Association des Professeurs de Mathématiques de l'Enseignement public
29, rue d'Ulm, Paris (5e)
Paris, Cc. 5.708-21

Les membres de l'Association (cotisation : 600 francs pour l'année scolaire) reçoivent gratuitement le *Bulletin* ainsi que les *Fascicules d'Enoncés.*

Cotisation pour les élèves des E.N.S. et les stagiaires des C.P.R. ...	300 F
Prix de l'abonnement au *Bulletin*	1 000 F
Prix du numéro	200 F

36e Année — No 179 *Octobre 1956*

BULLETIN de l'Association des Professeurs de Mathématiques de l'Enseignement Public

I. ÉTUDES

APPLICATIONS LINÉAIRES ET MATRICES *

Dans les conférences précédentes de ce cycle, les principales structures algébriques, celles de groupe, anneau, corps et celle d'espace vectoriel, ont été étudiées. Dans mes deux conférences consacrées aux applications linéaires d'un espace vectoriel dans un autre et à leur représentation par des matrices, nous rencontrerons des exemples variés de ces différentes notions.

Le sujet de mes deux conférences correspond exactement à ce qu'on nommait naguère la théorie des substitutions linéaires et, en toute honnêteté, l'optique moderne n'apporte pas ici de résultats substantiellement nouveaux. Ce que je vais dire, vous le savez tous. C'est l'économie de moyens, la puissance de généralisation qui fait le prix du point de vue moderne et de cette mutation de la règle du jeu mathématique qu'il impose.

Sous la pluie dense d'écritures et d'indices explicités de l'antique théorie des substitutions linéaires, on ne distinguait qu'assez péniblement les résultats fondamentaux et les idées directrices des démonstrations. Les déterminants et la théorie des équations linéaires fondée sur leur considération semblaient jouer un rôle essentiel.

Dans notre nouveau jeu, nous aimons écrire le moins possible, raisonner directement, intrinsèquement, sur les vecteurs ou les applications, distinguer clairement les grandes structures algébriques qui sous-tendent la théorie, et nous n'aimons plus les déterminants ; ils nous apparaissent comme un outil qui, lorsqu'on en fait trop usage, masque la simplicité des faits algébriques fondamentaux. Je ne les utiliserai pratiquement pas dans mes deux conférences, à la seule exception de la partie consacrée aux valeurs propres ; il est là aussi utile et possible de s'en passer, mais cela m'aurait entraîné trop loin. Je ne suppose pas non plus connus les résultats de la classique théorie des équations linéaires, et j'ai essayé de choisir les différentes démonstrations exposées de façon à mettre en lumière la souplesse et la richesse des modes de démonstrations de l'algèbre moderne.

La théorie que je suis amené à vous exposer vous choquera sans

(*) Cette *étude* est la rédaction, par l'auteur lui-même, des deux Conférences qu'il a prononcées, les 19 et 26 avril 1956, à l'Institut Henri-Poincaré, dans le cadre des conférences sur l'Algèbre organisées par la Société Mathématique de France, en accord avec l'A.P.M., à l'intention spéciale des professeurs.

doute plus qu'elle ne vous séduira parce que vous, comme moi, avons été initialement conditionnés à autre chose. Ne me reprochez pas trop d'être « abstrait » : il est bien difficile de savoir au juste ce que cette douce injure signifie et la dose de concret n'est pas proportionnelle à la quantité de craie éparse sur un tableau. Le concret est bien souvent précisément ce à quoi nous avons été conditionnés, les éléments de l'histoire de notre formation personnelle. Pour bien des mathématiciens de ma génération, l'algèbre moderne a paru abstraite jusqu'au moment où nous l'avons assimilée. A ceux de nos collègues plus jeunes qui ont — si j'ose dire — sucé dès l'enfance le lait des espaces vectoriels, ce sont eux qui apparaissent un splendide concret.

Rappel.

On rappelle qu'une structure d'*espace vectoriel* est définie sur un ensemble E, relativement à un corps commutatif K (corps des scalaires), par la donnée de deux lois de composition :

1° une loi de composition interne à E, notée additivement, définissant sur E une structure de groupe abélien ;

2° une loi de composition dite « le produit par un scalaire » qui à $\alpha \in K$, $\vec{x} \in E$ associe $\alpha\vec{x} \in E$ et qui jouit des propriétés suivantes :

$$1.\vec{x} = \vec{x} \qquad \alpha(\beta\vec{x}) = (\alpha\beta)\vec{x}$$

$$(\alpha + \beta)\vec{x} = \alpha\vec{x} + \beta\vec{x} \qquad \alpha(\vec{x} + \vec{y}) = \alpha\vec{x} + \alpha\vec{y}$$

quels que soient $\alpha, \beta \in K$, $\vec{x}, \vec{y} \in E$.

Les espaces vectoriels envisagés ici sont supposés *de dimension finie*, la dimension étant l'ordre maximum d'un système libre de vecteurs de E (ou système de vecteurs linéairement indépendants).

Un *sous-espace vectoriel* de E est une partie V de E telle que, quels que soient $\vec{x}, \vec{y} \in V$ et $\alpha \in K$, $\vec{x} + \vec{y}$ et $\alpha\vec{x}$ appartiennent à V. Les combinaisons linéaires des vecteurs d'un système de E sont les éléments d'un sous-espace de E, le sous-espace « engendré par le système ». Le *rang* du système de vecteurs est l'ordre maximum des systèmes libres de vecteurs que l'on peut extraire du système donné ; c'est aussi la dimension du sous-espace engendré par le système de vecteurs.

Etant donné un sous-espace V de E, il existe des sous-espaces U tels que tout vecteur de E soit, d'une manière et d'une seule, somme d'un vecteur de U et d'un vecteur de V. En particulier, $U \cap V = 0$; U est dit un sous-espace *supplémentaire* de V dans E.

I. Généralités sur les applications linéaires

1. *Notion d'application linéaire.*

a) Donnons-nous deux espaces vectoriels E et F, de dimensions respectives n et p, définis sur le même corps K de scalaires (par exemple le corps des réels ou le corps des complexes).

Définition : *Une application f de* E *dans* F *est dite linéaire si elle satisfait aux deux conditions suivantes :*

a) *quels que soient* $\vec{x}, \vec{y} \in E$ $\qquad f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$,

b) *quels que soient* $\vec{x} \in E$, $\alpha \in K$ $\qquad f(\alpha\vec{x}) = \alpha f(\vec{x})$.

Une application linéaire f de E *sur* F est dite un *homomorphisme* de E sur F ; si de plus f est *biunivoque*, f est dit un *isomorphisme* de E sur F.

Une forme linéaire sur E fournit un exemple d'application linéaire de E dans l'espace vectoriel à 1 dimension défini par K relativement à lui-même. Si la forme n'est pas identiquement nulle, on a même un homomorphisme de E sur K. La géométrie élémentaire met en jeu un grand nombre d'applications linéaires : étant donné dans l'espace de la géométrie élémentaire un trièdre *oxyz*, prenons pour E l'espace vectoriel de tous les vecteurs (libres), pour F celui des vecteurs du plan *xoy*, pour application f la projection parallèlement à *oz*. On a une application linéaire de E sur F.

De même, E restant le même, prenons F confondu avec E. L'homothétie de centre O, de rapport λ, définit une application linéaire biunivoque de E sur lui-même.

b) Les vecteurs de F images par f des vecteurs de E constituent une partie $f(E)$ de F qui *en est un sous-espace vectoriel*. En effet, si $(\vec{l}_i)$ est une base quelconque de E, au vecteur :

$$(1\text{-}1) \qquad \vec{x} = \sum_{i=1}^{n} x^i \vec{l}_i \qquad (x^i \in K)$$

correspond, f étant linéaire, le vecteur de F :

$$(1\text{-}2) \qquad f(\vec{x}) = \sum_{i=1}^{n} x^i f(\vec{l}_i).$$

Ainsi, tout élément de $f(E)$ s'exprime par une combinaison linéaire des $f(\vec{l}_i)$. Inversement, tout vecteur de F qui s'exprime par une telle combinaison linéaire avec des coefficients (x^i) est l'image par f de (1-1) et $f(E)$ est le sous-espace de F engendré par les $f(\vec{l}_i)$.

On appelle *rang de l'application linéaire f* la dimension de $f(E)$; ce rang r n'est autre que le rang du système des n vecteurs $f(\vec{l}_i)$ de F ; ainsi $r \leqslant n, p$.

c) Si l'application f de E sur $f(E)$ est *biunivoque*, $f(\vec{x}) = 0$ entraîne $\vec{x} = 0$. Si $f(\vec{x})$ défini par (1-2) est nul, $\vec{x}$ défini par (1-1) est nul et $x^i = 0$. Par suite, les $f(\vec{l}_i)$ forment un système libre et $r = n$, soit :

$$\dim. f(E) = \dim. E.$$

Inversement, considérons deux espaces E et F de *même dimension*. Il est aisé de construire des isomorphismes de l'un sur l'autre : si $(\vec{l}_i)$ est une base de E, $(\vec{f}_i)$ une base de F, appliquons $\vec{x} = \Sigma x^i \vec{l}_i \in E$ sur $f(\vec{x}) = \Sigma x^i \vec{f}_i \in F$. On vérifie immédiatement qu'on définit ainsi une application linéaire biunivoque f de E sur F, c'est-à-dire un isomorphisme. Ainsi, *pour que deux espaces vectoriels soient isomorphes, il faut et il suffit qu'ils aient même dimension.*

2. *Image d'un sous-espace vectoriel.*

a) Les applications linéaires f ont la propriété fondamentale de conserver la structure d'espace vectoriel. D'une façon précise, *si* V *est un sous-espace vectoriel de* E, *les vecteurs de* F *images par* f *de vecteurs de* V *constituent une partie* f(V) *de* F *qui en est un sous-espace vectoriel :* il suffit de restreindre f à V et d'appliquer le résultat que nous venons d'établir pour f(E). Si f est biunivoque, dim. V $=$ dim. f(V) ; le rang d'un système de vecteurs est conservé par isomorphisme.

Si W est un sous-espace de f(E), $\overset{-1}{f}$ (W) est l'ensemble des vecteurs de E dont l'image par f appartient à W. Il est clair que $\overset{-1}{f}$ (W) est encore un sous-espace de E ; en effet, si $\vec{x}, \vec{y} \in \overset{-1}{f}$ (W), $f(\vec{x}+\vec{y}) = f(\vec{x}) + f(\vec{y}) \in$ W et $\vec{x}+\vec{y} \in \overset{-1}{f}$ (W) ; de même si $\alpha \in$ K, $f(\alpha\vec{x}) = \alpha f(\vec{x}) \in$ W *et* $\alpha\vec{x} \in \overset{-1}{f}$ (W).

En particulier, si nous prenons W $= 0$, $\overset{-1}{f}$ (0) est un sous-espace de E.

b) Au sujet de ce sous-espace, nous allons établir le théorème suivant :

Théorème : *L'espace vectoriel* f(E) *est isomorphe à tout sous-espace de* E *supplémentaire de* $\overset{-1}{f}$ (0).

Soit L un sous-espace supplémentaire de $\overset{-1}{f}$ (0) dans E. A tout élément $\vec{z} \in$ L associons l'élément $f(\vec{z})$ de f(E). Il est aisé de voir que tout f(E) est ainsi atteint ; en effet, si $f(\vec{x})$ ($\vec{x} \in$ E) est un élément quelconque de f(E), considérons la décomposition :

$$\vec{x} = \vec{y} + \vec{z} \qquad (\vec{y} \in \overset{-1}{f} (0),\ \vec{z} \in \mathrm{L}).$$

On a :

$$f(\vec{x}) = f(\vec{y}) + f(\vec{z}) = f(\vec{z}),$$

puisque $f(\vec{y}) = 0$. L'application ainsi définie de L *sur* f(E) est biunivoque, puisque si $f(\vec{z}) = 0$, $\vec{z} \in \mathrm{L} \cap \overset{-1}{f}$ (0) donc est nul. Nous avons ainsi défini une application linéaire biunivoque de L sur f(E) et ces deux espaces sont isomorphes. On en déduit :

$$\text{dim. } \overset{-1}{f} (0) = n - r.$$

En particulier, pour que f soit un isomorphisme de E sur f(E), il faut et il suffit que $r = n$. Pour que f applique E *sur* F, il faut et il suffit que $r = p$. Pour que f soit un isomorphisme de E sur F, il faut et il suffit que $r = n = p$.

3. *L'espace vectoriel des applications de* E *dans* F.

A l'ensemble L(E, F) des applications linéaires de E dans F, nous pouvons étendre les lois de composition introduites sur l'ensemble des formes linéaires sur E. Introduisons sur L(E, F) les deux lois de composition suivantes :

1° Etant donnés $f, g \in$ L(E, F), nous pouvons leur associer l'application, notée $f + g$, définie par :

$$f + g : \quad \vec{x} \to f(\vec{x}) + g(\vec{x}) \quad \text{quel que soit } \vec{x} \in \mathrm{E}.$$

Il est clair que cette application est linéaire.

2° Etant donnés $\alpha \in$ K et $f \in$ L(E, F), nous pouvons leur associer l'application, notée αf, manifestement linéaire, définie par :

$$\alpha f : \quad \vec{x} \to f(\alpha\vec{x}) = \alpha f(\vec{x}) \quad \text{quel que soit } \vec{x} \in \mathrm{E}.$$

On vérifie aisément que l'ensemble de ces deux lois définit sur L(E, F) une structure d'espace vectoriel. Nous savons ainsi ajouter des applications linéaires et les multiplier par un scalaire. L'application nulle est celle qui à $\vec{x} \in$ E fait toujours correspondre $\vec{0} \in$ F ; l'application opposée de f est celle qui à tout $\vec{x}$ fait correspondre le vecteur opposé à $f(\vec{x})$.

4. *Produit de deux ou plusieurs applications linéaires.*

Considérons trois espaces vectoriels E, F, G de dimensions respectives n, p, q. Soit f une application linéaire de E dans F, g une application linéaire de F dans G. Considérons l'application de E dans G définie de la manière suivante : si $\vec{x} \in$ E, on obtient par f l'élément $\vec{y} = f(\vec{x})$ de F, puis par g l'élément $\vec{z} = g(\vec{y}) = g[f(\vec{x})]$ de G. Cette application de E dans G sera notée $g_\circ f$. Elle est manifestement linéaire, car :

$$g[f(\vec{x}_1 + \vec{x}_2)] = g[f(\vec{x}_1) + f(\vec{x}_2)] = g[f(\vec{x}_1)] + g[f(\vec{x}_2)]$$

$$g[f(\alpha\vec{x})] = g[\alpha f(\vec{x})] = \alpha g[f(\vec{x})]$$

pour $\alpha \in$ K, $\vec{x}$, $\vec{x}_2$, $\vec{x}_2 \in$ E.

L'application linéaire $g_\circ f$ est dite le *composé ou produit* des applications linéaires f et g. Ce produit s'étend immédiatement à plus de deux applications linéaires et est *associatif* d'après sa définition même. Il est *distributif* par rapport à l'addition des applications linéaires [dans L(E, F) ou L(F, G)].

II. Représentation par des matrices rectangulaires

5. *Représentation d'une application linéaire par une matrice.*

a) Reprenons les espaces vectoriels E et F de dimensions n et p et choisissons dans E une base $(\vec{l}_i)$. Cette base étant fixée, à toute application linéaire $f \in$ L(E, F) correspond un système de n vecteurs $f(\vec{l}_i)$ de F ; inversement, à tout système de n vecteurs (f_i) de F, faisons correspondre l'application f de E dans F qui applique :

$$\vec{x} = \sum_{i=1}^{n} x^i \vec{l}_i \in \mathrm{E}$$

sur le vecteur.

$$\vec{y} = \sum_{i=1}^{n} x^i \vec{f}_i \in F.$$

Cette application est linéaire et telle que $f(\vec{l}_i) = \vec{f}_i$.

Ainsi, une base étant choisie dans E, *nous avons défini une application biunivoque de l'ensemble* L(E, F) *des applications linéaires de* E *dans* F *sur l'ensemble des systèmes de n vecteurs de* F.

b) Pour définir les vecteurs d'un tel système, le plus commode est d'introduire leurs composantes dans une base de F. Etant donnée une base $(\vec{\varepsilon}_\mu)$ $(\mu = 1, 2, \dots p)$ de F, rapportons les $f(\vec{l}_i)$ à cette base :

$$(5-1) \quad \begin{cases} f(\vec{l}_1) = a_1^1 \vec{\varepsilon}_1 + a_1^2 \vec{\varepsilon}_2 + \dots + a_1^p \vec{\varepsilon}_p \\ f(\vec{l}_2) = a_2^1 \vec{\varepsilon}_1 + a_2^2 \vec{\varepsilon}_2 + \dots + a_2^p \vec{\varepsilon}_p \\ \dots\dots\dots\dots\dots\dots \\ f(\vec{l}_n) = a_n^1 \vec{\varepsilon}_1 + a_n^2 \vec{\varepsilon}_2 + \dots + a_n^p \vec{\varepsilon}_p \end{cases}$$

Nous sommes ainsi conduits à introduire les np nombres de K rangés en un tableau rectangulaire :

$$A = \begin{vmatrix} a_1^1 & a_1^2 & \dots & a_1^p \\ a_2^1 & a_2^2 & \dots & a_2^p \\ \dots & \dots & \dots & \dots \\ a_n^1 & a_n^2 & \dots & a_n^p \end{vmatrix} = \left(a_i^\mu\right) \qquad (i = 1, 2, \dots n \; ; \; \mu = 1, 2, \dots p).$$

A un tel tableau, désigné par la lettre A (ou une autre majuscule), on donne le nom de *matrice à n lignes et p colonnes* (ou matrice $n \times p$).

Ainsi, deux bases ayant été choisies l'une dans E, l'autre dans F, à toute application linéaire f correspond par (5-1) une matrice $n \times p$ qui est dite *représentative* de f dans les bases choisies. Inversement, étant donnée une matrice $n \times p$ à éléments de K, les éléments qui figurent dans une même ligne sont les composantes, dans la base de F, d'un vecteur $\vec{f}_i$; les n vecteurs de F ainsi obtenus définissent d'après *a*) une application linéaire f. Ainsi :

Théorème : *Une base de* E *et une base de* F *étant choisies, les formules* (5-1) *définissent une application biunivoque de* L(E, F) *sur l'ensemble des matrices* $n \times p$.

c) Considérons l'espace vectoriel K^p dont les éléments sont les suites ordonnées de p nombres de K et qui se trouve rapporté à une base canonique [la base (1, 0, ..., 0), (0, 1, ..., 0), etc...]. Etant donnée une matrice $n \times p$, soit A, chaque ligne de A définit un vecteur de K^p appelé un vecteur-ligne. On appelle *rang de la matrice* A *le rang du système des n vecteurs lignes.*

D'après le § 1, la donnée d'une base de F définit un isomorphisme entre K^p et F, dans lequel les vecteurs correspondants ont mêmes composantes. Pour un choix arbitraire de bases de E et F, le rang de A est égal au rang du système des n vecteurs de F définis par les vecteurs lignes et par suite est égal au rang de l'application f dont la matrice A est représentative. Ainsi, *toute matrice* A, *représentative d'une application linéaire f a même rang que cette application.*

De même qu'on calcule souvent sur les vecteurs par l'intermédiaire de leurs composantes ou sur les points par l'intermédiaire de leurs coordonnées, on conçoit qu'il est souvent commode de calculer sur les applications linéaires par l'intermédiaire de matrices représentatives. On aboutit ainsi à une véritable « géométrie analytique » des applications linéaires.

6. *Exemples de matrices représentatives.*

a) *Matrice-projection.*

Dans l'espace de la géométrie ordinaire, donnons-nous trois vecteurs unitaires $\vec{i}$, $\vec{j}$, $\vec{k}$, deux à deux orthogonaux. Prenons pour E l'espace vectoriel des vecteurs de la géométrie ordinaire, pour F le sous-espace engendré par $\vec{i}$ et $\vec{j}$, pour f la projection orthogonale de E sur F qui est une application linéaire de E sur F ; E étant rapporté à la base $(\vec{i}, \vec{j}, \vec{k})$, F à la base $(\vec{i}, \vec{j})$, on a :

$$\begin{cases} f(\vec{i}) = \vec{i} \\ f(\vec{j}) = \vec{j} \\ f(\vec{k}) = \vec{0} \end{cases} \qquad \text{d'où la mention représentative } A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

qui est de rang 2, dimension de $f(E) = F$.

b) *Matrice-homothétie.*

E restant le même, prenons F confondu avec E et pour f l'application linéaire de E sur lui-même défini par une homothétie de rapport λ ; E et F étant rapportés à la base $(\vec{i}, \vec{j}, \vec{k})$:

$$\begin{cases} f(\vec{i}) = \lambda\vec{i} \\ f(\vec{j}) = \lambda\vec{j} \\ f(\vec{k}) = \lambda\vec{k} \end{cases} \qquad \text{d'où la matrice représentative } A = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

qui est de rang 3.

7. *Algèbre des matrices rectangulaires : addition et multiplication par un scalaire.*

Désignons par M(n, p) l'ensemble des matrices $n \times p$ à éléments dans K. Nous nous proposons de définir sur M(n, p) une structure d'espace vectoriel relativement à K.

a) *Addition de deux éléments de* M(n, p). Soit A et B deux matrices $n \times p$ à éléments dans K. Si E et F sont deux espaces vectoriels relativement à K, de dimensions respectives n et p, et munis de bases, A et B sont les matrices représentatives de deux applications linéaires f et g de E dans F.

Etudions la matrice S représentative, pour ces bases $(\vec{l}_i)$ et $(\vec{\varepsilon}_\mu)$, de l'application $f+g$. Si $A=(a_i^\mu)$ et $B=(b_i^\mu)$, on a :

$$f(\vec{l}_i)=\sum_\mu a_i^\mu \vec{\varepsilon}_\mu \qquad g(\vec{l}_i)=\sum_\mu b_i^\mu \vec{\varepsilon}_\mu$$

et par suite : $$(f+g)(\vec{l}_i)=f(\vec{l}_i)+g(\vec{l}_i)=\sum_\mu a_i^\mu + b_i^\mu \vec{\varepsilon}_\mu$$

et S a pour éléments :

$$(7\text{-}1) \qquad s_i^\mu = a_i^\mu + b_i^\mu$$

A la matrice $n\times p$, S dont les éléments sont donnés par (7-1), nous donnerons le nom de somme des matrices A et B :

$$S = A + B.$$

La loi de composition ainsi définie doue $M(n, p)$ d'une structure de groupe abélien. La matrice O est celle dont tous les éléments sont nuls ; la matrice opposée à une matrice a pour éléments les opposés des éléments de cette matrice.

b) *Produit par un scalaire.* Les notations étant les mêmes et $\alpha\in K$, étudions la matrice P représentative de l'application αf. Il vient :

$$f(\vec{l}_i)=\sum_\mu a_i^\mu \vec{\varepsilon}_\mu \qquad \alpha f(\vec{l}_i)=\sum_\mu \alpha\, a_i^\mu \vec{\varepsilon}_\mu$$

et P a pour éléments :

$$(7\text{-}2) \qquad p_i^\mu = \alpha\, a_i^\mu$$

On vérifie immédiatement que les deux lois ainsi définies douent $M(n, p)$ d'une *structure d'espace vectoriel.*

c) Une base étant choisie dans E et une dans F, considérons l'application biunivoque Φ de $L(E, F)$ sur $M(n, p)$ qui à toute application linéaire $f\in L(E, F)$ fait correspondre sa matrice représentative $A=\Phi(f)$. D'après nos définitions de l'addition et du produit par un scalaire :

$$\Phi(f+g)=\Phi(f)+\Phi(g) \qquad \Phi(\alpha f)=\alpha\Phi(f).$$

Par suite, Φ *est un isomorphisme de l'espace vectoriel* $L(E, F)$ *sur l'espace vectoriel* $M(n, p)$.

Il est aisé de construire une base de $M(n, p)$: soit E_μ^i la matrice $n\times p$ dont tous les éléments sont nuls, à l'exception de l'élément e_i^μ qui vaut 1. Il est clair que si $A=a_i^\mu$.

$$A=\sum_{i,\mu} a_i^\mu E_\mu^i$$

Ainsi, toute matrice $n\times p$ s'exprime par une combinaison linéaire des E. De plus, pour la matrice O, tous les coefficients sont nécessairement nuls. Les E forment une base de $M(n, p)$ qui se compose de np éléments. Ainsi, *les espaces vectoriels isomorphes* $M(n, p)$ *et* $L(E, F)$ *ont la dimension np.*

8. *Produit de matrices rectangulaires.*

A partir de la composition ou produit d'applications linéaires, on peut définir dans certains cas le produit de matrices rectangulaires.

Soit E, F, G trois espaces vectoriels relativement à K, de dimensions respectives n, p, q, dans lesquels nous choisissons des bases respectives $(\vec{l}_i)$ $(\vec{\varepsilon}_\mu)$ $(\vec{\eta}_\lambda)$.

Si A est une matrice $n\times p$, B une matrice $p\times q$, A est représentative pour les bases choisies d'une application f de E dans F et B d'une application g de F dans G. Etudions la matrice D représentative de l'application $g\circ f$ de E dans G ; c'est une matrice $n\times q$. Si :

$$f(\vec{l}_i)=\sum_\mu a_i^\mu \vec{\varepsilon}_\mu \qquad g(\vec{\varepsilon}_\mu)=\sum_\lambda b_\mu^\lambda \vec{\eta}_\lambda$$

on a :

$$g[f(\vec{l}_i)]=\sum_\mu a_i^\mu g(\vec{\varepsilon}_\mu)=\sum_{\mu,\lambda} a_i^\mu b_\mu^\lambda \vec{\eta}_\lambda$$

soit :

$$g[f(\vec{l}_i)]=\sum_\lambda\Big(\sum_\mu b_\mu^\lambda a_i^\mu\Big)\vec{\eta}_\lambda$$

Ainsi, la matrice D a pour éléments :

$$(8\text{-}1) \qquad d_i^\lambda=\sum_\mu b_\mu^\lambda a_i^\mu.$$

La matrice D est dite le produit de B par A et est notée BA. On ne peut effectuer le produit BA que si le nombre des colonnes de A est égal au nombre des lignes de B, et ce produit s'effectue ligne de A par colonne de B. Les remarques suivantes sont évidentes :

a) Le produit des matrices est *associatif* : si A est $n\times p$, B est $p\times q$, C est $q\times r$, on a :

$$(CB)A = C(BA).$$

Cela résulte immédiatement du fait que cette associativité est vérifiée par les produits des applications linéaires que les différentes matrices représentent.

b) Le produit des matrices est *distributif* à droite et à gauche par rapport à l'addition : si A, A_1 et A_2 sont des matrices $n\times p$, B, B_1, B_2 des matrices $p\times q$:

$$B(A_1+A_2)=BA_1+BA_2 \qquad (B_1+B_2)A=B_1A+B_2A,$$

les mêmes relations étant vérifiées par les applications linéaires correspondantes.

c) A étant $n\times p$ et B $p\times q$, on peut effectuer BA, mais on ne peut effectuer AB que si $n=q$. S'il en est ainsi, BA est $n\times n$ et AB est $p\times p$. Même si en outre $n=p$, on voit facilement sur des exemples que AB et BA sont en général distincts. Le produit n'est pas commutatif.

9. *Exemples de produit.*

a) Considérons un plan rapporté à deux axes rectangulaires xoy de vecteurs unitaires $\vec{i}$ et $\vec{j}$; E, F, G étant confondus avec l'espace des vecteurs libres du plan, prenons pour f une rotation d'angle α et pour g une symétrie par rapport à ox. Il vient :

$$\begin{cases} f(\vec{i})=\vec{i}\cos\alpha+\vec{j}\sin\alpha \\ f(\vec{j})=\vec{i}\sin\alpha+\vec{j}\cos\alpha \end{cases} \quad \text{matrice représentative dans } (\vec{i},\vec{j}): \quad A=\begin{pmatrix}\cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha\end{pmatrix}$$

de même :

$$\begin{cases} g(\vec{i}) = \vec{i} \\ g(\vec{j}) = -\vec{j} \end{cases} \qquad \text{matrice représentative :} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

On a :

$$BA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & -\cos\alpha \end{pmatrix}$$

$$AB = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & -\cos\alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

Géométriquement, AB représente une symétrie par rapport à ox suivie d'une rotation d'angle α, c'est-à-dire une symétrie par rapport à la droite d'angle $\alpha/2$; BA représente la symétrie par rapport à la droite d'angle $-\alpha/2$.

b) Reprenons la projection (§ 6 *a*), mais en la considérant comme une application de E *dans* E. La matrice représentative dans $(\vec{i}, \vec{j}, \vec{k})$ est alors :

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

Géométriquement, $f \circ f = f$. Il en résulte $A.A = A$, ce qu'on peut vérifier immédiatement.

10. *Matrice représentative d'un vecteur par rapport à une base.*

Soit E un espace vectoriel relativement au corps K et considérons K comme un espace vectoriel à 1 dimension de base canonique définie par l'élément 1. A toute application linéaire ξ de K dans E correspond un vecteur $\vec{x} = \xi(1)$ de E. Inversement, à tout vecteur $\vec{x}$ de E correspond l'application linéaire ξ de K dans E :

$$\xi(1) = \alpha \rightarrow \alpha\vec{x} \qquad (\alpha \in K).$$

Nous définissons ainsi un isomorphisme entre E et L(K, E).

Choisissons dans E une base $(\vec{l}_i)$. On a :

$$\xi(1) = \vec{x} = x^1\vec{l}_1 + x^2\vec{l}_2 + \dots + x^n\vec{l}_n$$

et la matrice représentative de ξ dans les bases choisies est la matrice $1 \times n$.

$$X = (x^1 x^2 \dots x^n).$$

Nous disons que cette matrice est représentative du vecteur $\vec{x}$ dans la base $(\vec{l}_i)$.

Soit F un autre espace vectoriel relativement à K, f une application linéaire de E dans F. L'application $f \circ \xi$ est une application linéaire de K dans F qui correspond au vecteur $f(\vec{x})$. Si A est la matrice représentative de f relativement à $(\vec{l}_i)$ et une base $(\vec{\varepsilon}_\mu)$ de F, la matrice représentative de $f \circ \xi$ relativement à 1 et $(\vec{\varepsilon}_\mu)$ est AX. Cette matrice $1 \times p$ est représentative dans la base $(\vec{\varepsilon}_\mu)$ du vecteur $f(\vec{x})$ de F.

Nous sommes ainsi conduits à représenter les vecteurs d'un espace vectoriel par des matrices X à 1 ligne et autant de colonnes qu'il y a de

dimensions dans l'espace envisagé. Le vecteur image par une application linéaire est représenté par une matrice AX de même type, produit de la matrice représentant l'application linéaire par la matrice représentant le vecteur initial.

11. *Matrice représentative d'une forme linéaire par rapport à une base.*

Soit y^* une forme linéaire sur l'espace vectoriel F. Par définition, y^* n'est autre qu'une application linéaire de F dans K. Si $(\vec{\varepsilon}_\mu)$ est une base de F, 1 la base canonique de K, la matrice représentative de y^* se déduit des formules :

$$\begin{cases} y^*(\vec{\varepsilon}_1) = y^*_1 . 1 \\ y^*(\vec{\varepsilon}_2) = y^*_2 . 1 \\ \dots\dots \\ y^*(\vec{\varepsilon}_p) = y^*_p . 1 \end{cases} \qquad \text{soit } Y^* = \begin{vmatrix} y^*_1 \\ y^*_2 \\ \dots \\ y^*_p \end{vmatrix}$$

où Y^* est une matrice à *1 colonne*.

Soit f une application linéaire de E dans F. L'application $y^* \circ f$ est une application linéaire de E dans K, c'est-à-dire une forme linéaire sur E, ou un élément du dual E* de E. Si A est la matrice représentative de f relativement aux bases $(\vec{l}_i)$ de E et $(\vec{\varepsilon}_\mu)$ de F, la matrice représentative de $y^* \circ f$ relativement à (l_i) et 1 est Y^*A.

On a ainsi déduit de f une application linéaire f de* F* *dans* E* *définie par :*

$$f^* : y^* \in F^* \rightarrow y^* \circ f \in E^*.$$

C'est l'application duale de f. Si on la rapportait aux bases de F* et E* duales des bases introduites, il y aurait passage aux matrices transposées (échange des lignes et des colonnes) et retour au formalisme du § 10.

III. Opérateurs linéaires et matrices carrées

12. *Anneaux des opérateurs linéaires et des matrices carrées.*

a) Etant donné un espace vectoriel E de dimension n, on appelle *opérateur linéaire f* de E (ou *endomorphisme* de E) une *application linéaire de E dans lui-même.*

Toute la théorie précédente peut donc être appliquée aux opérateurs linéaires en supposant seulement F confondu avec E. Pour représenter un opérateur linéaire f par une matrice, il suffit ici de se donner *une* base $(\vec{l}_i)$ de E et de rapporter les vecteurs $f(\vec{l}_i)$ à cette base même :

$$f(\vec{l}_i) = \sum_{j=1}^{n} a_i^j \vec{l}_j$$

L'opérateur f est ainsi représenté relativement à la base $(\vec{l}_i)$ [la base $(\vec{\varepsilon}_\mu)$ du cas général est confondue avec $(\vec{l}_i)$] par la matrice $n \times n$ (matrice carrée) :

$$A = (a_i^j)$$

b) Soit L(E, E) l'ensemble des opérateurs linéaires de E ; L(E, E) se trouve naturellement doué d'une structure d'espace vectoriel. Mais ici il y a plus : le produit $g{\circ}f$ de deux opérateurs linéaires de E est *toujours défini* et est un opérateur linéaire de E. Cette loi de composition dans L(E, E) est associative, distributive à droite et à gauche par rapport à l'addition, mais elle n'est pas commutative. Enfin, il existe un opérateur linéaire remarquable, l'opérateur identité e, qui à tout vecteur de E fait correspondre le vecteur lui-même ; quel que soit $f \in$ L(E, E) :

$$f{\circ}e = e{\circ}f = f.$$

Ainsi, notre loi de composition admet un élément unité. Si $\alpha \in$ K, on a :

$$f{\circ}(\alpha e) = (\alpha e){\circ}f = \alpha f,$$

et la multiplication par le scalaire α peut être considérée comme un cas particulier du produit des opérateurs linéaires (produit à droite ou à gauche par αe).

Nous disposons ainsi sur L(E, E) de deux lois de composition :

une loi notée additivement douant L(E, E) d'une structure de groupe abélien ;

une loi notée $\circ$ satisfaisant aux axiomes des anneaux non commutatifs et admettant un élément unité.

THÉORÈME : *L'addition et le produit des opérateurs linéaires définissent sur* L(E, E) *une structure d'anneau non commutatif admettant un élément unité.*

c) Soit M(n, n) l'ensemble des matrices carrées $n \times n$; le produit de deux telles matrices A, B est toujours défini et $BA = (\sum_j b_j^k a_i^j)$. Une base $(\vec{l}_i)$ étant choisie dans E, considérons l'application biunivoque Φ de L(E, E) sur M(n, n) qui à tout opérateur linéaire f fait correspondre sa matrice représentative $A = \Phi(f)$. D'après la définition de l'addition et du produit des matrices :

$$\Phi(f + g) = \Phi(f) + \Phi(g) \qquad\qquad \Phi(g{\circ}f) = \Phi(g).\Phi(f),$$

et ce que nous venons d'établir sur L(E, E) est automatiquement valable pour M(n, n). En particulier, le produit admet pour élément unité la matrice I représentative de l'opérateur identité, soit :

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ . & . & . & . \\ 0 & 0 & & 1 \end{pmatrix} = (\delta_j^i).$$

Nous énoncerons :

THÉORÈME : *L'addition et le produit des matrices* $n \times n$ *définissent sur* M(n, n) *une structure d'anneau non commutatif, admettant un élément unité, et isomorphe à la structure d'anneau de* L(E, E).

Ces anneaux admettent des *diviseurs de zéro :* on entend par là que le produit de deux matrices peut être nul sans qu'aucune d'entre elles le soit, comme le montre l'exemple suivant :

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Enfin, il est aisé de définir la notion de *polynôme* à coefficients

dans K et à variable (et valeurs) dans l'un des anneaux des opérateurs ou des matrices carrées. Nous poserons :

$$A^0 = I \qquad A^1 = A \qquad A^2 = A.A \qquad A^p = A^{p-1}.A.$$

On a manifestement :

$$A^p.A^q = A^{p+q} = A^q.A^p.$$

Un monôme est ici un produit αA^p d'une puissance de A par $\alpha \in$ K. Un polynôme est une somme d'un nombre fini de monômes. A tout polynôme à coefficients et variable dans K :

$$P(\zeta) = \sum_{k=0}^{p} \alpha_k \zeta^k$$

on peut associer le polynôme de variable matricielle :

$$P(A) = \sum_{k=0}^{p} \alpha_k A^k \qquad (A \in M(n, n))$$

Si K est le corps des complexes, on peut mettre $P(\zeta)$ sous la forme :

$$P(\zeta) = \alpha_p \prod_{k=1}^{p} (\zeta - \mu_k)$$

On vérifie immédiatement par identification que :

$$P(A) = \alpha_p \prod_{k=1}^{p} (A - \mu_k I)$$

13. *Opérateurs linéaires réguliers et matrices régulières.*

a) Un opérateur linéaire f de E est dit *régulier* s'il définit une application linéaire de E *sur* E. Pour que f soit régulier, il faut et il suffit que f(E) coïncide avec E, c'est-à-dire que f soit de rang n. S'il en est ainsi, f est biunivoque et définit un isomorphisme de E sur E (§ 2), c'est-à-dire ce qu'on nomme un *automorphisme* de E. Un opérateur non régulier est dit *singulier.*

Le produit $g{\circ}f$ de deux opérateurs linéaires réguliers f et g est un opérateur linéaire lui-même manifestement régulier, puisqu'il applique E sur E. Inversement, si f et g sont deux opérateurs linéaires tels que $g{\circ}f$ soit *régulier*, f et g sont tous deux réguliers. En effet :

$$\text{rang } (g{\circ}f) = \dim g[f(E)] \leqslant \dim f(E) = \text{rang } f.$$

Comme rang $(g{\circ}f) = n$, on a rang $f = n$; f est donc régulier et $f(E) = E$; g(E) étant de dimension n, g est aussi régulier.

f étant régulier, donc biunivoque, à tout $\vec{y} \in$ E correspond un vecteur $\vec{x}$ et un seul tel que $\vec{y} = f(\vec{x})$; l'application biunivoque $\vec{y} \to \vec{x}$ de E sur E est l'application inverse $\overset{-1}{f}$ de f, et pour tout $\vec{x} \in$ E :

$$(13\text{-}1) \qquad \overset{-1}{f}[f(\vec{x})] = \vec{x},$$

ce qui caractérise l'application $\overset{-1}{f}$. Celle-ci est linéaire comme on le voit par un raisonnement identique à celui de § 2*a*. Ainsi, tout opérateur linéaire régulier f admet une application inverse $\overset{-1}{f}$ qui est aussi un opérateur linéaire régulier, l'opérateur inverse. Manifestement, l'inverse de $\overset{-1}{f}$ est f. La relation (13-1) et la relation obtenue par échange de f et $\overset{-1}{f}$ peuvent être traduites par :

(13-2) $$\overset{-1}{f} \circ f = f \circ \overset{-1}{f} = e.$$

Si deux opérateurs linéaires f et g sont tels que $g \circ f = e$, ils sont réguliers, et pour tout $\vec{x} \in E$:

$$g[f(\vec{x})] = \vec{x},$$

ce qui exprime que g est l'inverse de l'opérateur linéaire régulier f. Nous énoncerons :

THÉORÈME : *Le produit noté $\circ$ doue l'ensemble des opérateurs linéaires réguliers de* E *d'une structure de groupe : le groupe des automorphismes de* E.

Si f et g sont deux opérateurs linéaires réguliers, $g \circ f$ étant régulier admet un inverse qui n'est autre que $\overset{-1}{f} \circ \overset{-1}{g}$. Cette propriété valable dans tout groupe se vérifie immédiatement, car :

$$\overset{-1}{f} \circ \overset{-1}{g} \circ g \circ f = \overset{-1}{f} \circ f = e.$$

b) Toute matrice représentative d'un opérateur linéaire régulier est dite régulière. *Pour qu'une matrice $n \times n$ soit régulière,* il faut et il suffit qu'elle soit de rang n. Une base de E étant choisie, l'application Φ, qui à tout opérateur linéaire régulier fait correspondre sa matrice représentative, permet de traduire les résultats précédents concernant les opérateurs réguliers en résultats concernant les matrices régulières.

En particulier, à toute matrice $n \times n$ régulière A, on peut associer une matrice $n \times n$ unique $\overset{-1}{A}$ telle que, quelle que soit la matrice $1 \times n$ X :

$$\overset{-1}{A}\,AX = X,$$

c'est-à-dire telle que :

$$\overset{-1}{A}\,A = I,$$

Cette matrice qui est régulière est appelée l'inverse de A et satisfait aussi à :

$$A\overset{-1}{A} = I.$$

Nous énoncerons :

THÉORÈME : *Le produit des matrices $n \times n$ doue l'ensemble des matrices $n \times n$ régulières à éléments dans* K *d'une structure de groupe isomorphe au groupe des automorphismes d'un espace vectoriel sur* K *de dimension n. Ce groupe est appelé le groupe linéaire à n variables de* K : GL $(n,$ K$)$.

Si A et B sont deux matrices $n \times n$ régulières, BA régulière admet pour inverse $\overset{-1}{A}\,\overset{1}{B}$. On notera aussi que si A est régulière, chacune des égalités AB = 0 et BA = 0 entraîne B = 0. Car si $\overset{-1}{A}$ est l'inverse de A,

de AB = 0, on déduit $\overset{-1}{A}\,AB = 0$, soit B = 0 ; de même pour l'autre égalité.

[14. *Intervention des déterminants.*

a) A toute matrice carrée, $n \times n$, A, on peut associer son déterminant (à valeur dans K), D(A). Pour que ce déterminant soit nul, il faut et il suffit que les vecteurs lignes soient linéairement dépendants, c'est-à-dire que le rang de la matrice soit inférieur à n. Par suite, pour qu'une matrice soit régulière, il faut et il suffit que son déterminant soit $\neq 0$.

Si A et B sont deux matrices $n \times n$, la règle usuelle de multiplication des déterminants se traduit par la relation :

$$D(BA) = D(B) \, . \, D(A).$$

b) La théorie usuelle des systèmes de Cramer fournit immédiatement la valeur des éléments de la matrice $\overset{-1}{A}$ inverse d'une matrice régulière A. L'équation matricielle :

$$AX = Y$$

où Y ($1 \times n$) est donnée et X ($1 \times n$) inconnue se traduit par le système des n équations linéaires :

$$\sum_{i=1}^{n} a^j_i x^i = y^j \quad (j = 1, 2, \ldots n),$$

qui est un système de Cramer. Si α^i_j est le coefficient de développement dans D(A) de l'élément a^j_i, on sait que :

$$x^i = \frac{\sum_{j=1}^{n} \alpha^i_j y^j}{D(A)} = \sum_{j=1}^{n} \frac{\alpha^i_j}{A(D)} y^j = \sum_{j=1}^{n} b^i_j y^j$$

en posant :

(14-1) $$b^i_j = \frac{\alpha^i_j}{D(A)}$$

Si B = (b^i_j), on a la relation matricielle :

$$X = BY$$

et B est la matrice inverse de A].

15. *Utilisation des matrices carrées pour la détermination du rang d'une matrice quelconque.*

Etant donnée une matrice $n \times p$, A, on appelle matrice carrée d'ordre q $(q \leqslant n, p)$ extraite de A une matrice $q \times q$ obtenue en biffant $(n\text{-}q)$ lignes et $(p\text{-}q)$ colonnes de A. Le rang de la matrice A peut alors être déterminé à l'aide du théorème suivant :

THÉORÈME : *Le rang d'une matrice* A *est égal à l'ordre maximum des matrices carrées régulières que l'on peut extraire de* A.

Supposons A de rang r. Il nous faut montrer qu'il est impossible d'extraire de A une matrice carrée régulière d'ordre $s > r$, mais qu'il est possible d'en extraire au moins une d'ordre r.

A étant de rang r, il existe une relation linéaire au moins entre $s > r$ vecteurs-lignes arbitraires de A. Soit B une matrice carrée d'ordre s extraite de A. Ses lignes sont des parties de s lignes de A ; par suite, il existe entre elles une relation linéaire au moins et B ne peut être de rang s, donc ne peut être régulière.

Le second point est un peu plus long à établir : nous considérons l'espace vectoriel K^p muni de sa base canonique $(\vec{\varepsilon}_\mu)$.

Si A est de rang r, il existe un système libre de r vecteurs-lignes et nous pouvons supposer, pour simplifier les notations, qu'ils correspondent aux r premières lignes de A ; soit $(\vec{\lambda}_a)$ $(a = 1, ..., r)$ ces vecteurs, V_1 le sous-espace de K^p qu'ils engendrent.

Si $r < p$, il existe un $\vec{\varepsilon}_\mu$ au moins n'appartenant pas à V_1, donc linéairement indépendant des $\vec{\lambda}_a$; adjoignons-le aux λ et recommençons avec le nouveau système de vecteurs ainsi obtenu. On voit ainsi qu'on peut toujours choisir parmi les $(\vec{\varepsilon}_\mu)$ $(p-r)$ vecteurs tels qu'avec les $\vec{\lambda}_a$ ils définissent une base de K^p ; supposons que ce soit les $(p-r)$ derniers vecteurs $(\vec{\varepsilon}_{r+1} ... \vec{\varepsilon}_p)$, soit $(\vec{\varepsilon}_A)$ $(A = r + 1, ..., p)$.

Nous désignerons par V_2 le sous-espace de K^p engendré par les r vecteurs $(\vec{\varepsilon}_1, ..., \vec{\varepsilon}_r)$, soit $(\vec{\varepsilon}_a)$, et par V_3 le sous-espace engendré par les $(\vec{\varepsilon}_A)$. Il est clair que V_1 et V_2 sont tous deux supplémentaires de V_3.

Soit $\vec{x} = \sum_{\mu=1}^{p} x^\mu \vec{\varepsilon}_\mu \in V_1$; K^p étant somme directe de V_2 et V_3 :

$$\vec{x} = \vec{y} + \vec{z} \qquad (\vec{y} \in V_2,\ \vec{z} \in V_3),$$

et l'on a :

$$\vec{y} = \sum_{a=1}^{r} x^a \vec{\varepsilon}_a \qquad \vec{z} = \sum_{A=r+1}^{p} x^A \vec{\varepsilon}_A$$

L'application f de V_1 dans V_2 définie par $\vec{x} \to \vec{y}$ est manifestement linéaire. Elle est biunivoque puisque $\vec{y} = \vec{0}$ ne peut être obtenu que pour un vecteur $\vec{x}$ appartenant simultanément à V_1 et V_3, donc nul. Par suite, le rang de f est égal à la dimension r de V_1 et f est un *isomorphisme* de V_1 sur V_2.

Dans cet isomorphisme au vecteur :

$$\vec{\lambda}_a = \sum_{\mu=1}^{p} a_a^\mu \vec{\varepsilon}_\mu \quad (a = 1, ..., r)$$

correspond le vecteur :

$$f(\vec{\lambda}_a) = \sum_{b=1}^{r} a_a^b \vec{\varepsilon}_b \quad (a = 1, ..., r)$$

et la matrice $B = (a_a^b)$ représentative de l'isomorphisme f dans les

bases $(\vec{\lambda}_a)$ de V_1 et $(\vec{\varepsilon}_b)$ de V_2 est de rang r, donc régulière. C'est une matrice régulière d'ordre r extraite de A.

16. *Contravariance et matrices semblables.*

a) Soit g un opérateur linéaire régulier de E, $(\vec{l}_i)$ une base de E, P la matrice régulière représentative de g dans cette base. Les n vecteurs $\vec{\varepsilon}_i = g(\vec{l}_i)$ formant un système libre définissent une autre base de E. Inversement, étant donnée une base $(\vec{\varepsilon}_i)$ de E, il existe un opérateur linéaire g et un seul appliquant la base $(\vec{l}_i)$ sur la base $(\vec{\varepsilon}_i)$ (§ 1) : c'est l'opérateur manifestement régulier qui à $\vec{x}' = \Sigma x'^i \vec{l}_i$ fait correspondre $g(\vec{x}') = \Sigma x'^i \vec{\varepsilon}_i$.

Le vecteur $\vec{x} = g(\vec{x}')$, qui est un vecteur arbitraire de E, est défini dans la base $(\vec{\varepsilon}_i)$ par la matrice $1 \times n$:

$$X' = (x'^i).$$

Dans la base initiale (l_i), il est défini par la matrice :

$$X = PX'.$$

Ainsi, étant données deux bases $(\vec{l}_i)$ et $(\vec{\varepsilon}_i)$, telles que la seconde se déduit de la première par l'opérateur linéaire représenté dans la première par la matrice régulière P, un vecteur $\vec{x}$ défini dans la base initiale par la matrice X est défini dans la base finale par la matrice :

$$(16\text{-}1) \qquad X' = P^{-1} X.$$

C'est ce qu'on entend en disant que les composantes d'un vecteur se transforment de façon *contravariante*.

b) L'espace E étant muni de deux bases $(\vec{l}_i)$ et $(\vec{\varepsilon}_i)$, soit A une matrice $n \times n$. Elle représente, relativement à $(\vec{l}_i)$, un certain opérateur linéaire f. Si X représente par rapport à $(\vec{l}_i)$ un vecteur $\vec{x}$ de E, le vecteur $f(\vec{x})$ est représenté par :

$$Y = AX.$$

Relativement à $(\vec{\varepsilon}_i)$, $\vec{x}$ et $f(\vec{x})$ sont représentés respectivement par :

$$X' = P^{-1} X \qquad Y' = P^{-1} Y.$$

Par suite :

$$Y' = P^{-1} APX'$$

et la matrice représentative de f dans la base $(\vec{\varepsilon}_i)$ est :

$$B = P^{-1} AP \qquad [P \in GL(n, K)].$$

La matrice B est dite *semblable* à A dans le groupe linéaire ; A et B représentent le même opérateur linéaire par rapport à deux bases différentes. En particulier, étant donnée une permutation π de la suite des indices i, les matrices $A = (a_i^j)$ et $B = (a_{\pi(i)}^{\pi(j)})$ sont semblables puisqu'elles représentent le même opérateur linéaire par rapport aux bases $(\vec{l_i})$ et $(\vec{e}_{\pi(i)})$. Nous dirons dans ce cas que B se déduit de A par *permutation*.

Il est aisé de voir que la similitude est une *relation d'équivalence* entre éléments de $M(n, n)$: en particulier, la transitivité est immédiate ;

si $C = \overset{-1}{Q} BQ$ $[Q \in GL(n, n)]$ et $B = \overset{-1}{P} AP$,

$$C = \overset{-1}{Q}\,\overset{-1}{P} APQ = \overset{-1}{(PQ)} APQ$$

où PQ est régulière.

IV. Notions sur le spectre et la réduction des matrices carrées dans le groupe linéaire

17. *Le problème de la réduction.*

La similitude étant une relation d'équivalence dans $M(n, n)$, considérons l'ensemble des matrices semblables à une matrice donnée. Cet ensemble dont les éléments sont deux à deux semblables est dit une *classe d'équivalence* dans le groupe linéaire et $M(n, n)$ est ainsi partagé en classes d'équivalence.

Une classe étant connue dès qu'on en connaît une matrice, il est naturel de chercher parmi toutes les matrices d'une classe une matrice aussi simple que possible de telle sorte qu'à chaque classe on puisse associer d'une manière unique à une permutation près une matrice qui la caractérise. Cette matrice, semblable à toute matrice de la classe, est dite la *réduite* d'une matrice quelconque A de la classe.

La réduite fait intervenir des fonctions des éléments de A qui jouissent d'une propriété importante : elles prennent mêmes valeurs pour toutes les matrices semblables à A. Une telle fonction est dite un *invariant de la matrice* A *relativement au groupe linéaire*. Le problème de la réduction d'une matrice $n \times n$ dans le groupe linéaire est un assez difficile problème d'Algèbre ; nous ne le traiterons que dans un cas particulier.

18. *Matrices diagonales.*

L'un des exemples les plus simples de matrices carrées est fourni par les matrices diagonales : une matrice $n \times n$, A est dite diagonale si tous ses éléments sont nuls, à l'exception de ceux figurant dans la diagonale principale ; I par exemple est une matrice diagonale. Soit α_i l'élément figurant à la i^e place de la diagonale de la matrice diagonale A. Si $(\vec{l_i})$ est une base de E, A représente un opérateur linéaire f tel que :

$$f(\vec{l_i}) = \alpha_i \vec{l_i}.$$

Par suite, *chaque sous-espace de* E *de dimension* 1 *engendré par un* $\vec{l_i}$ *est invariant par* f. Nous écrirons $A = (\alpha_i)$.

Les matrices diagonales définissent un *sous-anneau commutatif* de $M(n, n)$. En effet, si $A = (\alpha_i)$ et $B = (\beta_i)$ sont deux matrices diagonales, elles représentent par rapport à $(\vec{l_i})$ deux opérateurs f et g tels que :

$$f(\vec{l_i}) = \alpha_i \vec{l_i} \qquad g(\vec{l_i}) = \beta_i \vec{l_i}$$

Par suite :

$$f(\vec{l_i}) + g(\vec{l_i}) = (\alpha_i + \beta_i)\vec{l_i} \qquad g[f(\vec{l_i})] = \beta_i\alpha_i \vec{l_i}$$

et $A + B = (\alpha_i + \beta_i)$, $BA = (\beta_i\alpha_i)$ sont des matrices diagonales, le produit BA étant manifestement commutatif.

Il est naturel de rechercher s'il est possible de prendre pour réduite une matrice diagonale. Cela est *impossible* en général, mais nous allons mettre en évidence un cas important où il en est ainsi.

19. *Vecteurs propres et valeurs propres.*

Dans toute la suite, nous supposons que K *est le corps des complexes* C.

a) Soit E un espace vectoriel relativement à C. Etant donné un opérateur linéaire f de E, cherchons les sous-espaces de dimension 1 invariants par f. Si $\vec{p} \neq 0$ appartient à un tel sous-espace, il existe un scalaire χ tel que :

(19-1) $\quad f(\vec{p}) = \chi\vec{p} \qquad \chi \in C$

ce qui peut s'écrire :

(19-2) $\quad (\chi e - f)\,(\vec{p}) = 0.$

S'il existe un tel vecteur $\vec{p}$, $(\chi e - f)$ ne peut être régulier. Inversement, si $(\chi e - f)$ est singulier, $(\chi e - f)^{-1}(0)$ est différent de zéro, et il existe des $\vec{p} \neq 0$ satisfaisant à (19-2) ou (19-1). Nous sommes ainsi amenés à la définition suivante :

Définition : *On appelle valeur propre d'un opérateur linéaire* f *toute valeur* $\chi \in C$ *telle que* $(\chi e - f)$ *soit singulier. Tout vecteur* $\vec{p} \neq 0$ *tel que* $f(\vec{p}) = \chi\vec{p}$ *est dit vecteur propre associé à la valeur propre* χ. *L'ensemble des valeurs propres de* f *est dit le spectre de* f.

b) [Adoptons dans E une base $(\vec{l_i})$ et soit A la matrice représentative de f ; $(\chi e - f)$ admet la matrice représentative $(\chi I - A)$. Pour que χ soit valeur propre de f, il faut et il suffit que $D(\chi I - A) = 0$, c'est-à-dire que χ soit un zéro de :

$$(19\text{-}3) \qquad \psi(\chi) = D(\chi I - A) = \begin{vmatrix} \chi - a^1_1 & -a^2_1 & \dots & -a^n_1 \\ -a^1_2 & \chi - a^2_2 & \dots & -a^n_2 \\ \cdot & \cdot & \cdot & \cdot \\ -a^1_n & -a^2_n & \dots & \chi - a^n_n \end{vmatrix}$$

Il importe de noter que, d'après sa définition même, le polynôme $\psi(\chi)$ de degré n est le même pour deux matrices semblables. En effet, si $B = \overset{-1}{P} AP$:

$$\chi I - B = \overset{-1}{P}(\chi I - A)P$$

et :

$$D(\chi I - B) = D(\chi I - A).$$

Au polynôme $\psi(\chi)$, on donne le nom de *polynôme caractéristique* de l'opérateur f ou de la matrice A. Ses zéros $\chi_1, \ldots, \chi_n$, distincts ou non, qui sont les valeurs propres de f, sont aussi dits les valeurs propres de A. Les différents coefficients de $\psi(\chi)$ sont des invariants de A dans le groupe linéaire. En particulier, les deux premiers termes de $\psi(\chi)$ s'écrivent manifestement :

$$\psi(\chi) = \chi^n - \left(\sum_i a_i^i\right)\chi^{n-1} + \ldots$$

Par suite, la fonction :

$$\text{tr. } A = \Sigma a_i^i = \chi_1 + \chi_2 + \ldots + \chi_n$$

est un invariant de A que l'on nomme la *trace* de cette matrice ou de l'opérateur f qu'elle représente dans une base $\vec{l}_i$].

20. *Réduction d'une matrice à valeurs propres distinctes.*

Soit A une matrice $n \times n$ à éléments dans C dont les valeurs propres $\chi_1, \ldots, \chi_n$ sont *toutes distinctes.* Dans l'espace vectoriel E sur C muni d'une base $(\vec{l}_i)$, A représente un opérateur f de valeurs propres (χ_i) que nous allons étudier. A chaque valeur propre χ_i correspond un vecteur propre $\vec{p}_i \neq 0$ au moins.

$$(20\text{-}1) \qquad f(\vec{p}_i) = \chi_i \vec{p}_i \qquad (i = 1, 2, \ldots, n).$$

Les n vecteurs propres $\vec{p}_i$ ainsi obtenus *forment un système libre si les χ_i sont distinctes.* En effet, supposons qu'il n'y ait que $q < n$ vecteurs $\vec{p}_i$ qui soient linéairement indépendants, par exemple les q premiers ; $\vec{p}_n$ peut alors s'exprimer d'une manière et d'une seule comme combinaison linéaire des p_μ $(\mu = 1, \ldots, q)$.

$$(20\text{-}2) \qquad \vec{p}_n = \sum_{\mu=1}^{p} \alpha_\mu \vec{p}_\mu \qquad (\alpha_\mu \text{ non tous nuls})$$

On en déduit :

$$f(\vec{p}_n) = \sum_\mu \alpha_\mu f(\vec{p}_\mu)$$

soit :

$$\chi_n \vec{p}_n = \sum_\mu \alpha_\mu \chi_\mu \vec{p}_\mu$$

Si $\chi_n = 0$, les χ_μ sont tous non nuls et il existerait une combinaison linéaire des $\vec{p}_\mu$ à coefficients non tous nuls qui serait égale à 0. On a donc $\chi_n \neq 0$ et

$$(20\text{-}3) \qquad \vec{p}_n = \sum_\mu \alpha_\mu \frac{\chi_\mu}{\chi_n} \vec{p}_\mu$$

qui doit coïncider avec (20-2) ; si $\alpha_\mu \neq 0$, on a donc $\chi_\mu = \chi_n$ et les valeurs propres ne seraient pas distinctes.

Le système libre $(\vec{p}_i)$ définit une base de E. Rapportons l'opérateur f à cette base ; d'après (20-1), f est représenté par la matrice diagonale :

$$R(A) = \begin{pmatrix} \chi_1 & 0 & \ldots & 0 \\ 0 & \chi_2 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & \chi_n \end{pmatrix}$$

qui est ainsi semblable à A.

Ainsi, toute matrice A à valeurs propres distinctes est semblable à une matrice diagonale dont les éléments sont les valeurs propres de A.

Dans le cas où les valeurs propres ne sont pas distinctes, les réduites ont des formes plus compliquées (réduites de Jordan).

Pour qu'une matrice B soit semblable à une matrice A, il est nécessaire qu'elles aient mêmes valeurs propres. Cette condition est suffisante si les valeurs propres sont distinctes, car les deux matrices sont alors semblables à R(A).

21. Théorème d'Hamilton-Cayley.

Indiquons une démonstration élémentaire d'une propriété importante du polynôme caractéristique $\psi(\chi)$ d'une matrice carrée A.

Soit $\alpha_j^i(\chi)$ le coefficient de développement dans $D(\chi I - A)$ de l'élément qui figure à la i^e ligne et à la j^e colonne. C'est un polynôme en χ de degré $(n - 1)$. Posons $C(\chi) = (\alpha_j^i(\chi))$; on a (§ 14) :

$$(\chi I - A)^{-1} = \frac{C(\chi)}{\psi(\chi)}$$

et par suite :

$$(21\text{-}1) \qquad C(\chi)(\chi I - A) = \psi(\chi) I$$

où $C(\chi)$ est un polynôme en χ à coefficients matriciels :

$$C(\chi) = \Gamma_1 \chi^{n-1} + \Gamma_2 \chi^{n-2} + \ldots + \Gamma_n \quad [\Gamma_i \in M(n, n)].$$

Si $\psi(\chi)$ s'écrit explicitement :

$$\psi(\chi) = \chi^n + c_1 \chi^{n-1} + \ldots + c_n,$$

on déduit de (21-1) par identification :

$$\begin{aligned} \Gamma_1 &= I \\ \Gamma_2 - \Gamma_1 A &= c_1 I \\ &\ldots \\ \Gamma_n - \Gamma_{n-1} A &= c_{n-1} I \\ -\Gamma_n A &= c_n I \end{aligned}$$

En multipliant les différents membres de ces relations successivement par $A^n, A^{n-1}, \ldots, A^0$ et ajoutant membre à membre, il vient :

$$\psi(A) = 0.$$

Théorème d'Hamilton-Cayley : *Si $\psi(\chi)$ désigne le polynôme caractéristique d'une matrice carrée A, on a :*

$$\psi(A) = 0.$$

André Lichnerowicz,
Professeur au Collège de France.