de dimension 2 et de valeur propre 1. Une involution de ce type commute avec les endomorphismes qui laissent invariants ces deux sous-espaces et avec ceux-là seulement.

Nous pouvons traduire ces résultats dans le langage de la géométrie élémentaire. Une origine étant choisie dans l'espace euclidien, l'ensemble des vecteurs géométriques issus de cette origine a la structure d'espace vectoriel de $\mathbf{R}^3$. Les sous-espaces de dimensions 1 et 2 de $\mathbf{R}^3$ correspondent aux droites et plans issus de l'origine. Nous voyons alors que :

— Une involution du type α) qui à tout vecteur donne pour image son opposé correspond à la symétrie par rapport à l'origine des coordonnées. Et nous pouvons énoncer : *Une symétrie par rapport à un point commute avec toute application linéaire de l'espace vectoriel relatif à ce point.*

— Une involution du type β) qui, étant donnés une droite et un plan fixes passant par l'origine, transforme tout vecteur en un vecteur ayant même composante sur la droite et composante opposée dans le plan, correspond à une symétrie par rapport à la droite parallèlement au plan. Et nous pouvons énoncer : *Une symétrie par rapport à un axe* D, *parallèle-*

ment à une direction de plan Π, *commute avec celles des applications linéaires de l'espace vectoriel relatif à un point de son axe qui laissent invariants cet axe et le plan de direction* Π *passant par ce point, et avec celles-là seulement.* (Exemples d'applications commutant avec la symétrie par rapport à D, parallèlement à Π : affinité sur le plan parallèle à Π, parallèlement à D ; symétrie par rapport à un plan passant par D, parallèlement à une direction de Π, etc...).

— Une involution de type γ), qui, étant donnés une droite et un plan fixes passant par l'origine, transforme tout vecteur en un vecteur ayant même composante sur le plan et composante opposée sur la droite, correspond à une symétrie par rapport au plan, parallèlement à la droite. Et nous pouvons énoncer : *Une symétrie par rapport à un plan* P *parallèlement à une direction de droite* Δ *commute avec celles des applications linéaires de l'espace vectoriel relatif à un point de* P *qui laissent invariants* P *et la droite de direction* Δ *passant par ce point, et avec celles-là seulement.* (Exemples d'applications commutant avec la symétrie par rapport à P, parallèlement à Δ : affinités parallèlement à P sur la parallèle à Δ ; symétrie par rapport à une droite de P issue de l'origine, parallèlement à un plan passant par Δ ; symétrie par rapport à un plan parallèle à Δ, issu de l'origine, parallèlement à une direction de D).

Exercice 41.

1° Supposons que : $$x_{p+1} = \sum_{i=1}^{p} \lambda_i x_i$$

et montrons que tous les vecteurs d'ordre supérieur s'exprimeront aussi comme combinaisons linéaires des x_i avec $1 \leqslant i \leqslant p$. Supposons qu'il en soit ainsi pour tous les vecteurs jusqu'à x_{p+k} :

$$x_{p+k} = \sum_{i=1}^{p} \mu_i x_i.$$

Il en résulte :

$$x_{p+k+1} = f(x_{p+k}) = \sum_{i=1}^{p} \mu_i f(x_i) = \sum_{i=1}^{p-1} \mu_i x_{i+1} + \mu_p x_{p+1}.$$

x_{p+1} étant combinaison linéaire des x_i $(1 \leqslant i \leqslant p)$, x_{p+k+1} l'est aussi ; les x_i $(1 \leqslant i \leqslant p)$ forment donc un système de générateurs minimum, donc une base, de F.

Pour que x_1 engendre E, il est nécessaire et suffisant que E et F aient même dimension, donc que $p = n$, ou que $x_1, x_2, ..., x_n$ soient linéairement indépendants.

2° Soient $\lambda_1, \lambda_2, ..., \lambda_n$ les éléments diagonaux de A :

$$x_1 = \Sigma\, \xi^i e_i \implies \begin{array}{l} x_2 = \Sigma\, \xi^i \lambda_i e_i \\ x_3 = \Sigma\, \xi^i (\lambda_i)^2 e_i \\ \cdots\cdots\cdots\cdots \\ x_n = \Sigma\, \xi^i (\lambda_i)^{n-1} e_i. \end{array}$$

Une condition d'indépendance du système de ces n vecteurs est que leur déterminant soit différent de zéro. Or, ce déterminant peut s'écrire :

$$\xi_1 \xi_2 \ldots \xi_n \begin{vmatrix} 1\ \lambda_1 \lambda_1^2 \ldots \lambda_1^{n-1} \\ 1\ \lambda_2 \lambda_2^2 \ldots \lambda_2^{n-1} \\ \cdots\cdots\cdots \\ 1\ \lambda_n \lambda_n^2 \ldots \lambda_n^{n-1} \end{vmatrix} = \xi_1 \xi_2 \ldots \xi_n \Pi (\lambda_i - \lambda_j)$$

le produit Π étant entendu à tous les couples d'indice i, j avec $i > j$ (déterminant de Vandermonde). Pour qu'il soit différent de zéro, il est nécessaire et suffisant que tous les λ_i, c'est-à-dire toutes les valeurs propres de f, soient distinctes.

3° Par rapport à $(x_1, x_2, ..., x_n)$, on trouve pour matrice de f :

$$A' = \begin{pmatrix} 0 & 0 & & 0 & a_1 \\ 1 & 0 & & \vdots & a_2 \\ 0 & 1 & & \vdots & \vdots \\ \vdots & 0 & \ddots & \vdots & a_i \\ \vdots & \vdots & \ddots & 0 & \vdots \\ \vdots & \vdots & & 1 & a_n \end{pmatrix}$$

si $x_{n+1} = \sum_{i=1}^{n} a_i x_i$.

Le polynôme caractéristique de f s'écrit donc :

$$P(\lambda) = \begin{vmatrix} -\lambda & 0 & 0 & \cdots & & a_1 \\ 1 & -\lambda & 0 & & & \vdots \\ 0 & 1 & -\lambda & & & \vdots \\ \vdots & 0 & 1 & -\lambda & & \vdots \\ \vdots & \vdots & 0 & 1 \ddots & & a_i \\ \vdots & \vdots & \vdots & 0 \ddots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & -\lambda & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 & a_n - \lambda \end{vmatrix}$$

Développons ce déterminant par rapport à la dernière colonne. Le mineur de a_1 vaut 1. Celui de a_i pour $1 < i < n$ vaut $(-\lambda)^{i-1}$. Enfin, celui de $a_n - \lambda$ vaut $(-\lambda)^{n-1}$. On trouve donc :

$$P(\lambda) = (-1)^{n+1} a_1 + \sum_{i=2}^{n-1} (-1)^{n+i} a_i (-\lambda)^{i-1} + (a_n - \lambda)(-\lambda)^{n-1}$$
$$= (-1)^n [\lambda^n - a_n \lambda^{n-1} \ldots - a_i \lambda^{i-1} \ldots - a_1].$$

La théorie générale nous apprend que $P(f) = 0$, c'est-à-dire que l'application $P(f)$ annule tous les vecteurs x_i. Calculons d'abord $P(f)x_1$ ou plutôt $(-1)^n P(f) x_1$ que nous désignerons par $g(x_1)$.

$$\begin{aligned} g(x_1) &= f^n(x_1) - a_n f^{n-1}(x_1) \ldots - a_2 f(x_1) - a_1 x_1 \\ &= f(x_n) - a_n x_n \ldots - a_2 x_2 - a_1 x_1 = 0. \end{aligned}$$

Supposons alors que l'on ait $g(x_i) = 0$ et cherchons $g(x_{i+1})$.

$$\begin{aligned} g(x_{i+1}) &= f^n(x_{i+1}) - a_n f^{n-1}(x_{i+1}) \ldots - a_{n-i+1} f^{n-i}(x_{i+1}) \\ &\quad - a_{n-i} f^{n-i-1}(x_{i+1}) \ldots - a_2 f(x_{i+1}) - a_1 x_{i+1} \\ &= f^{n+1}(x_i) - a_n f^n(x_i) \ldots - a_{n-i+1} f^{n-i+1}(x_i) \\ &\quad - a_{n-i} x_n \ldots - a_2 x_{i+2} - a_1 x_{i+1} \\ &= f(f^n(x_i) - a_n f^{n-1}(x_i) \ldots - a_{n-i+1} f^{n-i}(x_i)) \\ &\quad - a_{n-i} x_n \ldots - a_2 x_{i+2} - a_1 x_{i+1}. \end{aligned}$$

Mais, en vertu de l'hypothèse de récurrence, $g(x_i) = 0$, la quantité entre parenthèses, qui est le début du développement de $g(x_i)$, peut être remplacée par :

$$a_{n-i}f^{n-i-1}(x_i) + \dots + a_2 f(x_i) + a_1 x_i$$

qui est égale à :

$$a_{n-i}x_{n-1} + \dots + a_2 x_{i+1} + a_1 x_i$$

et dont l'image par f est :

$$a_{n-i}x_n + \dots + a_2 x_{i+2} + a_1 x_{i+1}.$$

On a donc bien $g(x_{i+1}) = 0$ et le théorème fondamental est vérifié.

Exercice 42.

Soient V et W deux variétés linéaires affines d'un espace affine E respectivement parallèles aux deux sous-espaces vectoriels H et J de l'espace associé T. Soit $H \cap J = L$.

Admettons que $V \cap W$ ne soit pas vide et soit $x_0 \in V \cap W$. Cherchons à caractériser tous les autres points $x \in V \cap W$. Ce sont ceux qui satisfont à :

$$\left.\begin{array}{l} x - x_0 \in H \\ x - x_0 \in J \end{array}\right\} \text{ ce qui équivaut à } x - x_0 \in L$$

L étant un sous-espace vectoriel de T, l'ensemble des x tels que $x - x_0 \in L$ est une variété affine parallèle à L. Par conséquent, ou bien l'intersection de deux variétés affines est vide, ou bien elle est une variété affine.

Si, au lieu de deux variétés affines, on considère un ensemble fini ou infini de variétés affines, le raisonnement précédent reste valable, l'intersection de l'ensemble des sous-espaces vectoriels parallèles aux variétés étant encore un sous-espace vectoriel.

Soit alors A une partie de E. L'intersection de toutes les variétés affines qui contiennent A est une variété affine V_0 qui contient A et est incluse dans toutes les autres. C'est la plus petite variété affine qui contienne A. Le barycentre de points de $A \subset V_0$ est un point de V_0.

Réciproquement, montrons que tout point de V_0 est barycentre de points de A. Nous savons que tout point de V_0 peut être obtenu comme barycentre de points fixes x_0, x_i $(i \in I)$, de V_0, pourvu que les vecteurs $t_{x_0x_i} = x_i - x_0$ constituent un système de générateurs de H_0 (sous-espace vectoriel auquel V_0 est parallèle). Or, nous pouvons trouver dans A un tel système de points x_i ; si, en effet, A n'en contenait pas, c'est que les vecteurs t_{x_0x}, x décrivant A, n'engendreraient pas H_0, mais un sous-espace inclus strictement dans H_0, et V_0 ne serait pas la plus petite variété affine contenant A. V_0 est donc bien l'ensemble des barycentres des points de A.

Exercice 43.

a) $b - a = c - d \iff t_{ab} = t_{dc}$.

Ajoutons t_{bd} aux deux membres de l'égalité. La somme vectorielle étant commutative, il vient :

$$t_{ab} + t_{bd} = t_{bd} + t_{dc} \iff t_{ad} = t_{bc} \iff d - a = c - b.$$

b) Nous avons défini le parallélogramme $abcd$ construit sur $t_1 = b - a$ et $t_2 = d - a$ par le fait que $c = a + t_1 + t_2$. Or, $a + t_1$ est le point b ; l'égalité de définition de c s'écrit donc $c = b + t_2$, donc $t_2 = c - b$; d'où :

$$c - b = d - a.$$

Réciproquement, si on part de $c - b = d - a$, on peut dire :

$$c = b + (d - a) = a + (b - a) + (d - a)$$

qui est la définition du cours.

c) Précisons d'abord que le milieu de deux points dans un espace affine est le barycentre de ces points affectés de coefficients égaux et donc égaux à 1/2 ; 1/2 est l'élément du corps qui, multiplié par deux, donne l'unité. Cet élément n'existe pas dans un corps de caractéristique 2 ; il existe au contraire dans tout autre corps puisqu'un corps de caractéristique différente de 2 admet pour corps premier **Q** ou $\mathbf{Z}_n$ avec n premier impair et que dans ces corps un tel élément existe. La notion de milieu a donc un sens dans tout espace affine sur un corps de caractéristique différente de 2, mais n'en a pas dans un corps de caractéristique 2.

Ceci posé, soit le parallélogramme de sommet a construit sur t_1 et t_2. Le milieu de ad est déduit de a par la translation $\frac{t_{ad}}{2}$, c'est donc $a + \frac{t_1 + t_2}{2}$.

Le milieu de bc, barycentre de b et c affectés des coefficients 1/2, est le point déduit de a par la translation $\frac{1}{2}\, t_1 + \frac{1}{2}\, t_2$. C'est bien $a + \frac{t_1}{2} + \frac{t_2}{2}$.

Réciproquement, dire que ac et bd ont même milieu c'est dire que :

$$\frac{a}{2} + \frac{c}{2} = \frac{b}{2} + \frac{d}{2}$$

ce qui peut s'écrire en prenant a pour origine des vecteurs :

$$\frac{1}{2}\, t_{ac} = \frac{1}{2}\, (t_{ab} + t_{ad}).$$

Soit $t_{ac} - t_{ab} = t_{ad}$, c'est-à-dire $c - b = d - a$.

d) Si K est de caractéristique 2, tout scalaire y est égal à son opposé, ce qui entraîne que tout vecteur de T est égal à son opposé :

$$\forall m, n \in E \qquad t_{mn} = t_{nm} \qquad (1).$$

Soit alors un parallélogramme a, b, c, d, ce qui se traduit par :

$$t_{ab} = t_{dc} \qquad (2)$$

et nous devons montrer que :

$$t_{\sigma(a)\sigma(b)} = t_{\sigma(c)\sigma(d)},$$

σ étant une permutation quelconque. On sait qu'une permutation peut toujours se décomposer en un produit de transpositions portant sur deux éléments consécutifs. Pour que la propriété soit établie, il suffit de vérifier que (2) reste vraie si on y échange a et b ou bien b et d ou bien d et c. La possibilité du deuxième de ces échanges a été établie en *a*) dans le cas général ; les deux autres résultent de (1).

Exercice 44.

Le résultat est évident à partir de l'une ou l'autre des définitions données :

a) L'image du barycentre d'un ensemble de points par la composée de deux applications affines sera le barycentre des images ; cette application composée est donc une application affine.

b) Si l'application u est affine c'est que :

$$x - a \longrightarrow u(x) - u(a)$$

est une application linéaire. Si u' est affine c'est que :

$$u(x) - u(a) \longrightarrow u'(u(x)) - u'(u(a))$$

est linéaire. La composée des deux applications est linéaire, donc :

$$x - a \longrightarrow u' \circ u(x) - u' \circ u(a)$$

est linéaire et $u' \circ u$ est affine.

Exercice 45.

Soit u une application affine de E dans F qui, de tout $x \in E$, donne l'image :

$$u(x) = u(a) + \varphi(t_{ax}),$$

φ étant une application linéaire. Une variété affine V parallèle au sous-espace vectoriel H de S est :

$$x_0 + H = \{ x = x_0 + t \; ; \; t \in H \} .$$

L'image de cette variété est l'ensemble des points :

$$\{ u(x) = u(x_0 + t) \; ; \; t \in H \} ,$$

or,

$$u(x_0) = u(a) + \varphi(t_{ax_0}),$$
$$u(x) = u(x_0 + t) = u(a) + \varphi(t_{ax}).$$

D'où par soustraction :

$$u(x) - u(x_0) = \varphi(t_{ax}) - \varphi(t_{ax_0}).$$

Mais φ étant linéaire :

$$\varphi(t_{ax}) - \varphi(t_{ax_0}) = \varphi(t_{x_0x}),$$

t_{x_0x} décrit H ; $\varphi(t_{x_0x})$ est l'image par φ d'un sous-espace vectoriel ; c'est donc un sous-espace vectoriel $\varphi(H)$ et on a :

$$u(V) = \{ u(x) \; ; \; x \in V \} = u(x_0) + \varphi(H).$$

L'image $u(V)$ *de* V *par l'application affine* u *est donc une variété affine de* F *parallèle à* $\varphi(H)$.

Soit maintenant W une variété affine de F parallèle à un sous-espace J de l'espace vectoriel T associé à F ; cherchons-en l'image réciproque. Si $W \cap u(E) = \emptyset$, l'image réciproque de W est vide. Supposons que $W \cap u(E) \neq \emptyset$ et choisissons x_0 appartenant à cette intersection. Il existe au moins un élément $y_0 \in E$ tel que $u(x_0) = y_0$. La variété W peut être mise sous la forme : $W = x_0 + J$. Pour qu'un autre x appartienne à $\overset{-1}{u}(W)$, il faut et il suffit que :

$$u(x) - u(x_0) \in J$$

ou

$$\varphi(t_{x_0x}) \in J.$$

L'ensemble des x répondant à la question est donc celui des points déduits de x_0 par l'ensemble des translations t appartenant à :

$$\overset{-1}{\varphi}(J) = H.$$

H, image réciproque d'un sous-espace vectoriel par une application linéaire, est un sous-espace vectoriel. L'ensemble des x répondant à la question est $x_0 + H$.

$\overset{-1}{u}(W)$ est une variété affine. L'image réciproque d'une variété affine par une application affine est vide ou bien est une variété affine.

Exercice 46.

D'abord, il est immédiat que $E \times F$ a une structure d'espace affine sur $S \times T$. En effet, du fait que E et F ont respectivement des structures affines sur S et sur T, il résulte que :

$$\forall (x, y)\, (x', y') \in (E \times F)^2 \quad \exists\, !(s, t) \in S \times T \qquad \begin{array}{l} x' = s(x) \\ y' = t(y) \end{array}$$

ce qu'on pourra noter :

$$\begin{array}{l} x' = x + s \\ y' = y + t \end{array}$$

Cherchons alors ce qu'on peut dire du graphe d'une application linéaire de E dans F. On sait qu'elle s'exprime sous la forme :

$$u(x) = u(a) + \varphi(x - a),$$

φ étant une application linéaire de S dans T. Le graphe est l'ensemble des points $X = (x, u(x))$ où x décrit tout E. Il contient le point $A = (a, u(a))$ et

$$\begin{cases} x = a + (x - a) \\ u(x) = u(a) + \varphi(x - a) \end{cases}$$

peut se noter $\qquad X = A + (x - a, \varphi(x - a)).$

On obtient l'ensemble des X en faisant décrire à $x - a$ tout l'espace vectoriel S. On peut donc dire que le graphe est l'ensemble :

$$A + \{ (t, \theta = \varphi(t)) \; ; \; t \in S \} .$$

L'ensemble $H = \{ (t, \theta) \; ; \; t \in S \}$ est un sous-espace vectoriel de $S \times T$. En effet, du fait de la linéarité de φ,

$$\left.\begin{array}{l} (t, \theta) \in H \\ (t', \theta') \in H \end{array}\right\} \Longrightarrow (t, \theta) + (t', \theta') = (t + t', \varphi(t) + \varphi(t')) = (t + t', \varphi(t + t')) \in H$$

$$(t, \theta) \in H \quad \Longrightarrow \quad \lambda(t, \theta) = (\lambda t, \lambda\varphi(t)) = (\lambda t, \varphi(\lambda t)) \in H.$$

Le graphe est donc une variété affine de $E \times F$ parallèle à H.

Réciproquement, nous allons considérer une application de E dans F dont le graphe soit une variété affine de $E \times F$. Mais, auparavant, observons que le graphe G d'une application quelconque est une partie de $E \times F$ telle que la restriction à G de la projection pr_E de $E \times F$ sur E (voir exercice 18), soit une bijection (pour qu'à tout x de E corresponde un $y \in F$ et un seul, donc un couple (x, y) et un seul). Pour abréger, nous dirons d'une telle partie G qu'elle se projette bijectivement sur E.

Nous considérerons donc une application u dont le graphe soit une variété affine qui se projette bijectivement sur E. Elle est de la forme :

$$A + H = (a, u(a)) + \{ (t, \varphi(t)) \; ; \; t \in S \}$$

et pour montrer que l'application u est affine, il faut seulement montrer que φ est linéaire. Or, H étant un sous-espace vectoriel,

$$\left.\begin{array}{l} (t, \varphi(t)) \in H \\ (t', \varphi(t')) \in H \end{array}\right\} \quad \Longrightarrow \quad (t + t', \varphi(t) + \varphi(t')) \in H.$$

Mais, puisque dans H, il existe un seul point dont la projection sur E soit $t + t'$, c'est le point $(t + t', \varphi(t + t'))$. On en conclut donc :

$$\varphi(t) + \varphi(t') = \varphi(t + t').$$

On montre de la même façon que $\varphi(\lambda t) = \lambda\varphi(t)$. Donc, φ est une application linéaire (*) et u est affine.

Nous pouvons donc énoncer :

Pour qu'une application de E *dans* F *soit affine, il faut et il suffit que son graphe soit une variété affine de* $E \times F$ ou, en tenant compte de la remarque générale faite ci-dessus :

Pour qu'une partie de $E \times F$ *soit le graphe d'une application affine de* E *dans* F, *il faut et il suffit que ce soit une variété affine se projetant bijectivement sur* E.

(*) On pouvait aussi raisonner ainsi. L'application $(t, \varphi(t)) \to t$, restriction à H de la projection pr_E du produit $E \times F$ sur E est une application linéaire (v. ex. 18) ; cette restriction est ici bijective.

L'application réciproque $t \to [t, \varphi(t)]$ est aussi linéaire et $(t, \varphi(t)) \to \varphi(t)$ restriction à H de la projection de $E \times F$ est aussi linéaire. L'application $t \to \varphi(t)$, qui peut être considérée comme la composée des deux précédentes est linéaire.

Exercice 47.

La recherche du barycentre de n points $x_1 x_2 \dots x_n$ est une opération qui peut être rendue associative. En effet, soient $x_1 \dots x_p$ affectés des coefficients $\alpha_i (1 \leqslant i \leqslant p)$ et $x_{p+1} \dots x_n$ affectés des coefficients $\beta_j (p+1 \leqslant j \leqslant n)$ avec $\Sigma\alpha_i + \Sigma\beta_j = 1$.

Cherchons le barycentre de $x_1 \dots x_p$ affectés des coefficients :

$$\alpha'_i = \frac{\alpha_i}{\Sigma\alpha_i},$$

$\Sigma\alpha'_i = 1$ et ce barycentre est le point y défini par :

$$t_{ay} = \Sigma\alpha'_i\, t_{ax_i}.$$

Puis prenons le barycentre de ce point affecté du coefficient $\Sigma\alpha_i$ et des points $x_{p+1} \dots x_n$ affectés de leurs coefficients ; c'est le point x défini par :

$$t_{ax} = (\Sigma\alpha_i)\, \Sigma\alpha'_i\, t_{ax_i} + \Sigma\beta_j t_{ax_j}.$$
$$= \Sigma\alpha_i t_{ax_i} + \Sigma\beta_j t_{ax_j}.$$

La somme vectorielle étant associative, ce barycentre ne diffère pas de celui des points $x_1 \dots x_n$ affectés de leurs coefficients.

Ceci posé, dans la recherche du barycentre des points $x_1 \dots x_n$, affectés de coefficients positifs, d'une partie A d'un espace affine, remplaçons x_1 et x_2 par leur barycentre déterminé comme il vient d'être indiqué ; puis cherchons le barycentre de x_3 et de ce barycentre et ainsi de suite..., tous les coefficients utilisés et calculés à partir des coefficients positifs initiaux étant eux-mêmes positifs.

Or, par définition du segment $x_1 x_2$, le barycentre de ces deux points affectés de coefficients positifs appartient au segment, donc à l'enveloppe convexe de A, le barycentre construit ensuite lui appartient pour la même raison... et finalement le barycentre des n points lui appartient.

L'enveloppe convexe contient donc l'ensemble $\mathcal{E}$ des barycentres des points de A pris en nombre fini quelconque et affectés de coefficients positifs. Nous allons montrer maintenant que cet ensemble $\mathcal{E}$ est convexe; comme il contient évidemment A, cela suffira à établir qu'il est l'enveloppe convexe.

Soient deux points a et b de $\mathcal{E}$. Si A est indexé par I, on peut écrire :

$$a = \sum_{i \in I} \alpha_i x_i \qquad b = \sum_{i \in I} \beta_i x_i,$$

où les α_i (resp. β_i) sont nuls, sauf pour un nombre fini d'indices, réels positifs et tels que $\Sigma\alpha_i = 1$ (resp. $\Sigma\beta_i = 1$). Un point x du segment $[a, b]$ s'écrit :

$$x = \lambda a + \mu b,$$

où λ et μ sont réels positifs de somme 1. On a donc :

$$x = \sum_{i \in I} (\lambda\alpha_i + \mu\beta_i) x_i.$$

Or, les coefficients $\lambda\alpha_i + \mu\beta_i$ sont nuls, sauf pour un nombre fini d'indices, pour lesquels ils sont réels positifs et de somme 1, car :

$$\Sigma(\lambda\alpha_i + \mu\beta_i) = \lambda\Sigma\alpha_i + \mu\Sigma\beta_i = \lambda + \mu = 1.$$

x est un barycentre positif de points de A, il appartient à $\mathcal{E}$.

Exercice 48.

a) f est symétrique. En effet, en vertu de (Q), avec $x = 0$, on a :

$$Q(y) + Q(-y) = 2\,Q(y),$$

d'où : $Q(y) = Q(-y)$. On en déduit :

$$f(x, y) - f(y, x) = \frac{Q(y-x) - Q(x-y)}{4} = 0.$$

b) f est additive par rapport à x, c'est-à-dire :

$$f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y).$$

On a, en effet :

$$Q(x_1 + x_2 + y) = 2\,Q(x_1 + y) + 2\,Q(x_2) - Q(x_1 - x_2 + y) \qquad (1)$$
$$Q(x_1 + x_2 - y) = 2\,Q(x_1 - y) + 2\,Q(x_2) - Q(x_2 - x_1 + y) \qquad (2)$$

Echangeant le rôle de x_1 et x_2, on obtient :

$$Q(x_1 + x_2 + y) = 2\,Q(x_2 + y) + 2\,Q(x_1) - Q(x_2 - x_1 + y) \qquad (3)$$
$$Q(x_1 + x_2 - y) = 2\,Q(x_2 - y) + 2\,Q(x_1) - Q(x_1 - x_2 + y) \qquad (4)$$

La combinaison $\dfrac{(1) - (2) + (3) - (4)}{2}$ donne :

$$Q(x_1 + x_2 + y) - Q(x_1 + x_2 - y)$$
$$= Q(x_1 + y) - Q(x_1 - y) + Q(x_2 + y) - Q(x_2 - y)$$

qui n'est autre que la formule à établir.

La symétrie de f et l'additivité par rapport à x entraînent l'additivité par rapport à y.

c) Prouvons enfin que : $\forall \lambda \in \mathbf{R} \qquad f(\lambda x, y) = \lambda f(x, y)$.

L'additivité par rapport à x entraîne (cf. *A.P.M. I*, exercice 61) :

$$\forall r \in \mathbf{Q} \qquad f(rx, y) = r f(x, y).$$

Mais $Q(\lambda x + y)$ étant continue en λ, $f(\lambda x, y)$ l'est aussi. Sa valeur étant $r f(x, y)$ pour $\lambda = r$ rationnel, est $\lambda f(x, y)$ pour λ réel quelconque. La symétrie implique $f(x, \lambda y) = \lambda f(x, y)$. L'application f est bien bilinéaire.

Exercice 49.

Soient x et y appartenant à la boule unité :

$$\|x\| \leqslant 1 \qquad \|y\| \leqslant 1.$$

On doit montrer que $\lambda x + (1 - \lambda) y$, avec $0 < \lambda < 1$, appartient aussi à la boule. Or,

$$\|\lambda x + (1 - \lambda) y\| \leqslant \lambda \|x\| + (1 - \lambda) \|y\| \leqslant \lambda + 1 - \lambda = 1.$$

Exercice 50.

Les solutions de l'équation différentielle sont de la forme :

$$y = e^{-ax}(\lambda \cos bx + \mu \sin bx).$$

Les deux solutions :

$$y = e^{-ax} \cos bx$$
$$y = e^{-ax} \sin bx$$

qui sont linéairement indépendantes, constituent une base de l'espace vectoriel sur **R** des solutions, base par rapport à laquelle les coordonnées d'une solution sont λ et μ.

$\{ y_1, y_2 \}$ est une forme bilinéaire associée à la forme quadratique :

$$\{ y, y \} = \frac{b}{\pi} \int_0^{\frac{2\pi}{b}} e^{2ax} y^2 \, dx$$

qui est bien définie positive puisqu'elle ne s'annule que si y est nul sur $\left[0, \frac{2\pi}{b}\right]$, ce qui, la fonction intégrée étant périodique de période $\frac{2\pi}{b}$, entraîne que y soit la fonction nulle. $\{ y_1, y_2 \}$ est donc bien un produit scalaire. Il vaut :

$$\{ y_1, y_2 \} = \frac{b}{\pi} \int_0^{\frac{2\pi}{b}} (\lambda_1 \cos bx + \mu_1 \sin bx)\ (\lambda_2 \cos bx + \mu_2 \sin bx)\ dx.$$

On trouve :

$$\{ y_1, y_2 \} = \lambda_1 \lambda_2 + \mu_1 \mu_2$$

et

$$\|y\| = \sqrt{\lambda^2 + \mu^2}.$$

Il en résulte que la base par rapport à laquelle les calculs ont été faits est une base orthonormée. Une suite de Cauchy d'éléments de cet espace vectoriel est une suite (y_n) de solutions qui satisfait à :

$$\forall \varepsilon \quad \exists \mathrm{N} \qquad n > \mathrm{N},\ p > \mathrm{N} \quad \Longrightarrow \quad \|y_n - y_p\| < \varepsilon,$$

or

$$\|y_n - y_p\| = \sqrt{(\lambda_n - \lambda_p)^2 + (\mu_n - \mu_p)^2}.$$

L'hypothèse entraîne :

$$|\lambda_n - \lambda_p| < \varepsilon \qquad |\mu_n - \mu_p| < \varepsilon.$$

C'est dire que (λ_n) et (μ_n) sont deux suites de Cauchy de nombres réels. Elles convergent donc respectivement vers deux réels l et m ; c'est-à-dire que ε' étant arbitrairement petit pour n assez grand :

$$|\lambda_n - l| < \varepsilon' \qquad |\mu_n - m| < \varepsilon',$$

ce qui implique : $\sqrt{(\lambda_n - l)^2 + (\mu_n - m)^2} < \varepsilon' \sqrt{2}$.

Or, le premier membre est la distance entre la solution (y_n) et la solution de coordonnées (l, m). La suite (y_n) converge donc vers cette solution. L'espace est donc complet. Il est isomorphe à $\mathbf{R}^2$, muni de la norme euclidienne, dont on a donc montré du même coup qu'il était complet.

Avec la deuxième définition, calculons :

$$[y_1\ y_2] = \int_0^{\infty} e^{-2ax} (\lambda_1 \cos bx + \mu_1 \sin bx)\ (\lambda_2 \cos bx + \mu_2 \sin bx)\ dx$$

$$= \frac{\lambda_1 \lambda_2 + \mu_1 \mu_2}{2} \int_0^{\infty} e^{-2ax} dx + \frac{\lambda_1 \lambda_2 - \mu_1 \mu_2}{2} \int_0^{\infty} e^{-2ax} \cos 2bx \, dx$$

$$+ \frac{\lambda_1 \mu_2 + \lambda_2 \mu_1}{2} \int_0^{\infty} e^{-2ax} \sin 2bx \, dx.$$

Les deux dernières intégrales sont respectivement les parties réelles et imaginaires de :

$$\int_0^{\infty} e^{-2ax} e^{2ibx} dx = \frac{1}{2} \frac{a + ib}{a^2 + b^2}$$

On trouve donc :

$$[y_1, y_2] = \frac{\lambda_1 \lambda_2 + \mu_1 \mu_2}{4a} + \frac{\lambda_1 \lambda_2 - \mu_1 \mu_2}{4} \cdot \frac{a}{a^2 + b^2} + \frac{\lambda_1 \mu_2 + \lambda_2 \mu_1}{4} \cdot \frac{b}{a^2 + b^2}$$

$$= \frac{\lambda_1 \lambda_2}{4} \cdot \frac{2a^2 + b^2}{a(a^2 + b^2)} + \frac{\mu_1 \mu_2}{4} \cdot \frac{b^2}{a(a^2 + b^2)} + \frac{\lambda_1 \mu_2 + \lambda_2 \mu_1}{4} \cdot \frac{ab}{a(a^2 + b^2)}$$

D'où :

$$[y, y] = \frac{1}{4a(a^2 + b^2)} \left((a\lambda + \mu b)^2 + \lambda^2 (a^2 + b^2) \right)$$

$[y, y]$ est donc bien une forme quadratique définie positive et la définition adoptée pour le produit scalaire était légitime.

$[y, y]$ peut être mis sous la forme :

$$[y, y] = \left(\frac{a\lambda + \mu b}{2\sqrt{a}\sqrt{a^2 + b^2}} \right)^2 + \left(\frac{\lambda}{2\sqrt{a}} \right)^2$$

Sous cette forme, on a mis en évidence une base orthogonale formée des deux solutions :

$$\mathrm{Y}_0\ (\lambda = 0,\ \mu = 1) \qquad \mathrm{Y}_1\ (\lambda = -b,\ \mu = a).$$

Pour avoir une base orthonormée, il suffit de diviser ces vecteurs par leur norme :

$$\|\mathrm{Y}_0\| = \frac{|b|}{2\sqrt{a}} \frac{1}{\sqrt{a^2 + b^2}} \qquad \|\mathrm{Y}_1\| = \frac{|b|}{2\sqrt{a}}$$

Les deux solutions qui forment une base orthonormée seront :

$$z_0 = \frac{2\sqrt{a}\ \sqrt{a^2 + b^2}}{b} e^{-ax} \sin bx \qquad z_1 = \frac{2\sqrt{a}}{b} e^{-ax} (-b \cos bx + a \sin bx).$$

Par rapport à cette base, la forme $[y, y]$ a pour valeur :

$$[y, y] = \xi_0^2 + \xi_1^2$$

si ξ_0 et ξ_1 sont les coordonnées de y $(y = \xi_0 z_0 + \xi_1 z_1)$.

A partir de là il n'y a rien à changer à ce que nous avons dit relativement aux suites de Cauchy.

Exercice 51.

a) Le sous-espace $x^{\perp}$, orthogonal à un vecteur propre x, est l'ensemble des vecteurs de l'espace orthogonaux à ce vecteur propre. L'orthogonalité est conservée par la transformation orthogonale u, d'autre part les vecteurs orthogonaux à un vecteur sont orthogonaux aux vecteurs colinéaires à ce vecteur.

$$\left(f(x, y) = 0 \quad \Longrightarrow \quad f(\lambda x, y) = 0 \right).$$

On peut donc dire :

$$y \perp x \iff u(y) \perp u(x) \iff u(y) \perp x$$

D'où

$$u(x^{\perp}) = x^{\perp}.$$

De même si H est un sous-espace invariant par u (c'est-à-dire si $u(\mathrm{H}) = \mathrm{H}$), le même raisonnement prouve que :

$$[u(\mathrm{H})]^{\perp} = u(\mathrm{H}^{\perp}).$$

b) Soient deux vecteurs propres x et y correspondant aux deux valeurs propres réelles différentes $+1$ et -1 respectivement ; on a :

$$u(x) = x \qquad u(y) = -y.$$

Donc, $\quad f(x, y) = f(u(x), u(y)) = f(x, -y) = -f(x, y),$

ce qui entraîne (l'espace vectoriel E considéré est sur $\mathbf{R}$ et non pas sur un corps de caractéristique 2) :

$$f(x, y) = 0 \text{ c'est-à-dire } x \perp y.$$

c) Supposons alors que u admette un vecteur propre a_1 de valeur propre $\lambda = 1$ pour fixer les idées. Nous pouvons choisir une base de E constituée de a_1 et d'une base du sous-espace orthogonal H_1. Nous pouvons toujours choisir a_1 unitaire et la base de H_1 orthonormée si bien que la base de E ainsi obtenue est orthonormée. Par rapport à cette base, la matrice M de u prend la forme :

$$\begin{pmatrix} 1 & 0 \\ 0 & (\mathrm{M}_1) \end{pmatrix}$$

M_1 étant la matrice orthogonale qui représente la restriction de u au sous-espace H_1 invariant par u, matrice dont le polynôme caractéristique a mêmes racines que celui de la matrice M à l'exception d'une racine 1.

Nous refaisons pour M_1 ce que nous avons fait pour M et si elle admet encore la valeur propre 1 et le vecteur a_2, nous mettons M sous la forme :

$$\begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & 0 \\ 0 & (\mathrm{M}_2) \end{pmatrix}$$

par rapport à une base orthonormée formée de a_1, a_2 et d'une base orthonormée de H_2, sous-espace orthogonal de celui engendré par a_1 et a_2, M_2 étant la matrice de la restriction de u à ce sous-espace H_2 ... et ainsi de suite...

Nous poursuivrons ce processus jusqu'à épuisement des racines du polynôme caractéristique. Si nous avons trouvé k fois la racine 1, l fois la racine -1, et si le polynôme caractéristique n'admet plus de racines réelles, si u est rapporté à la base orthonormée formée de $k + l$ vecteurs propres choisis unitaires et deux à deux orthogonaux et d'une base orthonormée du sous-espace H' orthogonal à celui qu'ils engendrent, sa matrice prendra la forme :

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \\ & & & & & & (\mathrm{M}') \end{pmatrix}$$

où M' est la matrice à valeurs propres imaginaires de la restriction de u à H'. Dans le cas où $k + l = n$, la matrice est diagonalisée.

d) Sous cette forme on voit que E est la somme directe de trois sous-

espaces deux à deux orthogonaux, tous trois invariants par u : H' et les deux sous-espaces E_1 et E_{-1} respectivement engendrés par les k vecteurs de valeur propre 1 et les l vecteurs de valeur propre -1. Ces deux sous-espaces sont formés de vecteurs propres ; dans le premier, ce sont des vecteurs propres de valeur 1 et la restriction de u à E_1 est l'identité ; dans le deuxième, ce sont des vecteurs propres de valeur -1 et la restriction de u à E_{-1} est une symétrie par rapport à l'origine.

En dehors de ces deux sous-espaces, u ne possède pas de vecteurs propres car un vecteur propre de valeur -1 (par exemple) doit être orthogonal à tous les vecteurs propres de valeur 1 d'après *b*), donc appartenir à $\mathrm{E}_{+1}^{\perp}$ qui est $\mathrm{E}_{-1} \oplus \mathrm{H}'$. Un vecteur de $\mathrm{E}_{-1} \oplus \mathrm{H}'$ qui serait propre de valeur propre -1, ayant pour composante dans E_{-1} un vecteur propre de valeur propre -1 devrait avoir aussi pour composante dans H' un vecteur propre, ce qui est impossible si cette composante n'est pas nulle, puisque H' ne possède pas de vecteurs propres. Tout vecteur propre de u, de valeur propre -1, appartient donc à E_{-1}. On montre de la même façon qu'un vecteur propre de valeur propre 1 appartient à E_1.

L'ensemble des vecteurs propres de u est donc formé de deux sous-espaces orthogonaux dont les dimensions sont égales à l'ordre de multiplicité des racines 1 et —1 dans le polynôme caractéristique.

Exercice 52.

Nous utilisons les notations du cours. Soit u une transformation orthogonale de E et $\tilde{u}$ la transformation correspondante de F. Soient λ et $\bar{\lambda}$ deux valeurs propres imaginaires conjuguées de $\tilde{u}$ et soient :

$$v = e_1 + i\, e_2 \qquad \bar{v} = e_1 - i\, e_2 \qquad (e_1 \text{ et } e_2 \text{ réels})$$

les deux vecteurs propres correspondants :

$$\tilde{u}(v) = \lambda v \qquad \tilde{u}(\bar{v}) = \bar{\lambda}\bar{v}.$$

Le sous-espace de F engendré par v et $\bar{v}$ est invariant par $\tilde{u}$. Les vecteurs

$$e_1 = \frac{v + \bar{v}}{2} \qquad e_2 = \frac{v - \bar{v}}{2i}$$

constituent une base réelle de ce sous-espace. [Pour s'en assurer il suffit de vérifier qu'ils ne peuvent être nuls, ce qui signifierait que v et $\bar{v}$ sont ou réels ou imaginaires purs. Or, nous savons qu'ils ne peuvent être réels ; d'autre part, v imaginaire pur entraînerait $\tilde{u}(v)$ imaginaire pur, mais par ailleurs $v = iv_1$ avec v_1 réel signifierait $\tilde{u}(v) = \lambda v = \lambda i\, v_1$ qui ne pourrait être imaginaire pur que si λ était réel, ce qui implique une contradiction]. Le sous-espace engendré par v et v' peut donc être rapporté à la base e_1, e_2.

Revenons maintenant au domaine réel. Le sous-espace engendré dans F par e_1 et e_2 a pour intersection avec E un sous-espace $\mathcal{E}$ invariant par u, qui est la restriction de $\tilde{u}$ à E. Ce sous-espace et son orthogonal constituent deux sous-espaces supplémentaires, tous deux invariants par u, et la restriction de u à chacun d'eux est une transformation orthogonale. $\mathcal{E}$, de dimension 2, est isomorphe à $\mathbf{R}^2$ et, u n'y ayant pas de valeur propre, réelle, sa restriction y est une rotation d'angle $\alpha \neq k\,\pi$.

Si u n'a aucune valeur propre réelle (ce qui exige que la dimension de E soit paire ; $E \approx \mathbf{R}^{2p}$), la restriction de u au sous-espace orthogonal à ć est aussi à valeurs propres imaginaires. Nous recommencerons pour lui ce que nous avons fait pour E et le décomposerons en deux sous-espaces orthogonaux invariants de dimension 2 et $2p - 4$. La restriction de u au premier sera une rotation d'angle $\alpha' \neq k\pi$. Et ainsi de suite...

Finalement E est décomposée en p sous-espaces E_i deux à deux orthogonaux tous isomorphes à $\mathbf{R}^2$, tous invariants par u, et la restriction de u à chacun d'eux est une rotation d'angle différent de $k\pi$.

Si on choisit dans chacun de ces sous-espaces une base orthonormée, on voit que la matrice de la transformation u à valeurs propres imaginaires est de la forme :

$$\begin{pmatrix} (M_1) & & & \\ & (M_2) & & \\ & & \ddots & \\ & & & (M_p) \end{pmatrix}$$

chacune des matrices M_i étant de la forme :

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

et la matrice de la transformation orthogonale la plus générale peut être mise par rapport à une base orthonormée correctement choisie sous la forme :

$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & (M_1) & \\ & & & & & & & \ddots \\ & & & & & & & & (M_k) \end{pmatrix}$$

Remarque. — Les vecteurs e_1 et e_2 sont orthogonaux. En effet :

$$f(e_1, e_2) = f\left(\frac{v + \bar{v}}{2}, \frac{v - \bar{v}}{2i}\right) = \frac{1}{4i} f(v + \bar{v}, v - \bar{v})$$

$$= \frac{1}{4i}\,[f(v, v) - f(\bar{v}, \bar{v}) + f(\bar{v}, v) - f(v, \bar{v})]$$

La forme étant symétrique, les deux derniers termes se suppriment.

D'autre part v et $\bar{v}$, vecteurs propres imaginaires, sont isotropes et les deux premiers termes sont nuls.

$$f(e_1, e_2) = 0$$

e_1 et e_2 fournissent donc une base orthogonale du sous-espace qu'ils engendrent et par rapport à elle la matrice de u prend la forme :

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Il est aisé de le vérifier directement ; en effet, λ et $\bar{\lambda}$ satisfont à :

$$f(v, \bar{v}) = \lambda\bar{\lambda} f(v, \bar{v}) \qquad (1).$$

Or, $\quad f(v, \bar{v}) = f(e_1, e_1) + f(e_2, e_2) + i f(e_2, e_1) - i f(e_1, e_2)$ $= f(e_1, e_1) + f(e_2, e_2)$ puisque f est symétrique. Cette quantité est strictement positive, donc l'égalité (1) exige $\lambda\bar{\lambda} = 1$ ou $|\lambda| = 1$. (Nous avons montré en passant que les *valeurs propres imaginaires sont de module 1*). On peut donc poser :

$$\lambda = e^{i\beta} \qquad \bar{\lambda} = e^{-i\beta} \qquad \beta \neq k\pi.$$

On trouve alors :

$$u(e_1) = \frac{e^{i\beta} v + e^{-i\beta} \bar{v}}{2} = e_1 \cos\beta - e_2 \sin\beta$$

$$u(e_2) = \frac{e^{i\beta} v - e^{-i\beta} \bar{v}}{2i} = e_1 \sin\beta + e_2 \cos\beta$$

Par rapport à (e_1, e_2), la matrice de u s'écrit :

$$\begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}$$

C'est celle de la rotation d'angle $\alpha = -\beta$.

Exercice 53.

Transformations orthogonales dans $\mathbf{R}^3$.

Il existe au moins une valeur propre réelle et une direction propre correspondant à cette valeur. *A priori*, nous pouvons considérer quatre cas suivant la valeur du déterminant Δ de la matrice et la valeur propre réelle. L'existence de transformations orthogonales correspondant effectivement à ces quatre cas sera prouvée *à posteriori* par la possibilité de construire une matrice orthogonale conforme à ces données.

La restriction de la transformation au sous-espace orthogonal au vecteur propre est une transformation orthogonale dont la nature (directe ou inverse) est déterminée par la condition que son déterminant Δ' vérifie $\Delta = \lambda\,\Delta'$. Nous avons donc les quatre cas :

$\Delta = 1$	$\lambda = 1$	$\Delta' = 1$
$\Delta = 1$	$\lambda = -1$	$\Delta' = -1$
$\Delta = -1$	$\lambda = 1$	$\Delta' = -1$
$\Delta = -1$	$\lambda = -1$	$\Delta' = 1$

En prenant pour la transformation dans le sous-espace orthogonal au vecteur propre tous les types possibles de transformations orthogonales dans $\mathbf{R}^2$, nous obtenons tous les types possibles dans $\mathbf{R}^3$.

1° $\Delta = 1 \quad \lambda = 1 \quad \Delta' = 1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

2° $\Delta = 1 \quad \lambda = -1 \quad \Delta' = -1$:

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

3° $\Delta = -1 \quad \lambda = 1 \quad \Delta' = -1$:

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

4° $\Delta = -1 \quad \lambda = -1 \quad \Delta' = 1$:

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

De ces 8 matrices, on voit que la 3e et la 4e sont identiques à l'ordre des vecteurs de la base près, de même que la 5e et la 7e. Nous retiendrons donc 6 types distincts :

— transformations positives :

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

- identité

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

- rotation autour d'un axe,
- une direction propre,
- un sous-espace orthogonal invariant.

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

- (cas particulier de la précédente), demi-tour autour d'un axe,
- une direction propre,
- un sous-espace orthogonal de vecteurs propres.

— transformations négatives :

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

- symétrie $\times$ rotation,
- une direction propre,
- un sous-espace orthogonal invariant.

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

- symétrie,
- une direction propre,
- un sous-espace orthogonal de vecteurs propres.

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

- symétrie par rapport à 0 (tous les vecteurs sont propres).

Transformations orthogonales dans $\mathbf{R}^4$.

Cette fois, nous pouvons n'avoir aucune valeur propre réelle. L'existence d'une transformation de ce type et sa forme ont été établies dans l'exercice 52.

Admettant l'existence d'une valeur propre réelle λ, nous raisonnerons comme pour $\mathbf{R}^3$ en classant les cas suivant la nature de la transformation et la valeur de λ et en considérant pour le sous-espace orthogonal à la direction propre associée à λ tous les types possibles de transformations orthogonales. Ceci nous amène à envisager 12 cas dont plusieurs sont identiques, à l'ordre des vecteurs de la base près. Finalement, on trouve :

1° Six types de transformations positives.

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & & \\ \sin\alpha & \cos\alpha & & \\ & & \cos\alpha' & -\sin\alpha' \\ & & \sin\alpha' & \cos\alpha' \end{pmatrix}$$

avec $\alpha \neq k\pi$, $\alpha' \neq k\pi$.

- produit de deux rotations,
- pas de vecteurs propres,
- deux sous-espaces $\approx \mathbf{R}^2$ invariants,

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\alpha & -\sin\alpha \\ & & \sin\alpha & \cos\alpha \end{pmatrix}$$

- rotation,
- un sous-espace $\approx \mathbf{R}^2$ de vecteurs propres,
- un autre sous-espace $\approx \mathbf{R}^2$ invariant,

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

- identité.

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

- symétrie par rapport à un point,
- tous les vecteurs sont propres,

$$\begin{pmatrix} -1 & & & \\ & -1 & & \\ & & \cos\alpha & -\sin\alpha \\ & & \sin\alpha & \cos\alpha \end{pmatrix}$$

- produit de deux rotations dont un demi-tour,
- un sous-espace de vecteurs propres $\approx \mathbf{R}^2$,
- un sous-espace $\approx \mathbf{R}^2$ invariant,

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

- demi-tour,
- deux sous-espaces de vecteurs propres $\approx \mathbf{R}^2$.

2° Trois types de transformations négatives.

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

- symétrie par rapport à une droite,
- un sous-espace $\approx \mathbf{R}^3$ de vecteurs propres,
- une direction propre,

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

- symétrie par rapport à un hyperplan,
- un sous-espace $\approx \mathbf{R}^3$ de vecteurs propres,
- une direction propre,

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \cos\alpha & -\sin\alpha \\ & & \sin\alpha & \cos\alpha \end{pmatrix}$$

- symétrie $\times$ rotation,
- deux directions propres,
- un sous-espace $\approx \mathbf{R}^2$ invariant.

On vérifie bien sur ces exemples les résultats généraux des exercices 51 et 52.

Exercice 54.

1) Un espace vectoriel sur un corps satisfait aux axiomes de la structure d'espace vectoriel sur tout sous-corps de ce corps. $E(\mathbf{C})$ peut donc être considéré comme espace vectoriel sur $\mathbf{R}$. Soit (a_k) une base de $E(\mathbf{C})$. Tout $x \in E(\mathbf{C})$ peut être mis sous la forme :

$$\begin{aligned} x &= \sum_k \lambda^k a_k & \lambda^k \in \mathbf{C} \\ &= \sum_k (r^k + is^k) \cdot a_k & r^k, s^k \in \mathbf{R} \\ &= \sum_k r^k a_k + s^k (ia_k). \end{aligned}$$

Sous cette forme, on voit que l'ensemble :

$$(a_k) \cup (ia_k)$$

constitue un système de générateurs de E(**R**). Montrons que cet ensemble est formé de vecteurs linéairement indépendants. Soient (r^k) et (s^k) tels que :

$$\sum_k r^k a_k + \sum_k s^k (ia_k) = 0.$$

Cette égalité peut s'écrire :

$$\sum_k (r^k + is^k) a_k = 0.$$

Mais les a_k constituant une base de E(**C**), ceci entraîne :

$$\forall k \qquad r^k + is^k = 0 \quad \text{ou} \quad r^k = s^k = 0$$

$(a_k) \cup (ia_k)$ est donc une base de E (**R**). *E(**R**) est de dimension 2n.*

2) Nous dirons qu'une application est **C**-linéaire (resp. **R**-linéaire) lorsqu'elle est linéaire pour les structures d'espaces vectoriels sur **C** (resp. sur **R**).

Toute application **C**-linéaire de **C** dans **C** est de la forme :

$$r + is = \lambda \in \mathbf{C} \longrightarrow \alpha\lambda \in \mathbf{C}$$

α étant un élément fixe de **C**, soit $\alpha = a + ib \quad (a, b \in \mathbf{R})$.

On aura :

$$\alpha\lambda = (a + ib)(r + is) = ar - bs + i(br + as).$$

Cette application équivaut donc à l'application de $\mathbf{R}^2$ dans $\mathbf{R}^2$.

$$(r, s) \in \mathbf{R}^2 \longrightarrow (ar - bs, br + as) \in \mathbf{R}^2$$

définie par la matrice :

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Ce n'est donc pas l'application **R**-linéaire la plus générale de $\mathbf{R}^2$ dans $\mathbf{R}^2$ définie par une matrice quelconque :

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \text{ soit } \qquad (r, s) \in \mathbf{R}^2 \longrightarrow (ar + cs, br + ds).$$

Le résultat peut aussi être énoncé géométriquement : les seules applications **R**-linéaires de $\mathbf{R}^2$ dans lui-même définies par des applications **C**-linéaires de **C** dans **C** sont les similitudes de centre l'origine.

3) Une forme linéaire sur E(**C**) est une application :

$$f : \mathrm{E}(\mathbf{C}) \longrightarrow \mathbf{C}$$

qui, à tout élément de E(**C**), qui peut aussi être considéré comme élément de E(**R**), fait correspondre un élément de **C**. On peut donc la considérer comme une application de :

$$\mathrm{E}(\mathbf{R}) \longrightarrow \mathbf{C},$$

elle satisfait à :

$$f(x_1 + x_2) = f(x_1) + f(x_2)$$
$$f(\lambda x_1) = \lambda f(x_1)$$

pour tout $\lambda \in \mathbf{C}$, donc pour tout $\lambda \in \mathbf{R}$. Elle est donc **R**-linéaire, et

$$\mathcal{L}(\mathrm{E}(\mathbf{C}), \mathbf{C}) \subset \mathcal{L}(\mathrm{E}(\mathbf{R}), \mathbf{C}). \qquad (1)$$

Soit réciproquement une application **R** linéaire :

$$f : \mathrm{E}(\mathbf{R}) \longrightarrow \mathbf{C},$$

elle satisfait à :

$$\forall x_1, x_2 \in \mathrm{E}(\mathbf{R}) \qquad f(x_1 + x_2) = f(x_1) + f(x_2)$$
$$\forall r \in \mathbf{R} \qquad f(rx) = rf(x). \qquad (2)$$

Les éléments de E(**R**) étant ceux de E(**C**), elle définit une application de E(**C**) dans **C** qui satisfait au premier axiome de la linéarité. Examinons ce que vaut :

$$f(\lambda x) = f[(r + is)x] \qquad r \in \mathbf{R} \qquad s \in \mathbf{R}$$
$$= f(rx) + f(isx)$$
$$= rf(x) + sf(ix)$$

en vertu des hypothèses (2). Pour que f soit **C**-linéaire, il est nécessaire et suffisant que ceci soit égal à :

$$(r + is) f(x) = rf(x) + isf(x)$$

donc que :

$$f(ix) = if(x). \qquad (3).$$

Pour qu'une application **R**-linéaire de E(**R**) dans **C** soit une forme linéaire sur E(**C**), il faut donc qu'elle vérifie la condition (3) pour tout x. Il est clair qu'il n'en est pas ainsi pour toute application f. Une application linéaire peut en effet être définie par les images des vecteurs de la base qu'on peut choisir arbitrairement. On peut donc trouver des applications **R**-linéaires telles que :

$$f(ia_k) \neq if(a_k).$$

La relation d'inclusion (1) est donc stricte.

Le résultat de la question précédente n'était d'ailleurs que le cas particulier de celui-ci pour E(**C**) = **C** et par conséquent E(**R**) = $\mathbf{R}^2$.

4) φ et ψ sont des applications de E(**C**) ou E(**R**) dans **R**.

$$f(x_1 + x_2) = \varphi(x_1 + x_2) + i\psi(x_1 + x_2)$$
$$f(x_1) + f(x_2) = \varphi(x_1) + i\psi(x_1) + \varphi(x_2) + i\psi(x_2).$$

Donc :

$$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2)$$
$$\psi(x_1 + x_2) = \psi(x_1) + \psi(x_2).$$

D'autre part, $\forall \lambda \in \mathbf{C}$, on a :

$$f(\lambda x_1) = \varphi(\lambda x_1) + i\psi(\lambda x_1)$$
$$\lambda f(x_1) = \lambda\varphi(x_1) + i\lambda\psi(x_1).$$

On en déduit, $\forall \lambda \in \mathbf{R}$:

$$\varphi(\lambda x_1) = \lambda\varphi(x)$$
$$\psi(\lambda x_1) = \lambda\psi(x_1).$$

φ et ψ sont bien des formes **R**-linéaires.

D'autre part :

$$f(ix) = if(x) = i\varphi(x) - \psi(x).$$

Mais, par définition de φ et ψ :

$$f(ix) = \varphi(ix) + i\psi(ix).$$

Donc :

$$\varphi(ix) = -\psi(x) \qquad \psi(ix) = \varphi(x).$$

5) Il en résulte que la seule application de E (**C**) dans **C** qui admette $\varphi(x)$ comme partie réelle de $f(x)$ est définie par :

$$f(x) = \varphi(x) - i\varphi(ix).$$

Reste à montrer que l'application f ainsi définie est **C**-linéaire.

$\varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2)$ entraîne immédiatement

$$f(x_1 + x_2) = f(x_1) + f(x_2).$$

D'autre part :

$$f(\lambda x) = f((r + is)x) = \varphi((r + is)x) - i\varphi((ir - s)x)$$
$$= \varphi(rx) + \varphi(isx) - i[\varphi(irx) - \varphi(sx)]$$
$$= (r + is)\varphi(x) + (s - ir)\varphi(ix)$$
$$= (r + is)(\varphi(x) - i\varphi(ix)) = \lambda f(x).$$

On a donc établi une bijection entre l'ensemble $\mathcal{L}$ (E(**C**), **C**) et l'ensemble $\mathcal{L}$ (E(**R**), **R**).

Il est immédiat d'établir que cette bijection respecte la somme. D'autre part, à $r\varphi$ ($r \in \mathbf{R}$), cette bijection fait correspondre rf. Les deux espaces vectoriels $\mathcal{L}$ (E(C), C) et $\mathcal{L}$ (E(R), R) sont donc *isomorphes en tant qu'espaces vectoriels sur* **R**.

$$\mathcal{L}(E(C), C) \approx \mathcal{L}(E(R), R).$$

6) Supposons que E, espace vectoriel sur **R**, soit aussi espace vectoriel sur **C**. À tout $x \in E$, faisons correspondre $ix \in E$. Cette correspondance (homothétie de rapport i) est un automorphisme. Soit u cet automorphisme. Il vérifie :

$$u^2(x) = -x.$$

Réciproquement, supposons qu'il existe un automorphisme de E (considéré comme espace vectoriel sur **R**) tel que :

$$\forall x \qquad u^2(x) = -x$$

et définissons la multiplication externe des éléments de E par ceux de **C** de la façon suivante :

$$\forall x \in E \qquad ix = u(x)$$

la multiplication par un complexe quelconque $a + ib$ étant définie par la condition que distributivité et associativité soient vérifiées.

$$(a + ib)x = ax + bu(x).$$

Montrons que cette multiplication donne à E une structure d'espace vectoriel sur **C**. Comme il est déjà un groupe additif, il reste à vérifier que la nouvelle multiplication externe vérifie les axiomes de la structure d'espace vectoriel.

$$\begin{aligned}
(r + is + r' + is')x &= (r + r')x + (s + s')u(x) \\
&= rx + su(x) + r'x + s'u(x) \\
&= (r + is)x + (r' + is')x \\
(r + is)(x + x') &= r(x + x') + su(x + x') \\
&= r(x + x') + s(u(x) + u(x')).
\end{aligned}$$

(puisque u est additive).

$$\begin{aligned}
&= (rx + is)x + (r + is)x' \\
(r' + is')[r + is)x] &= (r' + is')(rx + su(x)) \\
&= r'rx + r'su(x) + s'u(rx + su(x)) \\
&= r'r + r'su(x) + s'ru(x) + s'su^2(x) \\
&= (r'r - ss')x + (rs' + sr')u(x) \\
&= [rr' - ss' + i(rs' + sr')]x \\
&= (r + is)(r' + is')x.
\end{aligned}$$

Enfin, l'élément neutre du corps pour la multiplication, n'ayant pas changé quand on a remplacé **R** par **C**, $1x$ reste égal à x.

Exercice 55.

Soit E l'espace vectoriel sur **R** où est définie une forme quadratique fondamentale $f(x, y)$ définie positive ; soit g l'application de E dans E* telle que :

$$f(x, y) = < x, g(y) >.$$

Soit $\varphi(x, y)$ une forme symétrique et ψ l'application de E dans E* telle que :

$$\varphi(x, y) = < x, \psi(y) >.$$

Deux vecteurs x et y dont le couple annule une des deux formes sont orthogonaux ou conjugués par rapport à cette forme. Pour simplifier le langage, nous conviendrons ici de nommer orthogonaux ceux dont le couple annule f et conjugués ceux dont le couple annule φ.

Le problème de diagonaliser la matrice Ψ de ψ (c'est-à-dire la matrice de la forme φ) par rapport à une base orthonormée de E est donc celui de trouver un système de vecteurs de E qui soient à la fois deux à deux conjugués et deux à deux orthogonaux.

1) Pour un vecteur de ce système, l'hyperplan orthogonal et l'hyperplan conjugué sont confondus. Autrement dit, si y est un tel vecteur, il y a identité entre l'ensemble des x qui annule la forme $g(y)$ et l'ensemble des x qui annule la forme $\psi(y)$. Pour cela, il est nécessaire et suffisant que ces deux formes soient les mêmes à une constante multiplicative près, y doit donc être tel que :

$$\exists \lambda \in \mathbf{R} \qquad \lambda g(y) = \psi(y)$$

g étant de rang égal à la dimension de l'espace admet une application réciproque g^{-1}. Cherchons l'image par g^{-1} des deux membres. Nous trouverons :

$$\lambda y = (g^{-1} \circ \psi)(y).$$

Sous cette forme, y apparaît comme un vecteur propre de l'application linéaire $g^{-1} \circ \psi$, application de E dans lui-même. Pour que y puisse être un élément d'une base orthogonale et conjuguée par rapport à φ, il est nécessaire et suffisant qu'il soit vecteur propre de l'application $g^{-1} \circ \psi$, les autres éléments devant être choisis dans l'hyperplan orthogonal.

2) L'équation caractéristique de $g^{-1} \circ \psi$ a toutes ses racines réelles. En effet, supposons qu'elle admette une racine imaginaire λ d'imaginaire conjuguée $\bar{\lambda}$. On en déduirait l'existence de deux vecteurs imaginaires conjugués v et $\bar{v}$ de l'espace F introduit en (IX, 5, 5) (voir aussi exercice 52) :

$$(g^{-1} \circ \psi)(v) = \lambda v \qquad \text{ou} \qquad \lambda g(v) = \psi(v)$$
$$(g^{-1} \circ \psi)(\bar{v}) = \bar{\lambda}\bar{v} \qquad \text{ou} \qquad \bar{\lambda} g(\bar{v}) = \psi(\bar{v}).$$

Soient alors les quantités $\varphi(v, \bar{v})$ et $\varphi(\bar{v}, v)$ qui doivent être égales puisque φ est symétrique. Elles valent :

$$\varphi(v, \bar{v}) = < v, \psi(\bar{v}) > = \bar{\lambda} < v, g(\bar{v}) >.$$
$$\varphi(\bar{v}, v) = < \bar{v}, \psi(v) > = \lambda < \bar{v}, g(v) >,$$

Comparons les derniers membres. g étant réel, $g(\bar{v})$ est imaginaire conjuguée de $g(v)$. On est donc ramené à comparer :

$$< v, \overline{g(v)} > \qquad \text{et} \qquad < \bar{v}, g(v) >.$$

Ces deux quantités sont égales (il suffit de se reporter à la forme que prend la forme bilinéaire canonique quand l'espace et son dual sont rapportés à des bases duales pour voir qu'il revient au même de remplacer l'élément de l'espace ou l'élément du dual par son imaginaire conjugué).

Par ailleurs, nous savons (cf. exercice 52, remarque) que $\varphi(v, \bar{v})$ ne peut être nul. L'égalité entre $\varphi(v, \bar{v})$ et $\varphi(\bar{v}, v)$ exige donc que λ soit réel.

Soit alors λ_1 une racine réelle de l'équation caractéristique de $g^{-1} \circ \psi$ et soit a_1 un vecteur propre correspondant. Nous pouvons toujours le choisir de norme 1 (norme définie à partir de f). La restriction de φ à l'orthogonal de a_1, qui est en même temps le sous-espace conjugué de la direction de a_1, est encore une forme bilinéaire symétrique. On pourra donc encore trouver une racine réelle de l'équation caractéristique de $g^{-1} \circ \psi$ et choisir un vecteur propre a_2 et ainsi de suite. Par ce procédé, nous obtiendrons n vecteurs réels et, si on veut, de norme 1, tels que par rapport à eux les deux matrices de f et φ soient diagonalisées (et celle de f réduite à la matrice unité si on a choisi les vecteurs de norme 1).

Remarque I : Si la matrice de f avait déjà été diagonalisée par rapport à une base orthonormée, la matrice G de g serait la matrice unité. Ce sont alors les racines de l'équation caractéristique de ψ, c'est-à-dire les valeurs propres de la matrice Ψ de φ qu'on serait amené à chercher.

Remarque II : Nous venons de généraliser les propriétés bien connues : une conique a toujours au moins un couple de directions conjuguées rectangulaires. Une quadrique a toujours au moins un triplet de directions conjuguées deux à deux rectangulaires. Ces directions fournissent une base orthonormée de l'espace par rapport à laquelle l'ensemble des termes du 2ᵉ degré de l'équation de la conique ou de la quadrique se présente comme une somme de carrés.

Exercice 56.

Considérons $f(\lambda x + y, \lambda x + y)$, x et y étant deux éléments de E et λ un nombre réel.

$$f(\lambda x + y, \lambda x + y) = \lambda^2 f(x, x) + \lambda[f(x, y) + f(y, x)] + f(y, y).$$

Or : $\quad f(x, y) + f(y, x) = f(x, y) + \overline{f(x, y)} = 2\,\mathcal{R} f(x, y)$

d'où : $\quad f(\lambda x + y, \lambda x + y) = f(x, x)\lambda^2 + 2\,\mathcal{R} f(x, y)\lambda + f(y, y).$

Ce trinôme en λ doit être positif ou nul quel que soit λ. Ceci exige :

$$[\mathcal{R} f(x, y)]^2 \leqslant f(x, x)\, f(y, y) \qquad (1).$$

On peut dire alors :

$$\sqrt{f(x, x)} = 0 \implies x = 0 \text{ par hypothèse.}$$

$$\sqrt{f(\lambda x, \lambda x)} = \sqrt{|\lambda|^2 f(x, x)} = |\lambda|\,\sqrt{f(x, x)}.$$

Enfin :

$$f(x + y, x + y) = f(x, x) + f(y, y) + 2\,\mathcal{R}\, f(x, x)$$
$$\leqslant f(x, x) + f(y, y) + 2\sqrt{f(x, x)}\,\sqrt{f(y, y)}$$
$$\sqrt{f(x + y, x + y)} \leqslant \sqrt{f(x, x)} + \sqrt{f(y, y)}$$

soit l'inégalité triangulaire. $\sqrt{f(x, x)}$ a bien les propriétés d'une norme.

Pour remplacer, dans le premier membre de l'inégalité (1), la partie réelle par le module, cherchons à remplacer x par un nombre kx tel que le deuxième membre ne change pas et tel que $f(kx, y)$ soit réel. La première de ces conditions sera réalisée en prenant pour k un nombre de module 1, soit $e^{i\theta}$:

$$f(e^{i\theta} x, e^{i\theta} x) = f(x, x),$$
$$f(e^{i\theta} x, y) = e^{i\theta} f(x, y).$$

Il suffira de choisir θ opposé à l'argument de $f(x, y)$ pour que

$$\mathcal{R} f(e^{i\theta} x, y) = |f(x, y)|.$$

Appliquée à $e^{i\theta} x, y$, l'inégalité devient :

$$|f(x, y)|^2 < f(x, x)\; f(y, y).$$

ERRATA DU TOME I

Page 13 : Relation avec le complémentaire.

2ᵉ ligne :

$$\complement(A \cup B) = \complement A \cap \complement B.$$

3ᵉ ligne :

$$\complement(A \cap B) = \complement A \cup \complement B.$$

Page 22 :

Ligne 22 du texte :

lire $E \longrightarrow E/R \longrightarrow f(E)$

au lieu de $E \longrightarrow f(E) \longrightarrow F.$

Page 33 : Section 4.

6 lignes avant la fin :

lire $A^{-1} = \{ b\,;\ \exists a \in A \quad e = ab \}$

au lieu de $A^{-1} = \{ b\,;\ \exists a \in A \quad c = ab \}.$

Page 36 : *Numéroter* (1) l'égalité

$$\varphi_b \circ \varphi_a = \varphi_{ba}$$

de la ligne 17.

Page 39 :

8ᵉ ligne à partir du bas :

lire : (2) $xy^{-1} \in g \implies (xy^{-1})^{-1} = yx^{-1} \in g$

au lieu de : (2) $xy^{-1} \in g \implies (xy^{-1}) = yx^{-1} \in g.$

Page 68 : § 4, Section 1 : *deux dernières lignes :*

lire : Tout homomorphisme de corps est injectif ou nul.

au lieu de : Il n'y a pas d'homomorphisme de corps qui ne soit un isomorphisme.

Page 91 : Exercice 61.

Remplacer croissante par monotone.

Page 111 : Solution de l'exercice 19.

Cas $n = 6$. Les deux premières structures indiquées sont isomorphes ; en effet, le groupe multiplicatif des entiers modulo 7 est cyclique de générateur 3 (ou 5) et est l'image isomorphe du groupe additif des entiers modulo 6 par l'isomorphisme f défini par $f(p) = 3^p$. Il n'y a donc que deux structures distinctes de groupe à 6 éléments.

A noter que dans l'exercice 57, il est prouvé plus généralement que le groupe multiplicatif de tout corps commutatif fini est cyclique. Le groupe multiplicatif des entiers modulo p, avec p premier, est donc toujours isomorphe au groupe additif des entiers modulo $p - 1$ et admet autant de générateurs qu'il existe de nombres premiers avec $p - 1$ et inférieurs à $p - 1$ ($\varphi(p - 1)$, fonction d'Euler).

Page 138 :

ligne 21 :

lire $\mathcal{G}'' = x^{r_1}\mathcal{A} + x^{r_1}\mathbf{Z} + \mathcal{G}'$
au lieu de $\mathcal{G}'' = x^{r_1}\mathcal{A} + \mathcal{G}'$.

ligne 26 :

lire $x^{r_2} = x^{r_1}(\mathrm{P} + n) + \mathrm{Q} \qquad \mathrm{P} \in \mathcal{A} \qquad n \in \mathbf{Z}$
au lieu de $x^{r_2} = x^{r_1}\mathrm{P} + \mathrm{Q}$.

Page 152 : Exercice 61.

8e ligne :

lire : Si f est croissante
au lieu de : Puisque f est croissante.

Dans la remarque, 1re ligne :

lire : monotone
au lieu de : croissante.

ligne 14 :

lire $\lambda p + \mu q$
au lieu de $\lambda(p + \mu q)$.

Index terminologique

I.F.Q.A.-Cahors - 31.148
Dépôt légal : IV-1963

Association des Professeurs de Mathématiques
de l'Enseignement Public,
29, rue d'Ulm, Paris (5^e^)

EXTRAIT DES STATUTS

Article II. — L'Association a pour but l'étude des questions intéressant l'enseignement des Mathématiques et la défense des intérêts professionnels de ses membres. Elle institue ou encourage des réunions, des discussions, des enquêtes sur l'enseignement des Mathématiques en France ou à l'étranger...

L'A.P.M. est ouverte à tous les Collègues enseignant dans les Facultés, les Grandes Ecoles, les Lycées, classiques, modernes ou techniques, les Ecoles Nationales Professionnelles, les Ecoles Normales, les Collèges d'Enseignement Général ou les Collèges Techniques, les écoles du premier degré, et en général toute personne exerçant dans l'enseignement public.

COTISATION. — Elle comprend l'abonnement au Bulletin, ainsi que les fascicules d'énoncés.

Cotisation normale .... 15 F

Cotisation réduite (stagiaires C.P.R., élèves des E.N.S. et des I.P.E.S., jeunes gens accomplissant leur service militaire, retraités) .... 8 F

ABONNEMENT (personnes n'appartenant pas à l'Enseignement Public, bibliothèques, etc...) :

France et autres pays : 20 F

Le numéro : 4 F

MODE DE PAIEMENT. — Virement postal (adressé au centre de chèques du tireur) ou mandat-carte à l'adresse :

A.P.M., 29, rue d'Ulm - PARIS, 5^e^ - C.C.P. Paris 5708-21

RECOMMANDATIONS DU TRESORIER. — Indications à porter sur le talon du chèque : 1° Nom (en majuscules) et prénom. — 2° Adresse où doit être envoyé le Bulletin. — 3° Ancienne adresse en cas de changement. — 4° Nom de l'établissement où l'on exerce. — 5° Nom de l'établissement précédent en cas de mutation en fin d'année scolaire.

N. B. — Toute nouvelle adhésion demandée en cours d'année scolaire compte à partir du 1^er^ octobre précédent. Elle donne droit à tous les bulletins déjà parus au cours de l'année scolaire, sous réserve qu'ils ne soient pas épuisés.

Voir, page 2 de la couverture, la liste des brochures de l'A.P.M.