Cours, licence 1 ère année, 2 ème semestre.

Espaces vectoriels.

Jean-Robert Belliard

année 2012–13.

Table des matières

Chapitre 0

Introduction

Ce cours est destiné aux étudiants de première année deuxième semestre de la licence de Sciences et Techniques. Il s'adresse plus particulièrement aux étudiants qui souhaitent s'orienter vers une licence de mathématiques, ou bénéficier d'une formation plus solide en mathématiques que celle destinée aux autres mentions de licence.

Le contenu de ce cours, qui prolonge celui de l'unité algèbre du semestre 1, est tout à fait classique. On y introduit le vocabulaire de base de l'algèbre linéaire. Les connaissances en algèbre linéaire des étudiants seront approfondies tout au long de la licence de mathématiques. Pour l'essentiel il s'agit ici d'assimiler trois notions incontournables :

• La notion <u>d'espace vectoriel</u> : pour en donner un aperçu intuitif ce sont des généralisations de l'ensemble $\mathbb{R}^n$ (produit cartésien de $\mathbb{R}$ avec lui-même n fois), et des opérations linéaires sur cet ensemble. Plus précisément un espace vectoriel est défini à partir d'un ensemble E. Les éléments de E sont appelés vecteurs. On associe implicitement à E un corps de base noté $\mathbb{K}$ dont les éléments sont appelés les scalaires (pour $\mathbb{R}^n$ le corps $\mathbb{K}$ est égal à $\mathbb{R}$). On dit alors que E est un $\mathbb{K}$-espace vectoriel si on peut effectuer dans E des opérations $\mathbb{K}$-linéaires, c'est-à-dire si on sait définir le vecteur $\lambda.u + \mu.v$ à partir des données de deux vecteurs u et v et de deux scalaires λ et μ. Ces opérations linéaires doivent être compatibles avec les opérations du corps de base et sont soumises aux règles de calcul usuelles (voir définition 1.5). On peut ainsi mener tous les calculs linéaires dans E comme on en avait l'habitude lors des résolutions de systèmes pendant le premier semestre.

• La notion <u>d'applications linéaires</u> entre espaces vectoriels. Ce sont des applications entre deux espaces vectoriels, disons E et F, sur un même corps de base, disons $\mathbb{K}$. Pour être appelées linéaires, ces applications doivent être compatibles avec les opérations $\mathbb{K}$-linéaires. La définition de cette notion est très simple. Mais son utilisation est vraiment au cœur de l'algèbre linéaire. On peut même dire que toute l'algèbre consiste à manipuler des applications compatibles avec les structures algébriques (on parle de "morphismes" dans d'autres contextes).

• La <u>théorie de la dimension</u>. C'est le concept principal de ce cours, et ce serait un bon sous-titre pour l'unité “espaces vectoriels”. Si E est un $\mathbb{K}$ espace vectoriel le nombre $\dim(E)$ est ou bien un entier naturel ou bien égal à l'infini et s'appelle la dimension de E. Par exemple on a l'égalité $\dim(\mathbb{R}^n) = n$. La principale démonstration de ce cours sera de voir que la dimension de E (le nombre d'éléments d'une famille particulière de vecteurs appelée base) est intrinsèque à l'espace E. Ce qui signifie dépend de E mais pas de la base choisie. On démontrera aussi que du point de vue de l'algèbre linéaire deux espaces de même dimension sur un même corps sont “identiques” (on dit isomorphes). De sorte que la dimension est un système d'invariant complet de la classe d'isomorphisme des espaces vectoriels.

Ces concepts théoriques seront illustrés en pratique au moyen d'un outil de calcul avec lequel les étudiants sont normalement déjà familiarisés : les matrices. En fait tous les problèmes d'algèbre linéaire se ramènent à des résolutions de systèmes (linéaires). Et comme on l'a déjà vu en premier semestre la manière la plus efficace de mener les calculs sur de tels systèmes est d'utiliser les représentations matricielles. Ces deux points de vue (théorique et matriciel) s'enrichissent mutuellement. Tout l'art de l'algèbre linéaire consiste d'abord à maîtriser à la fois les aspects théoriques et matriciels et ensuite à basculer au moment opportun vers l'aspect le mieux adapté à la question courante !

Chapitre 1

Espaces vectoriels

1.1 Corps de base

Comme vu en introduction, la donnée d'un espace vectoriel sous-entend toujours celle d'un corps des scalaires. On commence donc par définir les corps.

Définition 1.1 *Soit E un ensemble. On appelle loi de composition interne sur E la donnée d'une application définie sur le produit cartésien $E \times E$ de E avec lui même et à valeur dans E. Pour (x, y) dans $E \times E$ l'usage est de noter $x + y$ ou $x * y$ ou $x.y$ ou $x \times y$ voire $x \perp y$ $\cdots$ l'image de (x, y) par cette application. Le résultat $x + y$ de l'opération interne est un élément de E.*

Des lois de composition internes particulières amènent à la notion de groupe :

Définition 1.2 *On appelle groupe (additif) un ensemble G muni d'une loi de composition interne disons $+$ tels que :*

1. *La loi $+$ est associative : $\forall a, b, c \in G$ $(a + b) + c = a + (b + c)$.*
2. *L'ensemble G est non vide et contient en particulier un élément e, appelé élément neutre pour $+$, et tel que : $\forall g \in G$ $e + g = g + e = g$.*
3. *Tout élément de g de G admet un inverse pour $+$, c'est-à-dire qu'il existe un $h \in G$ tel que : $h + g = g + h = e$, où e est l'élément neutre de G.*

La notation additive $+$ est en principe réservée aux groupes dits commutatifs dans lesquels la loi vérifie la propriété supplémentaire :

4. *La loi $+$ est commutative : $\forall a, b \in G$ $a + b = b + a$.*

Exemple : L'ensemble $\mathbb{Z}$ des entiers naturels $\mathbb{Z} = \{\cdots -2, -1, 0, 1, 2, 3 \cdots\}$ muni de l'addition habituelle

$$+: \quad \begin{aligned} \mathbb{Z} \times \mathbb{Z} &\longrightarrow \mathbb{Z} \\ (n, m) &\longmapsto n + m \ , \end{aligned}$$

est un groupe additif.

Pour des groupes généraux (non forcément commutatif) on ne parle pas de groupe additif et on privilégie d'autres notations que + pour indiquer les lois de composition internes. Pour des groupes additifs on préfère noter 0 plutôt que e l'élément neutre et on parle d'opposé de g pour l'inverse h de g qui vérifie $h + g = g + h = 0$.

Lorsqu'il existe, l'élément neutre pour une loi de composition interne $*$ est unique. En effet si e et e' sont deux éléments neutre pour $*$ alors on a $e * e' = e$ car e' est neutre mais aussi $e * e' = e'$ car e est neutre. Cela donne l'égalité $e' = e$.

Lorsqu'il existe, l'opposé de g pour une loi de composition interne + est unique. En effet si h et h' sont deux opposé de g alors on a $h' + g + h = h' + e = h'$ car h est un opposé de g. Mais on a aussi $h' + g + h = e + h = h$ car h' est un opposé de g. Cela donne l'égalité $h = h'$.

Dans un groupe additif l'usage est de noter $-g$ l'opposé de g pour la loi + (uniquement déterminé par g). Si la loi est notée multiplicativement on notera g^{-1} l'opposé de g.

On peut maintenant en venir à la notion de corps :

Définition 1.3 *On appelle corps (commutatif) la donnée d'un ensemble $\mathbb{K}$ et de deux loi de composition interne l'une additive notée + et l'autre multiplicative notée $\times$ sur $\mathbb{K}$ telles que :*

1. *L'ensemble $\mathbb{K}$ muni de l'addition + est un groupe (additif) pour lequel on note 0 l'élément neutre.*
2. *La multiplication est associative : $\forall x, y, z \in \mathbb{K},\ (x \times y) \times z = x \times (y \times z)$.*
3. *Il existe un neutre multiplicatif (l'élément unité du corps) noté 1 et tel que : $\forall x \in \mathbb{K},\ x \times 1 = 1 \times x = x$.*
4. *La multiplication est distributive à droite et à gauche par rapport à l'addition : $\forall x, y, z \in \mathbb{K},\ (x + y) \times z = x \times z + y \times z$ et $z \times (x + y) = z \times x + z \times y$.*
5. *La multiplication est commutative : $\forall x, y \in \mathbb{K},\ x \times y = y \times x$.*
6. *Tout élément non nul $x \neq 0$ de $\mathbb{K}$ admet un inverse pour la multiplication : $\forall x \in \mathbb{K} \setminus \{0\},\ \exists y \in \mathbb{K},\ x \times y = y \times x = 1$.*

Exemples :

- Les corps $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont supposés connus ici.
- Sur l'ensemble à deux éléments $\mathbb{F}_2 = \{0, 1\}$ on peut définir les opérations + et $\times$ comme suit $0+0 = 1+1 = 0$ et $1+0 = 0+1 = 1$; $0 \times 1 = 0 \times 0 = 1 \times 0 = 0$ et $1 \times 1 = 1$. Avec ces opérations $\mathbb{F}_2$ est le corps à deux éléments.
- Par acquis de conscience on rappelle comment on peut définir $\mathbb{C}$ à partir de $\mathbb{R}$. On appelle $\mathbb{C}$ l'ensemble (en bijection avec $\mathbb{R}^2$) dont les éléments z s'écrivent d'un et d'une seule manière $z = a + bi$ avec a et b dans $\mathbb{R}$. Par convention deux éléments $z = a + bi$ et $z' = a' + b'i$ de $\mathbb{C}$ sont égaux si et seulement si $a = a'$ et $b = b'$. On définit ensuite l'addition et la multiplication dans $\mathbb{C}$ à partir des opérations dans $\mathbb{R}$ et avec les formules :

$$(a + bi) + (a' + b'i) = (a + a') + (b + b')i$$

et

$$(a + bi) \times (a' + b'i) = (a \times a' - b \times b') + (a \times b' + a' \times b)i$$

Ces définitions permettent de voir les nombres réels comme des cas particulier de nombres complexes en identifiant $x \in \mathbb{R}$ au nombre complexe $x + 0\, i \in \mathbb{C}$. Les opérations entre nombres réels dans $\mathbb{R}$ ou dans $\mathbb{C}$ conduisent aux mêmes résultats. Cela conduit à voir le nombre “imaginaire pur” $i = 0 + 1 \times i \in \mathbb{C}$ comme une solution dans $\mathbb{C}$ à l'équation $i^2 = -1$, qui n'avait pas de solution dans $\mathbb{R}$. Pour démontrer qu'avec cette définition $\mathbb{C}$ est bien un corps on doit vérifier les propriétés 1 à 6 de la définition 1.3. Les propriétés 1 à 5 sont laissées au lecteur. On les vérifie en écrivant soigneusement les calculs et en utilisant les propriétés de $\mathbb{R}$. Ces calculs sont parfois long et doivent être fait avec soin, mais ils ne posent pas de vrai difficulté. Pour la propriété 6 on procède ainsi. On fixe $z = a + bi \in \mathbb{C}$ et on suppose $z \neq 0$. Alors a ou b au moins est non nul dans $\mathbb{R}$ et donc le module $|z| := a^2 + b^2$ est non nul dans $\mathbb{R}$ aussi. Mais par définition des opérations dans $\mathbb{C}$ on constate l'égalité $z\overline{z} = z(a - bi) = |z|$ de sorte que pour inverser z dans $\mathbb{C}$ il suffit d'inverser $|z|$ dans $\mathbb{R}$. Donc pour vérifier que $z = a + bi$ est inversible avec $a \neq 0$ ou $b \neq$ on peut simplement effectuer le calcul :

$$(a + bi) \times \left(\frac{a}{a^2 + b^2} + \frac{b}{a^2 + b^2} i \right) = 1 \ .$$

– À partir de l'ensemble $\mathbb{R}[T]$ des polynômes à coefficients réels en une indéterminée T, on peut définir l'ensemble des fractions rationnelles

$$\mathbb{R}(T) = \left\{ \frac{P(T)}{Q(T)} \;\middle|\; P(T), Q(T) \in \mathbb{R}[T], Q(T) \neq 0 \right\},$$

avec la convention

$$\frac{P}{Q} = \frac{R}{S} \iff PS = QR \ .$$

Cet ensemble qui a déjà été manipulé au premier semestre (réduction en éléments simples, par exemple) forme un corps avec les opérations usuelles sur les fractions.

1.2 Espaces vectoriels

La notion d'espace vectoriel est le cadre naturel pour des nombreuses applications en mathématiques ou à d'autres sciences. Par exemple l'espace plan $\mathbb{R}^2$ conduit à la géométrie d'Euclide, l'espace physique est usuellement modélisé par $\mathbb{R}^3$, l'espace-temps par $\mathbb{R}^4$, l'espace des états d'un ordinateur comprenant N bytes est $\mathbb{F}_2^N \cdots$

Pour la suite de ce cours on peut prendre, pour fixer les idées, $\mathbb{K} = \mathbb{Q}$ ou $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Il est néanmoins important de savoir que tout ce qui concerne l'algèbre linéaire reste valable, tel quel et avec les mêmes démonstrations, sur un corps $\mathbb{K}$ quelconque.

Définition 1.4 *Soit $\mathbb{K}$ un corps et soit E un ensemble (non vide). On appelle opération externe ou loi externe ou action de $\mathbb{K}$ sur E la donnée d'une application définie sur le produit cartésien $\mathbb{K} \times E$ et à valeur dans E.*

Définition 1.5 *Soit $\mathbb{K}$ un corps. On dit qu'un ensemble E est un* <u>*$\mathbb{K}$-espace vectoriel*</u> *lorsqu'il est muni d'une loi de composition interne commutative, notée + et d'une opération externe de $\mathbb{K}$, notée ., telle que :*

1. *$(E, +)$ est un groupe additif d'élément neutre noté 0_E ou 0.*
2. *Les lois + et . sont compatibles entre elles et avec les lois de la structure de corps de $\mathbb{K}$. C'est-à-dire qu'elles vérifient :*
   (a) $\forall \lambda \in \mathbb{K},\ \forall u, v \in E \quad \lambda.(u + v) = \lambda.u + \lambda.v$
   (b) $\forall \lambda, \mu \in \mathbb{K},\ \forall u \in E \quad (\lambda \times \mu).u = \lambda.(\mu.u)$
   (c) $\forall \lambda, \mu \in \mathbb{K},\ \forall u \in E \quad (\lambda + \mu).u = (\lambda.u) + (\mu.v)$
   (d) $\forall u \in E, \quad 1_{\mathbb{K}}.u = u$

Proposition 1.6 *Soit E un $\mathbb{K}$-espace vectoriel. Alors pour tout λ de $\mathbb{K}$ et tout u de E, on a :*

1. $\lambda.0_E = 0_E$.
2. $0.u = 0_E$
3. $(-1).u = -u$.

preuve : 1. Comme 0_E est neutre additif dans E, on a $0_E + 0_E = 0_E$. On en déduit $\lambda.(0_E + 0_E) = \lambda.0_E$ et donc avec la propriété 2a de la définition 1.5 on obtient $\lambda.0_E + \lambda.0_E = \lambda.0_E$. Pour conclure on ajoute de part et d'autre l'opposé de $-\lambda.0_E$ dans le groupe additif E et on obtient $\lambda.0_E = \lambda 0_E + \lambda.0_E - \lambda.0_E = \lambda.0_E - \lambda.0_E = 0_E$. □

2. Par la propriété 2c de la définition 1.5 on a $0.u + 0.u = (0 + 0).u = 0.u$. Si on ajoute l'opposé $-0.u$ à $0.u$ dans le groupe E de part et d'autre de ces équations on obtient $0_E = 0.u + 0.u - 0.u = 0.u$. □

3. Dans le groupe E, pour u fixé, l'opposé $-u$ est unique. En effet si $0_E = u+v = u+w$ alors on a aussi $w+u = 0_E$ comme + est commutatif et donc $w = w+u+v = v$. De sorte que pour montrer $(-1).u = -u$ il suffit de montrer que $u + (-1).u = 0_E$. Mais par la propriété 2d de la définition 1.5 on a $u + (-1).u = 1.u + (-1).u$. Et avec la propriété 2c de cette même définition on trouve $1.u + (-1).u = (1 + (-1)).u = 0.u$. De sorte que 3 suit de 2. □

Exemples : Voici des exemples (dont 2 sont absolument fondamentaux) d'espaces vectoriels :

1. $\{0_E\}$ l'espace vectoriel réduit à 0, sur n'importe quel corps $\mathbb{K}$.
2. Le corps $\mathbb{K}$ lui-même est un espace vectoriel sur $\mathbb{K}$ appelé droite vectorielle.
3. Pour tout entier naturel n le produit cartésien $E = \mathbb{K}^n$ de $\mathbb{K}$ avec lui-même n fois est un $\mathbb{K}$-espace vectoriel.
4. L'ensemble des polynômes $\mathbb{K}[X]$ en une indéterminée X et à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel.
5. L'ensemble $M_{n,p}(\mathbb{K})$ des matrices à n lignes, p colonnes et à coefficients dans $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel.

6. L'ensemble $\mathcal{F}(\mathbb{R},\mathbb{R})$ dont les éléments sont les applications de $\mathbb{R}$ dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.
7. L'ensemble $\mathcal{C}^0(\mathbb{R},\mathbb{R})$ dont les éléments sont les applications continues de $\mathbb{R}$ dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.
8. L'ensemble $\mathcal{C}^1(\mathbb{R},\mathbb{R})$ dont les éléments sont les applications dérivable à dérivée continue de $\mathbb{R}$ dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.
9. En général, si on se donne un ensemble X non vide et un $\mathbb{K}$-espace vectoriel E alors l'ensemble $\mathcal{F}(X,E)$ dont les éléments sont les applications de X dans E est un $\mathbb{K}$-espace vectoriel.

Preuve : On va définir les opérations et vérifier que les exemples 3 et 9 donnent bien un espace vectoriel. Le même travail pour les autres exemples peut être effectué en exercice.

L'exemple 3 : On fixe $n \geq 1$. Étant donné $x, y \in \mathbb{K}^n$ quelconques et $\lambda \in \mathbb{K}$ quelconque on doit d'abord définir $x+y$ et $\lambda.x$, puis vérifier que ces opérations satisfont les propriétés des définitions 1.2 et 1.5. Puisque $\mathbb{K}$ est un corps on peut définir les opérations coordonnées par coordonnées et utiliser la définition 1.3. On peut donc écrire $x = (x_1, \cdots, x_i, \cdots, x_n) = (x_i)_{i=1}^{i=n}$ et $y = (y_1, \cdots, y_i, \cdots, y_n) = (y_i)_{i=1}^{i=n}$ et poser :

$$x + y := (x_i + y_i)_{i=1}^{i=n} \qquad \text{et} \qquad \lambda.x := (\lambda x_i)_{i=1}^{i=n} \ .$$

On vérifie maintenant qu'avec ces deux lois $\mathbb{K}^n$ est bien un espace vectoriel sur $\mathbb{K}$. Pour ce on se donne des vecteurs x, y et z dans $\mathbb{K}^n$ de coordonnées respectives $x = (x_i)_i, y = (y_i)_i$ et $z = (z_i)_i$ et deux scalaires λ et μ dans $\mathbb{K}$. On a alors

1. associativité de $+$:

   $$(x+y)+z = (x_i+y_i)_{i=1}^{i=n} + z = ((x_i+y_i)+z_i)_{i=1}^{i=n} = (x_i+(y_i+z_i))_{i=1}^{i=n},$$

   parce que l'addition est associative dans $\mathbb{K}$. Ainsi

   $$(x+y)+z = x + (y_i+z_i)_{i=1}^{i=n} = x + (y+z).$$

   D'où l'associativité de l'addition $+$ dans $\mathbb{K}^n$.
2. On note $O_{\mathbb{K}^n} = (0, \cdots, 0, \cdots, 0)$ le vecteur dont toutes les coordonnées sont nulles. Alors pour tout $x = (x_i)_i$ dans $\mathbb{K}^n$ on a

   $$O_{\mathbb{K}^n} + x = (0 + x_i)_{i=1}^{i=n} = (x_i)_{i=1}^{i=n} = (x_i + 0)_{i=1}^{i=n} x + O_{\mathbb{K}^n} \ .$$

   Le vecteur $O_{\mathbb{K}^n}$ est donc un élément neutre pour l'addition $+$ dans $\mathbb{K}^n$.
3. Si $x = (x_i)$ est un élément quelconque de $\mathbb{K}^n$ alors le vecteur $-x = (-x_i)_{i=1}^{i=n}$ vérifie :

   $$x + (-x) = (x_i - x_i)_{i=1}^{i=n} = O_{\mathbb{K}^n} = (-x) + x \ .$$

   Donc le vecteur $-x$ est l'opposée de x pour la loi $+$ dans $\mathbb{K}^n$.

4. commutativité de + :

$$x + y = (x_i + y_i)_{i=1}^{i=n} = (y_i + x_i)_{i=1}^{i=n},$$

parce que l'addition est commutative dans $\mathbb{K}$. Ainsi

$$x + y = (y_i + x_i)_{i=1}^{i=n} = y + x \ .$$

Et l'addition + est commutative dans $\mathbb{K}^n$.

On a démontré que $\mathbb{K}^n$ muni de l'opération + forme un groupe commutatif. vérifions que les opérations + et . sont compatibles.

1. Distributivité de . par rapport à l'addition de $\mathbb{K}^n$:

$$\lambda.(x + y) = \lambda.(x_i + y_i)_{i=1}^{i=n} = (\lambda(x_i + y_i))_{i=1}^{i=n} = (\lambda x_i + \lambda y_i)_{i=1}^{i=n},$$

car la multiplication dans $\mathbb{K}$ est distributive par rapport à l'addition de $\mathbb{K}$. On en déduit :

$$\lambda.(x + y) = (\lambda x_i)_{i=1}^{i=n} + (\lambda y_i)_{i=1}^{i=n} = \lambda.(x_i)_{i=1}^{i=n} + \lambda.(y_i)_{i=1}^{i=n} = \lambda.x + \lambda.y \ .$$

Et la loi externe . est distributive par rapport à l'addition des vecteurs dans $\mathbb{K}^n$.

2. Compatibilité de . avec la multiplication dans $\mathbb{K}$:

$$(\lambda \times \mu).x = ((\lambda \times \mu) \times x_i)_{i=1}^{i=n} = (\lambda \times (\mu \times x_i))_{i=1}^{i=n} \ ,$$

car la multiplication dans $\mathbb{K}$ est associative. On en déduit :

$$(\lambda \times \mu).x = \lambda.(\mu \times x_i)_{i=1}^{i=n} = \lambda.(\mu.(x_i)_{i=1}^{i=n}) = \lambda.(\mu.x) \ .$$

Et les lois $\times$ et . sont bien compatibles.

3. Distributivité de . par rapport à l'addition dans $\mathbb{K}$:

$$(\lambda + \mu).x = ((\lambda + \mu)x_i)_{i=1}^{i=n} = (\lambda x_i + \mu x_i)_{i=1}^{i=n} \ ,$$

car la multiplication est distributive par rapport à l'addition dans $\mathbb{K}$. On en déduit : =

$$(\lambda + \mu).x = (\lambda x_i)_{i=1}^{i=n} + (\mu x_i)_{i=1}^{i=n} = \lambda.x + \mu.x \ .$$

Et la loi externe . est bien distributive par rapport à l'addition dans $\mathbb{K}$.

4. Compatibilité entre le neutre multiplicatif $1_{\mathbb{K}} \in \mathbb{K}$ et la loi externe . :

$$1_{\mathbb{K}}.x = (1_{\mathbb{K}} \times x_i)_{i=1}^{i=n} = (x_i)_{i=1}^{i=n} = x \ ,$$

car 1_K est le neutre multiplicatif dans $\mathbb{K}$.

Cela démontre que l'exemple 3 avec ces deux lois est bien un espace vectoriel sur $\mathbb{K}$. $\square$

Cet exemple est absolument fondamental parce qu'on verra par la suite que tous les espaces vectoriel (en dimension finie) sont isomorphes à un $\mathbb{K}^n$.

L'exemple 9 : Pour démontrer qu'un ensemble E est un espace vectoriel à ce stade du cours on n'a pas d'autre choix que de suivre les définitions. On procède donc en suivant le même plan de démonstration que pour l'exemple 3. C'est-à-dire on définit d'abord l'addition interne et la loi de composition externe sur $\mathcal{F}(X,E)$ puis on vérifie les huit propriétés des définitions 1.2 et 1.5. On fixe donc X un ensemble non vide, et E un $\mathbb{K}$-espace vectoriel. En particulier E contient son élément neutre additif 0_E, donc $E \neq \emptyset$ et $\mathcal{F}(X,E)$ contient l'application identiquement nulle $O_{\mathcal{F}(X,E)} = (\forall x \in X,\ x \mapsto 0_E)$ et n'est pas vide non plus. Dans toute cette démonstration on va s'appuyer sur la structure de $\mathbb{K}$-espace vectoriel de E. On se donne trois applications génériques $f, g, h \colon X \longrightarrow E$ dans $\mathcal{F}(X,E)$ et deux scalaires génériques λ, μ dans $\mathbb{K}$. On définit les opérations $+$ et $.$ sur $\mathcal{F}(X,E)$ point par point en se basant sur celles déjà définies dans E. Explicitement l'application $(f+g)$ est celle qui à tout x de X associe $f(x)+g(x)$, où l'addition est calculée dans E. L'application $\lambda.f$ est celle qui à tout x de E associe $\lambda.f(x)$, où le $.$ renvoie à l'opération externe $.$ sur E. Comme deux applications coïncident si et seulement si leurs images en tout point x de X sont égales, l'ensemble $\mathcal{F}(X,E)$ va hériter des huit propriétés des espaces vectoriel tout simplement parce qu'elles sont vraies dans E. On fixe un $x \in X$ générique. On a alors

1. associativité de $+$:

   $$((f+g)+h)(x) = (f+g)(x)+h(x) = (f(x)+g(x))+h(x) = f(x)+(g(x)+h(x))\ ,$$

   parce que l'addition est associative dans E. Ainsi

   $$((f+g)+h)(x) = f(x) + (g+h)(x) = (f+(g+h))(x)\ .$$

   On voit ainsi que les deux applications $f+(g+h)$ et $(f+g)+h$ coïncident en tout $x \in X$. Cela donne l'associativité de $+$ dans $\mathcal{F}(X,E)$.

2. On a déjà défini le neutre $O_{\mathcal{F}(X,E)}$ par la propriété $\forall x \in X,\ O_{\mathcal{F}(X,E)}(x) = 0_E$, où 0_E est le neutre dans E. Alors pour tout $x \in X$, comme 0_E est le neutre de E, on a :

   $$(O_{\mathcal{F}(X,E)} + f)(x) = (0_E + f(x)) = f(x) = f(x) + 0_E = (f + O_{\mathcal{F}(X,E)})(x)\ .$$

   Donc $O_{\mathcal{F}(X,E)}$ est un neutre pour l'addition $+$ sur $\mathcal{F}(X,E)$.

3. Soit $f \in \mathcal{F}(X,E)$ une application quelconque. On définit $\widetilde{f} \in \mathcal{F}(X,E)$ en posant pour tout x de X : $\widetilde{f}(x) = -f(x)$, où, pour x fixé, l'opposé $-f(x)$ est calculé dans E. On a alors, pour tout x dans X :

   $$(f+\widetilde{f})(x) = f(x) + (-f(x)) = 0_E = (-f(x)+f(x)) = (\widetilde{f}+f)(x).$$

   Cela donne l'identité entre applications $f+\widetilde{f} = O_{\mathcal{F}(X,E)} = \widetilde{f}+f$, et l'application $\widetilde{f}$ est bien un opposé à f relativement à la loi $+$ sur $\mathcal{F}(X,E)$.

4. commutativité de $+$:

   $$(f+g)(x) = f(x)+g(x) = g(x)+f(x)\ ,$$

   parce que l'addition est commutative dans $\mathbb{K}$. Ainsi

   $$(f+g)(x) = g(x)+f(x) = (g+f)(x)\ .$$

   Et l'addition $+$ est commutative dans $\mathcal{F}(X,E)$.

On a démontré que $\mathcal{F}(X, E)$ muni de l'opération $+$ forme un groupe commutatif. vérifions que les opérations $+$ et $.$ sont compatibles.

1. Distributivité de $.$ par rapport à l'addition de $\mathcal{F}(X, E)$:
$$(\lambda.(f+g))(x) = \lambda.((f+g)(x)) = \lambda.(f(x)+g(x)) = \lambda.f(x) + \lambda.g(x) \ ,$$
car la multiplication dans $\mathbb{K}$ est distributive par rapport à l'addition de E. On en déduit :
$$(\lambda.(f+g))(x) = (\lambda.f)(x) + (\lambda.g)(x) = (\lambda.f + \lambda.g)(x) \ .$$
Comme x est quelconque, cela donne l'identité entre applications $\lambda.(f+g) = (\lambda.f + \lambda.g)$ et la distributivité de $.$ par rapport à l'addition $+$ dans $\mathcal{F}(X, E)$.
2. Compatibilité de $.$ avec la multiplication dans $\mathbb{K}$:
$$((\lambda \times \mu).f)(x) = (\lambda \times \mu).f(x) = \lambda.(\mu.f(x)) \ ,$$
car la multiplication dans $\mathbb{K}$ est compatible avec la loi externe $.$ de E. On en déduit :
$$((\lambda \times \mu).f)(x) = \lambda.((\mu.f)(x)) = (\lambda.(\mu.f))(x) \ .$$
Comme x est quelconque cela donne l'identité entre applications $(\lambda \times \mu).f = \lambda.(\mu.f)$ et donc la compatibilité de la loi $\times$ avec la loi $.$ sur $\mathcal{F}(X, E)$. Dans les deux derniers points de cette démonstration on ne répètera pas l'argument qui fait passer des identité $u(x) = v(x)$ à l'identité $u = v$ qui a déjà été utilisé plusieurs fois.
3. Distributivité de $.$ par rapport à l'addition dans $\mathbb{K}$:
$$((\lambda + \mu).f)(x) = (\lambda + \mu).f(x) = \lambda.f(x) + \mu.f(x) \ ,$$
car l'opération externe $.$ est distributive par rapport à l'addition dans E. On en déduit : $=$
$$((\lambda + \mu).f)(x) = ((\lambda.f)(x) + (\mu.f)(x)) = (\lambda.f + \mu.f)(x) \ .$$
Et la loi externe $.$ de $\mathcal{F}(X, E)$ est bien distributive par rapport à l'addition dans $\mathbb{K}$.
4. Compatibilité entre le neutre multiplicatif $1_{\mathbb{K}} \in \mathbb{K}$ et la loi externe $.$:
$$(1_{\mathbb{K}}.f)(x) = 1_{\mathbb{K}} \times f(x) = f(x) \ ,$$
car 1_K est le neutre multiplicatif dans $\mathbb{K}$. Cela donne $1_{\mathbb{K}}.f = f$ et conclut cette démonstration. $\square$

Comme on a beaucoup de liberté pour les choix de X et de E, pour lequel on peut prendre dors et déjà $\mathbb{K}$ ou $\mathbb{K}^n$, cet exemple est extrêmement souple et général. Il est aussi très utile pour démontrer que d'autres espaces sont bien des espaces vectoriels, dès qu'on dispose de la notion de sous-espace (voir le chapitre 2 de ce cours). Les sept autres exemples d'espace vectoriel sont des variantes ou bien des sous-espaces de ces deux exemples fondamentaux. On ne vérifie pas en détail dans ce cours qu'ils forment bien des espaces vectoriels.

1.3 Applications linéaires

Définition 1.7 *Soient E et F deux $\mathbb{K}$-espaces vectoriels. Une application $f\colon E \longrightarrow F$ est dite linéaire si elle est compatible avec les structures de $\mathbb{K}$-espaces sur E et sur F. Autrement dit si f vérifie :*

1. $\forall \lambda \in \mathbb{K},\ \forall u \in E, \quad f(\lambda.u) = \lambda.f(u)$.
2. $\forall u, v \in E, \quad f(u+v) = f(u) + f(v)$.

Proposition 1.8 *Soit $f\colon E \longrightarrow F$ une application entre deux $\mathbb{K}$-espaces E et F. Alors f est linéaire si et seulement si pour tout $\lambda \in \mathbb{K}$ et tout u, v de E on a $f(\lambda.u + v) = \lambda.f(u) + f(v)$.*

Preuve : Soit $f\colon E \longrightarrow F$ une application entre deux $\mathbb{K}$-espaces vectoriels. On suppose que f vérifie la définition 1.7, et on se fixe un scalaire $\lambda \in \mathbb{K}$ et deux vecteurs u et v dans E. Alors on a $f(\lambda.u + v) = f(\lambda.u) + f(v)$ par le 1 de la définition 1.7. Et avec le 2 de cette même définition 1.7 on trouve $f(\lambda.u+v) = \lambda.f(u)+f(v)$. Cela démontre le sens direct de l'équivalence de la proposition 1.8. Réciproquement on suppose que pour tout $\lambda \in \mathbb{K}$ et tout u, v de E on ait $f(\lambda.u + v) = \lambda.f(u) + f(v)$. Alors pour tout u, v dans E en prenant $\lambda = 1_{\mathbb{K}}$ dans cette identité on arrive à

$$f(u+v) = f(1_{\mathbb{K}}.u + v) = 1_{\mathbb{K}}.f(u) + f(v) = f(u) + f(v)\ .$$

Cela montre déjà que f est additive. Pour la deuxième propriété on a besoin de voir que $f(0_E) = 0_F$. C'est une propriété qui intervient dans d'autres contexte et qui est suffisamment utile pour être isolée dans un lemme.

Lemme 1.9 *Soient E et F deux groupes additifs et soit $f\colon E \longrightarrow F$ une application additive entre E et F. Alors l'image $f(0_E)$ du neutre 0_E de E est égale au neutre 0_F de F :*

$$f(0_E) = 0_F$$

Preuve du lemme 1.9 : En fait on n'utilise même pas la commutativité des lois de groupe. Comme 0_E est un neutre pour la loi de groupe de E, on a $0_E+0_E = 0_E$. En appliquant f qui est compatible avec les lois de groupes on en déduit $f(0_E)+f(0_E) = f(0_E + 0_E) = f(0_E)$. Puis si on compose dans F avec l'opposé $-f(0_E)$ on arrive à $f(0_E) = f(0_E) + f(0_E) - f(0_E) = f(0_E) - f(0_E) = 0_F$. $\square$

On reprend la preuve de la proposition 1.8. Il reste à démontrer le point 1 de la définition 1.7, en supposant la propriété $f(\lambda.u + v) = \lambda.f(u) + f(v)$. On fixe un scalaire λ et un vecteur u et on prend $v = 0_E$ dans cette propriété. En utilisant le lemme, on obtient :

$$f(\lambda.u) = f(\lambda.u + 0_E) = \lambda.f(u) + f(0_E) = \lambda.f(u) + 0_F = \lambda.f(u)\ .$$

Cela démontre la proposition 1.8 $\square$

Exemples :

1. Si E est un $\mathbb{K}$-espace vectoriel et si $\lambda \in \mathbb{K}$ l'homothétie vectorielle de rapport λ définie par $x \mapsto \lambda.x$ est linéaire.
2. Si X est un ensemble non vide, si $a \in X$ est un élément fixé de X et si E est un $\mathbb{K}$-espace vectoriel alors l'application de spécialisation en a définie de $\mathcal{F}(X, E)$ dans E par $f \mapsto f(a)$ est linéaire.
3. Si $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ désigne le $\mathbb{R}$-espace vectoriel des applications continues de $\mathbb{R}$ dans $\mathbb{R}$, l'application de $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ dans $\mathbb{R}$ définie par $\phi(f) = \int_0^1 f(t)dt$ est linéaire.
4. Si $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ désigne le $\mathbb{R}$-espace vectoriel des applications continues de $\mathbb{R}$ dans $\mathbb{R}$, l'application de $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ dans $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ définie par
$$\phi(f)(x) = \int_0^x f(t)dt$$
est linéaire.
5. Si $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ désigne le $\mathbb{R}$-espace vectoriel des applications continues dérivables et à dérivées continues de $\mathbb{R}$ dans $\mathbb{R}$, l'application de $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ dans $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$ définie par $\phi(f) = f'$ est linéaire.
6. Si $A \in M_{n,m}(\mathbb{K})$, A définit une application linéaire ϕ_A de $\mathbb{K}^m$ dans $\mathbb{K}^n$ par $\phi_A(x_1, \ldots, x_m) = (y_1, \ldots, y_n)$ où
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$
On peut encore écrire : $\phi(x_1, \ldots, x_m) = (x_1, \ldots, x_m){}^tA$ où ${}^tA$ désigne la matrice transposée de A. Cette matrice est définie en inversant les lignes et les colonnes de A, c'est-à-dire que si le coefficient de la i-ième ligne et j-ième colonne de A est le scalaire $a_{i,j}$ alors $a_{i,j}$ est aussi le coefficient de la j-ième ligne et i-ième colonne de ${}^tA$.

Démonstration : La définition des espaces vectoriels et celle des lois sur $\mathcal{F}(X, E)$ sont faites pour que les applications des exemples 1 et 2 soient linéaires. Les application des exemples 3 et 4 sont linéaire par la linéarité de l'intégrale qui ne sera pas démontrée dans ce cours ; pas plus que la linéarité de la dérivation utilisée dans l'exemple 5 qui, elle aussi, est du ressort de l'UE fonctions et suites. On démontre en détail que l'application ϕ_A du dernier exemple est linéaire. On se donne une matrice $A = (a_{i,j}) \in M_{n,m}(\mathbb{K})$, où $a_{i,j}$ est le coefficient sur la i-ième ligne et j-ième colonne de A. Par définition du produit matriciel l'équation
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix},$$
équivaut à
$$\forall i,\ 1 \leq i \leq n,\ y_i = \sum_{j=1}^{j=m} a_{i,j} x_j$$

On se donne deux vecteurs $x = (x_j)_{j=1}^{j=m}$ et $x' = (x'_j)_{j=1}^{j=m}$ dans $\mathbb{K}^m$ et un scalaire $\lambda \in \mathbb{K}$. Alors on a $\lambda.x + x' = (\lambda x_j + x'_j)_{j=1}^{j=m}$ et donc

$$\begin{aligned}
\phi_A(\lambda.x + x') &= \begin{pmatrix} \sum_{j=1}^m a_{1,j}(\lambda x_j + x'_j) \\ \vdots \\ \sum_{j=1}^n a_{i,j}(\lambda x_j + x'_j) \\ \vdots \\ \sum_{j=1}^n a_{n,j}(\lambda x_j + x'_j) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1,j}\lambda x_j + \sum_{j=1}^m a_{1,j}x'_j \\ \vdots \\ \sum_{j=1}^m a_{i,j}\lambda x_j + \sum_{j=1}^m a_{i,j}x'_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}\lambda x_j + \sum_{j=1}^m a_{n,j}x'_j \end{pmatrix} \\
&= \begin{pmatrix} \sum_{j=1}^m a_{1,j}\lambda x_j \\ \vdots \\ \sum_{j=1}^m a_{i,j}\lambda x_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}\lambda x_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^m a_{1,j}x'_j \\ \vdots \\ \sum_{j=1}^m a_{i,j}x'_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}x'_j \end{pmatrix} \\
&= \lambda. \begin{pmatrix} \sum_{j=1}^m a_{1,j}x_j \\ \vdots \\ \sum_{j=1}^m a_{i,j}x_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}x_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^m a_{1,j}x'_j \\ \vdots \\ \sum_{j=1}^m a_{i,j}x'_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}x'_j \end{pmatrix} = \lambda.\phi_A(x) + \phi_A(x')
\end{aligned}$$

Cela donne la linéarité de l'application ϕ_A. On dit que ϕ_A est l'application linéaire associée à la matrice A. On verra qu'en dimension finie toute les applications linéaires sont de ce type.

Chapitre 2

Sous-espaces vectoriels

On conserve la notation $\mathbb{K}$ pour le corps des scalaires, qui est fixé. On allègera, de temps en temps, les énoncés et les notations en omettant la référence à $\mathbb{K}$.

2.1 Généralités

On commence par définir la notion de sous-espace.

Définition 2.1 *Soit E un $\mathbb{K}$-espace vectoriel. On dit que F est un* <u>*sous-espace vectoriel*</u> *de E lorsque :*

1. *F est un sous-ensemble de E : $F \subset E$.*
2. *F est non vide : $F \neq \emptyset$.*
3. *F est $\mathbb{K}$-linéairement stable : $\forall x, y \in F, \forall \lambda, \mu \in \mathbb{K}$, $\lambda.x + \mu.y \in F$*

Les sous-espaces sont aussi caractérisées comme suit :

Proposition 2.2 *Soit F un sous-ensemble d'un $\mathbb{K}$-espace vectoriel E. Alors F est un sous-espace de E si et seulement si il vérifie les deux conditions suivantes :*

1. *$F \neq \emptyset$*
2. *$\forall x, y \in F, \forall \lambda \in \mathbb{K}$, $\lambda.x + y \in F$.*

Preuve : On suppose que F est un sous-espace de E au sens de la définition 2.1. Alors $F \subset E$ et $F \neq \emptyset$. En outre si $x, y \in F$ et si $\lambda \in \mathbb{K}$ alors en prenant $\mu = 1_{\mathbb{K}}$ dans la formule 3 on obtient $\lambda.x + y = \lambda.x + 1.y \in F$. Cela démontre le sens direct de l'équivalence de la proposition 2.2. On suppose maintenant que F est un sous-ensemble non vide de E qui vérifie $\forall x, y \in F, \forall \lambda \in \mathbb{K}$, $\lambda.x+y \in F$. On vérifie d'abord que $0_E \in F$. Soit $x \in F$ dont l'existence est assuré parce que $F \neq \emptyset$. Alors on a $0_E = (-1).x+x \in F$. Ensuite si $\lambda, \mu \in \mathbb{K}$ et si $x, y \in F$, on a $\mu.y = \mu.y+0_E \in F$, puis $\lambda.x + (\mu.y) \in F$ en utilisant deux fois la propriété $\forall x, y \in F, \forall \lambda \in \mathbb{K}$, $\lambda.x + y \in F$. $\square$

Proposition 2.3 *Soit F un sous-espace d'un $\mathbb{K}$-espace vectoriel E. Alors F est un $\mathbb{K}$-espace vectoriel.*

Preuve : A priori la loi de composition interne $+$ de E est une application de $E \times E$ à valeur dans E. Si on prend un sous-ensemble non vide $F \subset E$ et si on restreint cette application à $F \times F$ alors on obtient une application $F \times F \longrightarrow E$. Mais en prenant $\lambda = \mu = 1$ dans la formule 3 de la définition 2.1 on voit que si F est un sous-espace vectoriel de E alors pour tout $x, y \in F$ on a $x + y \in F$. De sorte que cette restriction définit une loi de composition interne sur F. De la même manière (en prenant $y = 0_E \in F$ dans la formule 3 de la définition 2.1) on voit que si F est un sous-espace vectoriel de E, si $x \in F$ et si $\lambda \in \mathbb{K}$ alors $\lambda.x \in F$. Ainsi, si F est un sous-espace de E, alors l'opération externe de $\mathbb{K}$ sur E définit par restriction à $\mathbb{K} \times F$ une opération externe de $\mathbb{K}$ sur F et non une application $\mathbb{K} \times F \longrightarrow E$ comme c'est le cas pour des sous-ensemble F qui ne sont pas des sous-espaces. Ceci posé si F est un sous-espace vectoriel de E, alors il est muni par restriction des lois sur E d'une loi de composition externe et d'une opération externe de $\mathbb{K}$. Il faut ensuite démontrer que ces lois sur F vérifient les huit propriétés des définitions 1.2 et 1.5. Ces propriétés portent sur des scalaires λ, μ dans $\mathbb{K}$ et sur des vecteurs x, y, z. Hormis l'existence de l'élément neutre et l'existence d'un inverse dans F, les propriétés vont être vraies en particulier pour x, y, z dans F parce qu'elles sont déjà vraies pour x, y, z dans E. On démontre donc l'existence d'un élément neutre et d'un opposé dans F. D'abord si F est un sous-espace de E, alors F est non vide et contient un élément, disons $x \in F$. Ensuite, en prenant $\lambda = \mu = 0$ et $x = y$ dans la formule 3 de la définition 2.1 et en utilisant la proposition 1.6, on voit que $0_E = 0.x + 0.y \in F$. De cette façon $0_E = 0_F$ est bien un élément neutre pour $+$ dans F. Ensuite si $x \in F$, alors en utilisant la proposition 1.6 on obtient que son opposé $-x$ dans E vérifie $-x = (-1).x = (-1).x + 0_F \in F$. Cela démontre que $-x \in F$ est bien un opposé dans F à x. $\square$

Remarque : En conséquence de cette proposition, pour démontrer qu'un ensemble F est un espace vectoriel, il est souvent plus efficace de trouver un espace E parmi les exemples habituels (le plus souvent E est soit $\mathbb{K}^n$ soit $\mathcal{F}(X, V)$ pour un choix judicieux d'espace V et d'ensemble X), puis de constater que F est un sous-espace de E.

Exemples :

1. L'ensemble $\mathcal{C}^1(\mathbb{R}, \mathbb{R})$ est un sous-espace vectoriel de $\mathcal{C}^0(\mathbb{R}, \mathbb{R})$
2. L'ensemble des applications impaires de $\mathbb{R}$ dans $\mathbb{R}$ est un sous espace vectoriel du $\mathbb{R}$-espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$.
3. L'ensemble des applications constantes de $\mathbb{R}$ dans $\mathbb{R}$ est un sous espace vectoriel du $\mathbb{R}$-espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$.
4. L'ensemble des polynômes de degré inférieur ou égal à n est un sous-espace vectoriel du K-espace vectoriel des polynômes à coefficients dans K.
5. Une matrice $[a_{i,j}]$, carré d'ordre n et à coefficient dans $\mathbb{K}$, est dite <u>triangulaire supérieure</u> lorsque tous les coefficients en dessous de sa diagonale, $\overline{\text{c'est-à-dire}}$ les $a_{i,j}$ pour $i < j$, sont nuls. L'ensemble des matrices carrées d'ordre n, triangulaires supérieures est un sous-espace du K espace vectoriel des matrices carrées d'ordre n à coefficients dans K.

2.2 Noyau et image d'une application linéaire

Définition 2.4 *Soit $f\colon E \longrightarrow F$ une application linéaire. On appelle noyau de f, et on note* $\operatorname{Ker} f$*, le sous-ensemble de E suivant*

$$\operatorname{Ker} f = \{x \in E,\ f(x) = 0\}.$$

Autrement dit $\operatorname{Ker} f$ *est l'image réciproque de l'élément neutre 0_F de F. On appelle image de f, et on note* $\operatorname{Im}(f)$*, l'image ensembliste de f, qui est un sous-ensemble de F. Autrement dit*

$$\operatorname{Im}(f) = \{y \in F,\ \exists x \in E |\ f(x) = y\}.$$

Proposition 2.5 *Soit $f\colon E \longrightarrow F$ une application linéaire. Alors on a*

1. $\operatorname{Ker} f$ *est un sous-espace de E.*
2. $\operatorname{Im}(f)$ *est un sous-espace de F.*
3. *f est injective si et seulement si* $\operatorname{Ker} f = \{0\}$.
4. *f est surjective si et seulement si* $\operatorname{Im}(f) = F$.

Preuve :

1. Par le lemme 1.9, on a $0_E \in \operatorname{Ker}(f)$ et donc $\operatorname{Ker}(f) \neq \emptyset$. On se donne $\lambda \in \mathbb{K}$ et $x, y \in \operatorname{Ker}(f)$. Alors on a $f(x) = f(y) = 0_F$. Puis, par linéarité de f on en déduit $f(\lambda.x + y) = \lambda.f(x) + f(y) = \lambda.0_F + 0_F = 0_F$. Donc $(\lambda.x + y)$ est bien un élément de $\operatorname{Ker}(f)$. Cela démontre que $\operatorname{Ker}(f)$ est un sous-espace de E.
2. Par le lemme 1.9 on a $f(0_E) = 0_F$ et donc $0_F \in \operatorname{Im}(f)$. On se donne $\lambda \in \mathbb{K}$ et $y, y' \in \operatorname{Im}(f)$. Alors par définition de l'image d'une application il existe x et x' dans E tels que $f(x) = y$ et $f(x') = y'$. Il suit $f(\lambda.x + x') = \lambda.f(x) + f(x') = \lambda.y + y'$; et donc $(\lambda.y + y') \in \operatorname{Im}(f)$. Cela démontre que $\operatorname{Im}(f)$ est un sous-espace de F.
3. On suppose que f est injective. Avec le lemme 1.9 on sait que $0_E \in \operatorname{Ker}(f)$. Si $x \in \operatorname{Ker}(f)$ alors x vérifie $f(x) = 0_F = f(0_E)$. Puisque f est injective cela donne $x = 0_E$, et donc $\operatorname{Ker}(f)$ est réduit à $\{0_E\}$. Réciproquement, on suppose que $\operatorname{Ker}(f) = \{0_E\}$ et on se donne x, x' dans E tels que $f(x) = f(x')$. Alors $f(x - x') = f(x) - f(x') = 0_F$, et donc $x - x' \in \operatorname{Ker}(f) = \{0_E\}$. Cela force $x = x'$, et démontre l'injectivité de f.
4. est immédiat.

$\square$

Proposition 2.6 *Soit $f\colon E \longrightarrow F$ une application linéaire bijective. Alors la bijection réciproque $f^{(-1)}\colon F \longrightarrow E$ est aussi linéaire et on dit que f et $f^{(-1)}$ sont des isomorphismes réciproque.*

Preuve : Par définition des bijections réciproques on a l'équivalence $f^{(-1)}(y) = x \iff f(x) = y$. On se donne $\lambda \in \mathbb{K}$ et $y, y' \in F$. Puisque f est une bijection il existe un unique $x \in E$ et un unique $x' \in E$ tels que $f(x) = y$ et $f(x') = y'$. On a, par linéarité de f, les identités $f(\lambda.x + x') = \lambda.f(x) + f(x') = \lambda.y + y'$. Par définition des bijections réciproque cela donne aussi $\lambda.f^{(-1)}(y) + f^{(-1)}(y') = \lambda.x + x' = f^{(-1)}(\lambda.y + y')$. Ceci démontre la linéarité de $f^{(-1)}$. □

Définition 2.7 *Lorsqu'il existe un isomorphisme $f : E \longrightarrow F$ entre deux $\mathbb{K}$-espaces vectoriels E et F on dit qu'ils sont isomorphes.*

Remarque : Lorsque $f : E \longrightarrow F$ est un isomorphisme, alors, du point de vue de l'algèbre linéaire, on peut parfaitement identifier E et F. Toute question d'algèbre linéaire peut être traitée indifféremment dans E ou dans F et on pourra transposer le résultat d'un espace à l'autre en appliquant f ou $f^{(-1)}$.

2.3 Application

1. L'ensemble des solutions d'un système de p équations linéaires homogènes (second membre nul) à n inconnues est un sous-espace vectoriel de $\mathbb{K}^n$. En effet si A est la matrice du système, le système s'écrit sous la forme matricielle :

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

   L'application de $\mathbb{K}^n$ dans $\mathbb{K}^p$ $\phi(x_1, \ldots, x_n) = (y_1, \ldots, y_p)$ définie par

$$A. \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$$

   est linéaire et son noyau est l'ensemble cherché.
2. Si $\mathcal{C}^\infty(\mathbb{R})$ désigne l'ensemble des fonctions indéfiniment dérivables de $\mathbb{R}$ dans $\mathbb{R}$, $\mathcal{C}^\infty(\mathbb{R})$ est un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. On peut alors considérer l'application ϕ de $\mathcal{C}^\infty(\mathbb{R})$ dans lui-même définie par :

$$\phi(f) = f'' - f' + f$$

   On obtient alors une application linéaire dont le noyau est l'espace vectoriel des solution de l'équation différentielle linéaire du second ordre :

$$y'' - y' + y = 0.$$

3. Si E désigne l'ensemble des suites réelles, et ϕ l'application de E dans E définie pour $u = (u_n)_{n \in \mathbb{N}}$, $\phi(u) = (\phi(u)_n)_{n \in \mathbb{N}}$ avec

$$\forall n \in \mathbb{N}, \phi(u)_n = u_{n+2} - 3u_{n+1} + 2u_n.$$

E est un espace vectoriel, ϕ est linéaire, et le noyau de ϕ est formé des suites vérifiant la relation de récurrence :

$$\forall n \in \mathbb{N}, u_{n+2} - 3u_{n+1} + 2u_n = 0.$$

On obtient aussi un espace vectoriel sur le corps des réels.

2.4 Sous-espace vectoriel engendré par une partie de E

On commence par une proposition qui est utile pour démontrer la consistance de la définition abstraite des sous-espaces engendrés.

Proposition 2.8 *Soit E un $\mathbb{K}$ espace-vectoriel et soit $(F_i)_{i\in I}$ des sous-espaces de E indexés par un ensemble I. Alors l'intersection $\bigcap_{i\in I} F_i$ est un sous-espace de E.*

Preuve : Pour tout $i \in I$, comme les F_i sont des sous-espaces on a $0_E \in F_i$. Par définition de l'intersection cela donne $0_E \in \bigcap_{i\in I} F_i$ et cette intersection est non vide. On se donne $\lambda \in \mathbb{K}$ et $x, y \in \bigcap_{i\in I} F_i$. Alors pour tout i on a $x, y \in F_i$ et puisque F_i est un sous-espace. On obtient $\lambda.x + y \in F_i$, et avec la définition de l'intersection on en déduit $\lambda.x + y \in \bigcap_{i\in I} F_i$; qui est donc un sous-espace de E. $\square$

Définition 2.9 *Soit S une partie quelconque de E. Alors on appelle <u>sous-espace engendré</u> par S et on note $\langle S\rangle$ le plus petit sous-espace de E (au sens de l'inclusion des sous-espaces) qui contienne S.*

La proposition 2.8 assure l'existence de $\langle S\rangle$. En effet si on prend pour $(F_i)_{i\in I}$ la famille de tous les sous-espaces qui contienne S, alors $\bigcap_{i\in I} F_i$ est un sous-espace, contient S et est contenu dans tout sous-espace contenant S.

Exemples : On obtient de cette façon les sous-espaces suivant :

1. $\langle\emptyset\rangle = \{0_E\}$
2. $\langle F\rangle = F \iff F$ est un sous-espace de E.
3. Étant donné $u \in E$, avec $u \neq 0_E$ alors on a $\langle u\rangle = \mathbb{K}.u = \{\lambda.u, \lambda \in \mathbb{K}\}$.

La définition 2.9 est très utile pour certaines démonstrations théoriques, et c'est elle qui se prête le mieux à des généralisations dans des contextes plus subtils que la simple algèbre linéaire. Cependant elle ne permet pas d'appréhender concrètement le sous-espace vectoriel engendré, par exemple elle ne donne pas d'idée précise de la forme d'un élément générique de $\langle S\rangle$. Pour obtenir cette approche plus terre à terre, il faut utiliser la notion de combinaison linéaire.

Définition 2.10

1. *Soit $S = \{s_1, \cdots, s_n\}$ une partie finie d'un espace vectoriel E. On appelle <u>combinaison linéaire</u> sur S tout élément x de la forme $x = \sum_{i=1}^{i=n} \lambda_i s_i$.*
2. *Soit S une partie infinie d'un espace vectoriel E. On appelle <u>combinaison linéaire</u> sur S toute combinaison linéaire sur une partie finie $\mathcal{F} \subset S$.*

Exemples :

1. Si S est réduit à un élément u, une combinaison linéaire d'éléments de S n'est autre qu'un élément de la forme $\lambda \cdot u$.
2. Dans l'espace vectoriel des polynômes $\mathbb{R}[T]$, les combinaisons linéaires d'éléments de la partie $S = \{T^n, n \in \mathbb{N}\}$ sont exactement les polynômes.
3. Dans $\mathbb{R}^2$ on a toujours $(x, y) = x \cdot (1, 0) + y \cdot (0, 1)$ donc tout élément de $\mathbb{R}^2$ est combinaison linéaire d'éléments de $S = \{(1, 0), (0, 1)\}$.
4. Dans $M_2(K)$ on a
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
Tout élément de $M_2(K)$ est donc combinaison linéaire d'éléments de
$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$
5. Par convention une combinaison linéaire d'éléments du vide est égale au vecteur nul 0_E.

Cette notion de combinaison linéaire sur S permet de concrétiser l'espace vectoriel engendré par S, grâce à la proposition :

Proposition 2.11 *Soit S une partie d'une espace E. Alors l'ensemble F_S de toutes les combinaisons linéaires sur S est un sous-espace de E et vérifie $F_S = \langle S \rangle$.*

Preuve : Par convention F_S contient 0_E même si $S = \emptyset$, et donc $F_S \neq \emptyset$. On se donne $\lambda \in \mathbb{K}$ et c, c' deux éléments de F_S. Alors il existe deux partie finie T, T' de S et pour chaque $t \in T$ et chaque $t' \in T'$ des scalaires λ_t et $\mu_{t'}$ tels que

$$c = \sum_{t \in T} \lambda_t . t \qquad \text{et} \qquad c' = \sum_{t' \in T'} \mu_{t'} . t'$$

On peut écrire $T \bigcup T'$ comme la réunion disjointe $T \bigcup T' = T_1 \amalg T_2 \amalg (T \bigcap T')$ où $T_1 = T \setminus T'$ et $T_2 = T' \setminus T$; c'est-à-dire par définition $t \in T_1$ si et seulement si $t \in T$ et $t \notin T'$ et $t' \in T_2$ si et seulement si $t' \in T'$ et $t' \notin T$. Ceci posé on obtient

$$\begin{aligned}
\lambda . c + c' &= \lambda \sum_{t \in T} \lambda_t . t + \sum_{t' \in T'} \mu_{t'} . t' \\
&= \sum_{t \in T_1} (\lambda \lambda_t) . t + \sum_{t \in T \bigcap T'} (\lambda \lambda_t) . t + \sum_{t' \in T \bigcap T'} \mu_{t'} . t' + \sum_{t' \in T_2} \mu_{t'} t' \\
&= \sum_{t \in T_1} (\lambda \lambda_t) . t + \sum_{t \in T \bigcap T'} ((\lambda \lambda_t) + \mu_t) . t + \sum_{t' \in T_2} \mu_{t'} t' \ .
\end{aligned}$$

Et comme $T \bigcap T'$ est aussi une partie finie de S on voit que $\lambda . c + c'$ est encore une combinaison linéaire sur S et donc un élément de F_S. Cela démontre que F_S est un sous-espace vectoriel de E. Comme en plus $S \subset F_S$ on voit, par définition de $\langle S \rangle$ que

$\langle S \rangle \subset F_S$. Réciproquement on se donne une combinaison linéaire finie sur S, disons $c = \sum_{t \in T} \lambda_t.t$ indexée par une partie finie T de S. Alors puisque $S \subset \langle S \rangle$ tous les $t \in T$ sont aussi des éléments de $\langle S \rangle$. Et puisque $\langle S \rangle$ est un sous-espace de E on voit que $c \in \langle S \rangle$. Cela démontre l'inclusion $F_S \subset \langle S \rangle$ et conclut cette preuve. $\square$

Les sous-espaces vectoriels engendrés se comportent relativement bien vis-à-vis des images directes d'applications linéaires. En effet on a la proposition :

Proposition 2.12 *Soit $f \colon E \longrightarrow F$ une application linéaire et S une partie de E. Alors on a l'égalité*

$$f(\langle S \rangle) = \langle f(S) \rangle.$$

Preuve : On commence par énoncer et démontrer un lemme.

Lemme 2.13 *Soit $f \colon E \longrightarrow F$ une application linéaire, et soit E' un sous-espace de E. Alors $f(E')$ est un sous-espace de F.*

Preuve du lemme Par le lemme 1.9 on a $f(0_E) = 0_F$ et puisque $0_E \in E'$ on en déduit $0_F \in f(E')$. On se donne $\lambda \in \mathbb{K}$ et $y, y' \in f(E')$. Alors par définition de l'image d'une application il existe x et x' dans E' tels que $f(x) = y$ et $f(x') = y'$. Puisque E' est un sous-espace on a $\lambda.x + x' \in E'$ et il suit par linéarité de f les identités $f(\lambda.x + x') = \lambda.f(x) + f(x') = \lambda.y + y'$; et donc $(\lambda.y + y') \in f(E')$. Cela démontre que $f(E')$ est un sous-espace de F. $\square$

On reprend la preuve de la proposition. Puisque $\langle S \rangle$ est un sous-espace de F, le lemme démontre que $f(\langle S \rangle)$ est un sous-espace de F. Comme ce sous-espace contient $f(S)$ on en déduit l'inclusion $\langle f(S) \rangle \subset f(\langle S \rangle)$, car $\langle f(S) \rangle$ est le plus petit sous-espace de F contenant $f(S)$. Pour établir l'inclusion réciproque on utilise la caractérisation par combinaison linéaire de $\langle S \rangle$. On doit démontrer qu'un élément quelconque $y \in f(\langle S \rangle)$ est aussi un élément de $\langle f(S) \rangle$. Mais puisque $y \in f(\langle S \rangle)$, il existe $x \in \langle S \rangle$ avec $y = f(x)$. L'élément x est une combinaison linéaire finie sur S, disons $x = \sum_{t \in T} \lambda_{x,t} t$, pour une partie finie $T \subset S$. Il suit

$$y = f(x) = f\left(\sum_{t \in T} \lambda_{x,t} t\right) = \sum_{t \in T} \lambda_{x,t} f(t) \in \langle f(S) \rangle.$$

$\square$

Définition 2.14 *Soit S une partie d'un espace vectoriel E. On dit que S est une <u>partie génératrice</u> de E, ou aussi que S engendre E lorsque $\langle S \rangle = E$.*

Exemples :

1. La partie vide est génératrice de l'espace vectoriel nul $\{0\}$.
2. Si E est un espace vectoriel, E engendre E.
3. La partie $\{(1,0), (0,1)\}$ forme une partie génératrice de $\mathbb{R}^2$.
4. La partie $S = \{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$ est une partie génératrice de $M_2(\mathbb{K})$.

Chapitre 3

Bases d'un espace vectoriel

Pour définir la notion de base d'un espace vectoriel on est amené à introduire la distinction entre partie et famille d'un espace vectoriel. Lorsqu'elle est finie, une famille est une liste d'éléments de l'espace, ordonnée, avec éventuellement des répétitions. Autrement dit si x et y sont deux éléments distincts d'un ensemble alors la famille (x, y, y) correspond à la partie $\{x, y\}$ mais est distincte de la famille (x, x, y) et aussi de la famille (y, x, y).

3.1 Familles

Définition 3.1 *Soient E un ensemble et soit I un autre ensemble (appelé dans ce contexte ensemble d'indice). On appelle famille de E indexée par I la donnée d'une application $i \mapsto e_i$ définie sur I et à valeur dans E. L'élément e_i est appelés le i-ième élément de la famille. L'usage est de noter $(e_i)_{i\in I}$ avec des parenthèse pour distinguer entre une famille et la partie $\{e_i, i \in I\}$ qui lui correspond. Si $J \subset E$, la restriction de l'application $i \mapsto e_i$ définit une famille de E indexée par I et on dit alors que la famille $(e_i)_{i\in J}$ est une sous-famille de $(e_i)_{i\in I}$.*

Avec cette définition les parties de $\mathbb{R}^2$ $\{(1, 0); (0, 1)\}$ et $\{(0, 1); (1, 0)\}$ coïncident tandis que $((1, 0); (0, 1)) \neq ((0, 1); (1, 0))$.

Définition 3.2 *Soit E un espace vectoriel et soit $(e_i)_{i\in I}$ une famille de E. On appelle combinaison linéaire sur $(e_i)_{i\in I}$ tout élément x de E de la forme $x = \sum_{i\in I} \lambda_i.e_i$, où $(\lambda_i)_{i\in I}$ est une famille de $\mathbb{K}$ telle que les λ_i soit tous nuls sauf un nombre fini d'entre eux.*

On voit tout de suite sur les définitions que l'ensemble des combinaisons linéaires sur une famille $(e_i)_{i\in I}$ coïncide avec l'ensemble des combinaisons linéaires sur la partie $\{e_i, i \in I\}$ qui lui correspond. En effet, en ce qui concerne les répétitions, par exemple si $e_i = e_j$, alors $\lambda_i.e_i + \lambda_j.e_j = (\lambda_i + \lambda_j).e_i$.

Définition 3.3 *Soit E un espace vectoriel et soit $(e_i)_{i\in I}$ une famille de E. On dit que la famille $(e_i)_{i\in I}$ est génératrice de E lorsque tout élément x de E est une combinaison linéaire sur les $(e_i)_{i\in I}$.*

Proposition 3.4 *Soit E un espace vectoriel et soit $(e_i)_{i\in I}$ une famille de E. Alors la famille $(e_i)_{i\in I}$ est génératrice du sous-espace $\langle\{e_i, i \in I\}\rangle$ engendré par la partie $\{e_i, i \in I\}$ qui lui correspond. En particulier ce sous-espace est le plus petit sous-espace de E qui contient tous les e_i.*

Preuve : Par la proposition 2.11 le sous-espace $\langle\{e_i, i \in I\}\rangle$ est l'espace dont les éléments sont les combinaisons linéaires sur la partie $\{e_i, i \in I\}$. Or de telles combinaisons linéaires sont des combinaisons sur la famille $(e_i)_{i\in I}$. $\square$

Proposition 3.5 *Soit $S = (s_1, \cdots, s_n)$ une famille finie indexée par $\{1, \cdots, n\}$ du $\mathbb{K}$-espace vectoriel E. Alors l'application $\phi\colon \mathbb{K}^n \longrightarrow E$ définie par*

$$\phi(\lambda_1, \lambda_2, \cdots, \lambda_n) = \sum_{i=1}^{n} \lambda_i s_i$$

est linéaire et en outre l'image de ϕ est le sous-espace engendré par S : $\mathrm{Im}(\phi) = \langle S\rangle$.

Preuve : On se donne $\lambda \in \mathbb{K}$ et deux vecteurs $(\lambda_1, \cdots, \lambda_n)$ et $(\mu_1, \cdots, \mu_n)$ de $\mathbb{K}^n$. Alors on a $\lambda.(\lambda_1, \cdots, \lambda_n) + (\mu_1, \cdots, \mu_n) = (\lambda\lambda_1 + \mu_1, \cdots, \lambda\lambda_i + \mu_i, \cdots, \lambda\lambda_n + \mu_n)$. Il suit :

$$\begin{aligned}\phi(\lambda.(\lambda_1, \cdots, \lambda_n) + (\mu_1, \cdots, \mu_n)) &= \sum_{i=1}^{n}(\lambda\lambda_i + \mu_i)s_i = \lambda\sum_{i=1}^{n}\lambda_i s_i + \sum_{i=1}^{n}\mu_i s_i \\ &= \lambda.\phi((\lambda_1, \cdots, \lambda_n)) + \phi((\mu_1, \cdots, \mu_n)) \ .\end{aligned}$$

Cela démontre la linéarité de ϕ. L'image de ϕ est un sous-espace de E qui contient S, cette image contient donc $\langle S\rangle$. Réciproquement tout élément de l'image de ϕ est une combinaison linéaire sur S, d'où l'égalité $\mathrm{Im}(\phi) = \langle S\rangle$. $\square$

On voit ainsi que si S est une famille génératrice d'un espace vectoriel E, tout élément de E s'écrit comme combinaison linéaire d'éléments de S. On s'intéresse maintenant à l'unicité d'une telle écriture. Soit donc S une partie génératrice de E, on suppose qu'il existe un vecteur non nul v qui s'écrit de deux façon différentes comme combinaison linéaire d'éléments de S :

$$\exists F_1 = (\lambda_u)_{u\in S}, \exists F_2 = (\beta_u)_{u\in S}, F_1 \neq F_2,\ v = \sum_{u\in S}\lambda_u \cdot u = \sum_{u\in S}\beta_u \cdot u\ ,$$

où les β_u et les λ_u sont nuls sauf pour un nombre fini d'entre eux. Alors

$$0_E = \sum_{u\in S}(\lambda_u - \beta_u)\cdot u.$$

Les coefficients $\lambda_u - \beta_u$ n'étant pas tous nuls. Cela permet d'introduire la définition suivante :

Définition 3.6 *Soit E un K-espace vectoriel. Une famille $(u_i)_{i\in I}$ d'éléments de E est dite liée sur K si le vecteur nul 0_E est combinaison linéaire à coefficients non tous nuls d'éléments de $(u_i)_{i\in I}$. On dit alors qu'une telle relation est une relation de dépendance linéaire non triviale entre les éléments de$(u_i)_{i\in I}$, ou encore que $(u_i)_{i\in I}$ est $\mathbb{K}$-linéairement dépendante.*

Exemples :

1. La famille $(1,1),(1,0),(0,1)$ est liée par la relation :

$$(1,1)-(1,0)-(0,1)=(0,0).$$

2. La famille $(u_i)_{i\in\{1,2\}}$ vérifiant $u_1=u_2$ est donc liée par la relation

$$u_1-u_2=0_E.$$

Définition 3.7 *Soit E un K-espace vectoriel. Une famille $(u_i)_{i\in I}$ d'éléments de E est dite linéairement indépendante (sur K) si elle n'est pas K-linéairement dépendante, ou encore si pour toute famille $(\lambda_i)_{i\in I}$ d'éléments de $\mathbb{K}$ avec un nombre fini de λ_i non nuls on a l'équivalence :*

$$\sum_{i\in I}\lambda_i\cdot u_i=0_E \iff \forall i\in I, \lambda_i=0.$$

On dit encore que la famille $(u_i)_{i\in I}$ est libre (sur K) .

Exemples :

1. La famille correspondante au singleton $\{u\}\subset E$ est linéairement indépendante si et seulement si $u\neq 0_E$.
2. La famille $(e_i)_{i\in\{1,2\}}$ d'éléments de $\mathbb{R}^2$, définie par

$$e_1=(1,0),\ e_2=(0,1),$$

   est linéairement indépendante sur $\mathbb{R}$.
3. La famille $(E_i)_{i\in\{1,2,3,4\}}$ d'éléments de $M_2(\mathbb{R})$, définie par :

$$E_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ E_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},\ E_3=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ E_4=\begin{pmatrix}0&0\\0&1\end{pmatrix},$$

   est linéairement indépendante sur $\mathbb{R}$.

Proposition 3.8 *Soit $(u_i)_{i\in I}$ une famille linéairement indépendante de vecteurs de E. Alors l'application $i\mapsto u_i$ est injective.*

Preuve : Supposons qu'il existe $i\neq j$ dans I tels que $u_i=u_j$, alors $u_i-u_j=0_E$ est une relation de dépendance linéaire. $\square$

Proposition 3.9 *Une famille $(x_i)_{i\in I}$ est libre si et seulement si pour tout partie finie F non vide de I, la sous-famille $(x_i)_{i\in F}$ est libre.*

Preuve : En effet toute combinaison linéaire d'éléments de la famille $(x_i)_{i\in F}$ est une combinaison linéaire d'éléments de la famille $(x_i)_{i\in I}$, donc si la famille indexée par I est libre toutes les sous-familles et en particulier les sous-familles finies sont libres.

Réciproquement comme si on considère la combinaison linéaire :

$$\sum_{i\in I} \lambda_i \cdot u_i = 0_E$$

l'ensemble $F = \{i \in I, \lambda_i \neq 0\}$ est fini, et donc $(x_i)_{i\in F}$ étant libre on a pour tout i dans F,$\lambda_i = 0$, puis pour tout i dans I, $\lambda_i = 0$, et la famille $(x_i)_{i\in I}$ est libre. $\square$

Proposition 3.10 *Soit E un K-espace vectoriel et $(u_i)_{i\in I}$ une famille libre d'éléments de E, et F le sous-espace $< \{u_i, i \in I\} >$ engendré par cette famille. Tout élément u de F s'écrit alors de façon unique comme combinaison linéaire des éléments de la famille $(u_i)_{i\in I}$.*

Preuve : En effet par définition de F tout élément de F est une combinaison linéaire des éléments de la famille $(u_i)_{i\in I}$. Si un élément u de F s'écrit de deux façons comme combinaison linéaire des $(u_i)_{i\in I}$, $u = \sum_{i\in I} \lambda_i \cdot u_i$ et $u = \sum_{i\in I} \beta_i \cdot u_i$, alors $0_E = \sum_{i\in I}(\lambda_i - \beta_i) \cdot u_i$ et l'indépendance des u_i donne l'égalité $\lambda_i = \beta_i$ pour tout i dans I. $\square$

3.2 Bases d'un espace vectoriel de type fini

On commence par définir ce qu'est un espace vectoriel de type fini.

Définition 3.11 *On dit qu'un espace vectoriel E est de type fini sur son corps des scalaires $\mathbb{K}$ lorsqu'il existe une partie génératrice finie de E.*

Par exemple $\mathbb{R}^2 = \langle (1, 0); (0, 1)\rangle$ est de type fini sur $\mathbb{R}$ et

$$M_2(\mathbb{C}) = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle,$$

est de type fini sur $\mathbb{C}$

Définition 3.12 *Soit E un espace vectoriel. On dit qu'une famille $(e_i)_{i\in I}$ est une base de E lorsque $(e_i)_{i\in I}$ est à la fois libre et génératrice.*

Exemples :

1. La famille $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ est une base de $\mathbb{R}^3$, c'est la base naturelle (canonique) de $\mathbb{R}^3$.
2. La famille $(T^n)_{n\in\mathbb{N}}$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{R}[T]$.
3. La famille $(E_{ij})_{\substack{1\le i\le 2\\ 1\le j\le 2}}$, définie par $E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ et $E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, est une base du $\mathbb{R}$-espace vectoriel $M_2(\mathbb{R})$.

Définition 3.13 *Soit E un espace vectoriel sur $\mathbb{K}$.*

1. *On dit qu'une sous-famille $(e_i)_{i\in J}$ d'une famille $(e_i)_{i\in I}$ est une sous-famille stricte lorsque $I \neq J$.*
2. *Une famille libre $(e_i)_{i\in J}$ est dite maximale lorsque toute famille $(e_i)_{i\in I}$ dont $(e_i)_{i\in J}$ soit une sous-famille stricte est liée.*
3. *Une famille génératrice $(e_i)_{i\in I}$ est dite minimale lorsque toute sous-famille stricte $(e_i)_{i\in J}$ de $(e_i)_{i\in I}$ n'est pas génératrice.*

Proposition 3.14 *Soit E un espace vectoriel sur $\mathbb{K}$ et soit $(e_i)_{i\in I}$ une famille de E. On a les équivalences*

$$\begin{aligned}(e_i)_{i\in I} \text{ est une base de } E &\iff (e_i)_{i\in I} \text{ est une famille libre maximale de } E\\ &\iff (e_i)_{i\in I} \text{ est une famille génératrice minimale de } E\ .\end{aligned}$$

Preuve : On va démontrer la chaîne d'implications

$$\begin{aligned}(e_i)_{i\in I} \text{ est une base de } E &\overset{(1)}{\Rightarrow} (e_i)_{i\in I} \text{ est une famille libre maximale de } E\\ &\overset{(2)}{\Rightarrow} (e_i)_{i\in I} \text{ est une famille génératrice minimale de } E\\ &\overset{(3)}{\Rightarrow} (e_i)_{i\in I} \text{ est une base de } E\ .\end{aligned}$$

(1) On suppose que $(e_i)_{i\in I}$ est une base de E. Alors (e_i) est libre et il suffit de montrer que (e_i) est maximale pour cette propriété. Prenons J un ensemble contenant strictement I et une sur-famille $(e_j)_{j\in J}$. Il faut montrer que $(e_j)_{j\in J}$ est liée. On fixe un $j_0 \in J \setminus I$. Comme $(e_i)_{i\in I}$ est une base il existe une combinaison linéaire sur les $(e_i)_{i\in I}$ telle que $e_{j_0} = \sum_{i\in I} \lambda_i e_i$. Mais cette identité conduit à la relation non triviale

$$0_E = e_{j_0} + \sum_{i\in I} -\lambda_i e_i\ .$$

L'implication (1) est démontrée.

(2) On suppose que $(e_i)_{i\in I}$ est une famille libre maximale. Montrons d'abord que $(e_i)_{i\in I}$ est génératrice. Soit $x \in E$, et soit j tel que $j \notin I$. On pose $J = I \bigcup \{j\}$ et $e_j = x$. Alors la famille $(e_i)_{i\in J}$ est une sur-famille stricte de $(e_i)_{i\in I}$ et donc par maximalité de $(e_i)_{i\in I}$ cette sur-famille est liée par une relation de dépendance linéaire :

$$0_E = \lambda_j x + \sum_{i\in I} \lambda_i e_i\ .$$

Dans cette relation on a forcément $\lambda_j \neq 0$, sinon la famille $(e_i)_{i\in I}$ serait liée. On obtient donc :

$$x = \sum_{i\in I} -\frac{\lambda_i}{\lambda_j} e_i \in \langle e_i,\, i \in I\rangle\ .$$

Cela démontre que la famille $(e_i)_{i\in I}$ est génératrice. Vérifions que cette famille est minimale. Soit F une partie stricte de I et soit $k \in I \setminus F$. On doit démontrer que $(e_i)_{i\in F}$ n'engendre pas E. Mais comme la famille $(e_i)_{i\in F\bigcup\{k\}}$ est libre en tant que sous-famille de la famille libre $(e_i)_{i\in I}$ on a forcément $e_k \notin \langle e_i,\, i \in F\rangle$. L'implication (2) est démontrée.

(3) On suppose que $(e_i)_{i\in I}$ est une famille génératrice minimale de E. Il reste à démontrer que $(e_i)_{i\in I}$ est libre. Pour ce on démontre que toute relation de dépendance linéaire entre ces e_i est triviale. Prenons une telle relation, c'est-à-dire qu'on se fixe une partie finie $F \subset I$ et une famille de scalaires $(\lambda_i)_{i\in F}$ telle que

$$\sum_{i\in F} \lambda_i e_i = 0_E \ .$$

On va démontrer, par l'absurde, que tous les λ_i, $i \in F$ sont nuls. On suppose, en vue d'une contradiction, que pour un certain $j \in F$ on ait $\lambda_j \neq 0$. Mais alors on obtient e_j comme combinaison linéaire des autres e_i avec l'identité

$$e_j = \sum_{i\in F, i\neq j} -\frac{\lambda_i}{\lambda_j} e_i \ .$$

Donc la sous-famille $(e_i)_{i\in I, i\neq j}$ est encore génératrice de E puisque l'espace engendré par cette sous-famille contient le système générateur $(e_i)_{i\in I}$ tout entier. Et l'hypothèse (absurde) $\lambda_j \neq 0$ contredit la minimalité de $(e_i)_{i\in I}$. □

Théorème 3.15 *Soit E un espace vectoriel de type fini, alors E admet une base finie.*

Preuve : On part d'une famille génératrice finie $(e_i)_{i=1\cdots n}$. Si cette famille est libre alors c'est une base et le théorème est démontré dans ce cas. Si cette famille n'est pas libre, il existe une relation de dépendance linéaire $0_E = \sum_{i=1}^n \lambda_i e_i$ avec au moins un j tel que $\lambda_j \neq 0$. Dans ce second cas la sous-famille $(e_i)_{i\neq j}$ est encore génératrice puisque l'espace qu'elle engendre contient e_j et donc le système générateur initial. On a ainsi extrait du système générateur initial une famille génératrice indexée par $\{1, \cdots, n-1\}$. Si cette famille extraite est libre, alors elle forme une base et le théorème est démontré. Si cette famille n'est pas libre on peut en extraire une famille génératrice indexée par $\{1, \cdots, n-2\}$. À l'issue d'au plus n telles étapes on aura construit une base de E. Si $E = \{0\}$ il ne contient aucune famille libre, et on convient que $\emptyset$ est une base de E, si $E \neq \{0\}$ il contient des vecteurs non nul et une famille génératrice doit comporter au moins un vecteur non nul. □

Le lemme qui suit est la clé de la théorie de la dimension, et sa démonstration est probablement la plus délicate de ce cours. L'énoncé est incontournable, et on verra dans toute la suite de ce chapitre les applications qui en découlent.

Lemme 3.16 *Soit E un espace vectoriel de type fini, soit $g_1, \cdots, g_n$ une famille génératrice finie de E, et soit $f_1, \cdots, f_s$ une famille libre de E. Alors $s \leq n$.*

Preuve : On va démonter que toute famille $(f_1, \cdots, f_{n+1})$ indexée par l'intervalle entier $\{1, \cdots, n+1\}$ est liée. Cela donne en particulier l'implication $s > n \Rightarrow (f_1, \cdots, f_s)$ est liée. La contraposée de cette implication est ce qu'on veut démontrer :

$$(f_1, \cdots, f_s) \text{ est libre} \Rightarrow s \leq n \ .$$

On va procéder par récurrence sur n. On commence par initialiser la récurrence avec $n = 1$. Dans ce cas l'espace E est engendré par $g = g_1$. Si on se donne f_1 et f_2 dans E, alors il existe λ_1 et λ_2 dans $\mathbb{K}$ avec $f_i = \lambda_i g$ pour $i = 1$ et $i = 2$. Si $\lambda_1 = 0$ alors $(f_1, f_2) = (0_E, f_2)$ est liée par la relation $1.f_1 = 0_E$. Si $\lambda_1 \neq 0$, alors (f_1, f_2) est lié par la relation non triviale :

$$-\lambda_2.f_1 + \lambda_1.f_2 = -\lambda_2.(\lambda_1.g) + \lambda_1.(\lambda_2.g) = 0.g = 0_E \ .$$

Cela démontre bien que dans un espace engendré par un seul vecteur toute famille d'au moins deux vecteurs est liée. On passe à l'hérédité de l'hypothèse de récurrence. On suppose que pour un certain rang $k \geq 2$, dans tout espace engendré par $k - 1$ éléments, les familles de k vecteurs sont liées. On se place ensuite dans un espace E engendré par k vecteurs $(g_1, \cdots g_k)$ et on prend une famille $(f_1, \cdots, f_k, f_{k+1})$ de E. Puisque les g_i sont générateurs on peut écrire le vecteur f_j, pour tout j compris entre 1 et $k + 1$, comme une combinaison linéaire

$$f_j = \sum_{i=1}^{k} \lambda_{i,j} g_i \ .$$

Pour démontrer que les f_j sont liés on peut, sans perte de généralité, supposer que $f_1 \neq 0$, car sinon on la relation $1.f_1 = 0_E$ et la preuve est terminée. Puisque $f_1 \neq 0_E$ l'un des $\lambda_{i,1}$ est non nul pour au moins un i. Quitte à changer l'ordre des g_i on peut donc supposer le pivot $\lambda_{1,1}$ non nul :$\lambda_{1,1} \neq 0$. Soit F le sous-espace engendré par les $k - 1$ vecteurs $F = \langle g_2, \cdots, g_k \rangle$. On forme les k vecteurs de F notés v_j pour $j = 2, \cdots, k + 1$ définis par :

$$v_j = \lambda_{1,1} f_j - \lambda_{1,j} f_1 = \sum_{i=1}^{n} (\lambda_{1,1}\lambda_{i,j} - \lambda_{1,j}\lambda_{i,1}) g_i = \sum_{i=2}^{n} (\lambda_{1,1}\lambda_{i,j} - \lambda_{1,j}\lambda_{i,1}) g_i \ .$$

Par l'hypothèse de récurrence, ces k vecteurs dans un espace engendré par $k - 1$ vecteurs sont liés et on dispose d'une relation de dépendance linéaire non triviale :

$$0_E = \sum_{j=2}^{k+1} \alpha_j v_j \ ,$$

avec au moins un j tel que $\alpha_j \neq 0$. On en déduit alors :

$$0_E = \sum_{j=2}^{k+1} \alpha_j v_j = \sum_{j=2}^{j=k+1} \alpha_j (\lambda_{1,1} f_j - \lambda_{1,j} f_1) = (-\sum_{j=2}^{j=k+1} \alpha_j \lambda_{1,j}) f_1 + \sum_{j=2}^{j=k+1} \alpha_j \lambda_{1,1} f_j \ ,$$

qui est une relation de dépendance linéaire non triviale entre les f_j. Le principe de récurrence conclut la démonstration du lemme. $\square$

Théorème 3.17 *Soit E un espace vectoriel de type fini. Alors toutes les bases de E sont finies et ont le même nombre d'éléments.*

Preuve : Puisque E est de type fini il admet un système générateur fini, disons $\langle g_1, \cdots, g_n\rangle = E$. Par le lemme 3.16, tout système libre dans E comporte au plus n éléments. En particulier les bases de E, qui existent par le théorème 3.15, ont au plus n éléments. On se donne donc deux bases de E $(e_1, \cdots, e_s)$ et $(f_1, \cdots, f_t)$. On utilise le lemme 3.16 dans les deux directions : puisque $(e_1, \cdots, e_s)$ est générateur et que $(f_1, \cdots, f_t)$ est libre on a $t \le s$. Puisque $(f_1, \cdots, f_t)$ est générateur et que $(e_1, \cdots, e_s)$ est libre on a $s \le t$. Cela donne $s = t$ et conclut la preuve du théorème : toutes les bases de E ont $s = t$ éléments. $\square$

Définition 3.18 *Soit E un espace vectoriel de type fini, et soit $(e_1, \cdots, e_n)$ une base finie de E à n éléments. En vertu du théorème précédent l'entier n est uniquement déterminé par E et ne dépend pas du choix de la base (e_i). Cet entier n s'appelle la dimension de E et se note $n = \dim_{\mathbb{K}}(E)$ ou bien $\dim(E)$ si le contexte permet d'omettre la référence à $\mathbb{K}$.*

Exemples :

1. $\dim(\mathbb{R}^n) = n$.
2. $\dim(M_{m,n}(\mathbb{K})) = mn$.

Théorème 3.19 *Soit E un espace vectoriel de dimension finie. On se donne deux ensemble d'indices $F \subset I$ et une famille $(u_i)_{i\in I}$, qui engendre E et telle que $(u_i)_{i\in F}$ soit libre. Alors F est finie et il existe une partie finie G telle que $F \subset G \subset I$ et que $(u_i)_{i\in G}$ soit une base de E.*

Preuve : Le lemme 3.16 démontre que l'ensemble F est fini : c'est l'ensemble d'indice d'une famille libre dans un espace de type fini. Soit n la dimension (finie) de E. En vertu du lemme 3.16, toute famille libre comprenant n éléments est maximale : en effet les sur-familles strictes comporteront au moins $n+1$ éléments et seront donc liées. Avec le lemme 3.14, toute famille libre à n éléments est une base. En partant de $(u_i)_{i\in F}$ on construit récursivement une famille libre maximale qui sera donc une base. On pose $G_0 = F$. On construit $G_0, G_1, \cdots G_\ell$ tels que

1. $G_k \subset G_{k+1} \subset I$
2. $\#G_{k+1} = \#G_k + 1$
3. Pour tout k la famille $(u_i)_{i\in G_k}$ est libre.
4. On s'arrête au rang ℓ tel que $\#G_\ell = n$.

Si cette construction est possible alors $G = G_\ell$ convient et la preuve du théorème est terminée. Supposons G_k construit (pour $k = 0$ on prend $G_0 = F$). Si $\#G_k = n$ alors on s'arrête. Si $\#G_k < n$ alors la famille libre $(u_i)_{i\in G_k}$ n'est pas génératrice. Si $\langle u_i, i \in G_k\rangle$ contenait tout $\{u_i, i \in I\}$ alors il contiendrait l'espace engendré par $\{u_i,\ i \in I\}$ c'est-à-dire E. De sorte qu'il existe au moins un $j \in J$ avec $u_j \notin \langle u_i, i \in G_k\rangle$. On pose $G_{k+1} = \{j\} \cup G_k$. Les conditions 1 et 2 sont vérifiée, il suffit maintenant de démontrer que la famille $(u_i)_{i\in G_{k+1}}$ est libre. On part d'une relation linéaire

$$0_E = \lambda_j u_j + \sum_{i\in G_k} \lambda_i u_i \ .$$

Alors si $\lambda_j \neq 0$ on obtient u_j comme combinaison linéaire des $(u_i)_{i \in G_k}$, et cela contredit $u_j \notin \langle u_i, i \in G_k \rangle$. Donc $\lambda_j = 0$. De sorte que la relation ci-dessus est une relation entre les $(u_i)_{i \in G_k}$ c'est-à-dire est forcément triviale. $\square$

Corollaire 3.20 (Base incomplète) *Toute famille libre d'un espace vectoriel de dimension finie peut être complétée en une base.*

Preuve : Il suffit de prendre E tout entier comme sur-famille génératrice et qui contient la famille libre comme sous-famille. $\square$

3.3 Digression sur la dimension infinie

Si on ne suppose plus que l'espace vectoriel E est de type fini, on dispose encore du théorème suivant (admis) :

Théorème 3.21 *Soit E un $\mathbb{K}$-espace vectoriel.*

1. *E admet une base.*
2. *Soit I un ensemble, F une partie de I et $(u_i)_{i \in I}$ une famille génératrice de E telle que $(u_i)_{i \in F}$ soit libre, alors il existe une partie G de I contenant F telle que la famille $(u_i)_{i \in G}$ est une base.*
3. *Si $(u_i)_{i \in I}$ et $(v_j)_{j \in J}$ sont deux bases de E, il existe une bijection entre I et J.*

Ce théorème est une généralisation des résultats du paragraphe précédent. Il est admis et sa démonstration utilise l'axiome du choix.

3.4 Lien avec les applications linéaires

Proposition 3.22 *Soit E un espace vectoriel et soit $(e_1, \cdots, e_n)$ une famille de E. Il y a équivalence entre les trois affirmations suivantes :*

1. *$(e_1, \cdots, e_n)$ est une base de E.*
2. *Pour tout $x \in E$ il existe une unique famille de scalaires $(\lambda_1, \cdots, \lambda_n)$ telle que $x = \sum_{i=1}^{i=n} \lambda_i e_i$.*
3. *L'application $\mathbb{K}^n \longrightarrow E$ définie par $(x_1, \cdots, x_n) \mapsto \sum_{i=1}^{i=n} x_i e_i$ est un isomorphisme d'espaces vectoriels.*

Preuve : Par la proposition 3.10 l'affirmation 1 entraîne l'affirmation 2. Réciproquement si $(e_1, \cdots, e_n)$ vérifie l'affirmation 2, alors c'est une famille génératrice de E et l'unicité de l'écriture $0_E = \sum_{i=1}^{n} 0.e_i$ montre qu'elle est libre : l'équivalence entre 1 et 2 est démontrée. On suppose 3, et on note ϕ l'application $\phi \colon \mathbb{K}^n \longrightarrow E$ telle que $\phi((x_1, \cdots, x_n)) = \sum_{i=1}^{i=n} x_i e_i$. Alors comme ϕ est surjective, tout x de E s'écrit pour un antécédent $(x_1, \cdots, x_n)$: $x = \phi((x_i)) = \sum_{i=1}^{n} x_i e_i$ et la famille (e_i) est génératrice. De plus comme ϕ est injective on a $\sum_{i=1}^{n} x_i e_i = 0 \iff \forall i\ x_i = 0$, et donc la famille (e_i) est une base. On a démontré que 3 entraîne 1. Réciproquement on suppose que la famille (e_i) est un base. Alors par la proposition 3.5 l'application ϕ

du 3 est bien définie, linéaire et surjective. Montrons que ϕ est injective en calculant $\mathrm{Ker}(\phi)$. Soit $(x_1, \cdots, x_n) \in \mathrm{Ker}(\phi)$. Alors on a $0_E = \phi(x_1, \cdots, x_n) = \sum_{i=1}^n x_i e_i$. Et comme la famille (e_i) est libre cela donne $\forall i \ i = 0$ et donc le vecteur $(x_1, \cdots, x_n)$ est nul. Donc ϕ est un isomorphisme est 1 est équivalent à 3. $\square$

Terminologie :

1. Lorsque $(e_1, \cdots, e_n)$ est une base de E les scalaires $(\lambda_1, \cdots, \lambda_n)$ tels que $x = \sum_{i=1}^{i=n} \lambda_i e_i$ s'appellent les coordonnées de x relativement à la base $(e_i)_{i \in I}$.
2. Dans $\mathbb{K}^n$ on appelle base canonique la famille $(e_i)_{i \in \{1, \cdots n\}}$ définie par $e_i = (\delta_{i,1}, \cdots, \delta_{i,n})$ où $\delta_{i,j}$ est le symbole de Kronecker :

$$\delta_{i,j} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases} .$$

   Alors le vecteur $x = (x_1, \cdots, x_n)$ vérifie $x = \sum_{i=1}^{i=n} x_i e_i$ et ses coordonnées relativement à la base canonique sont bien ses coordonnées au sens usuel.

Remarque : Fixer une base de E de dimension n revient ainsi à fixer un isomorphisme entre $\mathbb{K}^n$ et E.

Corollaire 3.23

1. *Soit E un $\mathbb{K}$-espace vectoriel de dimension n. Alors E est isomorphe à $\mathbb{K}^n$.*
2. *Soient F et G, deux $\mathbb{K}$-espaces vectoriels de dimension n, alors F est isomorphe à G.*

Preuve : 1. On fixe une base $(e_1, \cdots, e_n)$ de E. La proposition 3.22 décrit un isomorphisme $\phi_{(e_i)}$, associé à ce choix de base, partant de $\mathbb{K}^n$ et d'image E.

2. On fixe $(f_i)_{i=1}^{i=n}$ et $(g_i)_{i=1}^{i=n}$ des bases respectives de F et G. On utilise les isomorphismes du 1 $\phi_{(f_i)} \colon \mathbb{K}^n \longrightarrow F$ et $\phi_{(g_i)} \colon \mathbb{K}^n \longrightarrow G$. Alors comme la composée de deux isomorphisme est un isomorphisme on obtient un isomorphisme entre F et G en prenant $\phi_{(g_i)} \circ \phi_{(f_i)}^{(-1)}$. $\square$

Proposition 3.24 *Soit $f \colon E \longrightarrow F$ une application linéaire et $(e_i)_{i \in I}$ une famille de E. Alors on a :*

1. *Si $(e_i)_{i \in I}$ engendre E, alors $(f(e_i))_{i \in I}$ engendre $\mathrm{Im}(f)$.*
2. *Si $(e_i)_{i \in I}$ engendre E, et si f est surjective alors $(f(e_i))_{i \in I}$ engendre F.*
3. *Si $(e_i)_{i \in I}$ est libre et si f est injective, alors $(f(e_i))_{i \in I}$ est libre.*
4. *Si $(e_i)_{i \in I}$ est une base de E et si f est un isomorphisme, alors $(f(e_i))_{i \in I}$ est une base de F.*

Preuve :

1. C'est une reformulation de la proposition 2.12.
2. C'est un cas particulier de 1.
3. On suppose f injective et $(e_i)_{i\in I}$ libre. On se donne une relation de dépendance linéaire entre les $f(e_i)_{i\in I}$, disons $0_F = \sum_{i\in I} \lambda_i f(e_i)$. Par linéarité de f on en déduit $0_F = f(\sum_{i\in I} \lambda_i e_i) = f(0_E)$ et comme f est injective on obtient une relation de dépendance linéaire entre les e_i : $\sum_{i\in I} \lambda_i e_i = 0_E$. Puisque les e_i sont libre tous les λ_i sont nuls et la seule relation de dépendance linéaire entre les $f(e_i)$ est la relation triviale.
4. C'est une conséquence de 2 et 3.

$\square$

Théorème 3.25 *Soit* $\mathbb{K}$ *un corps et soient* E *et* F *deux espaces vectoriels de dimension finie. Alors* E *et* F *sont isomorphes si et seulement si ils ont même dimension.*

Preuve : Si il existe un isomorphisme $f\colon E \xrightarrow{\sim} F$, par le 4 de la proposition 3.24 l'image d'une base de E est une base de F. Donc il y a autant d'éléments dans chacune des bases de E que dans chacune des bases de F. Réciproquement si E et F ont même dimension ils sont isomorphes par le 2 du corollaire 3.23 $\square$

Théorème 3.26 (Théorème du rang) *Soit* $f\colon E \longrightarrow F$ *une application linéaire. Alors on a*

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

Preuve : Soit $n = \dim(E)$. Comme $\mathrm{Ker}(f)$ est un sous-espace de E, toute base de $\mathrm{Ker}(f)$ est une famille libre dans E et donc se complète en une base de E par le corollaire 3.20. On se fixe une base $(e_1, \cdots, e_k)$ de $\mathrm{Ker}(f)$ et on la complète avec $(e_{k+1}, \cdots, e_n)$ pour obtenir une base de E. L'affirmation à démontrer revient à $\dim(\mathrm{Im}(f)) = n - k$ et il s'agit de trouver une base de $\mathrm{Im}(f)$ comprenant $n-k$ éléments. On va voir que $(f(e_{k+1}), \cdots, f(e_n))$ est une base de $\mathrm{Im}(f)$ et cela suffira à démontrer le théorème du rang. Par la proposition 3.24 la famille $f(e_i)_{i=1}^{i=n}$ engendre $\mathrm{Im}(f)$. Mais comme pour $i = 1, \cdots, k$ on a $f(e_i) = 0_F$, la famille $f(e_i)_{i=k+1}^{i=n}$ est encore génératrice de $\mathrm{Im}(f)$. Montrons maintenant que cette famille est libre. On part d'une relation de dépendance linéaire $0_F = \sum_{i=k+1}^{n} \lambda_i f(e_i)$. Par linéarité de f on en déduit $0_F = f(\sum_{i=k+1}^{n} \lambda_i e_i)$ et donc on arrive à l'appartenance $\sum_{i=k+1}^{n} \lambda e_i \in \mathrm{Ker}(f)$, et on peut calculer dans E. Maintenant la famille $(e_1, \cdots, e_k)$ est une base de $\mathrm{Ker}(f)$ et donc on peut trouver une combinaison linéaire sur cette base telle que $\sum_{i=1}^{k} \lambda_i e_i = \sum_{i=k+1}^{n} \lambda_i e_i$. Mais comme la famille $(e_i)_{i=1}^{i=n}$ est libre dans E tous les λ_i sont nuls et la seule relation entre les $(f(e_i))_{i=k+1}^{i=n}$ est la relation triviale. $\square$

Remarque : Le rang d'une application linéaire $f\colon E \longrightarrow F$ est la dimension $\mathrm{rg}(f) = \dim(\mathrm{Im}(f))$. Il a une interprétation matricielle compatible avec la terminologie "rang des matrices" rencontrée en semestre 1. Ce théorème de rang donne la formule $\mathrm{rg}(f) = \dim(E) - \dim(\mathrm{Ker}(f))$. En particulier $\mathrm{rg}(f)$ qui est dans sa définition attaché au sous-espace $\mathrm{Im}(f)$ de F peut se calculer en restant entièrement dans E.

Chapitre 4

Matrices

4.0 Rappels du semestre 1

Les notions de matrices, de produit matriciel et quelques propriétés ont été traités pendant l'ue d'algèbre du semestre starter. Pour fixer les idées, dans cette section, on rappelle les définitions et les propriétés utiles *sans démonstration.*

Définition 4.1 *Soit $\mathbb{K}$ un corps de scalaires, et soit n et m deux entiers.*

1. *On appelle matrice à n lignes et m colonnes la données d'une famille $(a_{i,j})$ d'éléments de $\mathbb{K}$ indexée par le produit cartésien $(i, j) \in \{1, \cdots, n\} \times \{1, \cdots m\}$. On présente habituellement ces données sous la formes d'un tableau d'éléments de $\mathbb{K}$ comportant n lignes et m colonnes comme ci-dessous pour $n = 2, m = 3$ par exemple :*
$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}$$
2. *L'élément $a_{i,j}$ s'appelle le coefficient de la i-ième ligne et j-ième colonne de la matrice $(a_{i,j})_{i,j}$. Le premier indice, usuellement noté i, correspond aux lignes du tableau, le deuxième, usuellement noté j, correspond aux colonnes.*
3. *L'ensemble dont les éléments sont toutes les matrices à n ligne et m colonnes s'appelle l'algèbre des matrices $n \times m$ à coefficient dans $\mathbb{K}$ et se note $M_{n,m}(\mathbb{K})$. Lorsque $n = m$ on parle de matrice carrée à n lignes et n colonnes et on note $M_n(\mathbb{K})$.*

Définition 4.2 *Soit $\mathbb{K}$ un corps de scalaires et soit n, m et p des entiers. Le produit matriciel est défini comme une application de $M_{n,m}(\mathbb{K}) \times M_{m,p}(\mathbb{K})$ et à valeur dans $M_{n,p}(\mathbb{K})$ avec les formules :*

$$(a_{i,j})\,(b_{j,l}) = (c_{i,l}) \quad \text{où} \quad c_{i,l} = \sum_{j=1}^{m} a_{i,j} b_{j,l}$$

Beaucoup de propriétés du produit matriciel ont dû être vues en algèbre au semestre 1. Il est distributif par rapport à l'addition des matrices : Lorsque tous les produits ont un sens on a $A(B + C) = AB + AC$ et $(B + C)A = BA + AC$. on prendra garde néanmoins à ce que ce produit n'est pas commutatif : En effet les produits AB et BA ne sont pas toujours simultanément défini. Lorsqu'ils sont tous deux définis le résultat n'appartient pas toujours à la même algèbre de matrice (les nombres de lignes et colonnes peuvent changer). Et même dans le cas de matrice carrées A et B de même taille n le plus souvent on a $AB \neq BA$. Par exemple :

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Par contre le produit matriciel est associatif. C'est-à-dire que les produits AB et $(AB)C$ sont définis si et seulement si les produits BC et $A(BC)$ sont définis et dans ce cas le résultat final coïncide : $(AB)C = A(BC)$.

4.1 Interprétation matricielle du pivot de Gauß

L'algorithme du pivot de Gauß de réduction des matrices a été utilisé en starter pour résoudre des systèmes linéaires et inverser des matrices. Cet algorithme extrêmement important et efficace possède beaucoup de variantes et beaucoup d'autres applications. C'est l'approche ultime pour toute question calculatoire en algèbre linéaire. On va voir ici comment s'en servir pour extraire une base et des relations linéaires à partir d'un système de générateur. Voici comment procéder sur un exemple.

Exemple : Dans $\mathbb{R}^2$ posons $e_1 = (1,0)$, $e_2 = (0,1)$, $u_1 = (1,2)$, $u_2 = (3,7)$, $u_3 = (4,1)$ et $u_4 = (2,1)$. Il s'agit d'extraire de la famille $(u_i)_{i\in\{1,2,3,4\}}$ une base du sous-espace $F =< (u_i)_{i\in\{1,2,3,4\}} >$. Pour cela on utilise la notation matricielle qui permet d'écrire les données :

$$(u_1, u_2, u_3, u_4) = (e_1, e_2) \cdot \begin{pmatrix} 1 & 3 & 4 & 2 \\ 2 & 7 & 1 & 1 \end{pmatrix}.$$

La réduction de la matrice $M = \begin{pmatrix} 1 & 3 & 4 & 2 \\ 2 & 7 & 1 & 1 \end{pmatrix}$ par la méthode du Pivot donne

$$\begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 4 & 2 \\ 2 & 7 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 25 & 11 \\ 0 & 1 & -7 & -3 \end{pmatrix}$$

Comme on va le voir dans la suite de cette section de cours, cela permet d'affirmer que u_1 et u_2 forment une base de $F = \mathbb{R}^2$ et aussi que

$$(u_3, u_4) = (u_1, u_2) \cdot \begin{pmatrix} 25 & 11 \\ -7 & -3 \end{pmatrix}$$

Soit $u_3 = 25 \cdot u_1 - 7 \cdot u_2$ et $u_4 = 11 \cdot u_1 - 3 \cdot u_2$.

Pour formaliser l'interprétation matricielle du pivot de Gauß on a besoin de quelques définitions et propriétés supplémentaires.

Définition 4.3 *Soit $n \geq 2$ un entier naturel.*

1. *On appelle* <u>*permutation*</u> *de* $\{1, \cdots, n\}$ *toute application bijective de* $\{1, \cdots, n\}$ *dans lui-même.*
2. *On appelle* <u>*groupe symétrique*</u> *sur* $\{1, \cdots, n\}$ *et on note* $\mathfrak{S}_n$ *l'ensemble de toute ces permutations de* $\{1, \cdots, n\}$.
3. *Étant donnée* $\sigma \in \mathfrak{S}_n$ *on appelle* <u>*matrice de permutation*</u> *associée à* σ *et on note* M_σ *la matrice de* $M_n(\mathbb{K})$ *dont les coefficients sont :*

$$M_\sigma = (m_{i,j}) \text{ avec } m_{i,j} = \begin{cases} 1 \text{ si } i = \sigma(j) \\ 0 \text{ si } i \neq \sigma(j) \end{cases}$$

Proposition 4.4 *Soit* n *un entier.*

1. *Une matrice* M *de* $M_n(\mathbb{K})$ *est la matrice d'une permutation si et seulement si tous les coefficients de* M *sont nuls sauf un et un seul dans chacune des lignes et chacune des colonnes de* M.
2. *Soit* $\sigma \in \mathfrak{S}_n$ *et soit* $C \in M_{m,n}(\mathbb{K})$ *une matrice comprenant un nombre arbitraire* m *de lignes et exactement* n *colonnes. On numérote* C_1 *à* C_n *les colonnes de* C *de sorte que* $C = (C_1, \cdots, C_i, \cdots, C_n)$. *Alors le produit* CM_σ *est défini et l'on a :*

$$(C_1, \cdots, C_i, \cdots, C_n)M_\sigma = (C_{\sigma(1)}, \cdots, C_{\sigma(i)}, \cdots, C_{\sigma(n)})$$

Preuve : 1. On suppose que M est la matrice $M = M_\sigma$ associée à une permutation $\sigma \in \mathfrak{S}_n$. Alors comme σ est une application pour toute colonne j il existe un et un seul i tel que $i = \sigma(j)$ donc une et une seule ligne i telle que $m_{i,j} = 1$ et on a par définition $m_{i,k} = 0$ pour les autres lignes k telles que $k \neq \sigma(j)$. Étant donnée une ligne l comme σ est bijective il existe un et un seul o tel que $l = \sigma(o)$. Cela donne une et une seule colonne o telle que $m_{l,o} = 1$ et on a par définition $m_{l,k} = 0$ pour tous les autres k tels que $l \neq \sigma(k)$. Les matrices de permutations ont donc bien un et seul coefficient égal à 1 dans chacune de leurs lignes et chacune de leur colonnes, et tous leurs autres coefficients sont nuls.

Supposons que $M = (m_{i,j})$ ait tous ces coefficients égal à 0 sauf un et un seul dans chacune de ces lignes et chacune de ces colonnes. Alors on peut définir une application $\sigma\colon \{1, \cdots, n\} \longrightarrow \{1, \cdots, n\}$ en prenant pour $\sigma(j)$ l'unique entier tel que $m_{\sigma(j),j} = 1$. Puisqu'on a en outre supposé que la matrice avait un et un seul coefficient égal à 1 dans chacune de ces lignes, cette application est une bijection et la matrice M est bien égale à la matrice de permutation associée à σ.

La formule 2 est une conséquence de la formule de multiplication des matrices "la règle lignes/colonnes". La colonne j du produit $(C_1, \cdots, C_i, \cdots, C_n)M_\sigma$ s'obtient en effectuant la somme $\sum_{i=1}^n m_{i,j}C_i$. Mais pour $i \neq \sigma(j)$ on a $m_{i,j} = 0$ et la j-ième colonne de ce produit matriciel vaut $\sum_{i=1}^n m_{i,j}C_i = m_{\sigma(j),j}C_{\sigma(j)} = C_{\sigma(j)}$. $\square$

Définition 4.5 *Soit* n *un entier.*

1. *Pour des entiers* i *et* j, *on appelle* <u>*symbole de Kronecker*</u> *et on note* $\delta_{i,j}$ *le scalaire* $\delta_{i,j} \in \mathbb{K}$ *tel que* $\delta_{i,j} = 0$ *si* $i \neq j$ *et* $\delta_{i,j} = 1$ *si* $i = j$.

2. *On appelle matrice identité d'ordre n et on note I_n la matrice de $M_n(\mathbb{K})$ définie $I_n = (\delta_{i,j})$, de sorte que tous les coefficients de I_n sont nuls sauf ceux de la diagonale égaux à 1. On remarque que si* $\mathrm{id} \in \mathfrak{S}_n$ *désigne la permutation identité* $\mathrm{id}(i) = i$, *alors on a* $I_n = M_{\mathrm{id}}$.
3. *On dit qu'une matrice $P \in M_n(\mathbb{K})$ est inversible lorsque qu'il existe une matrice $Q \in M_n(\mathbb{K})$ telle que $QP = I_n$.*
4. *Soit i un entier $i \leq n$ et soit λ un scalaire. On appelle i-ième matrice de dilatation de rapport $\lambda \in \mathbb{K}$ et on note $D_i(\lambda)$ la matrice de $M_n(\mathbb{K})$ définie par $D_i(\lambda) = (d_{k,l})$ avec $d_{k,l} = 0$ si $k \neq l$, et pour $l = k$ $d_{l,l} = 1$ si $l \neq i$ et $d_{i,i} = \lambda$.*
5. *Soit $i \neq j$ deux entiers de $\{1, \cdots, n\}$ et soit μ un scalaire. On appelle i, j-ième matrice de transvection de rapport μ et on note $T_{i,j}(\mu)$ la matrice définie par $T_{i,j}(\mu) = (t_{k,l})$ avec*

$$t_{k,l} = \begin{cases} 1 \text{ si } k = l \\ \mu \text{ si } k = i \text{ et } l = j \\ 0 \text{ sinon} \end{cases}$$

Proposition 4.6 *Soit n un entier.*

1. *Pour tout entier m et toute matrice $A \in M_{m,n}(\mathbb{K})$ on a $AI_n = A$.*
2. *Pour tout entier m et toute matrice $A \in M_{n,m}(\mathbb{K})$ on a $I_nA = A$.*
3. *Une matrice P est inversible si et seulement si il existe une matrice Q' telle que $PQ' = I_n$.*
4. *Si une matrice P inversible est fixée alors la matrice Q telle que $QP = I_n$ est unique, la matrice Q' telle que $PQ' = I_n$ est unique et on a $Q' = Q$. La matrice Q s'appelle la matrice inverse de P et se note P^{-1}.*
5. *Pour $\lambda \neq 0$ La matrice $D_i(\lambda)$ est inversible et son inverse est $D_i(\lambda^{-1})$. Pour tout $\mu \in \mathbb{K}$ la matrice $T_{i,j}(\mu)$ est inversible et son inverse est $T_{i,j}(-\mu)$.*

Preuve : 1 et 2 sont évidents. Les affirmations 3. et 4. sont admises. Une preuve de 5. a été esquissée en cours, mais pour l'essentiel ces formules sont admises. Ce sont des cas particuliers des formules d'opérations élémentaires de la proposition qui suit.

Proposition 4.7 *Soit $L \in M_{n,m}(\mathbb{K})$ une matrice une matrice à n lignes et un nombre arbitraire m de colonnes. On numérote $L_1, \cdots, L_n$ les lignes de L de telle sorte que*

$$L = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix}$$

Alors les produits $T_{i,j}(\mu)L$ et $D_i(\lambda)L$ sont définis et correspondent à des opérations élémentaires sur les lignes de L. Explicitement on a les formules :

1. si $i > j$ alors

$$T_{i,j}(\mu)L = T_{i,j}(\mu)\begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \\ L_n \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_j + \mu L_i \\ \vdots \\ L_n \end{pmatrix}$$

2. si $i < j$ alors

$$T_{i,j}(\mu)L = T_{i,j}(\mu)\begin{pmatrix} L_1 \\ \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ L_j + \mu L_i \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix}$$

3.

$$D_i(\lambda)L = D_i(\lambda)\begin{pmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_n \end{pmatrix} = \begin{pmatrix} L_1 \\ \vdots \\ \lambda L_i \\ \vdots \\ L_n \end{pmatrix}$$

Preuve : Ces formules ont été admises en cours faute de temps. Ce serait un bon exercice à partir de la définition du produit matriciel ou encore mieux à partir des formules de multiplication des matrices $E_{i,j} = (\delta_{i,k}\delta_{i,l})_{k,l}$ de la base canonique de $M_n(\mathbb{K})$. Il n'y a pas vraiment de difficulté dans cette démonstration et je préfère m'attarder sur des points moins calculatoires de ce cours. $\square$

Remarque : Dans ce cas il est plus utile de retenir le résultat plutôt que sa preuve. Pour bien retrouver ses formules, il faut retenir absolument que si on multiplie une matrice A à gauche par un une matrice élémentaire E et qu'on forme EA on modifie A par une opérations élémentaire sur ces lignes. Pour retrouver la forme générale de cette matrice E on peut simplement appliquer l'opération élémentaire à modéliser aux lignes de la matrice I_n, puisqu'on retrouve alors $EI_n = E$. Lorsqu'on multiplie à droite la matrice A par une matrice de permutation M_σ on permute les colonnes de A, suivant σ.

Théorème 4.8 *Soient n, m deux entiers et soit $A \in M_{n,m}(\mathbb{K})$ une matrice à n lignes et m colonnes. Alors il existe un entier r, appelé le rang de A, avec $r \leq n$ et $r \leq m$, une matrice inversible $P \in M_n(\mathbb{K})$, une permutation $\sigma \in \mathfrak{S}_m$ telle que si M_σ*

désigne la matrice de permutation la matrice produit PAM_σ soit réduite de rang r au sens de Gauß, c'est-à-dire soit de la forme :

$$PAM_\sigma = \begin{pmatrix} I_r & A_1 \\ \mathrm{O}_{n-r,r} & \mathrm{O}_{n-r,m-r} \end{pmatrix}.$$

Les matrices $\mathrm{O}_{n-r,r}\mathrm{O}_{n-r,m-r}$ sont les matrices nulles de tailles $n-r \times r$ et respectivement $n-r \times m-r$. La matrice A_1 comporte, elle, r lignes et $m-r$ colonnes.

Commentaires :

1. La matrice P inversible est obtenue comme un produit de $T_{i,j}(\mu)$ et de $D_i(\lambda)$ qui correspondent à des opérations sur les lignes de la matrice A bien choisies. Cela donne l'inversibilité de P puisque si R et S sont inversibles avec $R'R = I_n$ et $S'S = I_n$ alors RS est inversible avec $S'R'RS = S'I_nS = S'S = I_n$.
2. La matrice M_σ correspond à une permutation des colonnes de A.

Preuve : Compte-tenu des commentaires qui précèdent il suffit de montrer qu'on peut réduire au sens de Gauß toute matrice A en effectuant sur cette matrice des permutations de colonnes et des opérations élémentaires sur les lignes. Cette réduction est toujours possible comme cela a été vu sur des exemples en semestre starter. Comme la démonstration donne aussi une méthode de calcul on la rappelle ici. Dans la suite de cette démonstration on dira qu'une matrice B est équivalente à une matrice A et on notera $A \sim B$ si on obtient B en transformant A par des permutations de colonnes ou par des opérations élémentaires sur les lignes. Cette relation est une relation d'équivalence puisque ces opérations élémentaire sont inversibles (on applique l'opération élémentaire inverse sur les lignes et la permutation réciproque sur les colonnes). On doit démontrer que toute matrice est équivalente à une matrice réduite au sens de Gauß. Dans la suite de cette démonstration on utilisera des permutations de colonnes non précisée, et des opérations élémentaires sur les lignes. On notera $L_i \leftarrow \lambda L_i$ l'opération (dilatation) qui multiplie la i-ième ligne par le scalaire non nul λ, et pour $i \neq j$ on notera $L_i \leftarrow L_i + \mu L_j$ l'opération (transvection) qui remplace la i-ième ligne L_i par la combinaison linéaire $L_i + \mu L_j$, définie pour $\mu \in \mathbb{K}$.

Premier pas : nettoyage de la première colonne Si A est la matrice nulle alors elle est déjà réduite de rang $r = 0$ et l'algorithme de réduction s'arrête. Supposons maintenant A non nulle. Si la première colonne de A est nulle, on échange cette colonne avec une colonne non nulle ce qui est une opération élémentaire admise (permutation de colonne) ; de sorte qu'on peut supposer que la première colonne de A est différente de $O_{n,1}$. Notons $A = (a_{i,j})$ pour les coefficients de A. On sait déjà qu'il y a une ligne i telle que $a_{i,1} \neq 0$, mais on veut prendre $a_{1,1}$ comme pivot et on doit transformer A pour obtenir $a_{1,1} = 1$. On distingue deux cas :

1. **Premier cas $a_{1,1} \neq 0$:** Alors on peut diviser la première ligne de A par $a_{1,1}$ (cela revient à multiplier A à gauche par $D_1(a_{1,1}^{-1})$). La matrice A est donc équivalente à une matrice avec $a_{1,1} = 1$.

2. **Deuxième cas $a_{1,1} = 0$:** On dispose d'une ligne avec $a_{i,1} \neq 0$. Si on remplace la première ligne L_1 de A par la somme $L_1 + L_i$ de cette première ligne avec la i-ième on retrouve en position $(1,1)$ l'ancien coefficient $a_{i,1}$ qui est non nul. Cela nous ramène au premier cas.

Dans tous les cas A est soit nulle soit équivalente à une matrice de la forme

$$A \sim \begin{pmatrix} 1 & * \cdots * \\ * & * \cdots * \\ \vdots & \ddots \\ * & * \cdots * \end{pmatrix}$$

Pour terminer le premier pas de la réduction on utilise le coefficient 1 en position de pivot pour nettoyer la première colonne. On note C_1 cette première colonne et $a_{i,1}$ les coefficients de cette première colonne. On a

$$A \sim \begin{pmatrix} 1 & * \\ a_{2,1} & \\ \vdots & \vdots \\ a_{n,1} & * \end{pmatrix} .$$

Dans cette matrice on effectue les opérations $L_i \leftarrow L_i - a_{i,1}L_1$ pour $i = 2 \cdots n$ ce qui annule le premier coefficient de chaque ligne. On obtient ainsi

$$A \sim \begin{pmatrix} 1 & * \cdots * \\ 0 & * \cdots * \\ \vdots & \ddots \\ 0 & * \cdots * \end{pmatrix}$$

Cela conclut le premier pas qui consiste à nettoyer la première colonne de A.

Passage de la k à la $k+1$-ième colonne : Par récurrence on peut supposer que A est équivalente à une matrice de la forme :

$$A \sim \widetilde{A} = \begin{pmatrix} I_k & B \\ O_{n-k,k} & C \end{pmatrix} ,$$

où la matrice I_k est la matrice identité carré d'ordre k, $O_{n-k,k}$ est la matrice nulle avec $n-k$ lignes et k colonnes et B et C sont des matrices respectivement dans $M_{k,m-k}(\mathbb{K})$ et $M_{n-k,m-k}(\mathbb{K})$. Si la matrice C est nulle alors la matrice $\widetilde{A}$ est déjà réduite au sens de Gauß de rang $r = k$ et l'algorithme s'arrête. On peut donc supposer $C \neq O_{n-k,m-k}$. Pour continuer à réduire, on doit dans un premier temps égaliser à 1 le coefficient $a_{k+1,k+1}$ pour ensuite s'en servir de pivot. Notons $\widetilde{a}_{i,j}$ les coefficients de $\widetilde{A}$. La matrice C est non nulle on peut donc supposer, quitte à permuter les colonnes que la $k+1$-ième colonne de C est non nulle, c'est-à-dire qu'il existe un $i \geq k+1$ avec $\widetilde{a}_{i,k+1} \neq 0$. On distingue deux cas

1. **Premier cas** $\widetilde{a}_{k+1,k+1} \neq 0$ **:** Alors il suffit de faire l'opération élémentaire $L_{k+1} \leftarrow \widetilde{a}_{k+1,k+1}^{-1} L_{k+1}$ ce qui ne change pas les k premières colonnes de $\widetilde{A}$ et on voit que
$$\widetilde{A} \sim \begin{pmatrix} I_k & * & B' \\ O_{1,k} & 1 & * \\ O_{n-k-1,k} & * & C' \end{pmatrix}$$
2. **Deuxième cas** $\widetilde{a}_{k+1,k+1} = 0$ **:** Alors on dispose d'un entier $i > k+1$ avec $\widetilde{a}_{i,k+1} \neq 0$ et en appliquant l'opération élémentaire $L_{k+1} \leftarrow L_{k+1} + \widetilde{a}_{i,k+1}^{-1} L_i$ on ne change pas les k premières colonnes de $\widetilde{A}$, et on voit aussi que
$$\widetilde{A} \sim \begin{pmatrix} I_k & * & B' \\ O_{1,k} & 1 & * \\ O_{n-k-1,k} & * & C' \end{pmatrix}$$

Dans tous les cas on s'est ramené à une matrice $\widetilde{A}$ de la forme
$$A \sim \widetilde{A} = \begin{pmatrix} I_k & * & B' \\ O_{1,k} & 1 & * \\ O_{n-k-1,k} & * & C' \end{pmatrix}$$
Sur cette matrice la $k+1$-ième ligne commence avec k coefficient nuls et donc les opérations de la forme $L_i \leftarrow L_i + \lambda L_{k+1}$ pour $i \neq k+1$ perturbent les matrices B' et C' mais ne changent pas les k premières colonnes de $\widetilde{A}$. On note $\widetilde{a}_{i,k+1}$ les coefficients de la $k+1$-ième colonne de $\widetilde{A}$, avec $\widetilde{a}_{k+1,k+1} = 1$. Pour $i = 1, \cdots, n$ et $i \neq k+1$ on effectue sur cette matrice $\widetilde{A}$ les opérations $L_i \leftarrow L_i - \widetilde{a}_{i,k+1} L_{k+1}$. Ces opérations laissent inchangées les k premières colonnes de $\widetilde{A}$, inchangée la $k+1$-ième ligne de $\widetilde{A}$ et annulent tous les coefficients de la $k+1$-ième colonne sauf le diagonal. On arrive ainsi à une matrice
$$A \sim \widetilde{A} \sim \widetilde{A}' = \begin{pmatrix} I_{k+1} & B \\ O_{n-k-1,k+1} & C \end{pmatrix} ,$$
ce qui termine le $k+1$-ième pas.

De cette façon on termine la réduction de la matrice A en au plus $\min(n, m)$ pas comme ceux décrits à l'instant. $\square$

4.2 Extraction de base

On se place dans l'espace vectoriel $\mathbb{K}^n$ et on se donne m vecteurs $u_1, \cdots, u_m$. On pose $F = \langle u_1, \cdots, u_m \rangle$. Le pivot de Gauß dit comment choisir r vecteurs parmi les u_i pour obtenir une base de F et donne aussi les coordonnées des autres u_i relativement à cette base. Précisons. On écrit en colonne les coordonnées des u_i dans la base canonique $e_1, \cdots, e_n$ de $\mathbb{K}^n$, et on note $A \in M_{n,m}(\mathbb{K})$ la matrice ainsi obtenue. En notation matricielle cela donne :
$$(u_1, \cdots, u_i, \cdots, u_m) = (e_1, \cdots, e_j, \cdots, e_n)A$$

Après réduction de Gauß la matrice A sera devenue A' de la forme

$$A' = \begin{pmatrix} I_r & A_1 \\ \mathrm{O}_{n-r,r} & \mathrm{O}_{n-r,m-r} \end{pmatrix}.$$

Dans la matrice A' les r premiers vecteurs sont égaux aux r premiers vecteurs de la base canonique et les colonnes de la matrice A_1 contiennent les coordonnées dans cette base des $m-r$ autres vecteurs. Toutes ces affirmations se traduisent par des relations linéaires entre les *colonnes* de la matrice A'. Or les relations linéaires entre les colonnes d'une matrice *ne changent pas lorsqu'on effectue des opérations élémentaires sur les lignes de cette matrices*!

Par exemple sur 3 colonnes si on a $xC_1 + yC_2 + zC_3 = 0$ et si on transforme la matrice (C_1, C_2, C_3) avec l'opération $L_j \leftarrow L_j + \mu L_i$ alors pour $k \neq j$ dans la k-ième ligne on a les coefficients (inchangés) $(c_{k,1}, c_{k,2}, c_{k,3})$ qui vérifient encore $xc_{k,1} + yc_{k,2} + zc_{k,3} = 0$, et dans la j-ième ligne on a les coefficients

$$(c_{j,1} + \mu c_{i,1}, c_{j,2} + \mu c_{i,2}, c_{j,3} + \mu c_{i,3})$$

qui vérifient aussi

$$\begin{aligned} &x(c_{j,1} + \mu c_{i,1}) + y(c_{j,2} + \mu c_{i,2}) + z(c_{j,3} + \mu c_{i,3}) \\ &= xc_{j,1} + yc_{j,2} + zc_{j,3} + x\mu c_{i,1} + y\mu c_{i,2} + z\mu c_{i,3} \\ &= 0 + \mu(xc_{i,1} + yc_{i,2} + zc_{i,3}) = 0 \ . \end{aligned}$$

En d'autres termes si on repense à la signification des matrices A et A' on voit qu'après la permutation σ de l'algorithme de Gauß les r premiers vecteurs

$$(u_{\sigma(1)}, \cdots, u_{\sigma(r)})$$

forment une base de F et que la matrice A_1 vérifie

$$(u_{\sigma(r+1)}, \cdots, u_{\sigma(m)}) = (u_{\sigma(1)}, \cdots, u_{\sigma(r)})A_1.$$

Autrement dit les colonnes de A_1 donnent les coordonnées des autres vecteurs u_i relativement à la base qu'on a extraite.

4.3 Matrice d'une application linéaire

Proposition 4.9 *Soit E et F deux espaces vectoriels de dimension finies. On se fixe une base $B_E = (e_1, \ldots, e_n)$ de E. Une application linéaire $f\colon E \longrightarrow F$ est uniquement déterminée par les images $f(e_i)$ des vecteurs de la base.*

Preuve : Soit $g\colon E \longrightarrow F$ une autre application linéaire telle que pour tout $i = 1, \cdots, n$ on ait $g(e_i) = f(e_i)$. Alors si $x \in E$ est un vecteur quelconque on peut l'écrire $x = \sum_{i=1}^n x_i e_i$ pour des scalaires x_i et on a alors par linéarité de f et de g :

$$g(x) = g\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i g(e_i) = \sum_{i=1}^n x_i f(e_i) = f\left(\sum_{i=1}^n x_i e_i\right) = f(x) \ .$$

Donc $f = g$. $\square$

Comme les vecteurs $f(e_i)$ sont eux-mêmes uniquement déterminés par leurs coordonnées dans une base de F on est conduit à introduire la définition suivante.

Définition 4.10 *Soit E et F deux espaces vectoriels de dimension finie dont on se fixe une base $B_E = (e_1, \ldots, e_n)$ de E et une base $B_F = (u_1, \ldots, u_m)$ de F. Soit $f\colon E \longrightarrow F$ une application linéaire, on appelle matrice de f relative aux bases B_E et B_F et on note*

$$M = \mathrm{Mat}_{B_F,B_E}(f)$$

l'unique matrice M de $M_{m,n}(K)$ définie par :

$$(f(e_1, \cdots, f(e_n)) = (u_1, \ldots, u_m)M$$

Cette matrice permet de calculer les images de f grâce à la proposition suivante.

Proposition 4.11 *Si f admet pour matrice M dans les bases B_E, B_F et si $u = \sum_{i=1}^{n} x_i \cdot e_i$ alors $f(u) = \sum_{j=1}^{m} y_j \cdot u_j$ avec*

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = M \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Le calcul est immédiat en utilisant la notation matricielle :

$$u = (e_1, \ldots, e_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Par linéarité de f on obtient :

$$f(u) = (f(e_1), \ldots, f(e_n)) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

soit

$$f(u) = (u_1, \ldots, u_m) \cdot M \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Puis par unicité des coordonnées dans la base de F :

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = M \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

$\square$

Corollaire 4.12 *Soit E et F deux espaces vectoriels de dimension finie. Une application linéaire f de E dans F est entièrement déterminée par sa matrice relative au choix d'une base de E et d'une base de F.*

Preuve : C'est une reformulation de la proposition 4.11 qui précède. □

Proposition 4.13 *Si B_E est une base de E, B_F une base de F et B_G une base de G, et si $f\colon E \longrightarrow F$ et $g\colon F \longrightarrow G$ sont deux applications linéaires les matrices associées à f, à g et à $g \circ f$ sont liées par la relation de Chasles :*

$$\mathrm{Mat}_{B_G,B_E}(g \circ f) = \mathrm{Mat}_{B_G,B_F}(g)\,\mathrm{Mat}_{B_F,B_E}(f)$$

Preuve : Notons $(e_i)_{1\leq i\leq n}$, $(u_i)_{1\leq i\leq m}$ et $(v_i)_{1\leq i\leq p}$ les bases B_E, B_F et B_G. On a

$$(f(e_1),\dots,f(e_n)) = (u_1,\dots,u_m)\,\mathrm{Mat}_{B_F,B_E}(f)$$

et

$$(g(u_1),\dots,g(u_m)) = (v_1,\dots,v_p)\,\mathrm{Mat}_{B_G,B_F}(g)$$

D'où

$$\begin{array}{rcc} (g\circ f(e_1),\dots,g\circ f(e_n)) & = & (g(u_1),\dots,g(u_m))\cdot \mathrm{Mat}_{B_F,B_E}(f) \\ & = & (v_1,\dots,v_p)\,\mathrm{Mat}_{B_G,B_F}(g)\,\mathrm{Mat}_{B_F,B_E}(f) \end{array}.$$

Ce qui donne le résultat. □

Proposition 4.14 *Soit $L(E,F)$ l'espace vectoriel dont les éléments sont les applications linéaires d'un espace vectoriel de E dans un autre espace vectoriel F. On a* $\dim(L(E,F)) = \dim(E)\dim(F)$.

Preuve : Si on se fixe une base de E et de F l'application qui à une application linéaire associe sa matrice est un isomorphisme entre $L(E,F)$ et $M_{\dim(F),\dim(E)}(\mathbb{K})$ et on a déjà vu la formule $\dim(M_{n,m}(\mathbb{K})) = nm$. □

Définition 4.15 *Si M est une matrice de $M_{n,m}(K)$ le rang de M est la dimension de l'espace vectoriel engendré par les vecteurs colonnes de M dans K^m.*

Le rang de la matrice M s'obtient par la méthode du Pivot de Gauß, c'est l'entier r, obtenu lorsque l'on arrive à une matrice semi-réduite ou réduite.

Proposition 4.16 *Si f est une application entre deux espaces vectoriels de dimension finie et de base B_E et B_F, le rang de f est égal au rang de la matrice de f relative aux bases B_E et B_F.*

Preuve : En effet soit $B_E = (e_1,\dots,e_n)$ et $B_F = (u_1,\dots,u_m)$, la matrice M vérifie :

$$(f(e_1),\dots,f(e_n)) = (u_1,\dots,u_m)\cdot M.$$

L'application linéaire de $\mathbb{K}^m$ dans F définie par

$$\phi(y_1,\dots,y_m) = (u_1,\dots,u_m)\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

est bijective (existence et unicité de l'écriture dans une base), et fait correspondre au j-ième vecteur colonne de M le vecteur $f(e_j)$. Si on considère le sous-espace G de K^m engendré par les vecteurs colonnes de M, la restriction de ϕ à G est une bijection linéaire entre G et $f(E)$ ces deux espaces ont donc la même dimension. □

4.4 Changement de bases

Le but de ce paragraphe est l'étude de la matrice d'une application linéaire $f\colon E \longrightarrow F$, en fonction des bases choisies. On se donne $(e_1, \dots e_n)$ une base de E, $(u_1, \dots, u_m)$ une base de F, et la matrice $M = \mathrm{Mat}_{u,e}(f)$ de f relative à ces bases. Se donner une autre base de E revient à se donner une matrice carrée inversible d'ordre n, $P = \mathrm{Mat}_{e_i,e'_i}(\mathrm{Id}_E)$, souvent appelée matrice de passage de la base $(e'_1, \dots, e'_n)$ à la base $(e_1, \dots, e_n)$ où encore matrice de changement de bases. Cette matrice vérifie

$$(e'_1, \dots, e'_n) = (e_1, \dots, e_n)P.$$

De même on peut se donner une matrice carrée inversible $Q = \mathrm{Mat}_{u_i,u'_i}(\mathrm{Id}_F)$ d'ordre m donc une nouvelle base

$$(u'_m, \dots, u'_m) = (u_m, \dots, u_m)Q.$$

Pour déterminer la matrice de f relative à ces nouvelles bases, il suffit d'utiliser les relations de Chasles :

$$\mathrm{Mat}_{u',e'}(f) = \mathrm{Mat}_{u',e'}(\mathrm{Id}_F \circ f \circ \mathrm{Id}_E) = \mathrm{Mat}_{u',u}(\mathrm{Id}_F)\,\mathrm{Mat}_{u,e}(f)\,\mathrm{Mat}_{e,e'}(\mathrm{Id}_E)$$

Puisqu'on a en outre

$$I_m = \mathrm{Mat}_{u',u'}(\mathrm{Id}_F) = \mathrm{Mat}_{u',u}(\mathrm{Id}_F)\,\mathrm{Mat}_{u,u'}(\mathrm{Id}_F) = \mathrm{Mat}_{u',u}(\mathrm{Id}_F)Q$$

on en déduit $Q^{-1} = \mathrm{Mat}_{u',u}(\mathrm{Id}_F)$ et donc finalement la formule (classique)

$$Q^{-1}\,\mathrm{Mat}_{u,e}(f)P = \mathrm{Mat}_{u',e'}(f)$$

4.5 Matrices équivalentes

Définition 4.17 *Deux matrices A et B de $M_{n,m}(K)$ sont dites équivalentes, s'il existe une matrice inversible P de $M_m(K)$ et une matrice inversible Q de $M_n(K)$ vérifiant :*

$$B = Q^{-1}AP.$$

On obtient une relation d'équivalence sur $M_{n,m}(K)$, en effet :

- La relation est réflexive, si $Q = I_n$ et $P = I_m$ alors $A = Q^{-1}AP$, donc A est équivalente à B.
- La relation est symétrique : P et Q étant inversibles on obtient

$$B = Q^{-1}AP \Leftrightarrow A = QBP^{-1}$$

Or cela s'écrit encore $A = Q_1^{-1}BP_1$ avec $Q_1 = Q^{-1}$ et $P_1 = P^{-1}$. Si B est équivalente à A, alors A est équivalente à B.

- La relation est transitive : Si B est équivalente à A et C est équivalente à B alors il existe des matrices inversibles Q_1 et Q_2 dans $M_n(K)$ et des matrices inversibles P_1 et P_2 dans $M_m(K)$ vérifiant :

$$B = Q_1^{-1} A P_1, \text{ et } C = Q_2^{-1} B P_2.$$

Alors

$$C = Q_2^{-1} Q_1^{-1} A P_1 P P_2$$

En notant $Q = Q_1 Q_2$ et $P = P_1 P_2$, Q est une matrice inversible de $M_n(K)$, d'inverse $Q^{-1} = Q_2^{-1} Q_1^{-1}$ et P est une matrice inversible de $M_m(K)$, de plus on a

$$C = Q^{-1} A P,$$

ce qui démontre que C est équivalente à A.

Quand on a une relation d'équivalence, on peut ranger les éléments que l'on étudie par classes d'équivalences, et une question intéressante est : “Comment peut-on décrire une classe d'équivalence ?”.

Proposition 4.18 *Les classes de $M_{n,m}(K)$ pour la relation d'équivalence des matrices sont décrites par le rang : Deux matrices A et B de $M_{n,m}(K)$ sont équivalentes si et seulement si elles ont même rang. En particulier il y a exactement* $\inf\{n,m\}+1$ *classes d'équivalences pour cette relation.*

Preuve : En effet d'après la méthode du Pivot de Gauss, il existe une matrice inversible P de $M_n(K)$ et une matrice de permutation S de $M_m(K)$ telles que PMS soit réduite de la forme :

$$PMS = \begin{pmatrix} I_r & M_1 \\ 0 & 0 \end{pmatrix}$$

Alors

$$PMS \begin{pmatrix} I_r & -M_1 \\ 0 & I_{m-r} \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

Donc une matrice de rang r est équivalente à la matrice $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ et réciproquement.

Il y a donc autant de classes d'équivalences que de matrices de la forme $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, soit $\inf\{n,m\}$. $\square$

4.6 Matrices semblables

Définition 4.19 *Deux matrices A et B de $M_n(K)$ sont dites semblables, s'il existe une matrice inversible P de $M_m(K)$ vérifiant :*

$$B = P^{-1} A P.$$

On obtient encore une relation réflexive, symétrique et transitive, donc une relation d'équivalence, les classes d'équivalences obtenues s'appelles les classes de similitudes. L'étude des classes d'équivalences pour cette relation est plus compliquée que la précédente, en particulier il y a une infinité de classes.

Proposition 4.20 *Soit f une application d'un espace vectoriel E de dimension n dans lui-même, si $u_1, \ldots, u_n$ est une base de E on note M la matrice de f dans cette base :*

$$(f(u_1), \ldots, f(u_n)) = (u_1, \ldots, u_n)M.$$

La classe de similitude de M est l'ensemble des matrices de $M_n(K)$ représentant f lorsque l'on fait varier la base de E.

En effet si $(v_1, \ldots, v_n) = (u_1, \ldots, u_n)P$ est une nouvelle base (P inversible) la matrice de f dans cette nouvelle base est $P^{-1}MP$ puisque

$$(f(v_1), \ldots, f(v_n)) = (v_1, \ldots, v_n)P^{-1}MP.$$

En particulier la classe de similitude d'une matrice diagonale de la forme λI_n est réduite à cette matrice.

Exemples : On se place dans $M_2(\mathbb{R})$.

- La matrice $\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$ est semblable à $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.
- La matrice $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ n'est pas semblable à $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Chapitre 5

Somme directe d'espaces vectoriels, produits

5.1 Produit d'espaces vectoriels

Nous commençons par la définition générale, où le produit est indexé par un ensemble quelconque, même si nous porterons presque exclusivement notre intérêt sur le cas d'un produit indexé par un ensemble fini.

Définition 5.1 *Soient $(E_i)_{i\in I}$ une famille d' espaces vectoriels définis sur le corps $\mathbb{K}$ et indexée par un ensemble I. Le produit cartésien $\prod_{i\in I} E_i$ admet une structure d'espace vectoriel naturelle, l'addition étant donnée par :*

$$\forall (u_i)_{i\in I} \in \prod_{i\in I} E_i, \forall (v_i)_{i\in I} \in \prod_{i\in I} E_i, (u_i)_{i\in I} + (v_i)_{i\in I} = (u_i + v_i)_{i\in I}$$

et la loi externe étant donnée par

$$\forall (u_i)_{i\in I} \in \prod_{i\in I} E_i, \forall \lambda \in \mathbb{K}, \lambda.(u_i)_{i\in I} = (\lambda.u_i)_{i\in I}$$

Remarque : Cette définition généralise la définition de la structure d'espace vectoriel sur $\mathbb{K}^n$. L'ensemble $\mathbb{K}^n$ lui-même est égal au produit cartésien de n copies de K. On omet la démonstration (de routine à ce stade du cours) que $\prod_{i\in I} E_i$ est bien un espace vectoriel.

Proposition 5.2 *Les projections $p_j \colon \prod_{i\in I} E_i \longrightarrow E_j$ définies par $(u_i)_{i\in I} \mapsto u_j$ sont des applications linéaires, et pour toute famille $(f_i)_{i\in I}$ d'applications linéaires $f_i \colon G \longrightarrow E_i$, il existe une et une seule application linéaire $\phi \colon G \longrightarrow \prod_{i\in I} E_i$ telle que $f_i = p_i \circ \phi$.*

Preuve : La structure sur $\prod_i E_i$ est définie pour que les projections p_j soient linéaires. Précisément on a $p_j(\lambda.(u_i)_i + (v_i)_i) = p_j((\lambda.u_i + v_i)_i) = \lambda.u_j + v_j = \lambda.p_j((u_i)_i) + p_j((v_i)_i)$. D'où la linéarité des p_j. On fixe une famille d'applications linéaires (f_i) comme dans l'énoncé. On définit $\phi \colon G \longrightarrow \prod_i E_i$ avec la formule

$$\forall g \in G,\ \phi(g) = (f_i(g))_{i\in I}\ .$$

On vérifie sans difficultés la linéarité de ϕ, les conditions $f_i = p_i \circ \phi$ et l'unicité de g. $\square$

En particulier une application ϕ de G dans $\prod_{i\in I} E_i$ est linéaire, si et seulement si les applications $p_i \circ \phi$ de G dans E_i le sont.

Définition 5.3 *Soit E un espace vectoriel, et I un ensemble d'indice, si l'on prend $E_i = E$ pour tout i, on note alors le produit $\prod_{i\in I} E_i$ sous la forme E^I, si de plus $I = \{1, \dots, n\}$ on note $\prod_{i\in I} E_i$ sous la forme E^n.*

Exemples : On considère le corps de base $\mathbb{R}$.

1. L'ensemble des suites numériques $\mathbb{R}^{\mathbb{N}}$ est un $\mathbb{R}$ espace vectoriel. L'ensemble d'indice est $\mathbb{N}$ et $E_i = \mathbb{R}$.
2. L'espace $\mathbb{R}^n$ est le $\mathbb{R}$-espace vectoriel produit de n copies de $\mathbb{R}$.

Ainsi on justifie le fait que l'application $f\colon \mathbb{R} \longrightarrow \mathbb{R}^{\mathbb{N}}$ définie par $f(x) = (nx)_{n\in\mathbb{N}}$ est une application linéaire en vérifiant que pour tout entier n, l'application $f_n\colon \mathbb{R} \longrightarrow \mathbb{R}$ définie par $f_n(x) = nx$ est linéaire.

Proposition 5.4 *Si pour $i \in \{1, \dots, n\}$ l'espace vectoriel E_i est de dimension finie d_i alors $E_1 \times E_2 \cdots \times E_n$ est de dimension finie $d_1 + d_2 + \cdots + d_n$.*

Preuve : On construit aisément une base de $E_1 \times E_2 \cdots \times E_n$ à l'aide de la donnée d'une base de chaque E_i en prenant pour chaque entier positif j inférieur ou égal à n, les n-uplets dont la j-ième composante est un élément de la base de E_j, et les autres composantes sont nulles. $\square$

5.2 Somme directe d'espaces vectoriels

Commençons là encore par une définition générale :

Définition 5.5 *Soient $(E_i)_{i\in I}$ une famille d' espaces vectoriels définis sur le corps K et indexée par un ensemble I. Le sous-espace vectoriel de $\prod_{i\in I} E_i$ défini par*

$$\bigoplus_{i\in I} E_i = \{(u_i)_{i\in I} \in \prod_{i\in I} E_i | \{i \in I, u_i \neq 0_{E_i}\} \text{ est fini}\}$$

est appelé la somme directe externe des E_i. Un élément de $\bigoplus_{i\in I} E_i$ se note $\oplus_i u_i$. On a alors :

$$\forall \lambda \in K, \forall \oplus_i u_i \in \bigoplus_{i\in I} E_i, \lambda. \oplus_i u_i = \oplus_i \lambda.u_i,$$

$$\forall \oplus_i u_i \in \bigoplus_{i\in I} E_i, \forall \oplus_i v_i \in \bigoplus_{i\in I} E_i, \oplus_i u_i + \oplus_i v_i = \oplus_i(u_i + v_i).$$

Remarque :

1. Clairement $\bigoplus_i E_i$ est un sous-espace de $\prod_i E_i$, ce qui justifie cette définition.
2. Dès que I est fini on a par définition l'égalité $\prod_i E_i = \bigoplus_i E_i$. On ne perçoit la distinction entre somme et produit que pour des ensemble d'indice infini et en particulier pour des espaces vectoriels de dimensions infinies. C'est pour cette raison qu'on est sorti provisoirement de la théorie des espaces vectoriels de type fini dans ce dernier chapitre.

Proposition 5.6 *Pour chaque indice j, l'application linéaire $\phi_j \colon E_j \longrightarrow \bigoplus_{i\in I} E_i$ définie par $\phi_j(u) = (v_i)_{i\in I}$ avec $v_j = u$ et $v_i = 0_{E_i}$ si $i \neq j$ est une injection appelée injection canonique de E_j dans $\bigoplus_{i\in I} E_i$. Soit G un $\mathbb{K}$-espace vectoriel. Toute familles $(f_i)_{i\in I}$, $f_i \colon E_i \longrightarrow G$, d'applications linéaires de E_i dans G définit une unique application linéaire $\varphi\colon \bigoplus_{i\in I} E_i \longrightarrow G$ telle que pour tout i on ait $\varphi\circ\phi_i = f_i$.*

Preuve : L'application linéaire $\phi_j \colon E_j \longrightarrow \prod_i E_i$ est parfaitement définie, linéaire, et unique grâce à la proposition 5.2 appliquée aux applications Id_{E_j} et aux applications nulles $E_j \longrightarrow E_i$ pour $i \neq j$. Manifestement l'image de ϕ_j est contenue dans le sous-espace $\bigoplus_i E_i$. Cela définit (de manière unique) l'application linéaire ϕ_j qui de plus est injective.

Étant donné une famille d'applications linéaires (f_i) comme dans l'énoncé, on définit $\varphi(\oplus_i u_i) = \sum_i f_i(u_i)$. Comme les u_i sont tous nuls sauf un nombre fini les $f_i(u_i)$ aussi et la somme a un sens dans G. Pour terminer la preuve il faut s'assurer que cette application φ est linéaire, est telle que $\varphi \circ \phi_i = f_i$ et est unique avec cette propriété. Ce sont des vérifications de routine à ce stade du cours, laissées aux lecteurs. $\square$

5.3 Cas de deux espaces vectoriels

Définition 5.7 *Soit E un espace vectoriel, F et G deux sous-espaces de E.*

1. *On note $F + G$ le sous-espace de E engendré par $F \cup G$.*
2. *F et G sont dits en somme directe si*
$$F \cap G = \{0_E\}.$$
Par abus de langage on note alors $F \oplus G$ le sous-espace $F + G$ de E.
3. *F et G sont dits supplémentaires si $F \cap G = \{0_E\}$ et $F + G = E$ (et dans ce cas on peut noter par abus de langage $F \oplus G = E$).*

Les abus de notations de la définition précédente sont sans conséquence grâce à la proposition suivante :

Proposition 5.8 *Soit E un espace vectoriel, F et G deux sous-espaces de E.*

1. *On a $F + G = \{u + v, u \in F, v \in G\}$.*
2. *Si $F \cap G = \{0_E\}$, alors tout $x \in F + G$ s'écrit de manière unique $x = x_f + x_g$ avec $x_f \in F$ et $x_g \in G$ et on a un isomorphisme canonique entre $F + G$ et la somme directe externe $F \bigoplus G$ de la définition 5.5.*

3. *Si F et G sont supplémentaires dans E alors E est isomorphe à la somme directe externe $F \bigoplus G$ de la définition 5.5.*

Preuve :

1. Par définition un élément de $F + G = \langle F \cup G \rangle$ est une combinaison linéaire d'éléments de $F \cup G$, c'est-à-dire un élément de la forme $x = \sum \lambda_i f_i + \sum \mu_j g_j$ pour des f_i dans F, des g_j dans G et des scalaires λ_i, μ_j dans $\mathbb{K}$. Mais comme F et G sont des sous-espaces on a $u = \sum \lambda_i f_i \in F$ et $v = \sum \mu_j g_j \in G$ et cela donne $x = u + v \in \{u + v, u \in F, v \in G\}$ et l'inclusion
$$F + G \subset \{u + v, u \in F, v \in G\}.$$
L'inclusion réciproque est immédiate.
2. On suppose que $x \in E$ admet deux écritures $x = x_f + x_g = x'_f + x'_g$ avec $x_f, x'_f \in F$ et $x_g, x'_g \in G$. Alors on en tire $x_f - x'_f = x_g - x'_g \in F \cap G$. Si on suppose maintenant $F \cap G = \{0\}$ alors on obtient $x_f = x'_f$ et $x_g = x'_g$, d'où l'unicité de l'écriture avec cette hypothèse. L'application naturelle $(x_f, x_g) \mapsto x = x_f + x_g$ est dans ce cas un isomorphisme de la somme directe externe $F \bigoplus G$ au sens de la définition 5.5 dans $F + G$.
3. est simplement le cas particulier $F + G = E$ de 2.
□

Exemples

- L'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ est somme directe du sous-espace E^+ formé par les fonctions paires et E^- formé par les fonctions impaires.
- L'espace vectoriel des matrices carrées d'ordre n est somme directe du sous-espace F des matrices symétriques (${}^t\!A = A$) et du sous-espace G des matrices antisymétriques (${}^t\!A = -A$).

Proposition 5.9 *Soit E un K-espace vectoriel de dimension finie n, et F un sous-espace vectoriel de E. Si $u_1, \cdots, u_k$ est une base de F et $u_{k+1}, \cdots, u_n$ sont $n-k$ vecteurs de E permettant de compléter la base de F en une base de E, alors l'espace vectoriel G engendré par $u_{k+1}, \ldots, u_n$ est un supplémentaire de F. En particulier tout sous-espace F de E admet un (le plus souvent des) supplémentaire(s).*

Preuve : Comme $u_1, \cdots, u_n$ est un système générateur de E on a bien $E = F + G$. Supposons que $x = \sum_{i=1}^{i=k} \lambda_i u_i = \sum_{i=k+1}^{i=n} \lambda_i u_i \in F \cap G$. Alors il suit $0_E = \sum_{i=1}^{i=k} \lambda_i u_i + \sum_{i=k+1}^{i=n} -\lambda_i u_i$ et puisque les u_i forment une base cela conduit à $\forall i,\ \lambda_i = 0$ puis $x = 0$. On a démontré l'égalité $E = F \bigoplus G$. □

On remarque que seuls $\{0_E\}$ et E admettent un supplémentaire unique. En effet si F est un sous-espace de E distincts de $\{0_E\}$ et E, et si G est un supplémentaire de F, de base $u_1, \ldots, u_k$, alors pour tout u non nul dans F, $u_1, \ldots, u_{k-1}, u_k + u$ engendre un supplémentaire de F distinct de G.

Proposition 5.10 *Soit E un espace vectoriel et F et G des sous-espaces vectoriels de E. On a la formule de dimension :*
$$\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)\ .$$

Preuve : Soit F' un supplémentaire de $F \cap G$ dans F et G' un supplémentaire de $F \cap G$ dans G. On a alors $\dim F' = \dim F - \dim(F \cap G)$ et $\dim G' = \dim G - \dim F \cap G$. Avec la proposition 5.4 on est ramené à démontrer les deux égalités $F + G = F' \bigoplus G = F' \bigoplus((F \cap G) \bigoplus G')$ qui sous-entendent à chaque fois que les sommes sont directes et que les ensembles concernés sont égaux. Seule la première égalité est à vérifier puisque la seconde revient à $G = G' \bigoplus(F \cap G)$, c'est-à-dire à la définition de G'. On a $F' \subset F$ et donc $F' + G \subset F + G$. Réciproquement si $x = x_f + x_g \in F + G$ on peut écrire $x_f = x'_f + x'_g$ avec $x'_f \in F'$ et $x'_g \in F \cap G$. Mais alors $x'_g + x_g \in G$ et donc $x = x'_f + (x'_g + x_g) \in F' + G$. D'où l'égalité d'ensemble $F + G = F' + G$. Vérifions que $F' \cap G = \{0\}$. Soit $x \in F' \cap G$ Alors $x \in F' \cap G \cap F$ car $F' \subset F$. Mais F' est un supplémentaire dans F de $F \cap G$ et donc $x = 0$. On a démontré l'égalité $F + G = F' \bigoplus G$. □

5.4 Projections linéaires

Définition 5.11 *Soit E un $\mathbb{K}$-espace vectoriel, un projecteur de E est un endomorphisme p, $\mathbb{K}$-linéaire de E, vérifiant $p \circ p = p$.*

Proposition 5.12 *Si p est un projecteur de E alors on a $E = \mathrm{Ker}(p) \bigoplus \mathrm{Im}(p)$.*

Preuve : On suppose $p \circ p = p$ c'est-à-dire pour tout $x \in E$ $p(p(x)) = p(x)$. Soit $x \in \mathrm{Ker}(p) \cap \mathrm{Im}(p)$. Alors on a $x = p(y)$ pour un $y \in E$ et aussi $0 = p(x) = p(p(y)) = p(y) = x$. Cela démontre $\mathrm{Ker}(p) \cap \mathrm{Im}(p) = \{0\}$. D'autre part si $x \in E$ on peut écrire $x = x - p(x) + p(x)$. Et on a $p(x - p(x)) = p(x) - p(p(x)) = p(x) - p(x) = 0$. Donc $x - p(x) \in \mathrm{Ker}(p)$ et $x = (x - p(x)) + p(x) \in \mathrm{Ker}(p) + \mathrm{Im}(p)$. □

Exemples

1. L'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par la matrice
$$\begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}$$
est un projecteur de noyau $\mathrm{Ker}\, p = \mathbb{R} \cdot e_1$ et d'image $\mathbb{R} \cdot (e_2 + 2e_1)$.
2. L'application linéaire de $\mathbb{R}^3$ dans $\mathbb{R}^3$ définie par la matrice
$$\begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
est un projecteur de noyau $\mathrm{Ker}\, p = \mathbb{R} \cdot e_1$ et d'image $\mathbb{R} \cdot (e_2 + 2e_1) + \mathbb{R} \cdot e_3$.
3. Soit E l'espace vectoriel sur $\mathbb{R}$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. L'application $\phi \colon E \longrightarrow E$, définie par
$$\forall f \in E, \forall x \in \mathbb{R}, \phi(f) = \frac{f(x) + f(-x)}{2}$$
est un projecteur, son image est formé des fonctions paires et son noyau des fonction impaires.

4. Étant donné une décomposition en somme directe $E = F \bigoplus G$ on appelle projection sur G de direction F l'application linéaire $p\colon E \longrightarrow G$ définie à partir de l'unique écriture $x = x_f + x_g$ avec $x_f \in F$ et $x_g \in G$ et la formule $p(x_f + x_g) = x_g$.
   Dans ce cadre si on se donne une base $\mathcal{B}$ de E obtenue en complétant une base de F avec une base de G la matrice de p relativement à cette base est la matrice
   $$\begin{pmatrix} O_{\dim F, \dim F} & O_{\dim G, \dim F} \\ O_{\dim G, \dim F} & I_{\dim G} \end{pmatrix} ,$$
   où $O_{n,m}$ désigne la matrice nulle de $M_{n,m}(\mathbb{K})$ et I_r désigne la matrice identité de $M_r(\mathbb{K})$. En particulier on voit immédiatement sur la matrice l'identité $p \circ p = p$.