Algèbre linéaire II (MAT 3541)

Notes de cours réalisées par Damien Roy
avec l'assistance de Pierre Bel
et l'appui financier du fonds de l'Université d'Ottawa
pour le développement de matériel pédagogique en français

Automne 2009
Département de mathématiques et de statistique
Université d'Ottawa

Table des matières

Préface

Ces notes de cours ont été rédigées à l'intention des étudiants du cours d'algèbre linéaire II (MAT 3541) à l'Université d'Ottawa. Les trois premiers chapitres de ces notes présentent une révision de notions de base étudiées au cours préalable d'algèbre linéaire I (MAT 2541) : espaces vectoriels, applications linéaires et diagonalisation. Le reste du cours est divisé en trois parties. Les chapitres 4 à 8 constituent le coeur du cours. Ils culminent avec le théorème de structure pour les modules de type fini sur un anneau euclidien, un résultat général qui conduit à la fois à la forme canonique de Jordan des opérateurs sur un espace vectoriel complexe de dimension finie et aussi au théorème de structure pour les groupes abéliens de type fini. Le chapitre 9 concerne la dualité et le produit tensoriel. Tandis que le chapitre 10 traite des théorèmes spectraux pour les opérateurs sur un espace euclidien (après un rappel des notions vues au cours préalable MAT 2541). Par rapport au plan initial, il manque encore un onzième chapitre sur les espaces hermitiens qui sera ajouté plus tard. Les notes sont complétées par deux appendices. L'appendice A fournit un rappel des structures de base de l'algèbre qui sont utilisées dans le cours : groupes, anneaux et corps (et les morphismes entre ces objets). L'appendice B étudie quant à lui le déterminant des matrices à coefficients dans un anneau commutatif quelconque avec ses propriétés. En général, le déterminant est simplement défini sur un corps et le but de cet appendice est de faire voir que cette notion se généralise facilement à un anneau commutatif quelconque.

Ce cours fournit aux étudiants l'occasion d'apprendre la théorie des modules qui n'est pas couverte, faute de temps, dans les autres cours d'algèbre sous-gradués. Pour simplifier et rendre le cours plus facilement digestible, j'ai évité d'introduire les modules quotients. Cela pourrait être ajouté ultérieurement en exercice (comme défi pour les étudiants les plus motivés). D'après les commentaires que j'ai reçus, les étudiants apprécient la richesse des concepts qui sont présentés dans ce cours et qui leur ouvrent de nouveaux horizons.

Je remercie Pierre Bel qui m'a beaucoup aidé dans la dactylographie et la mise au point de ces notes de cours. Je remercie aussi le Rectorat aux Études de l'Université d'Ottawa pour son appui finacier à ce projet via le fonds pour le développement de matériel pédagogique en français. Ce projet n'aurait pas pu voir le jour sans leur appui. Enfin, je remercie mon épouse Laura Dutto pour son aide dans la révision des versions préliminaires de ces notes et les étudiants qui ont suivi ce cours avec moi à l'hiver et à l'automne 2009 pour leur participation et leur commentaires.

Damien Roy Ottawa, janvier 2010.

Chapitre 1

Retour sur les espaces vectoriels

Ce chapitre, comme les deux suivants, est consacré à un rappel des notions fondamentales d'algèbre linéaire apprises au cours précédent (MAT 2541). Ce premier chapitre concerne l'objet d'étude privilégié de l'algèbre linéaire, c'est-à-dire les espaces vectoriels. On rappelle ici les notions d'espace vectoriel, de sous-espace vectoriel, de base, de dimension, de coordonnées, et de somme directe. Le lecteur devra prêter une attention spéciale à la notion de somme directe puisqu'elle joue un rôle fondamental dans la suite.

1.1 La notion d'espace vectoriel

Définition 1.1.1. Un *espace vectoriel* sur un corps K est un ensemble V muni d'une addition et d'une multiplication par les éléments de K :

$$\begin{array}{ccc} V \times V & \longrightarrow & V \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \mathbf{u} + \mathbf{v} \end{array} \quad \text{et} \quad \begin{array}{ccc} K \times V & \longrightarrow & V \\ (a, \mathbf{v}) & \longmapsto & a\mathbf{v} \end{array}$$

qui satisfont les axiomes suivants :

$$\left.\begin{array}{ll} \textbf{EV1}. & \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \\ \textbf{EV2}. & \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \end{array}\right\} \text{pour tout choix de } \mathbf{u}, \mathbf{v}, \mathbf{w} \in V.$$

EV3. Il existe $\mathbf{0} \in V$ tel que $\mathbf{v} + \mathbf{0} = \mathbf{v}$ pour tout $\mathbf{v} \in V$.
EV4. Pour tout $\mathbf{v} \in V$, il existe $-\mathbf{v} \in V$ tel que $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

$$\left.\begin{array}{ll} \textbf{EV5}. & 1\mathbf{v} = \mathbf{v} \\ \textbf{EV6}. & a(b\mathbf{v}) = (ab)\mathbf{v} \\ \textbf{EV7}. & (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v} \\ \textbf{EV8}. & a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v} \end{array}\right\} \text{pour tout choix de } a, b \in K \text{ et de } \mathbf{u}, \mathbf{v} \in V.$$

Ces huit axiomes s'interprètent aisément. Les quatre premiers signifient que $(V, +)$ est un *groupe abélien* (cf. Appendice **A**). En particulier **EV1** et **EV2** impliquent que l'ordre dans

lequel on additionne les éléments $\mathbf{v}_1, \ldots, \mathbf{v}_n$ de V n'affecte par leur somme notée

$$\mathbf{v}_1 + \cdots + \mathbf{v}_n \quad \text{ou} \quad \sum_{i=1}^{n} \mathbf{v}_i.$$

La condition **EV3** détermine uniquement l'élément $\mathbf{0}$ de V. On l'appelle le *vecteur nul* de V, le terme "vecteur" étant le nom générique employé pour désigner un élément d'un espace vectoriel. De même, pour chaque $\mathbf{v} \in V$, il existe un et un seul vecteur $-\mathbf{v} \in V$ qui satisfait **EV4**. On l'appelle l'*inverse additif* de $\mathbf{v}$. L'existence de l'inverse additif permet de définir la soustraction dans V par

$$\mathbf{u} - \mathbf{v} := \mathbf{u} + (-\mathbf{v}).$$

Les axiomes **EV5** et **EV6** ne concernent que la multiplication par un scalaire tandis que les axiomes **EV7** et **EV8** lient les deux opérations. Ces derniers demandent que la multiplication par un scalaire soit distributive sur l'addition à gauche comme à droite (c'est-à-dire sur l'addition dans K, comme dans V). Ils impliquent les formules de distributivité générales :

$$\left(\sum_{i=1}^{n} a_i\right)\mathbf{v} = \sum_{i=1}^{n} a_i\mathbf{v} \quad \text{et} \quad a\sum_{i=1}^{n} \mathbf{v}_i = \sum_{i=1}^{n} a\mathbf{v}_i$$

quels que soient $a, a_1, \ldots, a_n \in K$ et $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n \in V$.

On reviendra sur ces axiomes quand on étudiera plus loin la notion de module sur un anneau commutatif qui généralise celle d'espace vectoriel sur un corps.

Exemple 1.1.2. Soit $n \in \mathbb{N}^*$. L'ensemble

$$K^n = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} ; a_1, a_2, \ldots, a_n \in K \right\}$$

des n-uplets d'éléments de K est un espace vectoriel sur K pour les opérations :

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix} \quad \text{et} \quad c\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} c\,a_1 \\ c\,a_2 \\ \vdots \\ c\,a_n \end{pmatrix}$$

En particulier, $K^1 = K$ est un espace vectoriel sur K.

Exemple 1.1.3. Plus généralement, soient $m, n \in \mathbb{N}^*$. L'ensemble

$$\mathrm{Mat}_{m\times n}(K) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} ; a_{11}, \ldots, a_{mn} \in K \right\}$$

des matrices $m \times n$ à coefficients dans K est un espace vectoriel sur K pour les opérations usuelles :

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \dots & a_{mn}+b_{mn} \end{pmatrix}$$

$$\text{et} \quad c \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} c\,a_{11} & \cdots & c\,a_{1n} \\ \vdots & & \vdots \\ c\,a_{m1} & \dots & c\,a_{mn} \end{pmatrix}.$$

Exemple 1.1.4. Soit X un ensemble quelconque. L'ensemble $\mathcal{F}(X, K)$ des fonctions de X dans K est un espace vectoriel sur K lorsqu'on définit la somme de deux fonctions $f\colon X \to K$ et $g\colon X \to K$ comme étant la fonction $f+g\colon X \to K$ donnée par

$$(f+g)(x) = f(x) + g(x) \quad \text{pour tout } x \in X,$$

et le produit de $f\colon X \to K$ par un scalaire $c \in K$ comme étant la fonction $c\,f\colon X \to K$ donnée par

$$(c\,f)(x) = c\,f(x) \quad \text{pour tout } x \in X.$$

Exemple 1.1.5. Enfin, si $V_1, \dots, V_n$ sont des espaces vectoriels sur K, leur produit cartésien

$$V_1 \times \cdots \times V_n = \{(\mathbf{v}_1, \dots, \mathbf{v}_n)\,;\, \mathbf{v}_1 \in V_1, \dots, \mathbf{v}_n \in V_n\}$$

est un espace vectoriel sur K pour les opérations

$$(\mathbf{v}_1, \dots, \mathbf{v}_n) + (\mathbf{w}_1, \dots, \mathbf{w}_n) = (\mathbf{v}_1 + \mathbf{w}_1, \dots, \mathbf{v}_n + \mathbf{w}_n),$$
$$c(\mathbf{v}_1, \dots, \mathbf{v}_n) = (c\mathbf{v}_1, \dots, c\mathbf{v}_n).$$

1.2 Sous-espaces vectoriels

Fixons un espace vectoriel V sur un corps K.

Définition 1.2.1. Un *sous-espace vectoriel* de V est un sous-ensemble U de V qui remplit les conditions suivantes :

SE1. $\mathbf{0} \in U$.
SE2. Si $\mathbf{u}, \mathbf{v} \in U$, alors $\mathbf{u} + \mathbf{v} \in U$.
SE3. Si $\mathbf{u} \in U$ et $c \in K$, alors $c\mathbf{u} \in U$.

Les conditions **SE2** et **SE3** signifient que U est stable sous l'addition dans V et stable sous la multiplication par les éléments de K. On obtient ainsi des opérations

$$\begin{array}{ccc} U \times U & \longrightarrow & U \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \mathbf{u} + \mathbf{v} \end{array} \quad \text{et} \quad \begin{array}{ccc} K \times U & \longrightarrow & U \\ (a, \mathbf{v}) & \longmapsto & a\mathbf{v} \end{array}$$

sur U et on montre que, pour ces opérations, U est lui-même un espace vectoriel. La condition **SE1** peut être remplacée par $U \neq \emptyset$ car, si U contient un élément $\mathbf{u}$, alors **SE3** implique que $0\,\mathbf{u} = \mathbf{0} \in U$. Néanmoins, il est généralement tout aussi simple de vérifier **SE1**. On a donc

Proposition 1.2.2. *Un sous-espace vectoriel U de V est un espace vectoriel sur K pour l'addition et la multiplication par un scalaire restreintes de V à U.*

Ainsi tous les sous-espaces vectoriels de V fournissent de nouveaux exemples d'espaces vectoriels. De plus, si U est un sous-espace de V, on peut aussi considérer les sous-espaces de U. Par contre, on montre que ceux-ci sont simplement les sous-espaces de V contenus dans U. Cela ne fournit pas de nouveaux exemples. En particulier la notion de sous-espace est transitive.

Proposition 1.2.3. *Si U est un sous-espace de V et si W est un sous-espace de U, alors W est un sous-espace de V.*

On peut aussi former la somme et l'intersection de sous-espaces de V :

Proposition 1.2.4. *Soient $U_1, \dots, U_n$ des sous-espaces de V. Les ensembles*

$$U_1 + \cdots + U_n = \{\mathbf{u}_1 + \cdots + \mathbf{u}_n \,;\, \mathbf{u}_1 \in U_1, \dots, \mathbf{u}_n \in U_n\}$$

$$U_1 \cap \cdots \cap U_n = \{\mathbf{u} \,;\, \mathbf{u} \in U_1, \dots, \mathbf{u} \in U_n\}$$

appelés respectivement la somme *et l'*intersection *de $U_1, \dots, U_n$ sont des sous-espaces de V.*

Exemple 1.2.5. Soient V_1 et V_2 des espaces vectoriels sur K. Alors

$$U_1 = V_1 \times \{\mathbf{0}\} = \{(\mathbf{v}_1, \mathbf{0}) \,;\, \mathbf{v}_1 \in V_1\} \quad \text{et} \quad U_2 = \{\mathbf{0}\} \times V_2 = \{(\mathbf{0}, \mathbf{v}_2) \,;\, \mathbf{v}_2 \in V_2\}$$

sont deux sous-espaces de $V_1 \times V_2$. On trouve que

$$U_1 + U_2 = V_1 \times V_2 \quad \text{et} \quad U_1 \cap U_2 = \{(\mathbf{0}, \mathbf{0})\}\,.$$

1.3 Bases

Dans cette section, on fixe un espace vectoriel V sur un corps K et des éléments $\mathbf{v}_1, \dots, \mathbf{v}_n$ de V.

Définition 1.3.1. On dit qu'un élément $\mathbf{v}$ de V est une *combinaison linéaire* de $\mathbf{v}_1, \dots, \mathbf{v}_n$ s'il existe $a_1, \dots, a_n \in K$ tels que

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n.$$

On désigne par

$$\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle_K = \{a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \,;\, a_1, \dots, a_n \in K\}$$

l'ensemble des combinaisons linéaires de $\mathbf{v}_1, \dots, \mathbf{v}_n$.

Les identités

$$0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n = \mathbf{0}$$
$$\sum_{i=1}^{n} a_i\mathbf{v}_i + \sum_{i=1}^{n} b_i\mathbf{v}_i = \sum_{i=1}^{n}(a_i + b_i)\mathbf{v}_i$$
$$c\sum_{i=1}^{n} a_i\mathbf{v}_i = \sum_{i=1}^{n}(c\,a_i)\mathbf{v}_i$$

montrent que $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K$ est un sous-espace de V. Ce sous-espace contient $\mathbf{v}_1, \ldots, \mathbf{v}_n$ car on a

$$\mathbf{v}_j = \sum_{i=1}^{n} \delta_{i,j}\mathbf{v}_i \quad \text{où} \quad \delta_{i,j} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

Enfin, tout sous-espace de V qui contient $\mathbf{v}_1, \ldots, \mathbf{v}_n$ contient aussi toutes les combinaisons linéaires de $\mathbf{v}_1, \ldots, \mathbf{v}_n$, c'est-à-dire qu'il contient l'ensemble $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K$ tout entier. Donc :

Proposition 1.3.2. *L'ensemble $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K$ des combinaisons linéaires de $\mathbf{v}_1, \ldots, \mathbf{v}_n$ est un sous-espace de V qui contient $\mathbf{v}_1, \ldots, \mathbf{v}_n$ et qui est contenu dans tout sous-espace de V contenant $\mathbf{v}_1, \ldots, \mathbf{v}_n$.*

On exprime cette propriété en disant que $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K$ est le plus petit sous-espace de V qui contient $\mathbf{v}_1, \ldots, \mathbf{v}_n$ (au sens de l'inclusion). On l'appelle le sous-espace de V *engendré* par $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Définition 1.3.3. On dit que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *engendrent* V ou que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est *un système de générateurs de* V si

$$\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K = V$$

On dit que V est un espace vectoriel *de type fini* s'il admet un système fini de générateurs.

Il est important de bien distinguer l'ensemble fini $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ formé de n éléments (en comptant les répétitions possibles) de l'ensemble $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K$ de leurs combinaisons linéaires qui est généralement infini.

Définition 1.3.4. On dit que les vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sont *linéairement indépendants* ou que l'ensemble $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est *linéairement indépendant* si le seul choix de $a_1, \ldots, a_n \in K$ tel que

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0} \tag{1.1}$$

est $a_1 = \cdots = a_n = 0$. Sinon, on dit que les vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sont *linéairement dépendants*.

En d'autres termes, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sont linéairement dépendants s'il existe $a_1, \ldots, a_n$ non tous nuls qui vérifient la condition (1.1). Une relation de la forme (1.1) avec $a_1, \ldots, a_n$ non tous nuls s'appelle une *relation de dépendance linéaire entre* $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Définition 1.3.5. On dit que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est une *base* de V si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est un système de générateurs linéairement indépendants de V.

L'importance de la notion de base tient dans le résultat suivant :

Proposition 1.3.6. *Supposons que* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *soit une base de* V*. Pour tout* $\mathbf{v} \in V$*, il existe un et un seul choix de scalaires* $a_1, \ldots, a_n \in K$ *tels que*

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n\,. \tag{1.2}$$

Preuve. Soit $\mathbf{v} \in V$. Comme les vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_n$ engendrent V, il existe $a_1, \ldots, a_n \in K$ qui satisfont (1.2). Si $a'_1, \ldots, a'_n \in K$ satisfont aussi $\mathbf{v} = a'_1\mathbf{v}_1 + \cdots + a'_n\mathbf{v}_n$, alors on a

$$\begin{aligned}\mathbf{0} &= (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) - (a'_1\mathbf{v}_1 + \cdots + a'_n\mathbf{v}_n)\\ &= (a_1 - a'_1)\mathbf{v}_1 + \cdots + (a_n - a'_n)\mathbf{v}_n.\end{aligned}$$

Comme les vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sont linéairement indépendants, cela entraîne que $a_1 - a'_1 = \cdots = a_n - a'_n = 0$, donc $a_1 = a'_1, \ldots, a_n = a'_n$. □

Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est une base de V, les scalaires $a_1, \ldots, a_n$ qui satisfont (1.2) s'appellent les *coordonnées* de $\mathbf{v}$ dans la base $\mathcal{B}$. On désigne par

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n$$

le vecteur-colonne formé par ces coordonnées. On notera que cela suppose qu'on a fixé un ordre sur les éléments de $\mathcal{B}$, autrement dit que $\mathcal{B}$ est un ensemble ordonné d'éléments de V. En principe, on devrait écrire $\mathcal{B}$ comme un n-uplet $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, mais la tradition veut qu'on conserve la notation ensembliste.

Exemple 1.3.7. Soit $n \in \mathbb{N}^*$, les vecteurs

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\ \ldots,\ \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

forment une base $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ de K^n appelée la *base canonique* de K^n. On a

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

donc

$$\left[\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}\right]_{\mathcal{E}} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \text{ pour tout } \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in K^n.$$

Le résultat suivant est fondamental :

Proposition 1.3.8. *Supposons que V soit un espace vectoriel de type fini. Alors*

(i) V possède une base et toutes les bases de V ont même cardinal.

(ii) Tout système de générateurs de V contient une base de V.

(iii) Tout ensemble linéairement indépendant d'éléments de V est contenu dans une base de V.

(iv) Tout sous-espace de V possède une base.

La cardinalité commune des bases de V s'appelle la *dimension* de V. On la note $\dim_K V$. Rappelons que, par convention, si $V = \{\mathbf{0}\}$ alors $\emptyset$ est une base de V et $\dim_K V = 0$. Si V est de dimension n, il découle des énoncés (ii) et (iii) de la proposition 1.3.8 que tout système de générateurs de V contient au moins n éléments et que tout ensemble linéairement indépendant de V contient au plus n éléments. On en déduit aussi que si U_1 et U_2 sont des sous-espaces de V de même dimension finie avec $U_1 \subseteq U_2$, alors $U_1 = U_2$.

On rappelle aussi le fait suivant.

Proposition 1.3.9. *Si U_1 et U_2 sont des sous-espaces de dimension finie de V, alors*

$$\dim_K(U_1 + U_2) + \dim_K(U_1 \cap U_2) = \dim_K(U_1) + \dim_K(U_2).$$

1.4 Somme directe

Dans cette section, on fixe un espace vectoriel V sur un corps K et des sous-espaces $V_1, \ldots, V_s$ de V.

Définition 1.4.1. On dit que la somme $V_1 + \cdots + V_s$ est *directe* si le seul choix de vecteurs $\mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_s \in V_s$ tels que

$$\mathbf{v}_1 + \cdots + \mathbf{v}_s = \mathbf{0}.$$

est $\mathbf{v}_1 = \cdots \mathbf{v}_s = \mathbf{0}$. On exprime cette condition en écrivant la somme des sous-espaces $V_1, \ldots, V_s$ sous la forme $V_1 \oplus \cdots \oplus V_s$ avec des cercles autour des signes d'addition.

On dit que V est la *somme directe* de $V_1, \ldots, V_s$ si $V = V_1 \oplus \cdots \oplus V_s$.

Exemple 1.4.2. Soit $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base d'un espace vectoriel V sur K. La proposition 1.3.6 montre que $V = \langle \mathbf{v}_1 \rangle_K \oplus \cdots \oplus \langle \mathbf{v}_n \rangle_K$.

Exemple 1.4.3. On a toujours $V = V \oplus \{\mathbf{0}\} = V \oplus \{\mathbf{0}\} \oplus \{\mathbf{0}\}$.

Ce dernier exemple montre que, dans une somme directe, on peut avoir un ou plusieurs facteurs réduits à $\{\mathbf{0}\}$. La proposition suivante justifie l'importance de cette notion.

Proposition 1.4.4. *L'espace vectoriel V est la somme directe des sous-espaces $V_1, \ldots, V_s$ si et seulement si, pour chaque $\mathbf{v} \in V$, il existe un et seul choix de vecteurs $\mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_s \in V_s$ tels que*

$$\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_s.$$

La preuve est semblable à celle de la proposition 1.3.6 et elle est laissée au lecteur. On démontre aussi le critère suivant (comparer à l'exercice 1.4.2).

Proposition 1.4.5. *La somme $V_1 + \cdots + V_s$ est directe si et seulement si*

$$V_i \cap \big(V_1 + \cdots + \widehat{V_i} + \cdots + V_s\big) = \{\mathbf{0}\} \ \textit{pour } i = 1, \ldots, s. \tag{1.3}$$

Dans cette condition, le chapeau au-dessus de V_i signifie que V_i est omis de la somme.

Preuve. Supposons d'abord que la condition (1.3) soit remplie. Si $\mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_s \in V_s$ satisfont $\mathbf{v}_1 + \cdots + \mathbf{v}_s = 0$, alors, pour $i = 1, \ldots, s$, on a

$$-\mathbf{v}_i = \mathbf{v}_1 + \cdots + \widehat{\mathbf{v}_i} + \cdots + \mathbf{v}_s \in V_i \cap \big(V_1 + \cdots + \widehat{V_i} + \cdots + V_s\big)$$

donc $-\mathbf{v}_i = \mathbf{0}$ et par suite $\mathbf{v}_i = \mathbf{0}$. Donc la somme $V_1 + \cdots + V_s$ est directe.

Réciproquement, supposons que cette somme soit directe, et fixons $i \in \{1, \ldots, s\}$. Si

$$\mathbf{v} \in V_i \cap \big(V_1 + \cdots + \widehat{V_i} + \cdots + V_s\big),$$

alors il existe $\mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_s \in V_s$ tels que

$$\mathbf{v} = -\mathbf{v}_i \quad \text{et} \quad \mathbf{v} = \mathbf{v}_1 + \cdots + \widehat{\mathbf{v}_i} + \cdots + \mathbf{v}_s.$$

En éliminant $\mathbf{v}$, on obtient que $-\mathbf{v}_i = \mathbf{v}_1 + \cdots + \widehat{\mathbf{v}_i} + \cdots + \mathbf{v}_s$, donc $\mathbf{v}_1 + \cdots + \mathbf{v}_s = \mathbf{0}$ et par suite $\mathbf{v}_1 = \cdots = \mathbf{v}_s = \mathbf{0}$. En particulier, on trouve $\mathbf{v} = 0$. Le choix de $\mathbf{v}$ étant arbitraire, cela montre que la condition (1.3) est satisfaite pour ce choix de i. □

En particulier, pour $s = 2$, la proposition 1.4.5 implique que $V = V_1 \oplus V_2$ si et seulement si $V = V_1 + V_2$ et $V_1 \cap V_2 = \{\mathbf{0}\}$. En appliquant la formule de la proposition 1.3.9, on en déduit :

Proposition 1.4.6. *Soient V_1 et V_2 des sous-espaces d'un espace vectoriel V de dimension finie. On a $V = V_1 \oplus V_2$ si et seulement si*

(i) $\dim_K V = \dim_K V_1 + \dim_K V_2$

(ii) $V_1 \cap V_2 = \{\mathbf{0}\}$.

Exemple 1.4.7. Soient $V = \mathbb{R}^3$,

$$V_1 = \left\langle \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\rangle_{\mathbb{R}} = \left\{ \begin{pmatrix} 2a \\ 0 \\ a \end{pmatrix} ; a \in \mathbb{R} \right\} \quad \text{et} \quad V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 ; x + y + z = 0 \right\}.$$

La proposition 1.3.2 montre que V_1 est un sous-espace de $\mathbb{R}^3$. Comme $\left\{\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}\right\}$ est une base de V_1, on a $\dim_{\mathbb{R}}(V_1) = 1$. Par ailleurs, comme V_2 est l'ensemble des solutions d'un système d'équations linéaires homogènes de rang 1 (avec une seule équation), V_2 est un sous-espace de $\mathbb{R}^3$ de dimension $\dim_{\mathbb{R}}(V_2) = 3 - (\text{rang du système}) = 2$. On trouve aussi que

$$\begin{pmatrix} 2a \\ 0 \\ a \end{pmatrix} \in V_2 \Longleftrightarrow 2a + 0 + a = 0 \Longleftrightarrow a = 0,$$

donc $V_1 \cap V_2 = \{\mathbf{0}\}$. Puisque $\dim_{\mathbb{R}} V_1 + \dim_{\mathbb{R}} V_2 = 3 = \dim_{\mathbb{R}} \mathbb{R}^3$, la proposition 1.4.6 montre que $\mathbb{R}^3 = V_1 \oplus V_2$.

Interprétation géométrique : L'ensemble V_1 représente une droite passant par l'origine dans $\mathbb{R}^3$ tandis que V_2 représente un plan passant par l'origine de vecteur normal $\mathbf{n} := (1, 1, 1)^t$. La droite V_1 n'est pas parallèle au plan V_2 car son vecteur directeur $(2, 0, 1)^t$ n'est pas perpendiculaire à $\mathbf{n}$. Dans cette situation, on voit que tout vecteur de l'espace se décompose de manière unique comme somme d'un vecteur de la droite et d'un vecteur du plan.

Le résultat suivant joue un rôle capital dans toute la suite.

Proposition 1.4.8. *Supposons que $V = V_1 \oplus \cdots \oplus V_s$ soit de dimension finie et soit $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \mathbf{v}_{i,2}, \ldots, \mathbf{v}_{i,n_i}\}$ une base de V_i pour $i = 1, \ldots, s$. Alors l'ensemble ordonné*

$$\mathcal{B} = \{\mathbf{v}_{1,1}, \ldots, \mathbf{v}_{1,n_1}, \mathbf{v}_{2,1}, \ldots, \mathbf{v}_{2,n_2}, \ldots\ldots, \mathbf{v}_{s,1}, \ldots, \mathbf{v}_{s,n_s}\}$$

obtenu par concaténation de $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_s$ est une base de V.

Dans la suite, on notera cette base sous la forme $\mathcal{B}_1 \amalg \mathcal{B}_2 \amalg \cdots \amalg \mathcal{B}_s$. Le symbole $\amalg$ représente "l'union disjointe".

Preuve. **1° On montre que $\mathcal{B}$ est un système de générateurs de V.**

Soit $\mathbf{v} \in V$. Puisque $V = V_1 \oplus \cdots \oplus V_s$, on peut écrire $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_s$ avec $\mathbf{v}_i \in V_i$ pour $i = 1, \ldots, s$. Puisque $\mathcal{B}_i$ est une base de V_i, on peut aussi écrire

$$\mathbf{v}_i = a_{i,1}\mathbf{v}_{i,1} + \cdots + a_{i,n_i}\mathbf{v}_{i,n_i}$$

avec $a_{i,1}, \ldots, a_{i,n_i} \in K$. On en déduit que

$$\mathbf{v} = a_{1,1}\mathbf{v}_{1,1} + \cdots + a_{1,n_1}\mathbf{v}_{1,n_i} + \cdots\cdots + a_{s,1}\mathbf{v}_{s,1} + \cdots + a_{s,n_s}\mathbf{v}_{s,n_s}.$$

Donc $\mathcal{B}$ engendre V.

2° On montre que $\mathcal{B}$ est linéairement indépendant.

Supposons que

$$a_{1,1}\mathbf{v}_{1,1} + \cdots + a_{1,n_1}\mathbf{v}_{1,n_i} + \cdots + a_{s,1}\mathbf{v}_{s,1} + \cdots + a_{s,n_s}\mathbf{v}_{s,n_s} = \mathbf{0}$$

pour des éléments $a_{1,1}, \ldots, a_{s,n_s} \in K$. En posant $\mathbf{v}_i = a_{i,1}\mathbf{v}_{i,1} + \cdots + a_{i,n_i}\mathbf{v}_{i,n_i}$ pour $i = 1, \ldots, s$, on obtient $\mathbf{v}_1 + \cdots + \mathbf{v}_s = \mathbf{0}$. Comme $\mathbf{v}_i \in V_i$ pour $i = 1, \ldots, s$ et que V est la somme directe de $V_1, \ldots, V_s$, cela implique $\mathbf{v}_i = \mathbf{0}$ pour $i = 1, \ldots, s$. On en déduit que $a_{i,1} = \cdots = a_{i,n_i} = 0$, car $\mathcal{B}_i$ est une base de V_i. Ainsi, on a $a_{1,1} = \cdots = a_{s,n_s} = 0$. Cela montre que $\mathcal{B}$ est linéairement indépendant. □

Corollaire 1.4.9. *Si $V = V_1 \oplus \cdots \oplus V_s$ est de dimension finie, alors*

$$\dim_K V = \dim_K(V_1) + \cdots + \dim_K(V_s).$$

Preuve. En utilisant la proposition 1.4.8 en conservant les mêmes notations, on a

$$\dim_K V = |\mathcal{B}| = \sum_{i=1}^{s} |\mathcal{B}_i| = \sum_{i=1}^{s} \dim_K(V_i).$$

□

Exemple 1.4.10. Dans l'exemple 1.4.7, on a $\mathbb{R}^3 = V_1 \oplus V_2$. On trouve que

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \text{et} \quad \mathcal{B}_2 = \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

sont respectivement des bases de V_1 et de V_2 (en effet, $\mathcal{B}_2$ est un sous-ensemble linéairement indépendant de V_2 formé de 2 éléments et $\dim_{\mathbb{R}}(V_2) = 2$, donc c'est une base de V_2). On en déduit que

$$\mathcal{B} = \mathcal{B}_1 \amalg \mathcal{B}_2 = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$$

est une base de $\mathbb{R}^3$.

Exercices.

1.4.1. Démontrer la proposition 1.4.4.

1.4.2. Soient $V_1, \ldots, V_s$ des sous-espaces d'un espace vectoriel V sur un corps K. Montrer que leur somme est directe si et seulement si

$$V_i \cap (V_{i+1} + \cdots + V_s) = \{\mathbf{0}\}$$

pour $i = 1, \ldots, s-1$.

1.4.3. Soient $V_1, \ldots, V_s$ des sous-espaces d'un espace vectoriel V sur un corps K. Supposons que leur somme soit directe.

(i) Montrer que, si U_i est un sous-espace de V_i pour $i = 1, \ldots, s$, alors la somme $U_1 + \cdots + U_s$ est directe.

(ii) Montrer que, si $\mathbf{v}_i$ est un élément non nul de V_i pour $i = 1, \ldots, s$, alors $\mathbf{v}_1, \ldots, \mathbf{v}_s$ sont linéairement indépendants sur K.

1.4.4. On considère les sous-espaces de $\mathbb{R}^4$ donnés par

$$V_1 = \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle_{\mathbb{R}} \quad \text{et} \quad V_2 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix} \right\rangle_{\mathbb{R}}.$$

Montrer que $\mathbb{R}^4 = V_1 \oplus V_2$.

1.4.5. Soient

$$\begin{aligned} V_1 &= \{\, f \colon \mathbb{Z} \to \mathbb{R} \,;\ \ f(-x) = f(x) \text{ pour tout } x \in \mathbb{Z} \,\} \\ \text{et} \quad V_2 &= \{\, f \colon \mathbb{Z} \to \mathbb{R} \,;\ \ f(-x) = -f(x) \text{ pour tout } x \in \mathbb{Z} \,\}. \end{aligned}$$

Montrer que V_1 et V_2 sont deux sous-espaces de $\mathcal{F}(\mathbb{Z}, \mathbb{R})$ et que $\mathcal{F}(\mathbb{Z}, \mathbb{R}) = V_1 \oplus V_2$.

1.4.6. Soient $V_1, V_2, \ldots, V_s$ des espaces vectoriels sur un même corps K. Montrer que

$$\begin{aligned} &V_1 \times V_2 \times \cdots \times V_s \\ &\quad = \big(V_1 \times \{\mathbf{0}\} \times \cdots \times \{\mathbf{0}\}\big) \oplus \big(\{\mathbf{0}\} \times V_2 \times \cdots \times \{\mathbf{0}\}\big) \oplus \cdots \oplus \big(\{\mathbf{0}\} \times \cdots \times \{\mathbf{0}\} \times V_s\big). \end{aligned}$$

Chapitre 2

Retour sur les applications linéaires

Le corps K des scalaires étant fixé, une application linéaire d'un espace vectoriel V dans un espace vectoriel W est une fonction qui "commute" à l'addition et la multiplication par les scalaires de K. Ce second chapitre de préliminaires passe en revue les notions d'application linéaire, noyau et image, et les opérations qu'on peut effectuer sur les applications linéaires (addition, multiplication par un scalaire et composition). On revoit ensuite les matrices des applications linéaires relatives à un choix de bases, comment elles se comportent face aux opérations mentionnées ci-dessus et aussi face à un changement de bases. On examine le cas particulier des opérateurs linéaires sur un espace vectoriel V, c'est-à-dire les applications linéaires de V dans lui-même. On prêtera spécialement attention à la notion de sous-espace invariant sous un opérateur linéaire qui joue un rôle fondamental dans la suite.

2.1 La notion d'application linéaire

Définition 2.1.1. Soient V et W des espaces vectoriels sur K. Une *application linéaire* de V dans W est une fonction $T\colon V \to W$ qui satisfait

AL1. $T(\mathbf{v}+\mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}')$ pour tout choix de $\mathbf{v}, \mathbf{v}' \in V$,

AL2. $T(c\mathbf{v}) = cT(\mathbf{v})$ pour tout $c \in K$ et tout $\mathbf{v} \in V$.

On rappelle les résultats suivants :

Proposition 2.1.2. *Soit $T\colon V \to W$ une application linéaire.*

(i) L'ensemble $\ker(T) := \{\mathbf{v}\,;\, T(\mathbf{v}) = \mathbf{0}\}$, *appelé* noyau *de T, est un sous-espace de V.*

(ii) L'ensemble $\mathrm{Im}(T) := \{T(\mathbf{v})\,;\, \mathbf{v} \in V\}$, *appelé* image *de T, est un sous-espace de W.*

(iii) L'application linéaire T est injective si et seulement si $\ker(T) = \{\mathbf{0}\}$. *Elle est surjective si et seulement si* $\mathrm{Im}(T) = W$.

(iv) Si V est de dimension finie, on a

$$\dim_K V = \dim_K(\ker(T)) + \dim_K(\mathrm{Im}(T)).$$

Exemple 2.1.3. **(Application linéaire associée à une matrice.)** Soit $M \in \mathrm{Mat}_{m\times n}(K)$. L'application

$$\begin{array}{rccc} T_M : & K^n & \longrightarrow & K^m \\ & X & \longmapsto & MX \end{array}$$

est linéaire (cela découle immédiatement des propriétés du produit matriciel). Son noyau est

$$\ker(T_M) := \{X \in K^n \,;\, MX = \mathbf{0}\}$$

c'est donc l'ensemble-solution du système homogène $MX = 0$. Si $C_1, \ldots, C_n$ désignent les colonnes de M, on trouve

$$\begin{aligned} \mathrm{Im}(T_M) &= \{MX \,;\, X \in K^n\} \\ &= \left\{(C_1 \cdots C_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;;\; x_1, \ldots, x_n \in K\right\} \\ &= \{x_1C_1 + \cdots + x_nC_n \,;\; x_1, \ldots, x_n \in K\} \\ &= \langle C_1, \ldots, C_n\rangle_K. \end{aligned}$$

Donc son image est l'espace-colonne de M. Comme la dimension de ce dernier est le rang de M, la formule (iv) de la proposition 2.1.2 redonne le résultat connu :

$$\dim\{X \in K^n \,;\, MX = \mathbf{0}\} = n - \mathrm{rang}(M).$$

Exemple 2.1.4. Soient V_1 et V_2 des espaces vectoriels sur K. La fonction

$$\begin{array}{rccc} \pi_1 : & V_1 \times V_2 & \longrightarrow & V_1 \\ & (\mathbf{v}_1, \mathbf{v}_2) & \longmapsto & \mathbf{v}_1 \end{array}$$

est une application linéaire car si $(\mathbf{v}_1, \mathbf{v}_2)$ et $(\mathbf{v}'_1, \mathbf{v}'_2)$ sont des éléments de $V_1 \times V_2$ et si $c \in K$, on trouve

$$\begin{aligned} \pi_1((\mathbf{v}_1, \mathbf{v}_2) + (\mathbf{v}'_1, \mathbf{v}'_2)) &= \pi_1(\mathbf{v}_1 + \mathbf{v}'_1, \mathbf{v}_2 + \mathbf{v}'_2) \\ &= \mathbf{v}_1 + \mathbf{v}'_1 \\ &= \pi_1(\mathbf{v}_1, \mathbf{v}_2) + \pi_1(\mathbf{v}'_1, \mathbf{v}'_2) \end{aligned}$$

et

$$\pi_1(c\,(\mathbf{v}_1, \mathbf{v}_2)) = \pi_1(c\,\mathbf{v}_1, c\,\mathbf{v}_2) = c\,\mathbf{v}_1 = c\,\pi_1(\mathbf{v}_1, \mathbf{v}_2).$$

On trouve aussi

$$\ker(\pi_1) = \{(\mathbf{v}_1, \mathbf{v}_2) \in V_1 \times V_2 \,;\, \mathbf{v}_1 = 0\} = \{(0, \mathbf{v}_2) \,;\, \mathbf{v}_2 \in V_2\} = \{\mathbf{0}\} \times V_2,$$

et $\mathrm{Im}(\pi_1) = V_1$ car $\mathbf{v}_1 = \pi_1(\mathbf{v}_1, \mathbf{0}) \in \mathrm{Im}(\pi_1)$ pour tout $\mathbf{v}_1 \in V_1$. Ainsi π_1 est surjective.

L'application $\pi_1 \colon V_1 \times V_2 \to V_1$ s'appelle la *projection sur le premier facteur*. On montre de même que l'application

$$\begin{array}{rccc} \pi_2 : & V_1 \times V_2 & \longrightarrow & V_2 \\ & (\mathbf{v}_1, \mathbf{v}_2) & \longmapsto & \mathbf{v}_2 \end{array}$$

appelée *projection sur le second facteur* est linéaire et surjective de noyau $V_1 \times \{\mathbf{0}\}$. De manière semblable, on montre encore que les applications

$$\begin{array}{rccc} i_1 : & V_1 & \longrightarrow & V_1 \times V_2 \\ & \mathbf{v}_1 & \longmapsto & (\mathbf{v}_1, \mathbf{0}) \end{array} \quad \text{et} \quad \begin{array}{rccc} i_2 : & V_2 & \longrightarrow & V_1 \times V_2 \\ & \mathbf{v}_2 & \longmapsto & (\mathbf{0}, \mathbf{v}_2) \end{array}$$

sont linéaires et injectives, d'images respectives $V_1 \times \{\mathbf{0}\}$ et $\{\mathbf{0}\} \times V_2$.

On conclut cette section avec le résultat suivant, crucial pour la suite, dont on verra au chapitre 9 qu'il se généralise aux applications dites "bilinéaires" (voir aussi l'appendice B pour le cas encore plus général des applications multilinéaires).

Théorème 2.1.5. *Soient V et W des espaces vectoriels sur K, avec $\dim_K(V) = n$. Soient $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, et $\mathbf{w}_1, \ldots, \mathbf{w}_n$ une suite de n éléments quelconques de W (pas nécessairement distincts). Alors, il existe une et une seule application linéaire $T\colon V \to W$ qui satisfasse $T(\mathbf{v}_i) = \mathbf{w}_i$ pour $i = 1, \ldots, n$.*

Preuve. **1° Unicité.** S'il existe une telle application linéaire T, alors

$$T\Big(\sum_{i=1}^{n} a_i \mathbf{v}_i\Big) = \sum_{i=1}^{n} a_i T(\mathbf{v}_i) = \sum_{i=1}^{n} a_i \mathbf{w}_i$$

pour tout choix de $a_1, \ldots, a_n \in K$. Cela détermine T de manière unique.

2° Existence. Considérons la fonction $T\colon V \to W$ définie par la formule

$$T\Big(\sum_{i=1}^{n} a_i \mathbf{v}_i\Big) = \sum_{i=1}^{n} a_i \mathbf{w}_i.$$

Elle satisfait $T(\mathbf{v}_i) = \mathbf{w}_i$ pour $i = 1, \ldots, n$. Il reste à montrer qu'elle est linéaire. On trouve

$$\begin{aligned} T\Big(\sum_{i=1}^{n} a_i \mathbf{v}_i + \sum_{j=1}^{n} a'_i \mathbf{v}_i\Big) &= T\Big(\sum_{i=1}^{n} (a_i + a'_i)\mathbf{v}_i\Big) \\ &= \sum_{i=1}^{n} (a_i + a'_i)\mathbf{w}_i \\ &= \sum_{i=1}^{n} a_i \mathbf{w}_i + \sum_{i=1}^{n} a'_i \mathbf{w}_i \\ &= T\Big(\sum_{i=1}^{n} a_i \mathbf{v}_i\Big) + T\Big(\sum_{j=1}^{n} a'_i \mathbf{v}_i\Big), \end{aligned}$$

pour tout choix de $a_1, \ldots, a_n, a'_1, \ldots, a'_n \in K$. On vérifie aussi que

$$T\Big(c \sum_{i=1}^{n} a_i \mathbf{v}_i\Big) = c\, T\Big(\sum_{i=1}^{n} a_i \mathbf{v}_i\Big),$$

pour tout $c \in K$. Donc T est bien une application linéaire. □

2.2 L'espace vectoriel $\mathcal{L}_K(V,W)$

Soient V et W des espaces vectoriels sur K. On désigne par $\mathcal{L}_K(V,W)$ l'ensemble des applications linéaires de V dans W. On rappelle d'abord que cet ensemble est naturellement muni d'une addition et d'une multiplication par les éléments de K.

Proposition 2.2.1. *Soient $T_1 : V \to W$ et $T_2 : V \to W$ des applications linéaires, et soit $c \in K$.*

(i) La fonction $T_1 + T_2 : V \to W$ donnée par

$$(T_1 + T_2)(\mathbf{v}) = T_1(\mathbf{v}) + T_2(\mathbf{v}) \quad \text{pour tout } \mathbf{v} \in V$$

est une application linéaire.

(ii) La fonction $cT_1 : V \to W$ donnée par

$$(cT_1)(\mathbf{v}) = cT_1(\mathbf{v}) \quad \text{pour tout } \mathbf{v} \in V$$

est aussi une application linéaire.

Preuve de (i). Pour tout choix de $\mathbf{v}_1, \mathbf{v}_2 \in V$ et de $a \in K$, on trouve en effet :

$$\begin{aligned}
(T_1+T_2)(\mathbf{v}_1+\mathbf{v}_2) &= T_1(\mathbf{v}_1+\mathbf{v}_2) + T_2(\mathbf{v}_1+\mathbf{v}_2) && \text{(définition de } T_1+T_2\text{)}\\
&= (T_1(\mathbf{v}_1)+T_1(\mathbf{v}_2)) + (T_2(\mathbf{v}_1)+T_2(\mathbf{v}_2)) && \text{(car } T_1 \text{ et } T_2 \text{ sont linéaires)}\\
&= (T_1(\mathbf{v}_1)+T_2(\mathbf{v}_1)) + (T_1(\mathbf{v}_2)+T_2(\mathbf{v}_2))\\
&= (T_1+T_2)(\mathbf{v}_1) + (T_1+T_2)(\mathbf{v}_2), && \text{(définition de } T_1+T_2\text{)}\\
(T_1+T_2)(a\,\mathbf{v}_1) &= T_1(a\,\mathbf{v}_1) + T_2(a\,\mathbf{v}_1) && \text{(définition de } T_1+T_2\text{)}\\
&= a\,T_1(\mathbf{v}_1) + a\,T_2(\mathbf{v}_1) && \text{(car } T_1 \text{ et } T_2 \text{ sont linéaires)}\\
&= a\,(T_1(\mathbf{v}_1)+T_2(\mathbf{v}_1))\\
&= a\,(T_1+T_2)(\mathbf{v}_1). && \text{(définition de } T_1+T_2\text{)}
\end{aligned}$$

Donc $T_1 + T_2$ est linéaire. La preuve que cT_1 est linéaire est semblable. □

Exemple 2.2.2. Soit V un espace vectoriel sur $\mathbb{Q}$, et soient $\pi_1 : V \times V \to V$ et $\pi_2 : V \times V \to V$ les applications linéaires définies, comme à l'exemple 2.1.4, par

$$\pi_1(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1 \quad \text{et} \quad \pi_2(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_2$$

pour tout $(\mathbf{v}_1, \mathbf{v}_2) \in V \times V$. Alors

1) $\pi_1 + \pi_2 : V \times V \to V$ est l'application linéaire donnée par

$$(\pi_1 + \pi_2)(\mathbf{v}_1, \mathbf{v}_2) = \pi_1(\mathbf{v}_1, \mathbf{v}_2) + \pi_2(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1 + \mathbf{v}_2,$$

2) $3\pi_1 : V \times V \to V$ est l'application linéaire donnée par

$$(3\pi_1)(\mathbf{v}_1, \mathbf{v}_2) = 3\pi_1(\mathbf{v}_1, \mathbf{v}_2) = 3\mathbf{v}_1.$$

En combinant addition et multiplication par un scalaire, on peut par exemple former

$$\begin{array}{rccc} 3\pi_1 + 5\pi_5 : & V \times V & \longrightarrow & V \\ & (\mathbf{v}_1, \mathbf{v}_2) & \longmapsto & 3\mathbf{v}_1 + 5\mathbf{v}_2. \end{array}$$

Théorème 2.2.3. *L'ensemble $\mathcal{L}_K(V,W)$ est un espace vectoriel sur K pour l'addition et la multiplication par un scalaire définies dans la proposition 2.2.1.*

Preuve. Il suffit de vérifier que ces opérations satisfont les 8 axiomes requis pour faire de $\mathcal{L}_K(V,W)$ un espace vectoriel sur K.

On note d'abord que la fonction $\mathcal{O}\colon V \to W$ donnée par

$$\mathcal{O}(\mathbf{v}) = \mathbf{0} \quad \text{pour tout } \mathbf{v} \in V$$

est une application linéaire (donc un élément de $\mathcal{L}_K(V,W)$). On note aussi que, si $T \in \mathcal{L}_K(V,W)$, alors la fonction $-T\colon V \to W$ définie par

$$(-T)(\mathbf{v}) = -T(\mathbf{v}) \quad \text{pour tout } \mathbf{v} \in V$$

est aussi une application linéaire (c'est $(-1)T$). Il suffit donc de montrer que, pour tout choix de $a, b \in K$ et de $T_1, T_2, T_3 \in \mathcal{L}_K(V,W)$, on a

1) $T_1 + (T_2 + T_3) = (T_1 + T_2) + T_3$
2) $T_1 + T_2 = T_2 + T_1$
3) $\mathcal{O} + T_1 = T_1$
4) $T_1 + (-T_1) = \mathcal{O}$
5) $1 \cdot T_1 = T_1$
6) $a(b\,T_1) = (ab)\,T_1$
7) $(a+b)\,T_1 = a\,T_1 + b\,T_1$
8) $a\,(T_1 + T_2) = a\,T_1 + b\,T_2$.

Montrons par exemple la condition 8). Pour établir que deux éléments de $\mathcal{L}_K(V,W)$ sont égaux, il faut montrer qu'ils prennent les mêmes valeurs en chaque point de leur domaine V. Dans la cas de 8), cela revient à vérifier que $(a\,(T_1 + T_2))(\mathbf{v}) = (a\,T_1 + a\,T_2)(\mathbf{v})$ pour tout $\mathbf{v} \in V$.

Soit $\mathbf{v} \in V$. On trouve

$$\begin{aligned} (a\,(T_1 + T_2))(\mathbf{v}) &= a(T_1 + T_2)(\mathbf{v}) \\ &= a(T_1(\mathbf{v}) + T_2(\mathbf{v})) \\ &= aT_1(\mathbf{v}) + aT_2(\mathbf{v}) \\ &= (aT_1)(\mathbf{v}) + (aT_2)(\mathbf{v}) \\ &= (aT_1 + aT_2)(\mathbf{v}). \end{aligned}$$

Donc, on a bien $a(T_1 + T_2) = a\,T_1 + a\,T_2$. □

Exemple 2.2.4. On sait que K est un espace vectoriel sur K. Donc, pour tout espace vectoriel V, l'ensemble $\mathcal{L}_K(V,K)$ des applications linéaires de V dans K est un espace vectoriel sur K. On l'appelle le *dual* de V et on le note V^*. Un élément de V^* s'appelle une *forme linéaire* sur V.

Exercices.

2.2.1. Compléter la preuve de la proposition 2.2.1 en montrant que cT_1 est une application linéaire.

2.2.2. Compléter la preuve du théorème 2.2.3 en montrant que les conditions 3), 6) et 7) sont remplies.

2.2.3. Soit n un entier positif. Montrer que, pour $i = 1, \ldots, n$, la fonction

$$\begin{array}{rccc} f_i : & K^n & \longrightarrow & K \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & x_i \end{array}$$

est un élément du dual $(K^n)^*$ de K, et que $\{f_1, \ldots, f_n\}$ est une base de $(K^n)^*$.

2.3 Composition

Soient U, V et W des espaces vectoriels sur un corps K. On rappelle que si $S\colon U \to V$ et $T\colon V \to W$ sont des applications linéaires, alors leur composée $T \circ S\colon U \to W$ donnée par

$$(T \circ S)(\mathbf{u}) = T(S(\mathbf{u})) \quad \text{pour tout } \mathbf{u} \in U$$

est aussi une application linéaire. Dans la suite, on utilisera beaucoup le fait que la composition est distributive sur l'addition et qu'elle "commute" à la multiplication par un scalaire de K comme le montre la proposition suivante.

Proposition 2.3.1. *Soient $S, S_1, S_2 \in \mathcal{L}_K(U,V)$ et $T, T_1, T_2 \in \mathcal{L}_K(V,W)$, et soit $c \in K$. On a*

(i) $(T_1 + T_2) \circ S = T_1 \circ S + T_2 \circ S$,

(ii) $T \circ (S_1 + S_2) = T \circ S_1 + T \circ S_2$,

(iii) $(cT) \circ S = T \circ (cS) = cT \circ S$.

Preuve de (i). Pour tout $\mathbf{u} \in U$, on a

$$\begin{aligned}((T_1 + T_2) \circ S)(\mathbf{u}) &= (T_1 + T_2)(S(\mathbf{u})) \\ &= T_1(S(\mathbf{u})) + T_2(S(\mathbf{u})) \\ &= (T_1 \circ S)(\mathbf{u}) + (T_2 \circ S)(\mathbf{u}) \\ &= (T_1 \circ S + T_2 \circ S)(\mathbf{u})\,.\end{aligned}$$

Les preuves de *(ii)* et de *(iii)* sont semblables. □

Soit $T\colon V \to W$ une application linéaire. On rappelle aussi que, si T est bijective, alors la fonction inverse $T^{-1}\colon W \to V$ caractérisée par

$$T^{-1}(\mathbf{w}) = \mathbf{u} \quad \Longleftrightarrow \quad T(\mathbf{u}) = \mathbf{w}$$

est aussi une application linéaire. Elle satisfait

$$T^{-1} \circ T = I_V \quad \text{et} \quad T \circ T^{-1} = I_W$$

où $I_V : \begin{array}{ccc} V & \to & V \\ \mathbf{v} & \mapsto & \mathbf{v} \end{array}$ et $I_W : \begin{array}{ccc} W & \to & W \\ \mathbf{w} & \mapsto & \mathbf{w} \end{array}$ désignent respectivement les applications identités de V et de W.

Définition 2.3.2. On dit qu'une application linéaire $T\colon V \to W$ est *inversible* s'il existe une application linéaire $S\colon W \to V$ telle que

$$S \circ T = I_V \quad \text{et} \quad T \circ S = I_W\,. \tag{2.1}$$

Les observations précédentes montrent que si T est bijective, alors elle est inversible. Par ailleurs, si T est inversible et si $S\colon W \to V$ satisfait (2.1), alors T est bijective et $S = T^{-1}$. Une application linéaire inversible n'est donc rien d'autre qu'une application linéaire bijective. On rappelle le fait suivant :

Proposition 2.3.3. *Supposons que $S\colon U \to V$ et $T\colon V \to W$ soient des applications linéaires inversibles. Alors U, V, W ont même dimension (finie ou infinie). De plus*

(i) $T^{-1}\colon W \to V$ est inversible et $(T^{-1})^{-1} = T$,

(ii) $T \circ S : U \to W$ est inversible et $(T \circ S)^{-1} = S^{-1} \circ T^{-1}$.

Une application linéaire inversible est aussi appelé *isomorphisme*. On dit que V est *isomorphe* à W et on écrit $V \simeq W$, s'il existe un isomorphisme de V dans W. On vérifie que

(i) $V \simeq V$,

(ii) $V \simeq W \iff W \simeq V$,

(iii) si $U \simeq V$ et $V \simeq W$, alors $U \simeq W$.

Il s'agit donc d'une relation d'équivalence sur la classe des espaces vectoriels sur K.

Le résultat suivant montre que, pour chaque entier $n \geq 1$, tout espace vectoriel de dimension n sur K est isomorphe à K^n.

Proposition 2.3.4. *Soit $n \in \mathbb{N}^*$, soit V un espace vectoriel sur K de dimension n, et soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V. L'application*

$$\begin{array}{rccc} \varphi : & V & \longrightarrow & K^n \\ & \mathbf{v} & \longmapsto & [\mathbf{v}]_{\mathcal{B}} \end{array}$$

est un isomorphisme d'espaces vectoriels sur K.

Preuve. Pour le prouver, on montre d'abord que φ est linéaire (exercice). Le fait que φ est bijective découle directement de la proposition 1.3.6. □

Exercices.

2.3.1. Démontrer la partie (iii) de la proposition 2.3.1.

2.3.2. Soient V et W des espaces vectoriels sur un corps K et soit $T\colon V \to W$ une application linéaire. Montrer que la fonction $T^*\colon W^* \to V^*$ donnée par $T^*(f) = f \circ T$ pour tout $f \in W^*$ est une application linéaire (voir l'exemple 2.2.4 pour les définitions de V^* et W^*). On dit que T^* est *l'application linéaire duale de T*.

2.4 Matrices des applications linéaires

Dans cette section, U, V et W dénotent des espaces vectoriels sur un même corps K.

Proposition 2.4.1. *Soit $T\colon V \to W$ une application linéaire. Supposons que $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ soit une base de V et que $\mathcal{D} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ soit une base de W. Alors, il existe une et une seule matrice de $\mathrm{Mat}_{m\times n}(K)$, notée $[T]_{\mathcal{D}}^{\mathcal{B}}$, telle que*

$$[T(\mathbf{v})]_{\mathcal{D}} = [T]_{\mathcal{D}}^{\mathcal{B}}\,[\mathbf{v}]_{\mathcal{B}} \quad \textit{pour tout} \quad \mathbf{v} \in V.$$

Cette matrice est donnée par

$$[T]_{\mathcal{D}}^{\mathcal{B}} = \Big([T(\mathbf{v}_1]_{\mathcal{D}} \cdots [T(\mathbf{v}_n]_{\mathcal{D}}\Big),$$

c'est-à-dire que la j-ième colonne de $[T]_{\mathcal{D}}^{\mathcal{B}}$ est $[T(\mathbf{v}_j)]_{\mathcal{D}}$.

Essayer d'en retrouver la preuve sans regarder l'argument donné ci-dessous. C'est un bon exercice !

Preuve. **1° Existence :** Soit $\mathbf{v} \in V$. Écrivons

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Par définition, cela signifie que $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$, donc

$$T(\mathbf{v}) = a_1T(\mathbf{v}_1) + \cdots + a_nT(\mathbf{v}_n).$$

Comme l'application $\begin{array}{ccc} W & \to & K^n \\ \mathbf{w} & \mapsto & [\mathbf{w}]_{\mathcal{D}} \end{array}$ est linéaire (voir la proposition 2.3.4), on en déduit que

$$\begin{aligned} [T(\mathbf{v})]_{\mathcal{D}} &= a_1[T(\mathbf{v}_1)]_{\mathcal{D}} + \cdots + a_n[T(\mathbf{v}_n)]_{\mathcal{D}} \\ &= \Big([T(\mathbf{v}_1)]_{\mathcal{D}} \cdots [T(\mathbf{v}_n)]_{\mathcal{D}}\Big) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \\ &= [\,T\,]_{\mathcal{D}}^{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}. \end{aligned}$$

2° Unicité : Si une matrice $M \in \mathrm{Mat}_{m\times n}(K)$ satisfait $[T(\mathbf{v})]_{\mathcal{D}} = M[\mathbf{v}]_{\mathcal{B}}$ pour tout $\mathbf{v} \in V$, alors, pour $j = 1, \ldots, n$, on obtient

$$\begin{aligned} [T(\mathbf{v}_j)]_{\mathcal{D}} = M[\mathbf{v}_j]_{\mathcal{B}} = M \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} &\leftarrow j\text{-ième ligne} \\ = j\text{-ième colonne de } M. & \end{aligned}$$

Donc $M = [\,T\,]_{\mathcal{D}}^{\mathcal{B}}$. □

Exemple 2.4.2. Soit $M \in \mathrm{Mat}_{m\times n}(K)$ et soit

$$\begin{array}{rccc} T_M : & K^n & \longrightarrow & K^n \\ & X & \longmapsto & M\,X \end{array}$$

l'application linéaire associée à la matrice M (voir l'exemple 2.1.3). Soient $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base canonique de K^n et $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ celle de K^m. Pour tout $X \in K^n$ et tout $Y \in K^m$, on a

$$[X]_{\mathcal{E}} = X \quad \text{et} \quad [Y]_{\mathcal{F}} = Y$$

(voir l'exemple 1.3.7), donc

$$[T_M(X)]_{\mathcal{F}} = [MX]_{\mathcal{F}} = MX = M[X]_{\mathcal{E}}.$$

On en déduit que

$$[T_M]^{\mathcal{E}}_{\mathcal{F}} = M,$$

c'est-à-dire que M est la matrice de T_M relative aux bases canoniques de K^n et de K^m.

Théorème 2.4.3. *Avec les notations de la proposition 2.4.1, l'application*

$$\begin{array}{ccc} \mathcal{L}_K(V,W) & \longrightarrow & \mathrm{Mat}_{m\times n}(K) \\ T & \longmapsto & [\,T\,]^{\mathcal{B}}_{\mathcal{D}} \end{array}$$

est un isomorphisme d'espaces vectoriels sur K.

Preuve. **1°** On montre que cette application est linéaire. Pour cela, on choisit $T_1, T_2 \in \mathcal{L}_K(V,W)$ et $c \in K$ quelconques. On doit vérifier que

1) $[T_1 + T_2]^{\mathcal{B}}_{\mathcal{D}} = [T_1]^{\mathcal{B}}_{\mathcal{D}} + [T_2]^{\mathcal{B}}_{\mathcal{D}}$,
2) $[cT_1]^{\mathcal{B}}_{\mathcal{D}} = c\,[T_1]^{\mathcal{B}}_{\mathcal{D}}$.

Pour démontrer 1), on note que, pour tout $\mathbf{v} \in V$, on a

$$\begin{aligned} [(T_1 + T_2)(\mathbf{v})]_{\mathcal{D}} &= [T_1(\mathbf{v}) + T_2(\mathbf{v})]_{\mathcal{D}} \\ &= [T_1(\mathbf{v})]_{\mathcal{D}} + [T_2(\mathbf{v})]_{\mathcal{D}} \\ &= [T_1]^{\mathcal{B}}_{\mathcal{D}}[\mathbf{v}]_{\mathcal{B}} + [T_1]^{\mathcal{B}}_{\mathcal{D}}[\mathbf{v}]_{\mathcal{B}} \\ &= ([T_1]^{\mathcal{B}}_{\mathcal{D}} + [T_2]^{\mathcal{B}}_{\mathcal{D}})[\mathbf{v}]_{\mathcal{B}}. \end{aligned}$$

Alors 1) suit en vertu de l'unicité de la matrice associée à $T_1 + T_2$. La preuve de 2) est semblable.

2° Il reste à montrer que l'application est bijective. La proposition 2.4.1 montre déjà qu'elle est injective. Pour montrer qu'elle est surjective, on choisit une matrice arbitraire $M \in \mathrm{Mat}_{m\times n}(K)$. On doit montrer qu'il existe $T \in \mathcal{L}_K(V,W)$ telle que $[\,T\,]^{\mathcal{B}}_{\mathcal{D}} = M$. Pour y parvenir, on considère l'application linéaire

$$\begin{array}{rccc} T_M : & K^n & \longrightarrow & K^n \\ & X & \longmapsto & M\,X \end{array}$$

associée à M. La proposition 2.3.4 montre que les applications

$$\begin{array}{rccc} \varphi : & V & \longrightarrow & K^n \\ & \mathbf{v} & \longmapsto & [\mathbf{v}]_{\mathcal{B}} \end{array} \qquad \text{et} \qquad \begin{array}{rccc} \psi : & W & \longrightarrow & K^m \\ & \mathbf{w} & \longmapsto & [\mathbf{w}]_{\mathcal{D}} \end{array}$$

sont des isomorphismes. Alors la composée

$$T = \psi^{-1} \circ T_M \circ \varphi \colon V \longrightarrow W$$

est une application linéaire, c'est-à-dire un élément de $\mathcal{L}_K(V,W)$. Pour tout $\mathbf{v} \in V$, on a

$$[T(\mathbf{v})]_{\mathcal{D}} = \psi(T(\mathbf{v})) = (\psi \circ T)(\mathbf{v}) = (T_M \circ \varphi)(\mathbf{v}) = T_M(\varphi(\mathbf{v})) = T_M([\mathbf{v}]_{\mathcal{B}}) = M[\mathbf{v}]_{\mathcal{B}},$$

donc $M = [\,T\,]^{\mathcal{B}}_{\mathcal{D}}$, comme requis. □

Corollaire 2.4.4. *Si* $\dim_K(V) = n$ *et* $\dim_K(W) = m$, *alors* $\dim_K \mathcal{L}_K(V,W) = mn$.

Preuve. Comme $\mathcal{L}_K(V,W) \simeq \mathrm{Mat}_{m\times n}(K)$, on obtient

$$\dim_K \mathcal{L}_K(V,W) = \dim_K \mathrm{Mat}_{m\times n}(K) = mn.$$

□

Exemple 2.4.5. Si V est un espace vectoriel de dimension n, il découle du corollaire 2.4.4 que son dual $V^* = \mathcal{L}_K(V,K)$ est de dimension $n \cdot 1 = n$ (voir l'exemple 2.2.4 pour la notion de dual de V). De même l'ensemble $\mathcal{L}_K(V,V)$ des applications de V dans lui-même est de dimension $n \cdot n = n^2$.

Proposition 2.4.6. *Soient* $S\colon U \to V$ *et* $T\colon V \to W$ *des applications linéaires. Supposons que* U, V, W *soient de dimension finie, et que* $\mathcal{A}$, $\mathcal{B}$ *et* $\mathcal{D}$ *soient des bases de* U, V *et* W *respectivement. Alors on a*

$$[T \circ S]^{\mathcal{A}}_{\mathcal{D}} = [T]^{\mathcal{B}}_{\mathcal{D}}\,[S]^{\mathcal{A}}_{\mathcal{B}}\,.$$

Preuve. Pour tout $\mathbf{u} \in U$, on a

$$[T \circ S(\mathbf{u})]_{\mathcal{D}} = [T(S(\mathbf{u}))]_{\mathcal{D}} = [\,T\,]^{\mathcal{B}}_{\mathcal{D}}[S(\mathbf{u})]_{\mathcal{B}} = [\,T\,]^{\mathcal{B}}_{\mathcal{D}}\left([S]^{\mathcal{A}}_{\mathcal{B}}\,[\mathbf{u}]_{\mathcal{A}}\right) = \left([\,T\,]^{\mathcal{B}}_{\mathcal{D}}\,[S]^{\mathcal{A}}_{\mathcal{B}}\right)[\mathbf{u}]_{\mathcal{A}}.$$

□

Corollaire 2.4.7. *Soit* $T\colon V \to W$ *une application linéaire. Supposons que* V *et* W *aient même dimension finie* n *et que* $\mathcal{B}$ *et* $\mathcal{D}$ *soient des bases respectives de* V *et* W. *Alors* T *est inversible si et seulement si la matrice* $[\,T\,]^{\mathcal{B}}_{\mathcal{D}} \in \mathrm{Mat}_{n\times n}(K)$ *est inversible, auquel cas on a*

$$[T^{-1}]^{\mathcal{D}}_{\mathcal{B}} = \left([\,T\,]^{\mathcal{B}}_{\mathcal{D}}\right)^{-1}.$$

Preuve. Si T est inversible, on a

$$T \circ T^{-1} = I_W \quad \text{et} \quad T^{-1} \circ T = I_V,$$

donc

$$\begin{array}{rrcccccc} & I_n & = & [\,I_W\,]^{\mathcal{D}}_{\mathcal{D}} & = & [T \circ T^{-1}]^{\mathcal{D}}_{\mathcal{D}} & = & [\,T\,]^{\mathcal{B}}_{\mathcal{D}}[T^{-1}]^{\mathcal{D}}_{\mathcal{B}} \\ \text{et} & I_n & = & [\,I_V\,]^{\mathcal{B}}_{\mathcal{B}} & = & [T^{-1} \circ T]^{\mathcal{B}}_{\mathcal{B}} & = & [T^{-1}]^{\mathcal{D}}_{\mathcal{B}}[\,T\,]^{\mathcal{B}}_{\mathcal{D}}. \end{array}$$

Par suite $[\,T\,]^{\mathcal{B}}_{\mathcal{D}}$ est inversible et son inverse est $[T^{-1}]^{\mathcal{D}}_{\mathcal{B}}$.

Réciproquement, si $[\,T\,]^{\mathcal{B}}_{\mathcal{D}}$ est inversible, il existe une application linéaire $S\colon W \to V$ telle que $[S]^{\mathcal{D}}_{\mathcal{B}} = \left([\,T\,]^{\mathcal{B}}_{\mathcal{D}}\right)^{-1}$. On obtient

$$\begin{array}{rrcccccc} & [T \circ S]^{\mathcal{D}}_{\mathcal{D}} & = & [\,T\,]^{\mathcal{B}}_{\mathcal{D}}\,[\,S\,]^{\mathcal{D}}_{\mathcal{B}} & = & I_n & = & [\,I_W\,]^{\mathcal{D}}_{\mathcal{D}} \\ \text{et} & [S \circ T]^{\mathcal{B}}_{\mathcal{B}} & = & [\,S\,]^{\mathcal{D}}_{\mathcal{B}}\,[\,T\,]^{\mathcal{B}}_{\mathcal{D}} & = & I_n & = & [\,I_V\,]^{\mathcal{B}}_{\mathcal{B}}, \end{array}$$

donc $T \circ S = I_W$ et $S \circ T = I_V$. Par suite T est inversible et son inverse est $T^{-1} = S$. □

2.5 Changements de coordonnées

On rappelle d'abord la notion de matrice de changement de coordonnées.

Proposition 2.5.1. *Soient $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ et $\mathcal{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ deux bases d'un même espace vectoriel V de dimension n sur K. Il existe une et une seule matrice inversible $P \in \mathrm{Mat}_{n\times n}(K)$ telle que*

$$[\mathbf{v}]_{\mathcal{B}'} = P[\mathbf{v}]_{\mathcal{B}}$$

pour tout $\mathbf{v} \in V$. Cette matrice est donnée par

$$P = [\, I_V \,]^{\mathcal{B}}_{\mathcal{B}'} = \Big([\mathbf{v}_1]_{\mathcal{B}'} \cdots [\mathbf{v}_n]_{\mathcal{B}'}\Big)$$

c'est à dire que la j-ième colonne de P est $[\mathbf{v}_j]_{\mathcal{B}'}$ pour $j = 1, \ldots, n$. Son inverse est

$$P^{-1} = [\, I_V \,]^{\mathcal{B}'}_{\mathcal{B}} = \Big([\mathbf{v}'_1]_{\mathcal{B}} \cdots [\mathbf{v}'_n]_{\mathcal{B}}\Big).$$

Preuve. Puisque l'application identité I_V de V est linéaire, la proposition 2.4.1 montre qu'il existe une et une seule matrice $P \in \mathrm{Mat}_{n\times n}(K)$ telle que

$$[\mathbf{v}]_{\mathcal{B}'} = [\, I_V(\mathbf{v}) \,]_{\mathcal{B}'} = P\,[\mathbf{v}]_{\mathcal{B}}$$

pour tout $\mathbf{v} \in V$. Cette matrice est

$$P = [\, I_V \,]^{\mathcal{B}}_{\mathcal{B}'} = \Big([\, I_V(\mathbf{v}_1)]_{\mathcal{B}'} \cdots [\, I_V(\mathbf{v}_n)]_{\mathcal{B}'}\Big) = \Big([\mathbf{v}_1]_{\mathcal{B}'} \cdots [\mathbf{v}_n]_{\mathcal{B}'}\Big).$$

De plus, comme I_V est inversible, égale à son propre inverse, le corollaire 2.4.7 donne

$$P^{-1} = [\, I_V^{-1} \,]^{\mathcal{B}'}_{\mathcal{B}} = [\, I_V \,]^{\mathcal{B}'}_{\mathcal{B}}.$$

□

La matrice P s'appelle la *matrice de passage des coordonnées en base $\mathcal{B}$ aux coordonnées en base $\mathcal{B}'$*. Pour abréger, on dit aussi que c'est la *matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$*.

Exemple 2.5.2. L'ensemble $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$ est une base de $\mathbb{R}^2$ car sa cardinalité est $|\mathcal{B}| = 2 = \dim_{\mathbb{R}} \mathbb{R}^2$ et que $\mathcal{B}$ est formé de vecteurs linéairement indépendants. Pour la même raison $\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ est aussi une base de $\mathbb{R}^2$. On trouve que

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = -2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + 3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 \\ 3 \end{pmatrix} = -3 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

donc

$$[\, I_{\mathbb{R}^2}]^{\mathcal{B}}_{\mathcal{B}'} = \left([\begin{pmatrix} 1 \\ 2 \end{pmatrix}]_{\mathcal{B}'}, [\begin{pmatrix} 1 \\ 3 \end{pmatrix}]_{\mathcal{B}'} \right) = \begin{pmatrix} -2 & -3 \\ 3 & 4 \end{pmatrix}$$

et par suite

$$[\mathbf{v}]_{\mathcal{B}'} = \begin{pmatrix} -2 & -3 \\ 3 & 4 \end{pmatrix} [\mathbf{v}]_{\mathcal{B}}$$

pour tout $\mathbf{v} \in \mathbb{R}^2$.

On peut aussi faire le calcul par l'intermédiaire de la base canonique $\mathcal{E} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ de $\mathbb{R}^2$. Comme $I_{\mathbb{R}^2} = I_{\mathbb{R}^2} \circ I_{\mathbb{R}^2}$, on trouve

$$\begin{aligned} [\, I_{\mathbb{R}^2}]_{\mathcal{B}'}^{\mathcal{B}} &= [\, I_{\mathbb{R}^2}]_{\mathcal{B}'}^{\mathcal{E}} \, [\, I_{\mathbb{R}^2}]_{\mathcal{E}}^{\mathcal{B}} \\ &= \left([\, I_{\mathbb{R}^2}]_{\mathcal{E}}^{\mathcal{B}'} \right)^{-1} \left([\, I_{\mathbb{R}^2}]_{\mathcal{E}}^{\mathcal{B}} \right) \\ &= \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} -2 & -3 \\ 3 & 4 \end{pmatrix} \end{aligned}$$

La seconde démarche employée dans l'exemple ci-dessus se généralise aisément. Cela donne :

Proposition 2.5.3. *Si $\mathcal{B}, \mathcal{B}'$ et $\mathcal{B}''$ sont trois bases d'un même espace vectoriel V, on a*

$$[\, I_V \,]_{\mathcal{B}''}^{\mathcal{B}} = [\, I_V \,]_{\mathcal{B}''}^{\mathcal{B}'} \, [\, I_V \,]_{\mathcal{B}'}^{\mathcal{B}}.$$

Les matrices de changement de coordonnées permettent aussi de relier entre elles les matrices d'une application linéaire $T \colon V \to W$ par rapport à divers choix de bases pour V et pour W.

Proposition 2.5.4. *Soit $T \colon V \to W$ une application linéaire. Supposons que $\mathcal{B}$ et $\mathcal{B}'$ soient des bases de V et que $\mathcal{D}$ et $\mathcal{D}'$ soient des bases de W. Alors on a*

$$[\, T \,]_{\mathcal{D}'}^{\mathcal{B}'} = [\, I_W \,]_{\mathcal{D}'}^{\mathcal{D}} \, [\, T \,]_{\mathcal{D}}^{\mathcal{B}} \, [\, I_V \,]_{\mathcal{B}}^{\mathcal{B}'}.$$

Preuve. Pour le voir, il suffit d'écrire $T = I_W \circ T \circ I_V$ et d'appliquer la proposition 2.4.6. □

On conclut avec le rappel du résultat suivant dont la preuve est laissée en exercice, et qui fournit un complément utile à la proposition 2.5.1.

Proposition 2.5.5. *Soit V un espace vectoriel sur K de dimension finie n, soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, et soit $\mathcal{B}' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ un ensemble de n éléments de V. Alors $\mathcal{B}'$ est une base de V si et seulement si la matrice*

$$P = \left([\mathbf{v}'_1]_{\mathcal{B}} \cdots [\mathbf{v}'_n]_{\mathcal{B}} \right)$$

est inversible.

2.6 Endomorphismes et sous-espaces invariants

Une application linéaire $T\colon V \to V$ d'un espace vectoriel V dans lui-même s'appelle un *endomorphisme* de V, ou encore un *opérateur linéaire* sur V. On désigne par

$$\mathrm{End}_K(V) := \mathcal{L}_K(V, V)$$

l'ensemble des endomorphismes de V. En vertu du théorème 2.2.3, c'est un espace vectoriel sur K.

Supposons que V soit de dimension finie n et que $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ soit une base de V. Pour tout opérateur linéaire $T\colon V \to V$, on écrit simplement $[\,T\,]_{\mathcal{B}}$ pour désigner la matrice $[\,T\,]^{\mathcal{B}}_{\mathcal{B}}$. Avec cette convention, on obtient

$$[T(\mathbf{v})]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}\,[\mathbf{v}]_{\mathcal{B}} \quad \text{pour tout } \mathbf{v} \in V.$$

Le théorème 2.4.3 nous apprend qu'on a un isomorphisme d'espaces vectoriels

$$\begin{array}{rcl} \mathrm{End}_K(V) & \longrightarrow & \mathrm{Mat}_{n\times n}(K)\,. \\ T & \longmapsto & [\,T\,]_{\mathcal{B}} \end{array}$$

Cela implique :

$$\dim_K(\mathrm{End}_K(V)) = \dim_K(\mathrm{Mat}_{n\times n}(K)) = n^2.$$

De plus, si T_1 et T_2 sont deux opérateurs linéaires sur V et si $c \in K$, alors on a :

$$[T_1 + T_2]_{\mathcal{B}} = [T_1]_{\mathcal{B}} + [T_2]_{\mathcal{B}} \quad \text{et} \quad [c\,T_1]_{\mathcal{B}} = c\,[T_1]_{\mathcal{B}}.$$

Par ailleurs, la composée $T_1 \circ T_2\colon V \to V$ est aussi un opérateur linéaire sur V et la proposition 2.4.6 donne

$$[T_1 \circ T_2]_{\mathcal{B}} = [T_1]_{\mathcal{B}}\,[T_2]_{\mathcal{B}}.$$

Enfin, les propositions 2.5.1 et 2.5.4 nous apprennent que, si $\mathcal{B}'$ est une autre base de V, alors, pour tout $T \in \mathrm{End}_K(V)$, on a

$$[\,T\,]_{\mathcal{B}'} = P^{-1}\,[\,T\,]_{\mathcal{B}}\,P$$

où $P = [\,I_V\,]^{\mathcal{B}'}_{\mathcal{B}}$.

Comme on l'a expliqué dans l'introduction, un des objectifs principaux de ce cours est de donner une description la plus simple possible d'un endomorphisme d'un espace vectoriel. Cette description passe par la notion suivante :

Définition 2.6.1. Soit $T \in \mathrm{End}_K(V)$. On dit qu'un sous-espace U de V est T-*invariant* si

$$T(\mathbf{u}) \in U \quad \text{pour tout } \mathbf{u} \in U.$$

Le résultat suivant est laissé en exercice :

Proposition 2.6.2. *Soit $T \in \mathrm{End}_K(V)$ et soient $U_1, \ldots, U_s$ des sous-espaces T-invariants de V. Alors $U_1 + \cdots + U_s$ et $U_1 \cap \cdots \cap U_s$ sont aussi T-invariants.*

On montre encore :

Proposition 2.6.3. *Soit $T \in \mathrm{End}_K(V)$ et soit U un sous-espace T-invariant de V. La fonction*

$$\begin{array}{rccc} T|_U : & U & \longrightarrow & U \\ & \mathbf{u} & \longmapsto & T(\mathbf{u}) \end{array}$$

est une application linéaire.

On dit que $T|_U$ est la *restriction de T à U*. La proposition 2.6.3 nous apprend que c'est un opérateur linéaire sur U.

Preuve. Pour tout choix de $\mathbf{u}_1, \mathbf{u}_2 \in U$ et de $c \in K$, on a

$$T|_U(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2) = T|_U(\mathbf{u}_1) + T|_U(\mathbf{u}_2),$$

$$T|_U(c\,\mathbf{u}_1) = T(c\,\mathbf{u}_1) = cT(\mathbf{u}_1) = c\,T|_U(\mathbf{u}_1).$$

Donc $T|_U : U \to U$ est bien linéaire. □

Enfin, le critère suivant est utile pour déterminer si un sous-espace de U de V est invariant sous un endomorphisme de V.

Proposition 2.6.4. *Soit $T \in \mathrm{End}_K(V)$ et soit $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ un système de générateurs d'un sous-espace U de V. Alors U est T-invariant si et seulement si $T(\mathbf{u}_i) \in U$ pour $i = 1, \ldots, m$.*

Preuve. Par définition, si U est T-invariant, on a $T(\mathbf{u}_1), \ldots, T(\mathbf{u}_m) \in U$. Réciproquement, supposons que $T(\mathbf{u}_i) \in U$ pour $i = 1, \ldots, m$ et choisissons un élément quelconque $\mathbf{u}$ de U. Comme $U = \langle \mathbf{u}_1, \ldots, \mathbf{u}_m \rangle_K$, il existe $a_1, \ldots, a_m \in K$ tels que $\mathbf{u} = a_1\,\mathbf{u}_1 + \cdots + a_m \mathbf{u}_m$. On en déduit

$$T(\mathbf{u}) = a_1\,T(\mathbf{u}_1) + \cdots + a_m T(\mathbf{u}_m) \in U.$$

Donc U est T-invariant. □

Exemple 2.6.5. Soit $\mathcal{C}_\infty(\mathbb{R})$ l'ensemble des fonctions $f : \mathbb{R} \to \mathbb{R}$ qui sont indéfiniment différentiables. La dérivation

$$\begin{array}{rccc} D : & \mathcal{C}_\infty(\mathbb{R}) & \longrightarrow & \mathcal{C}_\infty(\mathbb{R}) \\ & f & \longmapsto & f' \end{array}$$

est un opérateur $\mathbb{R}$-linéaire sur $\mathcal{C}_\infty(\mathbb{R})$. Soit U le sous-espace de $\mathcal{C}_\infty(\mathbb{R})$ engendré par les fonctions f_1, f_2 et f_3 données par

$$f_1(x) = e^{2x}, \quad f_2(x) = x\,e^{2x}, \quad f_3(x) = x^2\,e^{2x}.$$

On trouve, pour tout $x \in \mathbb{R}$,

$$\begin{aligned} f_1'(x) &= 2e^x = 2f_1(x), \\ f_2'(x) &= e^{2x} + 2x\,e^{2x} = f_1(x) + 2f_2(x), \\ f_3'(x) &= 2x\,e^{2x} + 2x^2\,e^{2x} = 2f_2(x) + 2f_3(x). \end{aligned}$$

Donc

$$D(f_1) = 2\,f_1, \quad D(f_2) = f_1 + 2\,f_2, \quad D(f_3) = 2\,f_2 + 2\,f_3 \tag{2.2}$$

appartiennent à U et par suite U est un sous-espace D-invariant de $\mathcal{C}_\infty(\mathbb{R})$.

Les fonctions f_1, f_2 et f_3 sont linéairement indépendantes, car si $a_1, a_2, a_3 \in \mathbb{R}$ satisfont $a_1\,f_1 + a_2\,f_2 + a_3\,f_3 = 0$, alors on a

$$a_1\,e^{2x} + a_2 x\,e^{2x} + a_3 x^2\,e^{2x} = 0 \quad \text{pour tout } x \in \mathbb{R}.$$

Comme $e^{2x} \neq 0$ quel que soit $x \in \mathbb{R}$, on en déduit (en divisant par e^{2x})

$$a_1 + a_2 x + a_3 x^2 = 0 \quad \text{pour tout } x \in \mathbb{R}$$

et par suite $a_1 = a_2 = a_3 = 0$ (un polynôme non nul de degré 2 possède au plus 2 racines). Donc $\mathcal{B} = \{f_1, f_2, f_3\}$ est une base de U et les formules (2.2) livrent

$$[D|_U]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

En combinant les considérations précédentes à celles de la section 1.4, on obtient :

Théorème 2.6.6. *Soit $T\colon V \longrightarrow V$ un opérateur linéaire sur un espace vectoriel V de dimension finie. Supposons que V soit la somme directe de sous-espaces T-invariants $V_1, \ldots, V_s$. Soit $\mathcal{B}_i = \{\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,n_i}\}$ une base de V_i pour $i = 1, \ldots, s$. Alors*

$$\mathcal{B} = \mathcal{B}_1 \amalg \cdots \amalg \mathcal{B}_s = \{\mathbf{v}_{1,1}, \ldots, \mathbf{v}_{1,n_1}, \ldots\ldots, \mathbf{v}_{s,1}, \ldots, \mathbf{v}_{s,n_s}\}$$

est une base de V et on a

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_{V_1}]_{\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & [T|_{V_2}]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & [T|_{V_s}]_{\mathcal{B}_s} \end{pmatrix}, \tag{2.3}$$

c'est-à-dire que la matrice de T relative à la base $\mathcal{B}$ est diagonale par blocs, le i-ième bloc de sa diagonale étant la matrice $[T|_{V_i}]_{\mathcal{B}_i}$ de la restriction de T à V_i relative à la base $\mathcal{B}_i$.

Le fait que $\mathcal{B}$ soit une base de V découle de la proposition 1.4.8. On démontre la formule (2.3) par récurrence sur s. Pour $s = 1$, il n'y a rien à démontrer. Pour $s = 2$, il est plus facile de changer de notation. On est ainsi amener à démontrer :

Lemme 2.6.7. *Soit $T\colon V \to V$ un opérateur linéaire sur un espace vectoriel V de dimension finie. Supposons que $V = U \oplus W$ où U et W sont des sous-espaces T-invariants. Soient $\mathcal{A} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ une base de U et $\mathcal{D} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ une base de W. Alors*

$$\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{w}_1, \ldots, \mathbf{w}_n\}$$

est une base de V et

$$[\,T\,]_{\mathcal{B}} = \begin{pmatrix} [T|_U]_{\mathcal{A}} & 0 \\ 0 & [T|_W]_{\mathcal{D}} \end{pmatrix}.$$

Preuve. Écrivons $[T|_U]_{\mathcal{A}} = (a_{ij})$ et $[T|_W]_{\mathcal{D}} = (b_{ij})$. Alors, pour $j = 1, \ldots, m$, on a :

$$T(\mathbf{u}_j) = T|_U(\mathbf{u}_j) = a_{1j}\mathbf{u}_1 + \cdots + a_{mj}\mathbf{u}_m \quad \Longrightarrow \quad [T(\mathbf{u}_j)]_{\mathcal{B}} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

De même, pour $j = 1, \ldots, n$, on a :

$$T(\mathbf{w}_j) = T|_W(\mathbf{w}_j) = b_{1j}\mathbf{w}_1 + \cdots + b_{nj}\mathbf{w}_n \quad \Longrightarrow \quad [T(\mathbf{w}_j)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}.$$

Par suite,

$$\begin{aligned} [\,T\,]_{\mathcal{B}} &= \Big([T(\mathbf{u}_1)]_{\mathcal{B}}, \ldots, [T(\mathbf{u}_m)]_{\mathcal{B}}, [T(\mathbf{w}_1)]_{\mathcal{B}}, \ldots, [T(\mathbf{w}_n)]_{\mathcal{B}}\Big) \\ &= \begin{pmatrix} a_{11} & \cdots & a_{1m} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n1} & \cdots & b_{nn} \end{pmatrix} \\ &= \begin{pmatrix} [T|_U]_{\mathcal{A}} & 0 \\ 0 & [T|_W]_{\mathcal{D}} \end{pmatrix} \end{aligned}$$

□

Preuve du théorème 2.6.6. Comme on l'a expliqué plus tôt, il suffit de démontrer (2.3). Le cas où $s = 1$ est clair. Supposons maintenant $s \leq 2$ et que le résultat soit vrai pour les valeurs inférieures de s. Les hypothèses du lemme 2.6.7 sont remplies pour le choix de

$$U := V_1, \quad \mathcal{A} := \mathcal{B}_1, \quad W := V_2 \oplus \cdots \oplus V_s \quad \text{et} \quad \mathcal{D} := \mathcal{B}_2 \amalg \cdots \amalg \mathcal{B}_s.$$

On a donc

$$[\,T\,]_{\mathcal{B}} = \begin{pmatrix} [T|_{V_1}]_{\mathcal{B}_1} & 0 \\ 0 & [T|_W]_{\mathcal{D}} \end{pmatrix}$$

Comme $W = V_2 \oplus \cdots \oplus V_s$ est somme directe de $s-1$ sous-espaces $T|_W$-invariants, l'hypothèse de récurrence donne

$$[\,T|_W\,]_{\mathcal{D}} = \begin{pmatrix} [T|_{V_2}]_{\mathcal{B}_2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & [T|_{V_s}]_{\mathcal{B}_s} \end{pmatrix},$$

car $(T|_W)|_{V_i} = T|_{V_i}$ pour $i = 2, \ldots, s$. La conclusion suit. □

Exercices.

2.6.1. Démontrer la proposition 2.6.2.

2.6.2. Soit V un espace vectoriel sur un corps K et soient S et T des opérateurs linéaires sur V qui commutent entre eux, c'est-à-dire tels que $S \circ T = T \circ S$. Montrer que $\ker(S)$ et $\mathrm{Im}(S)$ sont des sous-espaces T-invariants de V.

2.6.3. Soit D l'opérateur de dérivation sur l'espace $C_\infty(\mathbb{R})$ des fonctions indéfiniment différentiables $f\colon \mathbb{R} \to \mathbb{R}$. On considère le sous-espace vectoriel $U = \langle\, 1, x, \ldots, x^m \,\rangle_{\mathbb{R}}$ de $C_\infty(\mathbb{R})$ pour un entier $m \geq 0$.

(i) Montrer que U est D-invariant.
(ii) Montrer que $\mathcal{B} = \{1, x, \ldots, x^m\}$ est une base de U.
(iii) Calculer $[\,D|_U\,]_{\mathcal{B}}$.

2.6.4. Soit D comme dans l'exercice 2.6.3, et soit $U = \langle\, \cos(3x), \sin(3x), x\cos(3x), x\sin(3x) \,\rangle_{\mathbb{R}}$.

(i) Montrer que U est D-invariant.
(ii) Montrer que $\mathcal{B} = \{\cos(3x), \sin(3x), x\cos(3x), x\sin(3x)\}$ est une base de U.
(iii) Calculer $[\,D|_U\,]_{\mathcal{B}}$.

2.6.5. Soit V un espace vectoriel de dimension 4 sur $\mathbb{Q}$, soit $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ une base de V, et soit $T\colon V \to V$ l'application linéaire déterminée par les conditions

$$\begin{aligned} T(\mathbf{v}_1) &= \mathbf{v}_1 + \mathbf{v}_2 + 2\mathbf{v}_3, & T(\mathbf{v}_2) &= \mathbf{v}_1 + \mathbf{v}_2 + 2\mathbf{v}_4, \\ T(\mathbf{v}_3) &= \mathbf{v}_1 - 3\mathbf{v}_2 + \mathbf{v}_3 - \mathbf{v}_4, & T(\mathbf{v}_4) &= -3\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 + \mathbf{v}_4. \end{aligned}$$

(i) Montrer que $\mathcal{B}_1 = \{\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_3 + \mathbf{v}_4\}$ et $\mathcal{B}_2 = \{\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_3 - \mathbf{v}_4\}$ sont les bases respectives de sous-espaces T-invariants V_1 et V_2 de V tels que $V = V_1 \oplus V_2$.
(ii) Calculer $[\,T|_{V_1}\,]_{\mathcal{B}_1}$ et $[\,T|_{V_2}\,]_{\mathcal{B}_2}$, et en déduire $[\,T\,]_{\mathcal{B}}$ où $\mathcal{B} = \mathcal{B}_1 \coprod \mathcal{B}_2$.

2.6.6. Soit $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire dont la matrice relative à la base canonique $\mathcal{E}$ de $\mathbb{R}^3$ est

$$[\,T\,]_{\mathcal{E}} = \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}.$$

(i) Montrer que $\mathcal{B}_1 = \{(1,1,1)^t\}$ et $\mathcal{B}_2 = \{(1,-1,0)^t, (0,1,-1)^t\}$ sont les bases respectives de sous-espaces T-invariants V_1 et V_2 de $\mathbb{R}^3$ tels que $\mathbb{R}^3 = V_1 \oplus V_2$.
(ii) Calculer $[\,T|_{V_1}\,]_{\mathcal{B}_1}$ et $[\,T|_{V_2}\,]_{\mathcal{B}_2}$, et en déduire $[\,T\,]_{\mathcal{B}}$ où $\mathcal{B} = \mathcal{B}_1 \coprod \mathcal{B}_2$.

2.6.7. Soit V un espace vectoriel de dimension finie sur un corps K, soit $T\colon V \to V$ un endomorphisme de V, et soit $n = \dim_K(V)$. Supposons qu'il existe un entier $m \geq 0$ et un vecteur $\mathbf{v} \in V$ tels que $V = \langle\, \mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots, T^m(\mathbf{v}) \,\rangle_K$. Montrer que $\mathcal{B} = \{\mathbf{v}, T(\mathbf{v}), \ldots, T^{n-1}(\mathbf{v})\}$ est une base de V et qu'il existe $a_0, \ldots, a_{n-1} \in K$ tels que

$$[\,T\,]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \ldots & 0 & a_0 \\ 1 & 0 & \ldots & 0 & a_1 \\ 0 & 1 & \ldots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & a_{n-1} \end{pmatrix}.$$

Note. On dit qu'un espace vectoriel V avec la propriété ci-dessus est *T-cyclique*, ou tout simplement *cyclique*.

Les trois derniers exercices proposent des exemples de la situation générale présentée à l'exercice 2.6.7.

2.6.8. Soit $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire dont la matrice en base canonique $\mathcal{E}$ est

$$[\,T\,]_{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}.$$

(i) Posons $\mathbf{v} = (1,1)^t$. Montrer que $\mathcal{B} = \{\mathbf{v}, T(\mathbf{v})\}$ forme une base de $\mathbb{R}^2$.
(ii) Déterminer $[T^2(\mathbf{v})]_{\mathcal{B}}$.
(iii) Calculer $[\,T\,]_{\mathcal{B}}$.

2.6.9. Soit V un espace vectoriel de dimension 3 sur $\mathbb{Q}$, soit $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ une base de V, et soit $T\colon V \to V$ l'application linéaire déterminée par les conditions

$$T(\mathbf{v}_1) = \mathbf{v}_1 + \mathbf{v}_3, \quad T(\mathbf{v}_2) = \mathbf{v}_1 - \mathbf{v}_2 \quad \text{et} \quad T(\mathbf{v}_3) = \mathbf{v}_2.$$

(i) Donner $[\,T\,]_{\mathcal{A}}$.
(ii) Montrer que $\mathcal{B} = \{\mathbf{v}_1, T(\mathbf{v}_1), T^2(\mathbf{v}_1)\}$ forme une base de V.

(iii) Calculer $[\,T\,]_{\mathcal{B}}$.

2.6.10. Soit $U = \langle\, 1, e^x, e^{2x} \,\rangle_{\mathbb{R}} \subset \mathcal{C}_\infty(\mathbb{R})$ et soit D la restriction à U de l'opérateur de dérivation sur $\mathcal{C}_\infty(\mathbb{R})$. On admettra que $\mathcal{A} = \{1, e^x, e^{2x}\}$ est une base de V.

(i) Soit $f(x) = 1 + e^x + e^{2x}$. Montrer que $\mathcal{B} = \{\, f,\, D(f),\, D^2(f) \,\}$ est une base de V.

(ii) Calculer $[\,D\,]_{\mathcal{B}}$.

Chapitre 3

Retour sur la diagonalisation

Dans ce chapitre, on rappelle comment déterminer si un opérateur linéaire ou une matrice carrée est diagonalisable et si oui comment procéder à sa diagonalisation.

3.1 Déterminants et matrices semblables

Soit K un corps. On définit le *déterminant* d'une matrice carrée $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(K)$ par

$$\det(A) = \sum_{(j_1,\dots,j_n)\in S_n} \epsilon(j_1,\dots,j_n) a_{1j_1}\cdots a_{nj_n}$$

où S_n désigne l'ensemble des permutations de $(1,2,\dots,n)$ et où, pour une permutation $(j_1,\dots,j_n) \in S_n$, l'expression $\epsilon(j_1,\dots,j_n)$ représente la *signature* de $(j_1,\dots,j_n)$ donnée par

$$\epsilon(j_1,\dots,j_n) = (-1)^{N(j_1,\dots,j_n)}$$

où $N(j_1,\dots,j_n)$ désigne le nombre de couples d'indices (k,l) avec $1 \leq k < l \leq n$ et $j_k > j_l$ (appelé le *nombre d'inversions de* $(j_1,\dots,j_n)$).

Rappelons que le déterminant jouit des propriétés suivantes :

Théorème 3.1.1. *Soit $n \in \mathbb{N}^*$ et soient $A, B \in \mathrm{Mat}_{n\times n}(K)$. Alors on a :*

(i) $\det(I) = 1$,

(ii) $\det(A^t) = \det(A)$,

(iii) $\det(AB) = \det(A)\det(B)$,

où I désigne la matrice identité $n \times n$ et A^t la transposée de A. De plus, la matrice A est inversible si et seulement si $\det(A) \neq 0$, *auquel cas on a*

(iv) $\det(A^{-1}) = \det(A)^{-1}$.

En fait, la formule pour le déterminant s'applique à toute matrice carrée à coefficients dans un anneau commutatif. On trouvera dans l'appendice B une démonstration du théorème 3.1.1 dans ce cadre général, avec K remplacé par un anneau commutatif quelconque.

Les propriétés (*iii*) et (*iv*) de "multiplicativité" du déterminant entraînent le résultat suivant.

Proposition 3.1.2. *Soit T un endomorphisme d'un espace vectoriel V de dimension finie, et soient $\mathcal{B}$, $\mathcal{B}'$ deux bases de V. Alors on a*

$$\det[T]_{\mathcal{B}} = \det[T]_{\mathcal{B}'}$$

Preuve. On a $[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P$ où $P = [I_V]_{\mathcal{B}}^{\mathcal{B}'}$, donc

$$\begin{aligned} \det[T]_{\mathcal{B}'} &= \det(P^{-1}[T]_{\mathcal{B}}P) \\ &= \det(P^{-1})\det[T]_{\mathcal{B}}\det(P) \\ &= \det(P)^{-1}\det[T]_{\mathcal{B}}\det(P) \\ &= \det[T]_{\mathcal{B}}. \end{aligned}$$

□

Cela permet de formuler :

Définition 3.1.3. Soit $T\colon V \to V$ un opérateur linéaire sur un espace vectoriel V de dimension finie. On définit le *déterminant* de T par

$$\det(T) = \det[T]_{\mathcal{B}}$$

où $\mathcal{B}$ désigne une base quelconque de V.

Le théorème 3.1.1 permet d'étudier le comportement du déterminant sous la composition des opérateurs linéaires :

Théorème 3.1.4. *Soient S et T des opérateurs linéaires sur un espace vectoriel V de dimension finie. On a :*

(i) $\det(I_V) = 1$,

(ii) $\det(S \circ T) = \det(S)\det(T)$.

De plus, T est inversible si et seulement si $\det(T) \neq 0$, *auquel cas on a*

(iii) $\det(T^{-1}) = \det(T)^{-1}$.

Preuve. Soit $\mathcal{B}$ une base de V. Comme $[I_V]_{\mathcal{B}} = I$ est la matrice identité $n \times n$, on a $\det(I_V) = \det(I) = 1$. Cela démontre (*i*). Pour (*iii*), on note, d'après le corollaire 2.4.7, que T est inversible si et seulement si sa matrice $[T]_{\mathcal{B}}$ est inversible et qu'alors

$$[T^{-1}]_{\mathcal{B}} = [T]_{\mathcal{B}}^{-1}.$$

On en déduit (*iii*). La relation (*ii*) est laissée en exercice. □

Le théorème 3.1.4 illustre le phénomène indiqué dans l'introduction que, le plus souvent dans le contexte de ce cours, un résultat sur les matrices admet un pendant en termes d'opérateurs linéaires et vice-versa. On conclut avec la notion suivante :

Définition 3.1.5. Soit $n \in \mathbb{N}^*$ et soient $A, B \in \mathrm{Mat}_{n\times n}(K)$ des matrices carrées de même format $n \times n$. On dit que A est *semblable* à B (ou *conjuguée* à B), s'il existe une matrice inversible $P \in \mathrm{Mat}_{n\times n}(K)$, telle que $B = P^{-1}\,A\,P$. On écrit alors $A \sim B$.

Exemple 3.1.6. Si $T\colon V \to V$ est un opérateur linéaire sur un espace vectoriel V de dimension finie et si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de V, alors on a

$$[\,T\,]_{\mathcal{B}'} = P^{-1}\,[\,T\,]_{\mathcal{B}}P$$

où $P = [\,I_V\,]_{\mathcal{B}}^{\mathcal{B}'}$, donc $[\,T\,]_{\mathcal{B}'}$ et $[\,T\,]_{\mathcal{B}}$ sont semblables.

Plus précisément, on peut montrer :

Proposition 3.1.7. *Soit $T\colon V \to V$ un opérateur linéaire sur un espace vectoriel V de dimension finie, et soit $\mathcal{B}$ une base de V. Une matrice M est semblable à $[\,T\,]_{\mathcal{B}}$ si et seulement s'il existe une base de $\mathcal{B}'$ de V telle que $M = [\,T\,]_{\mathcal{B}'}$.*

Enfin le théorème 3.1.1 fournit une condition nécessaire pour que deux matrices soient semblables :

Proposition 3.1.8. *Si A et B sont deux matrices semblables, alors* $\det(A) = \det(B)$.

La preuve est laissée en exercice. La proposition 3.1.2 peut être vue comme un corollaire de ce résultat joint à l'exemple 3.1.6.

Exercices.

3.1.1. Montrer que la similitude $\sim$ est une relation d'équivalence sur $\mathrm{Mat}_{n\times n}(K)$.

3.1.2. Démontrer la proposition 3.1.7.

3.1.3. Démontrer la proposition 3.1.8.

3.2 Diagonalisation des opérateurs

Dans tout ce paragraphe, on fixe un espace vectoriel V de dimension finie n sur un corps K et un opérateur linéaire $T\colon V \to V$. On dit que :

- T est *diagonalisable* s'il existe une base $\mathcal{B}$ de V telle que $[T]_{\mathcal{B}}$ soit une matrice diagonale,
- un élément λ de K est une *valeur propre* de T s'il existe un vecteur $\mathbf{v}$ de V *non nul* tel que
$$T(\mathbf{v}) = \lambda \mathbf{v}\,,$$
- un vecteur $\mathbf{v}$ non nul qui satisfait la relation ci-dessus pour un élément λ de K s'appelle un *vecteur propre* de T ou plus précisément un *vecteur propre de T pour la valeur propre λ*.

On rapelle aussi que si $\lambda \in K$ est une valeur propre de T alors l'ensemble

$$E_\lambda := \{\mathbf{v} \in V\,;\, T(\mathbf{v}) = \lambda \mathbf{v}\}$$

(encore noté $E_\lambda(T)$) est un sous-espace de V car, pour $\mathbf{v} \in V$, on a

$$T(\mathbf{v}) = \lambda \mathbf{v} \quad \Longleftrightarrow \quad (T - \lambda I)(\mathbf{v}) = \mathbf{0}$$

donc

$$E_\lambda = \ker(T - \lambda I).$$

On dit que E_λ est l'*espace propre de T pour la valeur propre λ*. Les vecteurs propres de T pour cette valeur propre sont les éléments non nuls de E_λ.

Pour chaque valeur propre $\lambda \in K$, l'espace propre E_λ est un sous-espace T-invariant de V car, si $\mathbf{v} \in E_\lambda$, on a $T(\mathbf{v}) = \lambda \mathbf{v} \in E_\lambda$. On note aussi qu'un élément $\mathbf{v}$ de V est un vecteur propre de T si et seulement si $\langle \mathbf{v} \rangle_K$ est un sous-espace T-invariant de dimension 1 de V (exercice).

Proposition 3.2.1. *L'opérateur T est diagonalisable si et seulement si V admet une base constituée de vecteurs propres de T.*

Preuve. Soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, et soient $\lambda_1, \ldots, \lambda_n \in K$. On a

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \Longleftrightarrow \quad T(\mathbf{v}_1) = \lambda_1 \mathbf{v}_1\,,\ \ldots\ ,\ T(\mathbf{v}_n) = \lambda_n \mathbf{v}_n.$$

La conclusion suit. □

Le point crucial de cette étude est le résultat suivant :

Proposition 3.2.2. *Soient $\lambda_1, \ldots, \lambda_s \in K$ des valeurs propres distinctes de T. La somme $E_{\lambda_1} + \cdots + E_{\lambda_s}$ est directe.*

Preuve. En vertu de la définition 1.4.1, il s'agit de montrer que le seul choix possible de vecteurs $\mathbf{v}_1 \in E_{\lambda_1}, \ldots, \mathbf{v}_s \in E_{\lambda_s}$ tels que

$$\mathbf{v}_1 + \cdots + \mathbf{v}_s = \mathbf{0} \tag{3.1}$$

est $\mathbf{v}_1 = \cdots = \mathbf{v}_s = \mathbf{0}$. Pour cela, on possède par récurrence sur s, en notant d'abord que, pour $s = 1$, c'est immédiat.

Supposons que $s \geq 2$ et que la somme $E_{\lambda_1} + \cdots + E_{\lambda_{s-1}}$ soit directe. Choisissons des vecteurs $\mathbf{v}_1 \in E_{\lambda_1}, \ldots, \mathbf{v}_s \in E_{\lambda_s}$ satisfaisant (3.1) et appliquons T aux deux membres de cette égalité. Comme $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ pour $i = 1, \ldots, s$ et que $T(\mathbf{0}) = \mathbf{0}$, on obtient

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_s \mathbf{v}_s = \mathbf{0}. \tag{3.2}$$

En soustrayant de (3.2) l'égalité (3.1) multipliée par λ_s, on trouve

$$(\lambda_1 - \lambda_s)\mathbf{v}_1 + \cdots + (\lambda_{s-1} - \lambda_s)\mathbf{v}_{s-1} = \mathbf{0}.$$

Comme $(\lambda_i - \lambda_s)\mathbf{v}_i \in E_{\lambda_i}$ pour $i = 1, \ldots, s-1$ et que, par hypothèse, la somme $E_{\lambda_1} + \cdots + E_{\lambda_{s-1}}$ est directe, cela implique que

$$(\lambda_1 - \lambda_s)\mathbf{v}_1 = \cdots = (\lambda_{s-1} - \lambda_s)\mathbf{v}_{s-1} = \mathbf{0}.$$

Enfin, comme $\lambda_s \neq \lambda_i$ pour $i = 1, \ldots, s-1$, on conclut que $\mathbf{v}_1 = \cdots = \mathbf{v}_{s-1} = \mathbf{0}$, puis, grâce à (3.1), que $\mathbf{v}_s = \mathbf{0}$. Donc la somme $E_{\lambda_1} + \cdots + E_{\lambda_s}$ est directe. □

On en déduit :

Théorème 3.2.3. *L'opérateur T possède au plus $n = \dim_K(V)$ valeurs propres. Soient $\lambda_1, \ldots, \lambda_s$ les valeurs propres distinctes de T dans K. Alors T est diagonalisable si et seulement si*

$$\dim_K(V) = \dim_K(E_{\lambda_1}) + \cdots + \dim_K(E_{\lambda_s}).$$

Si tel est le cas, on obtient une base $\mathcal{B}$ de V constituée de vecteurs propres de T en concaténant des bases de $E_{\lambda_1}, \ldots, E_{\lambda_s}$ et alors

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & 0 & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_s & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_s \end{pmatrix}$$

chaque λ_i étant répété un nombre de fois égal à $\dim_K(E_{\lambda_i})$.

Preuve. Supposons d'abord que $\lambda_1, \ldots, \lambda_s$ soient des valeurs propres distinctes de T (pas nécessairement toutes). Alors, comme la somme $E_{\lambda_1} + \cdots + E_{\lambda_s}$ est directe, on trouve

$$\dim_K(E_{\lambda_1} + \cdots + E_{\lambda_s}) = \dim_K(E_{\lambda_1}) + \cdots + \dim_K(E_{\lambda_s}).$$

Comme $E_{\lambda_1} + \cdots + E_{\lambda_s} \subseteq V$ et que $\dim_K(E_{\lambda_i}) \geq 1$ pour $i = 1, \ldots, s$, on en déduit que

$$n = \dim_K(V) \geq \dim_K(E_{\lambda_1}) + \cdots + \dim_K(E_{\lambda_s}) \geq s.$$

Cela démontre la première affirmation du théorème : T possède au plus n valeurs propres.

Supposons maintenant que $\lambda_1, \ldots, \lambda_s$ soient toutes les valeurs propres distinctes de T. En vertu des propositions 3.2.1, 3.2.2 et du corollaire 1.4.9, on en déduit que

$$\begin{aligned} T \text{ est diagonalisable} &\iff V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_s} \\ &\iff n = \dim_K(E_{\lambda_1}) + \cdots + \dim_K(E_{\lambda_s}). \end{aligned}$$

Cela démontre la seconde assertion du théorème. La troisième et dernière découle du théorème 2.6.6. □

Pour pouvoir en pratique *diagonaliser* T, c'est-à-dire déterminer une base $\mathcal{B}$ de V telle que $[T]_{\mathcal{B}}$ soit diagonale, si une telle base existe, il reste à pouvoir déterminer les valeurs propres de T et les espaces propres correspondants. Comme $E_\lambda = \ker(T - \lambda I)$ pour chaque valeur propre λ, le problème se réduit à la détermination des valeurs propres de T. Pour ce faire, on s'appuie sur le résultat suivant :

Proposition 3.2.4. *Soit $\mathcal{B}$ une base de V. Le polynôme*

$$\mathrm{car}_T(x) := \det(x\, I - [T]_{\mathcal{B}}) \in K[x]$$

ne dépend pas du choix de $\mathcal{B}$. Les valeurs propres de T sont les racines de ce polynôme dans K.

Avant de passer à la démonstration de ce résultat, on note d'abord que l'expression $\det(x\, I - [T]_{\mathcal{B}})$ est bien un élément de $K[x]$, c'est-à-dire un polynôme en x à coefficients dans K. Pour le voir, écrivons $[T]_{\mathcal{B}} = (a_{ij})$. Alors on a

$$\begin{aligned} \det(x\, I - [T]_{\mathcal{B}}) &= \det \begin{pmatrix} x - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x - a_{nn} \end{pmatrix} \\ &= (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}) + (\text{termes de degré} \leq n - 2) \\ &= x^n - (a_{11} + \cdots + a_{nn}) x^{n-1} + \cdots \end{aligned} \tag{3.3}$$

comme le montre la formule pour le développement du déterminant rappelée à la section 3.1. Ce polynôme s'appelle le *polynôme caractéristique de T*, d'où la notation $\mathrm{car}_T(X)$.

Preuve de la proposition 3.2.4. Soit $\mathcal{B}'$ une autre base de V. On sait que

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}\,P$$

où $P = [I]_{\mathcal{B}}^{\mathcal{B}'}$. On en déduit que

$$xI - [T]_{\mathcal{B}'} = xI - P^{-1}[T]_{\mathcal{B}}\,P = P^{-1}(xI - [T]_{\mathcal{B}})\,P,$$

donc

$$\begin{aligned}\det(xI - [T]_{\mathcal{B}'}) &= \det\left(P^{-1}(xI - [T]_{\mathcal{B}})\,P\right)\\ &= \det(P)^{-1}\det(xI - [T]_{\mathcal{B}})\det(P)\\ &= \det(xI - [T]_{\mathcal{B}})\end{aligned} \tag{3.4}$$

Cela démontre que le polynôme $\det(xI - [T]_{\mathcal{B}})$ est indépendant du choix de $\mathcal{B}$. Notez que les calculs (3.4) font intervenir des matrices à coefficients polynomiaux dans $K[x]$, et la multiplicativité du déterminant pour un produit de telles matrices. Cela est justifié dans l'appendice B. Si le corps K est infini, on peut contourner cette difficulté en identifiant les polnômes de $K[x]$ aux fonctions polynomiales de K dans K (voir la section 4.2).

Enfin soit $\lambda \in K$. Par définition, λ est une valeur propre de T si et seulement si

$$\begin{aligned}\ker(T - \lambda I) \neq \{\mathbf{0}\} &\iff \text{il existe } \mathbf{v} \in V \setminus \{\mathbf{0}\} \text{ tel que } (T - \lambda\, I)\mathbf{v} = \mathbf{0}\,,\\ &\iff \text{il existe } X \in K^n \setminus \{\mathbf{0}\} \text{ tel que } [T - \lambda\, I]_{\mathcal{B}}X = \mathbf{0}\,,\\ &\iff \det([T - \lambda\, I]_{\mathcal{B}}) = 0\,.\end{aligned}$$

Puisque $[T - \lambda I]_{\mathcal{B}} = [T]_{\mathcal{B}} - \lambda[I]_{\mathcal{B}} = [T]_{\mathcal{B}} - \lambda I = -(\lambda I - [T]_{\mathcal{B}})$, on a aussi

$$\det([T - \lambda I]_{\mathcal{B}}) = (-1)^n \mathrm{car}_T(\lambda)\,.$$

Donc les valeurs propres de T sont les racines de $\mathrm{car}_T(x)$. □

Corollaire 3.2.5. *Soit $\mathcal{B}$ une base de V. Écrivons $[T]_{\mathcal{B}} = (a_{ij})$. Alors le nombre*

$$\mathrm{trace}(T) := a_{11} + a_{22} + \cdots + a_{nn}$$

appelé la trace *de T ne dépend pas du choix de $\mathcal{B}$. On a*

$$\mathrm{car}_T(x) = x^n - \mathrm{trace}(T)\,x^{n-1} + \cdots + (-1)^n \det(T).$$

Preuve. Les calculs (3.3) montrent que le coefficient de x^{n-1} dans $\det(x\,I - [T]_{\mathcal{B}})$ est égal à $-(a_{11} + \cdots + a_{nn})$. Comme $\mathrm{car}_T(x) = \det(xI - [T]_{\mathcal{B}})$ ne dépend pas du choix de $\mathcal{B}$, la somme $a_{11} + \cdots + a_{nn}$ n'en dépend pas non plus. Cela démontre la première assertion du corollaire. Pour la seconde, il suffit de montrer que le coefficient constant de $\mathrm{car}_T(x)$ est $(-1)^n \det(T)$. Or ce coefficient est

$$\mathrm{car}_T(0) = \det(-[T]_{\mathcal{B}}) = (-1)^n \det([T]_{\mathcal{B}}) = (-1)^n \det(T).$$

□

Exemple 3.2.6. Avec les notations de l'exemple 2.6.5, on pose $T = D|_U$ où U est le sous-espace de $\mathcal{C}_\infty(\mathbb{R})$ engendré par les fonctions $f_1(x) = e^{2x}$, $f_2(x) = x\,e^{2x}$ et $f_3(x) = x^2e^{2x}$, et où D est la dérivation sur $\mathcal{C}_\infty(\mathbb{R})$. On a vu que $\mathcal{B} = \{f_1, f_2, f_3\}$ est une base de U et que

$$[\,T\,]_{\mathcal{B}} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

Le polynôme caractéristique de T est donc

$$\mathrm{car}_T(x) = \det(xI - [\,T\,]_{\mathcal{B}}) = \begin{vmatrix} x-2 & -1 & 0 \\ 0 & x-2 & -2 \\ 0 & 0 & x-2 \end{vmatrix} = (x-2)^3.$$

Comme $x = 2$ est la seule racine de $(x-2)^3$ dans $\mathbb{R}$, on en déduit que 2 est la seule valeur propre de T. L'espace propre de T pour cette valeur propre est

$$E_2 = \ker(T - 2I).$$

Soit $f \in U$. On a

$$f \in E_2 \iff (T-2I)(f) = 0 \iff [\,T-2I\,]_{\mathcal{B}}[f]_{\mathcal{B}} = 0 \iff ([\,T\,]_{\mathcal{B}} - 2I)[f]_{\mathcal{B}} = 0$$

$$\iff \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}[f]_{\mathcal{B}} = 0 \iff [f]_{\mathcal{B}} = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix} \text{ avec } t \in \mathbb{R}$$

$$\iff f = t\,f_1 \quad \text{avec } t \in \mathbb{R}\,.$$

Donc $E_2 = \langle f_1 \rangle_{\mathbb{R}}$ est de dimension 1. Comme

$$\dim_{\mathbb{R}}(E_2) = 1 \neq 3 = \dim_{\mathbb{R}}(U),$$

le théorème 3.2.3 nous apprend que T n'est pas diagonalisable.

Exercices.

3.2.1. Soit V un espace vectoriel sur un corps K, soit $T\colon V \to V$ un opérateur linéaire sur V, et soient $\mathbf{v}_1, \ldots, \mathbf{v}_s$ des vecteurs propres de T pour des valeurs propres distinctes $\lambda_1, \ldots, \lambda_s \in K$. Montrer que $\mathbf{v}_1, \ldots, \mathbf{v}_s$ sont linéairement indépendants sur K.

Indice. On peut par exemple combiner la proposition 3.2.2 à l'exercice 1.4.3 (ii).

3.2.2. Soit $V = \langle\, f_0, f_1, \ldots, f_n \,\rangle_{\mathbb{R}}$ le sous-espace de $\mathcal{C}_\infty(\mathbb{R})$ engendré par les fonctions $f_i(x) = e^{ix}$ pour $i = 0, \ldots, n$.

(i) Montrer que V est invariant sous l'opérateur D de dérivation dans $\mathcal{C}_\infty(\mathbb{R})$ et que, pour $i = 0, \ldots, n$, la fonction f_i est un vecteur propre de D pour la valeur propre i.
(ii) En utilisant l'exercice 3.2.1, en déduire que $\mathcal{B} = \{f_0, f_1, \ldots, f_n\}$ est une base de V.
(iii) Calculer $[\, D|_V \,]_{\mathcal{B}}$.

3.2.3. On considère le sous-espace $\mathcal{P}_2(\mathbb{R})$ de $\mathcal{C}_\infty(\mathbb{R})$ donné par $\mathcal{P}_2(\mathbb{R}) = \langle\, 1, x, x^2 \,\rangle_{\mathbb{R}}$. Dans chaque cas, déterminer si l'application linéaire donnée est diagonalisable et, si elle l'est, construire une base dans laquelle sa matrice soit diagonale.
(i) $T\colon \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ donnée par $T(p(x)) = p'(x)$.
(ii) $T\colon \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ donnée par $T(p(x)) = p(2x)$.

3.2.4. Soit $T\colon \mathrm{Mat}_{2\times 2}(\mathbb{R}) \to \mathrm{Mat}_{2\times 2}(\mathbb{R})$ l'application linéaire donnée par $T(A) = A^t$.
(i) Déterminer une base des espaces propres de T pour les valeurs propres 1 et -1.
(ii) En déduire une base $\mathcal{B}$ de $\mathrm{Mat}_{2\times 2}(\mathbb{R})$ telle que $[\, T \,]_{\mathcal{B}}$ soit diagonale et donner $[\, T \,]_{\mathcal{B}}$.

3.2.5. Soit $S\colon \mathbb{R}^3 \to \mathbb{R}^3$ la symétrie par rapport au plan d'équation $x + y + z = 0$.
(i) Décrire géométriquement les espaces propres de S et donner une base de chacun.
(ii) Déterminer une base $\mathcal{B}$ de $\mathbb{R}^3$ relative à laquelle la matrice de S est diagonale et donner $[\, S \,]_{\mathcal{B}}$.
(iii) En déduire la matrice de T dans la base canonique.

3.3 Diagonalisation des matrices

Comme on l'a mentionné déjà, les résultats sur les opérateurs se transposent aux matrices et vice versa. Le problème de la diagonalisation en fournit un bon exemple.

Dans tout ce paragraphe, on fixe un entier $n \geq 1$.

- On dit qu'une matrice $A \in \mathrm{Mat}_{n\times n}(K)$ est *diagonalisable* s'il existe une matrice inversible $P \in \mathrm{Mat}_{n\times n}(K)$ telle que $P^{-1}\, A\, P$ soit une matrice diagonale.
- On dit qu'un élément λ de K est une *valeur propre* de A s'il existe un vecteur-colonne $X \in K^n$ non nul tel que $AX = \lambda X$.
- Un vecteur-colonne $X \in K^n$ non nul qui satisfait la relation $AX = \lambda X$ pour un scalaire $\lambda \in K$ s'appelle un *vecteur propre de A pour la valeur propre λ*.
- Si λ est une valeur propre de A, alors l'ensemble

$$\begin{aligned} E_\lambda(A) &:= \{X \in K^n\,;\, A\,X = \lambda\, X\} \\ &= \{X \in K^n\,;\, (A - \lambda I)\, X = \mathbf{0}\} \end{aligned}$$

est un sous-espace de K^n appelé l'*espace propre de A pour la valeur propre λ*.

- Le polynôme
$$\mathrm{car}_A(x) := \det(x\,I - A) \in K[x]$$
s'appelle le *polynôme caractéristique* de A.

Rappelons qu'à tout $A \in \mathrm{Mat}_{n\times n}(K)$, on associe l'opérateur linéaire $T_A\colon K^n \to K^n$ donné par
$$T_A(X) = A\,X \quad \text{pour tout} \quad X \in K^n.$$
Alors on voit que les définitions ci-dessus coïncident avec les notions correspondantes pour l'opérateur T_A. Plus précisément, on note que :

Proposition 3.3.1. *Soit $A \in \mathrm{Mat}_{n\times n}(K)$.*

(i) Les valeurs propres de T_A coïncident avec celles de A.

(ii) Si λ est une valeur propre de T_A, alors les vecteurs propres de T_A pour cette valeur propre sont les mêmes que ceux de A et on a $E_\lambda(T_A) = E_\lambda(A)$.

(iii) $\mathrm{car}_{T_A}(x) = \mathrm{car}_A(x)$.

Preuve. Les énoncés (i) et (ii) découlent directement des définitions tandis que (iii) découle du fait que $[\,T_A\,]_{\mathcal{E}} = A$ où $\mathcal{E}$ représente la base canonique de K^n. □

On note aussi :

Proposition 3.3.2. *Soit $A \in \mathrm{Mat}_{n\times n}(K)$. Les conditions suivantes sont équivalentes :*

(i) T_A est diagonalisable,

(ii) il existe une base de K^n constituée de vecteurs propres de A,

(iii) A est diagonalisable.

De plus, si ces conditions sont remplies et si $\mathcal{B} = \{X_1, \ldots, X_n\}$ est une base de K^n constituée de vecteurs propres de A, alors la matrice $P = (X_1 \cdots X_n)$ est inversible et
$$P^{-1}\,A\,P = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$
où λ_i est la valeur propre de A associée à X_i pour $i = 1, \ldots, n$.

Preuve. Si T_A est diagonalisable, la proposition 3.2.1 montre qu'il existe une base $\mathcal{B} = \{X_1, \ldots, X_n\}$ de K^n constituée de vecteurs propres de T_A, ou de A c'est pareil. Pour une telle base, la matrice $P = [\,I\,]_{\mathcal{E}}^{\mathcal{B}} = (X_1 \cdots X_n)$ de passage de la base $\mathcal{B}$ à la base canonique $\mathcal{E}$ est inversible et on trouve que
$$P^{-1}\,A\,P = P^{-1}[\,T_A\,]_{\mathcal{E}}P = [\,T_A\,]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

où λ_i est la valeur propre de A associée à X_i pour $i = 1, \ldots, n$. Cela démontre les implications $(i) \Rightarrow (ii) \Rightarrow (iii)$ ainsi que la dernière affirmation de la proposition. Pour conclure il suffit de montrer l'implication $(iii) \Rightarrow (i)$.

Supposons donc que A soit diagonalisable et soit $P \in \mathrm{Mat}_{n\times n}(K)$ une matrice inversible telle que $P^{-1}AP$ soit diagonale. Comme P est inversible, ses colonnes forment une base $\mathcal{B}$ de K^n et, pour ce choix de base, on a $P = [\,I\,]_{\mathcal{E}}^{\mathcal{B}}$. Alors $[\,T_A\,]_{\mathcal{B}} = P^{-1}[\,T_A\,]_{\mathcal{E}}P = P^{-1}AP$ est diagonale, ce qui montre bien que T_A est diagonalisable. □

En combinant le théorème 3.2.3 aux deux propositions précédentes, on obtient

Théorème 3.3.3. *Soit $A \in \mathrm{Mat}_{n\times n}(K)$. Alors A possède au plus n valeurs propres dans K. Soient $\lambda_1, \ldots, \lambda_s$ les valeurs propres distinctes de A dans K. Alors A est diagonalisable si et seulement si*

$$n = \dim_K(E_{\lambda_1}(A)) + \cdots + \dim_K(E_{\lambda_s}(A))\,.$$

Si tel est le cas, on obtient une base $\mathcal{B} = \{X_1, \ldots, X_n\}$ de K^n constituée de vecteurs propres de A en concaténant des bases de $E_{\lambda_1}(A), \ldots, E_{\lambda_s}(A)$. Alors la matrice $P = (X_1 \cdots X_n)$ dont les colonnes sont $X_1, \ldots, X_n$ est inversible et

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & 0 & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_s & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_s \end{pmatrix}$$

chaque λ_i étant répété un nombre de fois égal à $\dim_K(E_{\lambda_i}(A))$.

Exemple 3.3.4. Soit $A = \begin{pmatrix} \bar{1} & \bar{2} & \bar{1} \\ \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{2} & \bar{1} \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{F}_3)$ où $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$ représente le corps fini à 3 éléments. On se donne le programme suivant :

1° Déterminer les valeurs propres de A dans $\mathbb{F}_3$.

2° Pour chacune d'elles, déterminer une base de l'espace propre correspondant.

3° Déterminer si A est diagonalisable. Si oui, former une matrice inversible P telle que $P^{-1}AP$ soit diagonale et donner cette matrice diagonale.

Solution: 1° On trouve

$$\mathrm{car}_A(x) = \begin{vmatrix} x-\bar{1} & -\bar{2} & -\bar{1} \\ \bar{0} & x & \bar{0} \\ -\bar{1} & -\bar{2} & x-\bar{1} \end{vmatrix} = x \begin{vmatrix} x-\bar{1} & -\bar{1} \\ -\bar{1} & x-\bar{1} \end{vmatrix} \quad \text{(développement par rapport à la deuxième ligne)}$$

$$= x(x^2 + x) = x^2(x + \bar{1}),$$

donc les valeurs propres de A sont $\bar{0}$ et $-\bar{1} = \bar{2}$.

2° On a

$$E_{\bar{0}}(A) = \{X \in \mathbb{F}_3^3\,;\, (A - \bar{0}I)X = \mathbf{0}\} = \{X \in \mathbb{F}_3^3\,;\, AX = \mathbf{0}\}.$$

En d'autres termes $E_{\bar{0}}(A)$ est l'ensemble-solution dans $\mathbb{F}_3^3$ du système d'équations linéaires homogènes $AX = \mathbf{0}$. Ce dernier ne change pas si on applique à la matrice A du système une suite d'opérations élémentaires de lignes :

$$A = \begin{pmatrix} \bar{1} & \bar{2} & \bar{1} \\ \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{2} & \bar{1} \end{pmatrix} \sim \begin{pmatrix} \bar{1} & \bar{2} & \bar{1} \\ \bar{0} & \bar{0} & \bar{0} \\ \bar{0} & \bar{0} & \bar{0} \end{pmatrix} \leftarrow L_3 - L_1$$

Donc, $E_{\bar{0}}(A)$ est l'ensemble-solution de l'équation

$$x + \bar{2}y + z = \bar{0}\,.$$

En posant $y = s$ et $z = t$, on trouve $x = s - t$, donc

$$E_{\bar{0}}(A) = \left\{ \begin{pmatrix} s-t \\ s \\ t \end{pmatrix} ;\, s,t \in \mathbb{F}_3^3 \right\} = \left\langle \begin{pmatrix} \bar{1} \\ \bar{1} \\ \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{2} \\ \bar{0} \\ \bar{1} \end{pmatrix} \right\rangle_{\mathbb{F}_3}.$$

Ainsi $\mathcal{B}_1 = \left\{ \begin{pmatrix} \bar{1} \\ \bar{1} \\ \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{2} \\ \bar{0} \\ \bar{1} \end{pmatrix} \right\}$ est une base de $E_0(A)$ (il est clair que les deux vecteurs-colonnes sont linéairement indépendants).

De même, on a :

$$E_{\bar{2}}(A) = \{X \in \mathbb{F}_3^3\,;\, (A - \bar{2}I)X = \mathbf{0}\} = \{X \in \mathbb{F}_3^3\,;\, (A + I)X = \mathbf{0}\}\,.$$

On trouve

$$\begin{aligned} A + I &= \begin{pmatrix} \bar{2} & \bar{2} & \bar{1} \\ \bar{0} & \bar{1} & \bar{0} \\ \bar{1} & \bar{2} & \bar{2} \end{pmatrix} \sim \begin{pmatrix} \bar{1} & \bar{1} & \bar{2} \\ \bar{0} & \bar{1} & \bar{0} \\ \bar{1} & \bar{2} & \bar{2} \end{pmatrix} \leftarrow \bar{2}L_1 \sim \begin{pmatrix} \bar{1} & \bar{1} & \bar{2} \\ \bar{0} & \bar{1} & \bar{0} \\ \bar{0} & \bar{1} & \bar{0} \end{pmatrix} \leftarrow L_3 - L_1 \\ &\sim \begin{pmatrix} \bar{1} & \bar{0} & \bar{2} \\ \bar{0} & \bar{1} & \bar{0} \\ \bar{0} & \bar{0} & \bar{0} \end{pmatrix} \begin{matrix} \leftarrow L_1 - L_2 \\ \\ \leftarrow L_3 - L_2 \end{matrix} \end{aligned}$$

Donc, $E_{\bar{2}}(A)$ est l'ensemble-solution du système

$$\begin{cases} x & & + \bar{2}z &= \bar{0} \\ & y & &= \bar{0}. \end{cases}$$

Ici, z est la seule variable indépendante. On pose $z = t$. Alors $x = -\bar{2}t = t$ et $y = \bar{0}$, donc

$$E_{\bar{2}}(A) = \left\{ \begin{pmatrix} t \\ \bar{0} \\ t \end{pmatrix} ; t \in \mathbb{F}_3 \right\} = \left\langle \begin{pmatrix} \bar{1} \\ \bar{0} \\ \bar{1} \end{pmatrix} \right\rangle_{\mathbb{F}_3}$$

et par suite $\mathcal{B}_2 = \left\{ \begin{pmatrix} \bar{1} \\ \bar{0} \\ \bar{1} \end{pmatrix} \right\}$ est une base de $E_{\bar{2}}(A)$.

3° Comme A est une matrice 3×3 et que

$$\dim_{\mathbb{F}_3} E_{\bar{0}}(A) + \dim_{\mathbb{F}_3} E_{\bar{2}}(A) = 3,$$

on conclut que A est diagonalisable, et que

$$\mathcal{B} = \mathcal{B}_1 \amalg \mathcal{B}_2 = \left\{ \begin{pmatrix} \bar{1} \\ \bar{1} \\ \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{2} \\ \bar{0} \\ \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} \\ \bar{0} \\ \bar{1} \end{pmatrix} \right\}$$

est une base de $\mathbb{F}_3^3$ constituée de vecteurs propres de A pour les valeurs propres respectives $\bar{0}, \bar{0}, \bar{2}$. Alors la matrice

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

est inversible et le produit $P^{-1}AP$ est une matrice diagonale de diagonale égale à $(\bar{0}, \bar{0}, \bar{2})$:

$$P^{-1}AP = \begin{pmatrix} \bar{0} & \bar{0} & \bar{0} \\ \bar{0} & \bar{0} & \bar{0} \\ \bar{0} & \bar{0} & \bar{2} \end{pmatrix}.$$

Exercices.

3.3.1. Dans chaque cas, déterminer les valeurs propres réelles de la matrice $A \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$ donnée et, pour chacune de ces valeurs propres, donner une base de l'espace propre correspondant. Si A est diagonalisable, déterminer une matrice inversible $P \in \mathrm{Mat}_{3\times 3}(\mathbb{R})$ telle que $P^{-1}AP$ soit diagonale, et donner $P^{-1}AP$.

$$\text{(i)} \quad A = \begin{pmatrix} -2 & 0 & 5 \\ 0 & 5 & 0 \\ -4 & 0 & 7 \end{pmatrix}, \qquad \text{(ii)} \quad A = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 6 & -6 \\ 1 & 2 & -1 \end{pmatrix}.$$

3.3.2. La matrice $A = \begin{pmatrix} 1 & -2 \\ 4 & -3 \end{pmatrix}$ n'est pas diagonalisable dans les réels. Par contre, elle l'est dans les complexes. Déterminer une matrice inversible P à coefficients complexes telle que $P^{-1}AP$ soit diagonale, et donner $P^{-1}AP$.

3.3.3. Soit $\mathbb{F}_2 = \{\bar{0}, \bar{1}\}$ le corps à 2 éléments et soit $A = \begin{pmatrix} \bar{1} & \bar{1} & \bar{1} \\ \bar{1} & \bar{0} & \bar{0} \\ \bar{0} & \bar{0} & \bar{1} \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{F}_2)$. Déterminer les valeurs propres de A dans $\mathbb{F}_2$ et, pour chacune d'elles, donner une base de l'espace propre correspondant. Si A est diagonalisable, déterminer une matrice inversible $P \in \mathrm{Mat}_{3\times 3}(\mathbb{F}_2)$ telle que $P^{-1}AP$ soit diagonale, et donner $P^{-1}AP$.

Chapitre 4

Polynômes, opérateurs linéaires et matrices

Soit V un espace vectoriel de dimension finie n sur un corps K. Dans ce chapitre, on montre que l'ensemble $\mathrm{End}_K(V)$ des opérateurs linéaires sur V est un anneau sous l'addition et la composition, isomorphe à l'anneau $\mathrm{Mat}_{n\times n}(K)$ des matrices $n \times n$ à coefficients dans K. On construit l'anneau $K[x]$ des polynômes en une indéterminée x sur le corps K. Pour chaque opérateur linéaire $T \in \mathrm{End}_K(V)$ et chaque matrice $A \in \mathrm{Mat}_{n\times n}(K)$, on définit des homomorphismes d'anneaux

$$\begin{array}{ccc} K[x] & \longrightarrow & \mathrm{End}_K(V) \\ P(x) & \longmapsto & P(T) \end{array} \qquad \text{et} \qquad \begin{array}{ccc} K[x] & \longrightarrow & \mathrm{Mat}_{n\times n}(K) \\ P(x) & \longmapsto & P(A) \end{array}$$

qui appliquent respectivement x sur T et x sur A. Le lien entre eux est que, si $\mathcal{B}$ est une base de V, alors $[P(T)]_{\mathcal{B}} = P([\,T\,]_{\mathcal{B}})$.

4.1 L'anneau des opérateurs linéaires

On renvoie le lecteur à l'appendice A pour un rappel des notions de base relatives aux anneaux. On montre d'abord :

Théorème 4.1.1. *Soit V un espace vectoriel de dimension n sur un corps K et soit $\mathcal{B}$ une base de V. L'ensemble $\mathrm{End}_K(V)$ des opérateurs linéaires sur V est un anneau sous l'addition et la composition. De plus, l'application*

$$\begin{array}{rccc} \varphi : & \mathrm{End}_K(V) & \longrightarrow & \mathrm{Mat}_{n\times n}(K) \\ & T & \longmapsto & [\,T\,]_{\mathcal{B}} \end{array}$$

est un isomorphisme d'anneaux.

Preuve. On sait déjà, d'après le théorème 2.2.3, que l'ensemble $\mathrm{End}_K(V) = \mathcal{L}_K(V,V)$ est un espace vectoriel sur K. En particulier, c'est un groupe abélien pour l'addition. On sait aussi

que la composée de deux opérateurs linéaires $S\colon V \to V$ et $T\colon V \to V$ est un opérateur linéaire $S \circ T\colon V \to V$, donc un élément de $\mathrm{End}_K(V)$. Comme la composition est associative et que l'application identité $I_V \in \mathrm{End}_K(V)$ est élément neutre pour cette opération, on conclut que $\mathrm{End}_K(V)$ est un monoïde pour la composition. Enfin la proposition 2.3.1 montre que si $R, S, T \in \mathrm{End}_K(V)$, alors

$$(R+S)\circ T = R\circ T + S\circ T \quad \text{et} \quad R\circ(S+T) = R\circ S + R\circ T.$$

Donc la composition est distributive sur l'addition dans $\mathrm{End}_K(V)$ et par suite, $\mathrm{End}_K(V)$ est un anneau.

Par ailleurs, le théorème 2.4.3 montre que l'application φ est un isomorphisme d'espaces vectoriels sur K. En particulier, c'est un isomorphisme de groupes abéliens sous l'addition. Enfin la proposition 2.4.6 montre que si $S, T \in \mathrm{End}_K(V)$, alors

$$\varphi(S\circ T) = [\,S\circ T\,]_{\mathcal{B}} = [\,S\,]_{\mathcal{B}}[\,T\,]_{\mathcal{B}} = \varphi(S)\varphi(T)\,.$$

Comme $\varphi(I_V) = I$, on conclut que φ est aussi un isomorphisme d'anneaux. □

Exercices.

4.1.1. Soit V un espace vectoriel sur un corps K et soit U un sous-espace de V. Montrer que l'ensemble des applications linéaires $T\colon V \to V$ telles que $T(\mathbf{u}) \in U$ pour tout $\mathbf{u} \in U$ est un sous-anneau de $\mathrm{End}_K(V)$.

4.1.2. Soit $A = \{a + b\sqrt{2}\,;\, a, b \in \mathbb{Q}\}$.

(i) Montrer que A est un sous-anneau de $\mathbb{R}$ et que $\mathcal{B} = \{1, \sqrt{2}\}$ est une base de A vu comme espace vectoriel sur $\mathbb{Q}$.

(ii) Soit $\alpha = a + b\sqrt{2} \in A$. Montrer que l'application $m_\alpha\colon A \to A$ définie par $m_\alpha(\beta) = \alpha\beta$ pour tout $\beta \in A$ est une application $\mathbb{Q}$-linéaire et calculer $[m_\alpha]_{\mathcal{B}}$.

(iii) Montrer que l'application $\varphi\colon A \to \mathrm{Mat}_{2\times 2}(\mathbb{Q})$ donnée par $\varphi(\alpha) = [m_\alpha]_{\mathcal{B}}$ est un homomorphisme d'anneaux injectif, et décrire son image.

4.2 L'anneau des polynômes

Proposition 4.2.1. *Soit A un anneau commutatif et soit x un symbole en dehors de A. Il existe un anneau commutatif, noté $A[x]$, contenant A comme sous-anneau et x comme élément tel que*

(i) tout élément de $A[x]$ s'écrive $a_0 + a_1\,x + \cdots + a_n\,x^n$ pour un entier $n \geq 0$ et des éléments $a_0, a_1, \ldots, a_n$ de A ;

(ii) *si* $a_0 + a_1\,x + \cdots + a_n\,x^n = 0$ *pour un entier* $n \geq 0$ *et des éléments* $a_0, a_1, \ldots, a_n$ *de* A, *alors* $a_0 = a_1 = \cdots = a_n = 0$.

Cet anneau $A[x]$ s'appelle l'*anneau des polynômes en l'indéterminée* x *sur* A. La preuve de ce résultat est un peu technique. L'étudiant pourra l'omettre en première lecture. Par contre, il est important de bien savoir manipuler les éléments de cet anneau, appelés *polynômes en* x *à coefficients dans* A.

Supposons l'existence de cet anneau $A[x]$ et soient

$$p = a_0 + a_1\,x + \cdots + a_m\,x^m \quad \text{et} \quad q = b_0 + b_1\,x + \cdots + b_n\,x^n$$

deux de ses éléments. On note d'abord que la condition (ii) implique

$$p = q \quad \Longleftrightarrow \quad \begin{cases} a_i = b_i & \text{pour} \quad i = 0, 1, \ldots, \min(m,n), \\ a_i = 0 & \text{pour} \quad i > n, \\ b_i = 0 & \text{pour} \quad i > m. \end{cases}$$

En définissant $a_i = 0$ pour chaque $i > m$ et $b_i = 0$ pour chaque $i > n$, on obtient aussi

$$p + q = \sum_{i=0}^{\max(n,m)} a_i x^i + \sum_{i=0}^{\max(n,m)} b_i x^i = \sum_{i=0}^{\max(n,m)} (a_i + b_i)\,x^i,$$

$$p \cdot q = \left(\sum_{i=0}^{m} a_i x^i\right) \cdot \left(\sum_{i=0}^{n} b_i x^i\right) = \sum_{i=0}^{m}\sum_{j=0}^{n} a_i b_j\,x^{i+j} = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i b_{k-i}\right) x^k.$$

Exemple 4.2.2. Sur le corps $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$ à 3 éléments, on a

$$\begin{aligned} (\bar{1} + x^2 + \bar{2}x^3) + (x + \bar{2}x^2) &= \bar{1} + x + \bar{2}x^3, \\ (\bar{1} + x^2 + \bar{2}x^3)(x + \bar{2}x^2) &= x + \bar{2}x^2 + x^3 + x^4 + x^5. \end{aligned}$$

Preuve de la proposition 4.2.1. Soit S l'ensemble des suites $(a_0, a_1, a_2, \ldots)$ d'éléments de A dont toutes les composantes sont nulles à partir d'un certain rang. On définit une "addition" sur S en posant

$$(a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots).$$

Le résultat est bien un élément de S, car si $a_i = b_i = 0$ pour tout $i > n$, alors $a_i + b_i = 0$ pour tout $i > n$. On vérifie sans peine que cette opération est associative et commutative, qu'elle admet pour élément neutre la suite $(0, 0, 0, \ldots)$, et qu'un élément quelconque $(a_0, a_1, a_2, \ldots)$ de S admet pour inverse additif la suite $(-a_0, -a_1, -a_2, \ldots)$. Donc S est un groupe abélien sous cette opération.

On définit aussi une "multiplication" sur S en posant

$$(a_0, a_1, a_2, \ldots)(b_0, b_1, b_2, \ldots) = (t_0, t_1, t_2, \ldots) \quad \text{où} \quad t_k = \sum_{i=0}^{k} a_i b_{k-i}.$$

Le résultat est un élément de S, car si on suppose encore $a_i = b_i = 0$ pour tout $i > n$, alors, pour tout entier $k > 2n$, chacun des produits $a_i b_{k-i}$ avec $i = 0, \dots, k$ est nul (un des indices i ou $k-i$ étant forcément $> n$) et par suite $t_k = 0$. Il est facile de voir que la suite $(1, 0, 0, \dots)$ est un élément neutre pour cette opération.

Pour conclure que S est un anneau sous ces opérations, il reste à montrer que la multiplication est associative et distributive sur l'addition. En d'autres termes, on doit montrer que si

$$p = (a_0, a_1, a_2, \dots), \quad q = (b_0, b_1, b_2, \dots) \quad \text{et} \quad r = (c_0, c_1, c_2, \dots)$$

sont trois éléments de S, alors

1) $(pq)r = p(qr)$
2) $(p+q)r = p\,r + q\,r$
3) $p(q+r) = p\,q + p\,r$.

Les trois égalités se démontrent sensiblement de la même manière : il suffit de montrer que les suites de part et d'autre de l'égalité possèdent le même élément d'indice k pour chaque $k \geq 0$. L'égalité 1) est la plus délicate à vérifier car chaque membre de celle-ci implique deux produits. On trouve que l'élément d'indice k de $(pq)r$ est

$$((pq)r)_k = \sum_{j=0}^{k} (pq)_j c_{k-j} = \sum_{j=0}^{k} \left(\sum_{i=0}^{j} a_i b_{j-i} \right) c_{k-j} = \sum_{j=0}^{k} \sum_{i=0}^{j} a_i b_{j-i} c_{k-j}$$

et que celui de $p(qr)$ est

$$(p(qr))_k = \sum_{i=0}^{k} a_i (qr)_{k-i} = \sum_{i=0}^{k} a_i \left(\sum_{j=0}^{k-i} b_j c_{k-i-j} \right) = \sum_{i=0}^{k} \sum_{j=0}^{k-i} a_i b_j c_{k-i-j}.$$

Dans les deux cas, on trouve qu'il s'agit de la somme de tous les produits $a_i b_j c_l$ avec $i, j, l \geq 0$ et $i + j + l = k$, donc

$$((pq)r)_k = \sum_{i+j+l=k} a_i b_j c_l = (p(qr))_k.$$

Cela démontre 1). Les preuves de 2) et 3) sont plus simples et laissées en exercice.

Pour tout choix de $a, b \in A$, on trouve que

$$(a, 0, 0, \dots) + (b, 0, 0, \dots) = (a + b, 0, 0, \dots),$$
$$(a, 0, 0, \dots)(b, 0, 0, \dots) = (ab, 0, 0, \dots).$$

Donc l'application $\iota \colon A \to S$ donné par

$$\iota(a) = (a, 0, 0, \dots)$$

est un homorphisme d'anneaux injectif. A partir de maintenant, on identifie A à son image sous ι. Alors A devient un sous-anneau de S et on trouve

$$a(b_0, b_1, \dots) = (a\,b_0, a\,b_1, a\,b_2, \dots)$$

pour tout $a \in A$ et toute suite $(b_0, b_1, b_2, \dots) \in S$.

Posons $X = (0, 1, 0, 0, \dots)$. On vérifie sans peine par récurrence que, pour tout entier $i \geq 1$, la puissance i-ième X^i de X est la suite dont tous les éléments sont nuls sauf celui d'indice i qui est égal à 1 :

$$X^i = (0, \dots, 0, \overset{i}{\overset{\vee}{1}}, 0, \dots).$$

On en déduit :

1° Pour toute suite $(a_0, a_1, a_2, \dots)$ de S, en choisissant un entier $n \geq 0$ tel que $a_i = 0$ pour tout $i > n$, on a :

$$(a_0, a_1, a_2, \dots) = \sum_{i=0}^{n} (0, \dots, 0, \overset{i}{\overset{\vee}{a_i}}, 0, \dots) = \sum_{i=0}^{n} a_i \, (0, \dots, 0, \overset{i}{\overset{\vee}{1}}, 0, \dots) = \sum_{i=0}^{n} a_i \, X^i.$$

2° Par ailleurs, si n est un entier ≥ 0 et si $a_0, a_1, \dots, a_n \in A$ satisfont

$$a_0 + a_1 \, X + \cdots + a_n \, X^n = 0,$$

alors

$$0 = \sum_{i=0}^{n} a_i \, (0, \dots, 0, \overset{i}{\overset{\vee}{1}}, 0, \dots) = \sum_{i=0}^{n} (0, \dots, 0, \overset{i}{\overset{\vee}{a_i}}, 0, \dots) = (a_0, a_1, \dots, a_n, 0, 0, \dots),$$

et par suite $a_0 = a_1 = \cdots = a_n = 0$.

Ces deux dernières observations montrent que l'anneau S possède les deux propriétés requises par la proposition sauf qu'au lieu de l'élément x, on a la suite $X = (0, 1, 0, \dots)$. Cela ne pose pas de réel problème : il suffit de remplacer X par x dans S et aussi dans les tables d'addition et de multiplication de S. □

Chaque polynôme $p \in A[x]$ induit une fonction de A dans A. Avant de décrire cette construction, on note d'abord le fait suivant dont la preuve est laissée en exercice :

Proposition 4.2.3. *Soient X un ensemble et A un anneau commutatif. L'ensemble $\mathcal{F}(X, A)$ des fonctions de X dans A est un anneau commutatif pour l'addition et la multiplication définies de la manière suivante. Si $f, g \in \mathcal{F}(X, A)$ alors $f + g$ et $f\,g$ sont les fonctions de X dans A données par*

$$(f + g)(x) = f(x) + g(x) \quad \textit{et} \quad (f\,g)(x) = f(x)\,g(x)$$

pour tout $x \in X$.

On montre aussi :

Proposition 4.2.4. *Soit A un sous-anneau d'un anneau commutatif C et soit $c \in C$. L'application*

$$\begin{array}{rcl} A[x] & \longrightarrow & C \\ \displaystyle\sum_{i=0}^{n} a_i x^i & \longmapsto & \displaystyle\sum_{i=0}^{n} a_i c^i \end{array} \tag{4.1}$$

est un homomorphisme d'anneaux.

Cet homomorphisme est appelé l'*évaluation en c* et l'image d'un polynôme p sous cet homomorphisme est notée $p(c)$.

Preuve de la proposition 4.2.4. Soient

$$p = \sum_{i=0}^{n} a_i x^i \quad \text{et} \quad q = \sum_{i=0}^{m} b_i x^i$$

des éléments de $A[x]$. Définissons $a_i = 0$ pour tout $i > n$ et $b_i = 0$ pour tout $i > m$. Alors on a :

$$p + q = \sum_{i=0}^{\max(n,m)} (a_i + b_i)\, x^i \quad \text{et} \quad p\, q = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) x^k.$$

On trouve

$$p(c) + q(c) = \sum_{i=0}^{m} a_i c^i + \sum_{i=0}^{n} b_i c^i = \sum_{i=0}^{\max(n,m)} (a_i + b_i)\, c^i = (p+q)(c)$$

et

$$p(c)q(c) = \left(\sum_{i=0}^{m} a_i c^i \right) \left(\sum_{i=0}^{n} b_i c^i \right) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_i b_j\, c^{i+j} = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) c^k = (pq)(c).$$

Ces dernières égalités jointes au fait que l'image de $1 \in A[x]$ est 1 montrent que l'application (4.1) est bien un homomorphisme d'anneaux. □

En particulier, si on prend $C = A[x]$ et $c = x$, l'application 4.1 devient l'application identité :

$$\begin{array}{rcl} A[x] & \longrightarrow & A[x] \\ \displaystyle p = \sum_{i=0}^{n} a_i x^i & \longmapsto & \displaystyle p(x) = \sum_{i=0}^{n} a_i x^i \end{array}$$

Pour cette raison, on convient généralement d'écrire $p(x)$ pour désigner un polynôme de $A[x]$. C'est ce que l'on fera le plus souvent à partir de maintenant.

Proposition 4.2.5. *Soit A un anneau commutatif. Pour chaque $p \in A[x]$, désignons par $\Phi(p) : A \to A$ la fonction donnée par $\Phi(p)(a) = p(a)$ pour tout $a \in A$. L'application $\Phi : A[x] \to \mathcal{F}(A, A)$ ainsi définie est un homomorphisme d'anneaux de noyau*

$$\ker(\Phi) = \{p \in A[x] \,;\, p(a) = 0 \ \textit{pour tout } a \in A\}.$$

On dit que $\Phi(p)$ est la *fonction polynomiale de A dans A induite par p*. L'image de Φ dans $\mathcal{F}(A, A)$ s'appelle l'*anneau des fonctions polynomiales de A dans A*. De manière plus concise, on peut définir Φ de la manière suivante :

$$\begin{array}{rccl} \Phi \,:\, A[x] & \longrightarrow & \mathcal{F}(A,A) \\ p & \longmapsto & \Phi(p)\colon A \longrightarrow A \\ & & \qquad a \longmapsto p(a) \end{array}$$

Preuve. Soient $p, q \in A[x]$. Pour tout $a \in A$, on trouve

$$\begin{aligned} \Phi(p+q)(a) &= (p+q)(a) = p(a) + q(a) = \Phi(p)(a) + \Phi(q)(a), \\ \Phi(p\,q)(a) &= (p\,q)(a) = p(a)\,q(a) = \Phi(p)(a) \cdot \Phi(q)(a), \end{aligned}$$

donc $\Phi(p+q) = \Phi(p) + \Phi(q)$ et $\Phi(p\,q) = \Phi(p)\Phi(q)$. Comme $\Phi(1)$ est la fonction constante égale à 1 (l'élément neutre de $\mathcal{F}(A, A)$ pour le produit), on conclut que Φ est un homomorphisme d'anneaux. La dernière assertion concernant son noyau découle des définitions. □

Exemple 4.2.6. Si $A = \mathbb{R}$, l'application $\Phi\colon \mathbb{R}[x] \to \mathcal{F}(\mathbb{R}, \mathbb{R})$ définie par la proposition 4.2.5 est injective, car on sait qu'un polynôme non nul de $\mathbb{R}[x]$ admet un nombre fini de racines réelles et par suite $\ker \Phi = \{\mathbf{0}\}$. Dans ce cas, Φ induit un isomorphisme entre l'anneau $\mathbb{R}[x]$ des polynômes à coefficients dans $\mathbb{R}$ et son image dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'anneau des fonctions polynomiales de $\mathbb{R}$ dans $\mathbb{R}$.

Exemple 4.2.7. Posons $A = \mathbb{F}_2 = \{\bar{0}, \bar{1}\}$, le corps à deux éléments. L'anneau $\mathcal{F}(\mathbb{F}_2, \mathbb{F}_2)$ des fonctions de $\mathbb{F}_2$ dans lui-même comporte 4 éléments tandis $\mathbb{F}_2[x]$ contient une infinité de polynômes. Dans ce cas, l'application Φ n'est pas injective. Plus précisément, le polynôme $p(x) = x^2 - x$ satisfait $p(\bar{0}) = \bar{0}$ et $p(\bar{1}) = \bar{0}$, donc $p(x) \in \ker \Phi$.

Exercices.

4.2.1. Démontrer la proposition 4.2.3.

4.2.2. Montrer que l'application $\Phi\colon A[x] \to \mathcal{F}(A, A)$ de la proposition 4.2.5 n'est pas injective si A est un anneau fini.

4.3 Évaluation en un opérateur linéaire ou en une matrice

Soit V un espace vectoriel sur un corps K et soit $T\colon V \to V$ un opérateur linéaire sur V. Pour tout polynôme

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x],$$

on pose

$$P(T) = a_0 I + a_1 T + \cdots + a_n T^n \in \mathrm{End}_K(V),$$

où I désigne l'application identité de V.

Proposition 4.3.1. *Avec les notations ci-dessus, l'application*

$$\begin{array}{ccc} K[x] & \longrightarrow & \mathrm{End}_K(V) \\ p(x) & \longmapsto & p(T) \end{array}$$

est un homomorphisme d'anneaux. Son image, notée $K[T]$, est le sous-anneau commutatif de $\mathrm{End}_K(V)$ donné par

$$K[T] = \{a_0 I + a_1 T + \cdots + a_n T^n \,;\, n \in \mathbb{N}^*, a_0, \ldots, a_n \in K\}.$$

Cet homomorphisme est appelé l'*évaluation en T*.

Preuve. Par définition, l'image du polynôme constant 1 est l'élément neutre I de $\mathrm{End}_K(V)$. Pour conclure que l'évaluation en T est un homomorphisme d'anneaux, il reste à montrer que si $p, q \in K[x]$ sont des polynômes quelconques, alors

$$(p+q)(T) = p(T) + q(T) \quad \text{et} \quad (pq)(T) = p(T) \circ q(T).$$

La preuve pour la somme est laissée en exercice. Celle pour le produit utilise le fait que, pour tout $a \in K$ et tout $S \in \mathrm{End}_K(V)$, on a

$$(a\, I) \circ S = S \circ (a\, I) = a\, S.$$

Écrivons

$$p = \sum_{i=0}^{n} a_i\, x^i \quad \text{et} \quad q = \sum_{j=0}^{m} b_j\, x^j.$$

On trouve

$$\begin{aligned}
p(T)\circ q(T) &= \Big(a_0 I + \sum_{i=1}^{n} a_i\, T^i\Big) \circ \Big(b_0 I + \sum_{j=1}^{m} b_j\, T^i\Big) \\
&= (a_0 I)\circ(b_0 I) + (a_0 I)\circ\Big(\sum_{j=1}^{m} b_j\, T^j\Big) + \Big(\sum_{i=1}^{n} a_i\, T^i\Big)\circ(b_0 I) + \Big(\sum_{i=1}^{n} a_i\, T^i\Big)\circ\Big(\sum_{j=1}^{m} b_j\, T^j\Big) \\
&= a_0 b_0 I + \sum_{j=1}^{m} a_0 b_j\, T^j + \sum_{i=1}^{n} a_i b_0\, T^i + \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j\, T^{i+j} \\
&= (a_0 b_0) I + \sum_{k=1}^{n+m} \Big(\sum_{i+j=k} a_i b_j\Big) T^k \\
&= (pq)(T).
\end{aligned}$$

Donc l'évaluation en T est bien un homomorphisme d'anneaux. Le fait que son image soit un sous-anneau commutatif de $\mathrm{End}_K(V)$ découle du fait que $K[x]$ est commutatif. □

Exemple 4.3.2. Soit D l'opérateur de dérivation sur l'espace vectoriel $\mathcal{C}_\infty(\mathbb{R})$ des fonctions $f\colon \mathbb{R}\to\mathbb{R}$ qui sont indéfiniment différentiables. Posons

$$p(x) = x^2 - 1 \in \mathbb{R}[x] \quad \text{et} \quad q(x) = 2x + 1 \in \mathbb{R}[x].$$

Pour toute fonction $f \in \mathcal{C}_\infty(\mathbb{R})$, on trouve

$$\begin{aligned}
p(D)(f) &= (D^2 - I)(f) = f'' - f, \\
q(D)(f) &= (2D + I)(f) = 2f' + f.
\end{aligned}$$

D'après la théorie, on a $p(D)\circ q(D) = (pq)(D)$ pour tout choix de polynômes $p, q \in \mathbb{R}[X]$. Vérifions-le pour le choix de p et q ci-dessus. Cela revient à montrer que

$$(p(D)\circ q(D))(f) = ((pq)(D))(f)$$

pour toute fonction $f \in \mathcal{C}_\infty(\mathbb{R})$. On trouve

$$\begin{aligned}
\big(p(D)\circ q(D)\big)(f) &= p(D)\big(q(D)(f)\big) \\
&= p(D)(2f' + f) \\
&= (2f' + f)'' - (2f' + f) \\
&= 2f''' + f'' - 2f' - f.
\end{aligned}$$

Comme $pq = (x^2-1)(2x+1) = 2x^3 + x^2 - 2x - 1$, on a aussi

$$((pq)(D))(f) = (2D^3 + D^2 - 2D - I)(f) = 2f''' + f'' - 2f' - f$$

comme annoncé.

Exemple 4.3.3. Soit $\mathcal{F}(\mathbb{Z},\mathbb{R})$ l'espace vectoriel des fonctions $f\colon \mathbb{Z}\to\mathbb{R}$. On définit un opérateur linéaire T sur cet espace vectoriel de la manière suivante. Si $f\colon \mathbb{Z}\to\mathbb{R}$ est une fonction quelconque, alors $T(f)\colon \mathbb{Z}\to\mathbb{R}$ est la fonction donnée par

$$\big(T(f)\big)(i) = f(i+1) \quad \text{pour tout } i\in\mathbb{Z}.$$

On trouve

$$\begin{aligned}\big(T^2(f)\big)(i) &= \big(T(T(f))\big)(i) = (T(f))(i+1) = f(i+2),\\ \big(T^3(f)\big)(i) &= \big(T(T^2(f))\big)(i) = (T^2(f))(i+1) = f(i+3),\end{aligned}$$

et ainsi de suite... En général,

$$\big(T^k(f)\big)(i) = f(i+k) \quad \text{pour} \quad k=1,2,3,\ldots$$

On en déduit que si $p(x) = a_0 + a_1\,x + \cdots + a_n\,x^n \in \mathbb{R}[x]$, alors $p(T)(f)$ est la fonction de $\mathbb{Z}$ dans $\mathbb{R}$ donnée, pour tout $i\in\mathbb{Z}$, par

$$\begin{aligned}(p(T)(f))(i) &= ((a_0 I + a_1\,T + \cdots + a_n\,T^n)(f))(i)\\ &= (a_0 f + a_1\,T(f) + \cdots + a_n\,T^n(f))(i)\\ &= a_0 f(i) + a_1 T(f)(i) + \cdots + a_n T^n(f)(i)\\ &= a_0 f(i) + a_1 f(i+1) + \cdots + a_n f(i+n)\,.\end{aligned}$$

Soit $A\in \mathrm{Mat}_{n\times n}(K)$ une matrice carrée $n\times n$ à coefficients dans un corps K. Pour tout polynôme

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x],$$

on pose

$$p(A) = a_0 I + a_1 A + \cdots + a_n A^n \in \mathrm{Mat}_{n\times n}(K),$$

où I désigne la matrice identité $n\times n$.

Proposition 4.3.4. *Avec les notations ci-dessus, l'application*

$$\begin{array}{ccc} K[x] & \longrightarrow & \mathrm{Mat}_{n\times n}(K)\\ p(x) & \longmapsto & p(A)\end{array}$$

est un homomorphisme d'anneaux. Son image, notée $K[A]$, est le sous-anneau commutatif de $\mathrm{Mat}_{n\times n}(K)$ donné par

$$K[A] = \{a_0 I + a_1 A + \cdots + a_n A^n\,;\, n\in\mathbb{N},\, a_0,\ldots,a_n\in K\}.$$

Cet homomorphisme est appelé l'*évaluation en A*. La preuve de cette proposition est semblable à celle de la proposition 4.3.1 et laissée en exercice.

Exemple 4.3.5. Soit $A = \begin{pmatrix} 1 & 1\\ 0 & 1\end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{Q})$ et soit $p(x) = x^3 + 2x - 1 \in \mathbb{Q}[x]$. On trouve

$$p(A) = A^3 + 2A - I = \begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix} + 2\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} - \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}2 & 5\\ 0 & 2\end{pmatrix}.$$

On conclut avec le résultat suivant :

Théorème 4.3.6. *Soit V un espace vectoriel de dimension finie n, soit $\mathcal{B}$ une base de V, et soit $T\colon V \to V$ un opérateur linéaire sur V. Pour tout polynôme $p(x) \in K[x]$, on a*

$$[p(T)]_{\mathcal{B}} = p([\,T\,]_{\mathcal{B}}).$$

Preuve. Écrivons $p(x) = a_0 + a_1\,x + \cdots + a_n\,x^n$. On trouve

$$\begin{aligned}[p(T)]_{\mathcal{B}} &= [a_0 I + a_1 T + \cdots + a_n T^n]_{\mathcal{B}} \\ &= a_0 I + a_1 [T]_{\mathcal{B}} + \cdots + a_n [T^n]_{\mathcal{B}} \\ &= a_0 I + a_1 [T]_{\mathcal{B}} + \cdots + a_n ([T]_{\mathcal{B}})^n \\ &= p([\,T\,]_{\mathcal{B}}).\end{aligned}$$

□

Exemple 4.3.7. Soit $V = \langle \cos(2t), \sin(2t) \rangle_{\mathbb{R}} \subseteq \mathcal{C}_\infty(\mathbb{R})$ et soit D l'opérateur de dérivation sur $\mathcal{C}_\infty(\mathbb{R})$. Le sous-espace V est stable sous D car

$$D(\cos(2t)) = -2\sin(2t) \in V \quad \text{et} \quad D(\sin(2t)) = 2\cos(2t) \in V. \tag{4.2}$$

Pour des raisons de simplicité, désignons par D la restriction $D|_V$ de D à V. On note que

$$\mathcal{B} = \{\cos(2t), \sin(2t)\}$$

est une base de V (exercice). Alors les formules (4.2) donnent

$$[\,D\,]_{\mathcal{B}} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$

Par suite, on a

$$[p(D)]_{\mathcal{B}} = p\Big(\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}\Big)$$

pour tout $p(x) \in \mathbb{R}[x]$. En particulier, pour $p(x) = x^2 - 4$, on trouve $p([\,D\,]_{\mathcal{B}}) = 0$, donc $p(D) = 0$.

Exercices.

4.3.1. Démontrer l'égalité $(p+q)(T) = p(T) + q(T)$ laissée en exercice dans la preuve de la proposition 4.3.1.

4.3.2. Soit $R_{\pi/4}\colon \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $\pi/4$ autour de l'origine dans $\mathbb{R}^2$, et soit $p(x) = x^2 - 1 \in \mathbb{R}[x]$. Calculer $p(R_{\pi/4})(1,2)$.

4.3.3. Soit $R_{2\pi/3}\colon \mathbb{R}^2 \to \mathbb{R}^2$ la rotation d'angle $2\pi/3$ autour de l'origine dans $\mathbb{R}^2$, et soit $p(x) = x^2 + x + 1 \in \mathbb{R}[x]$. Montrer que $p(R_{2\pi/3}) = 0$.

4.3.4. Soit D la restriction à $V = \langle\, e^x, xe^x, x^2e^x \,\rangle_{\mathbb{R}}$ de l'opérateur de différentiation dans $\mathcal{C}_\infty(\mathbb{R})$.

(i) Montrer que $\mathcal{B} = \{e^x, xe^x, x^2e^x\}$ est une base de V et déterminer $[\,D\,]_{\mathcal{B}}$.

(ii) Soit $p(x) = x^2 - 1$ et $q(x) = x - 1$. Calculer $[p(D)]_{\mathcal{B}}$ et $[q(D)]_{\mathcal{B}}$. Est-ce que ces matrices commutent ? Pourquoi ?

4.3.5. Soit V un espace vectoriel sur $\mathbb{C}$ de dimension 3, soit $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ une base de V, et soit $T\colon V \to V$ l'application linéaire déterminée par les conditions

$$T(\mathbf{v}_1) = i\mathbf{v}_2 + \mathbf{v}_3, \quad T(\mathbf{v}_2) = i\mathbf{v}_3 + \mathbf{v}_1 \quad \text{et} \quad T(\mathbf{v}_3) = i\mathbf{v}_1 + \mathbf{v}_2.$$

(i) Calculer $[p(T)]_{\mathcal{B}}$ où $p(x) = x^2 + (1+i)x + i$.

(ii) Calculer $p(T)(\mathbf{v}_1)$.

Chapitre 5

Factorisation unique dans les anneaux euclidiens

L'objectif de ce chapitre est de montrer que tout polynôme non constant à coefficients dans un corps K s'écrit comme produit de polynômes *irréductibles* de $K[x]$ et que cette factorisation est unique à l'ordre des termes près et à multiplication près de chaque facteur par un élément non nul de K. On va montrer que cette propriété s'étend en fait à une classe assez large d'anneaux qu'on appelle les *anneaux euclidiens*. Cette classe inclut les anneaux de polynômes $K[x]$ sur un corps K, l'anneau $\mathbb{Z}$ des entiers ordinaires et l'anneau des entiers de Gauss pour ne citer que les exemples les plus importants. La propriété centrale des anneaux euclidiens est que tout idéal d'un tel anneau est *principal*, c'est-à-dire engendré par un seul élément.

5.1 Divisibilité dans les anneaux intègres

Définition 5.1.1. Un *anneau intègre* est un anneau commutatif A avec $0 \neq 1$ qui possède la propriété que $ab \neq 0$ pour tout choix de $a, b \in A$ avec $a \neq 0$ et $b \neq 0$.

Exemple 5.1.2. L'anneau $\mathbb{Z}$ est intègre. Tout sous-anneau d'un corps K est intègre.

On verra plus loin que l'anneau $K[x]$ est intègre quel que soit le corps K.

Définition 5.1.3. Soit A un anneau intègre et soient a, b avec $a \neq 0$. On dit que a *divise* b (dans A) ou que b est *divisible* par a s'il existe $c \in A$ tel que $b = c\,a$. On note cette condition $a \,|\, b$. Un *diviseur* de b (dans A) est un élément non nul de A qui divise b.

Exemple 5.1.4. Dans $\mathbb{Z}$, on a $-2 \,|\, 6$, car $-6 = (-2)(-3)$. Par contre 5 ne divise pas 6 dans $\mathbb{Z}$, mais 5 divise 6 dans $\mathbb{Q}$.

Les propriétés suivantes sont laissées en exercices :

Proposition 5.1.5. *Soit A un anneau intègre et soient $a, b, c \in A$ avec $a \neq 0$.*

(i) $1 \mid b$

(ii) *Si* $a \mid b$, *alors* $a \mid bc$.

(iii) *Si* $a \mid b$ *et* $a \mid c$, *alors* $a \mid (b + c)$.

(iv) *Si* $a \mid b$, $b \neq 0$ *et* $b \mid c$, *alors* $a \mid c$.

(v) *Si* $b \neq 0$ *et* $ab \mid ac$, *alors* $b \mid c$.

Définition 5.1.6. Soit A un anneau intègre et soient $a, b \in A \setminus \{0\}$. On dit que a et b sont *associés* ou que a est *associé à* b si $a \mid b$ et $b \mid a$. On note cette condition $a \sim b$.

Exemple 5.1.7. Dans $\mathbb{Z}$, deux entiers non nuls a et b sont associés si et seulement si $a = \pm b$. En particulier, tout entier non nul a est associé à un et un seul entier positif, à savoir $|a|$.

Lemme 5.1.8. *Soit A un anneau intègre et soit $a \in A \setminus \{0\}$. Les conditions suivantes sont équivalentes :*

(i) $a \sim 1$,

(ii) $a \mid 1$,

(iii) a *est inversible.*

Preuve. Comme $1 \mid a$, les conditions (i) et (ii) sont équivalentes. Comme A est commutatif, les conditions (ii) et (iii) sont aussi équivalentes. □

Définition 5.1.9. Les éléments non nuls d'un anneau intègre A qui satisfont les conditions équivalentes du lemme 5.1.8 sont appelés les *unités* de A. Leur ensemble est noté A^*.

Proposition 5.1.10. *L'ensemble A^* des unités d'un anneau intègre A est stable sous le produit dans A et, pour cette opération, c'est un groupe abélien.*

Pour la preuve, voir l'appendice A.

Exemple 5.1.11. Le groupe des unités de $\mathbb{Z}$ est $\mathbb{Z}^* = \{1, -1\}$. Le groupe des unités d'un corps K est $K^* = K \setminus \{0\}$. On l'appelle aussi le *groupe multiplicatif de K*.

Proposition 5.1.12. *Soit A un anneau intègre et soient $a, b \in A \setminus \{0\}$. Alors $a \sim b$ si et seulement si $a = u\,b$ pour une unité u de A.*

Preuve. Supposons d'abord que $a \sim b$. Alors il existe $u, v \in A$ tels que $a = u\,b$ et $b = v\,a$. On en déduit que $a = u\,v\,a$, donc $(1 - u\,v)a = 0$. Comme A est intègre et que $a \neq 0$, cela implique $1 - u\,v = 0$, donc $u\,v = 1$ et par suite $u \in A^*$.

Réciproquement, supposons que $a = u\,b$ avec $u \in A^*$. Comme $u \in A^*$, il existe $v \in A$ tel que $v\,u = 1$. On en déduit que $b = v\,u\,b = v\,a$. Des égalités $a = u\,b$ et $b = v\,a$, on tire que $b \mid a$ et $a \mid b$, donc $a \sim b$. □

Définition 5.1.13. On dit qu'un élément p d'un anneau intègre A est *irréductible* (dans A), si $p \neq 0$, si $p \notin A^*$ et si les seuls diviseurs de p (dans A) sont les éléments de A associés à 1 ou à p.

Exemple 5.1.14. Les éléments irréductibles de $\mathbb{Z}$ sont de la forme $\pm p$ où p est un nombre premier.

Définition 5.1.15. Soient a, b des éléments d'un anneau intègre A, avec $a \neq 0$ ou $b \neq 0$. On dit qu'un élément non nul d de A est un *plus grand commun diviseur* (pgcd) *de a et de b* si

(i) $d \mid a$ et $d \mid b$,

(ii) pour tout $c \in A \setminus \{0\}$ tel que $c \mid a$ et $c \mid b$, on a $c \mid d$.

On écrit alors $d = \mathrm{pgcd}(a, b)$.

On note que le pgcd n'existe pas toujours et que, s'il existe, il n'est pas unique en général. Plus précisément si d est un pgcd de a et de b, alors les pgcd de a et de b sont les éléments de A associés à d (voir l'exercice 5.1.3). Pour cette raison, on écrit souvent $d \sim \mathrm{pgcd}(a, b)$ pour dire que d est un pgcd de a et de b.

Exercices.

5.1.1. Démontrer la proposition 5.1.5.

Note. La partie (v) utilise de manière cruciale le fait que A est un anneau intègre.

5.1.2. Soit A un anneau intègre. Montrer que la relation $\sim$ de la définition 5.1.6 est une relation d'équivalence sur $A \setminus \{0\}$.

5.1.3. Soit A un anneau intègre, et soient $a, b \in A \setminus \{0\}$. Supposons que d et d' soient des pgcd de a et de b. Montrer que $d \sim d'$.

5.1.4. Soit A un anneau intègre, et soient p et q des éléments non nuls de A avec $p \sim q$. Montrer que p est irréductible si et seulement si q est irréductible.

5.1.5. Soit A un anneau intègre, et soient p et q des éléments irréductibles de A avec $p|q$. Montrer que $p \sim q$.

5.2 Divisibilité en termes d'idéaux

Soit A un anneau *commutatif*. On rappelle qu'un *idéal* de A est un sous-ensemble I de A tel que

(i) $0 \in I$,

(ii) si $r, s \in I$, alors $r + s \in I$,

(iii) si $a \in A$ et $r \in I$, alors $ar \in I$.

Exemple 5.2.1. Les ensembles $\{0\}$ et A sont des idéaux de A.

Le résultat suivant est laissé en exercice.

Lemme 5.2.2. *Soient $r_1, \ldots, r_m$ des éléments d'un anneau commutatif A. L'ensemble*

$$\{a_1 r_1 + \cdots + a_m r_m \,;\, a_1, \ldots, a_m \in A\}$$

est le plus petit idéal de A qui contienne $r_1, \ldots, r_m$.

Définition 5.2.3. L'idéal construit au lemme 5.2.2 s'appelle l'*idéal de A engendré par $r_1, \ldots, r_m$* et on le note $(r_1, \ldots, r_m)$. On dit qu'un idéal de A est *principal* s'il est engendré par un seul élément. L'idéal principal engendré par r est

$$(r) = \{a\, r \,;\, a \in A\}.$$

Exemple 5.2.4. L'idéal de $\mathbb{Z}$ engendré par 2 est l'ensemble $(2) = \{2\, a \,;\, a \in \mathbb{Z}\}$ des entiers pairs.

Exemple 5.2.5. Dans $\mathbb{Z}$, on a $1 \in (5, 18)$ car $1 = -7 \cdot 5 + 2 \cdot 18$, donc $\mathbb{Z} = (1) \subseteq (5, 18)$ et par suite $(5, 18) = (1) = \mathbb{Z}$.

Soit A un anneau intègre et soit $a \in A \setminus \{0\}$. En vertu des définitions, pour tout élément b de A, on a

$$\boxed{a \,|\, b \iff b \in (a) \iff (b) \subseteq (a).}$$

On en déduit que

$$a \sim b \iff (a) = (b) \qquad \text{et} \qquad a \in A^* \iff (a) = A.$$

Exercices.

5.2.1. Démontrer le lemme 5.2.2.

5.2.2. Soient I et J des idéaux d'un anneau commutatif A.

(i) Montrer que

$$I + J := \{r + s \,;\, r \in I,\, s \in J\}$$

est le plus petit idéal de A qui contienne I et J.

(ii) Si $I = (a)$ et $J = (b)$, montrer que $I + J = (a, b)$.

5.2.3. Soient $a_1, \ldots, a_s$ des éléments non tous nuls d'un anneau intègre A. Supposons qu'il existe $d \in A$ tel que

$$(d) = (a_1, \ldots, a_s).$$

Montrer que

(i) d n'est pas nul et divise $a_1, \ldots, a_s$;
(ii) si un élément non nul c de A divise $a_1, \ldots, a_s$, alors $c|d$.

Note. Un élément d de A qui remplit les conditions (i) et (ii) s'appelle un *plus grand commun diviseur de* $a_1, \ldots, a_s$ (en abrégé, on écrit *pgcd*).

5.2.4. Soient $a_1, \ldots, a_s$ des éléments non nuls d'un anneau intègre A. Supposons qu'il existe $m \in A$ tel que

$$(m) = (a_1) \cap \cdots \cap (a_s).$$

Montrer que
(i) m n'est pas nul et $a_1, \ldots, a_s$ divisent m ;
(ii) si un élément nul c de A est divisible par $a_1, \ldots, a_s$, alors $m|c$.

Note. Un élément m de A qui remplit les conditions (i) et (ii) s'appelle un *plus petit commun multiple de* $a_1, \ldots, a_s$ (en abrégé, on écrit *ppcm*).

5.3 Division euclidienne pour les polynômes

Soit K un corps. Tout polynôme non nul $p(x) \in K[x]$ s'écrit sous la forme

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

pour un entier $n \geq 0$ et des éléments $a_0, a_1, \ldots, a_n$ de K avec $a_n \neq 0$. Cet entier n est appelé le *degré* de $p(x)$ et noté $\deg(p(x))$. Le coefficient a_n est appelé le *coefficient dominant de* $p(x)$. On dit que $p(x)$ est *unitaire* si son coefficient dominant est 1.

Le degré du polynôme 0 n'est pas défini (certains auteurs posent $\deg(0) = -\infty$, mais nous n'adopterons pas cette convention dans ce cours).

Proposition 5.3.1. *Soient $p(x), q(x) \in K[x] \setminus \{0\}$. Alors, on a $p(x)q(x) \neq 0$ et*

$$\deg(p(x)q(x)) = \deg(p(x)) + \deg(q(x)).$$

Preuve. Écrivons $p(x) = a_0 + \cdots + a_n x^n$ et $q(x) = b_0 + \cdots + b_m x^m$ avec $a_n \neq 0$ et $b_m \neq 0$. Alors le produit

$$p(x)q(x) = a_0 b_0 + \cdots + a_n b_m x^{n+m}$$

a pour coefficient dominant $a_n b_m \neq 0$ et par suite

$$\deg(p(x)q(x)) = n + m = \deg(p(x)) + \deg(q(x)).$$

□

Corollaire 5.3.2.
(i) $K[x]$ est un anneau intègre.
(ii) $K[x]^* = K^* = K \setminus \{0\}$.

(iii) Tout polynôme $p(x) \in K[x] \setminus \{0\}$ est associé à un et un seul polynôme unitaire de $K[x]$.

Preuve. L'affirmation (i) découle directement de la proposition précédente. Pour démontrer (ii), on note d'abord que si $p(x) \in K[x]^*$, alors il existe $q(x) \in K[x]$, tel que $p(x)q(x) = 1$. En appliquant la proposition, on obtient

$$\deg(p(x)) + \deg(q(x)) = \deg(1) = 0,$$

donc $\deg(p(x)) = 0$ et par suite $p(x) \in K \setminus \{0\} = K^*$. Cela montre que $K[x]^* \subseteq K^*$. Comme l'inclusion $K^* \subseteq K[x]^*$ est immédiate, on obtient $K[x]^* = K^*$. L'assertion (iii) est laissée en exercice. □

La preuve du résultat suivant est elle aussi laissée en exercice :

Proposition 5.3.3. *Soient $p_1(x), p_2(x) \in K[x]$ avec $p_2(x) \neq 0$. On peut écrire*

$$p_1(x) = q(x)p_2(x) + r(x)$$

pour un et un seul choix de polynômes $q(x) \in K[x]$ et $r(x) \in K[x]$ satisfaisant

$$r(x) = 0 \quad ou \quad \deg(r(x)) < \deg(p_2(x)).$$

Ce fait est appelé *division euclidienne*. On le démontre, pour $p_2(x)$ fixé, par récurrence sur le degré de $p_1(x)$ (en supposant $p_1(x) \neq 0$). Les polynômes $q(x)$ et $r(x)$ sont appelés respectivement le *quotient* et le *reste* de la division de $p_1(x)$ par $p_2(x)$.

Exemple 5.3.4. Pour $K = \mathbb{Q}$, $p_1(x) = x^3 - 5x^2 + 8x - 4$ et $p_2(x) = x^2 - 2x + 1$, on trouve que

$$p_1(x) = (x - 3)\, p_2(x) + (x - 1)$$

avec $\deg(x - 1) = 1 < \deg(p_2(x)) = 2$.

Exercices.

5.3.1. Démontrer la partie (iii) du corollaire 3.2.

5.3.2. Soit $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$ le corps à 3 éléments et soient

$$p_1(x) = x^5 - x^2 + \bar{2}x + \bar{1},\ p_2(x) = x^3 - \bar{2}x + \bar{1} \in \mathbb{F}_3[x].$$

Déterminer le quotient et le reste de la division de $p_1(x)$ par $p_2(x)$.

5.4 Anneaux euclidiens

Définition 5.4.1. Un *anneau euclidien* est un anneau intègre A muni d'une fonction

$$\varphi : A \setminus \{0\} \to \mathbb{N}$$

qui jouit des propriétés suivantes :

(i) pour tout choix de $a, b \in A \setminus \{0\}$ avec $a \mid b$, on a $\varphi(a) \leq \varphi(b)$,

(ii) pour tout choix de $a, b \in A$ avec $b \neq 0$, on peut écrire

$$a = q\,b + r$$

avec $q, r \in A$ satisfaisant $r = 0$ ou $\varphi(r) < \varphi(b)$.

Une fonction φ qui possèdent ces propriétés est appelée *stathme euclidien*. La condition (i) est souvent omise. Nous l'avons incluse car elle permet de simplifier certains arguments et parce qu'elle est remplie dans les cas les plus importants pour nous.

Exemple 5.4.2. L'anneau $\mathbb{Z}$ est euclidien pour la valeur absolue

$$\begin{array}{rccc} \varphi : & \mathbb{Z} & \longrightarrow & \mathbb{N}\,. \\ & a & \longmapsto & |a| \end{array}$$

Pour tout corps K, l'anneau $K[x]$ est euclidien pour le degré

$$\begin{array}{rccc} \deg : & K[x] \setminus \{0\} & \longrightarrow & \mathbb{N}\,. \\ & p(x) & \longmapsto & \deg(p(x)) \end{array}$$

Dans la suite, on fixe un anneau euclidien A et une fonction $\varphi\ A \setminus \{0\} \to \mathbb{N}$ comme dans la définition 5.4.1.

Théorème 5.4.3. *Tout idéal de A est principal.*

Preuve. Si $I = \{0\}$, alors $I = (0)$ est engendré par 0. Supposons maintenant que $I \neq \{0\}$. Alors l'ensemble $I \setminus \{0\}$ n'est pas vide et par suite il contient un élément a pour lequel $\varphi(a)$ est minimal. On va montrer que $I = (a)$.

On note d'abord que $(a) \subseteq I$ puisque $a \in I$. Pour montrer l'inclusion réciproque, choisissons $b \in I$. Puisque A est euclidien, on peut écrire

$$b = q\,a + r$$

avec $q, r \in A$ qui satisfont $r = 0$ ou $\varphi(r) < \varphi(a)$. Puisque

$$r = b - q\,a = b + (-q)a \in I$$

(car $a, b \in I$), le choix de a implique $\varphi(r) \geq \varphi(a)$ si $r \neq 0$. On doit donc nécessairement avoir $r = 0$ et par suite $b = q\,a \in (a)$. Le choix de b étant arbitraire, cela montre que $I \subseteq (a)$. Donc $I = (a)$, comme annoncé. □

Pour un idéal engendré par deux éléments, on peut déterminer un générateur en utilisant l'algorithme suivant :

Algorithme euclidien 5.4.4. *Soient $a_1, a_2 \in A$ avec $a_2 \neq 0$. Par division euclidienne répétée, on écrit*

$$\begin{cases} a_1 = q_1 a_2 + a_3 & \text{avec } a_3 \neq 0 \text{ et } \varphi(a_3) < \varphi(a_2), \\ a_2 = q_2 a_3 + a_4 & \text{avec } a_4 \neq 0 \text{ et } \varphi(a_4) < \varphi(a_3), \\ \quad \dots & \dots \\ a_{k-2} = q_{k-2} a_{k-1} + a_k & \text{avec } a_k \neq 0 \text{ et } \varphi(a_k) < \varphi(a_{k-1}), \\ a_{k-1} = q_{k-1} a_k. & \end{cases}$$

Alors on a $(a_1, a_2) = (a_k)$.

Une schéma de preuve de ce résultat est proposé dans l'exercice 5.4.1.

Corollaire 5.4.5. *Soient $a, b \in A$ avec $a \neq 0$ ou $b \neq 0$. Alors a et b admettent un pgcd dans A. Plus précisément, un élément d de A est un pgcd de a et de b si et seulement si $(d) = (a, b)$.*

Preuve. Soit d un générateur de (a, b). Comme $a \neq 0$ ou $b \neq 0$, on a $(a, b) \neq (0)$ et par suite $d \neq 0$. Puisque $a, b \in (a, b) = (d)$, on note d'abord que d divise a et b. Par ailleurs, si $c \in A \setminus \{0\}$ divise a et b, alors $a, b \in (c)$, donc $d \in (a, b) \subseteq (c)$ et par suite d divise c. Cela montre que d est un pgcd de a et de b.

Si d' est un autre pgcd de a et de b, on a $d \sim d'$ (voir l'exercice 5.1.3), et par suite $(d') = (d) = (a, b)$. □

Le corollaire 5.4.5 implique en particulier que tout pgcd d'éléments a et b de A avec $a \neq 0$ ou $b \neq 0$ s'écrit sous la forme

$$d = r\,a + s\,b$$

avec $r, s \in A$ puisqu'un tel élément appartient à l'idéal (a, b).

Exemple 5.4.6. Calculer le pgcd de 15 et 24 et l'écrire sous forme $15r + 24s$ avec $r, s \in \mathbb{Z}$.

Solution: On applique l'algorithme euclidien 5.4.4 pour déterminer le générateur de $(15, 24)$. Par division euclidienne répétée, on trouve

$$24 = 15 + 9 \tag{5.1}$$
$$15 = 9 + 6 \tag{5.2}$$
$$9 = 6 + 3 \tag{5.3}$$
$$6 = 2 \cdot 3 \tag{5.4}$$

donc $\mathrm{pgcd}(15, 24) = 3$. Pour déterminer r et s, on remonte simplement la chaîne de calculs (5.1) à (5.4) à partir de la fin. On trouve ainsi

$$\begin{aligned} 3 &= 9 - 6 && \text{grâce à (5.3)},\\ &= 9 - (15 - 9) = 2 \cdot 9 - 15 && \text{grâce à (5.2)},\\ &= 2(24 - 15) - 15 = 2 \cdot 24 - 3 \cdot 15 && \text{grâce à (5.1)}. \end{aligned}$$

En conclusion

$$\mathrm{pgcd}(15, 24) = 3 = (-3) \cdot 15 + 2 \cdot 24.$$

Exemple 5.4.7. Calculer le pgcd de

$$p_1(x) = x^4 + 3x^2 + x - 1 \quad \text{et} \quad p_2(x) = x^2 - 1$$

dans $\mathbb{Q}[x]$ et l'écrire sous la forme $a_1(x)\, p_1(x) + a_2(x)\, p_2(x)$ avec $a_1(x), a_2(x) \in \mathbb{Q}[x]$.

Solution: Par division répétée, on trouve

$$p_1(x) = (x^2 + 4)p_2(x) + (x + 3) \tag{5.5}$$
$$p_2(x) = (x - 3)(x + 3) + 8 \tag{5.6}$$
$$x + 3 = (1/8)(x + 3) \cdot 8 \tag{5.7}$$

donc $\mathrm{pgcd}(p_1(x), p_2(x)) = 8$ (ou 1). En reprenant les calculs depuis la fin, on trouve

$$\begin{aligned} 8 &= p_2(x) - (x - 3)(x + 3) && \text{grâce à (5.6)},\\ &= p_2(x) - (x - 3)(p_1(x) - (x^2 + 4)p_2(x)) && \text{grâce à (5.5)},\\ &= (3 - x)\, p_1(x) + (x^3 - 3\, x^2 + 4\, x - 11)\, p_2(x)\,. \end{aligned}$$

Une autre conséquence du théorème 5.4.3 est le résultat suivant qui joue un rôle capital dans la suite.

Théorème 5.4.8. *Soit p un élément irréductible de A et soient $a, b \in A$. Supposons que $p \mid ab$. Alors on a $p \mid a$ ou $p \mid b$.*

Preuve. Supposons que $p \nmid a$. On doit montrer que $p \mid b$. Pour ce faire, on choisit d tel que $(p, a) = (d)$. Comme $d \mid p$ et que p est irréductible, on a $d \sim 1$ ou $d \sim p$. Comme on a aussi $d \mid a$ et que $p \nmid a$, la possibilité que $d \sim p$ est exclue. Donc d est une unité de A et par suite

$$1 \in (d) = (p, a).$$

Cela signifie qu'on peut écrire $1 = r\, p + s\, a$ avec $r, s \in A$. Alors $b = (r\, p + s\, a)b = r\, b\, p + s\, a\, b$ est divisible par p, car $p \mid a\, b$. □

Par récurrence, ce résultat s'étend à un produit quelconque d'éléments de A :

Corollaire 5.4.9. *Soit p un élément irréductible de A et soient $a_1, \ldots, a_s \in A$. Supposons que $p \mid a_1 \cdots a_s$. Alors p divise au moins un des éléments $a_1, \ldots, a_s$.*

Exercices.

5.4.1. Soient a_1 et a_2 des éléments d'un anneau euclidien A, avec $a_2 \neq 0$.
(i) Montrer que l'algorithme 5.4.4 termine après un nombre fini d'étapes.
(ii) Dans les notations de cet algorithme, montrer que $(a_i, a_{i+1}) = (a_{i+1}, a_{i+2})$ pour chaque $i = 1, \ldots, k-2$.
(iii) Dans les mêmes notations, montrer que $(a_{k-1}, a_k) = (a_k)$.
(iv) Conclure que $(a_1, a_2) = (a_k)$.

5.4.2. Déterminer le pgcd de $f(x) = x^3 + x - 2$ et de $g(x) = x^5 - x^4 + 2x^2 - x - 1$ dans $\mathbb{Q}[x]$ et l'exprimer comme une combinaison linéaire de $f(x)$ et de $g(x)$.

5.4.3. Déterminer le pgcd de $f(x) = x^3 + \bar{1}$ et de $g(x) = x^5 + x^4 + x + \bar{1}$ dans $\mathbb{F}_2[x]$ et l'exprimer comme une combinaison linéaire de $f(x)$ et de $g(x)$.

5.4.4. Déterminer le pgcd de 114 et de 45 dans $\mathbb{Z}$ et l'exprimer comme une combinaison linéaire de 114 et de 45.

5.4.5. Soient $a_1, \ldots, a_s$ des éléments non tous nuls d'un anneau euclidien A.
(i) Montrer que $a_1, \ldots, a_s$ admettent un pgcd dans A au sens de l'exercice 5.2.3.
(ii) Si $s \geq 3$, montrer que $\mathrm{pgcd}(a_1, \ldots, a_s) = \mathrm{pgcd}(a_1, \mathrm{pgcd}(a_2, \ldots, a_s)$.

5.4.6. Soient $a_1, \ldots, a_s$ des éléments non nuls d'un anneau euclidien A.
(i) Montrer que $a_1, \ldots, a_s$ admettent un ppcm dans A au sens de l'exercice 5.2.4.
(ii) Si $s \geq 3$, montrer que $\mathrm{ppcm}(a_1, \ldots, a_s) = \mathrm{ppcm}(a_1, \mathrm{ppcm}(a_2, \ldots, a_s)$.

5.4.7. Déterminer le pgcd de 42, 63 et 140 dans $\mathbb{Z}$, et l'exprimer comme combinaison linéaire de ces trois entiers (voir l'exercice 5.4.5).

5.4.8. Déterminer le pgcd de $x^4 - 1$, $x^6 - 1$ et $x^3 + x^2 + x + 1$ dans $\mathbb{Q}[x]$, et l'exprimer comme combinaison linéaire de ces trois polynômes (voir l'exercice 5.4.5).

5.5 Le théorème de factorisation unique

On fixe encore un anneau euclidien A et une fonction $\varphi : A \setminus \{0\} \to \mathbb{N}$ comme dans la définition 5.4.1. Les résultats de la section précédente s'appuient seulement sur la seconde propriété de φ requise par la définition 5.4.1. Le résultat suivant utilise la première propriété.

Proposition 5.5.1. *Soient $a, b \in A \setminus \{0\}$ avec $a \,|\, b$. Alors on a*

$$\varphi(a) \leq \varphi(b)$$

avec égalité si et seulement si $a \sim b$.

Preuve. Puisque $a \,|\, b$, on a $\varphi(a) \leq \varphi(b)$ comme requis dans la définition 5.4.1.

Si $a \sim b$, on a aussi $b \,|\, a$, donc $\varphi(b) \leq \varphi(a)$ et par suite $\varphi(a) = \varphi(b)$.

Réciproquement, supposons que $\varphi(a) = \varphi(b)$. On peut écrire

$$a = q\,b + r$$

avec $q, r \in A$ qui satisfont $r = 0$ ou $\varphi(r) < \varphi(b)$. Comme $a \,|\, b$, on trouve que a divise $r = a - q\,b$, donc $r = 0$ ou $\varphi(a) \leq \varphi(r)$. Comme $\varphi(a) = \varphi(b)$, la seule possibilité est que $r = 0$. Cela signifie que $a = q\,b$, donc $b \,|\, a$ et par conséquent $a \sim b$. □

En appliquant ce résultat avec $a = 1$, on en déduit

Corollaire 5.5.2. *Soit $b \in A \setminus \{0\}$. On a $\varphi(b) \geq \varphi(1)$ avec égalité si et seulement si $b \in A^*$.*

On est maintenant en mesure de démontrer le résultat principal de ce chapitre.

Théorème 5.5.3 (théorème de factorisation unique). *Soit $a \in A \setminus \{0\}$. Il existe une unité $u \in A^*$, un entier $r \geq 0$ et des éléments irréductibles $p_1, \ldots, p_r$ de A tels que*

$$a = u\,p_1 \cdots p_r. \tag{5.8}$$

Pour toute autre factorisation de a comme produit

$$a = u'\,q_1 \cdots q_s \tag{5.9}$$

d'une unité $u' \in A^*$ et d'éléments irréductibles $q_1, \ldots, q_s$ de A, on a $s = r$ et il existe une permutation $(i_1, i_2, \ldots, i_r)$ de $(1, 2, \ldots, r)$ telle que

$$q_1 \sim p_{i_1}\,,\; q_2 \sim p_{i_2}\,,\; \ldots\,,\; q_r \sim p_{i_r}.$$

Preuve. **1° Existence :** On démontre d'abord l'existence d'une factorisation de la forme (5.8) par récurrence sur $\varphi(a)$.

Si $\varphi(a) \leq \varphi(1)$, le corollaire 5.5.2 nous apprend que $\varphi(a) = \varphi(1)$ et que $a \in A^*$. Dans ce cas, l'égalité (5.8) est remplie pour le choix de $u = a$ et $r = 0$.

Supposons maintenant que $\varphi(a) > \varphi(1)$. Parmi les diviseurs p de a avec $\varphi(p) > \varphi(1)$, on en choisit un pour lequel $\varphi(p)$ est minimal. Si q est un diviseur de cet élément p alors $q \,|\, a$ et par suite on a $\varphi(q) = \varphi(1)$ ou $\varphi(q) \geq \varphi(p)$. En vertu de la proposition 5.5.1, cela signifie que $q \sim 1$ ou $q \sim p$. Cela étant vrai pour tout diviseur q de p, on conclut que p est irréductible. Écrivons $a = b\,p$ avec $b \in A$. Comme $b \,|\, a$ et $b \not\sim a$, la proposition 5.5.1 donne $\varphi(b) < \varphi(a)$. Par récurrence, on peut supposer que b admet une factorisation du type (5.8). Alors $a = b\,p$ admet aussi une telle factorisation (avec le facteur p en plus).

2° Unicité : Supposons que a admette aussi la factorisation (5.9), on obtient

$$u\,p_1 \cdots p_r = u'\,q_1 \cdots q_s \tag{5.10}$$

où u, u' sont des unités et $p_1, \ldots, p_r, q_1, \cdots, q_s$ des éléments irréductibles de A. Si $r = 0$, le membre de gauche de cette égalité est réduit à u et, puisqu'aucun élément irréductible de A ne peut diviser une unité, cela implique $s = 0$ comme requis. Supposons maintenant que $r \geq 1$. Comme p_1 divise $u\, p_1 \cdots p_r$, il divise aussi $u'\, q_1 \cdots q_s$. Puisque p_1 est irréductible, cela implique, en vertu du corollaire 5.4.9, que p_1 divise au moins un des facteurs $q_1, \ldots, q_s$ (car $p_1 \nmid u'$. Quitte à permuter $q_1, \ldots, q_s$ si nécessaire, on peut supposer que $p_1 \,|\, q_1$. Alors on a $q_1 = u_1 p_1$ pour une unité $u_1 \in A^*$, et l'égalité (5.10) livre

$$u\, p_2 \cdots p_r = u''\, q_2 \cdots q_s$$

où $u'' = u' u_1 \in A^*$. Puisque le membre de gauche de cette nouvelle égalité contient un facteur irréductible de moins, on peut supposer, par récurrence, que cela implique $r - 1 = s - 1$ et $q_2 \sim p_{i_2}, \ldots, q_r \sim p_{i_r}$ pour une permutation $(i_2, \ldots, i_r)$ de $(2, \ldots, r)$. Alors $r = s$ et $q_1 \sim p_1$, $q_2 \sim p_{i_2}$, $\ldots$, $q_r \sim p_{i_r}$ comme requis. □

Puisque le groupe des unités de $\mathbb{Z}$ est $\mathbb{Z}^* = \{1, -1\}$ et que tout élément irréductible de $\mathbb{Z}$ est associé à un et un seul nombre premier p, on en déduit :

Corollaire 5.5.4. *Tout entier a non nul s'écrit sous la forme*

$$a = \pm p_1 \cdots p_r$$

pour un choix de signe $\pm$, un entier $r \geq 0$ et des nombres premiers $p_1, \ldots, p_r$. Cette écriture est unique à une permutation près de $p_1, \ldots, p_r$.

Si $r = 0$, il faut interpréter le produit $\pm p_1 \cdots p_r$ comme étant ± 1.

Puisque qu'on peut ordonner les nombres premiers par ordre de grandeur, on en déduit encore :

Corollaire 5.5.5. *Pour tout entier non nul a, il existe un et un seul choix de signe $\pm$, d'entier $s \geq 0$, de nombres premiers $p_1 < p_2 < \cdots < p_s$ et d'entiers positifs $e_1, e_2, \ldots, e_s$ tels que*

$$a = \pm p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s}.$$

De même, si K est un corps, on sait que $K[x]^* = K^*$ et que tout polynôme irréductible de $K[x]$ est associé à un et un seul polynôme irréductible unitaire. Par contre, il n'y a pas d'ordre naturel sur $K[x]$. On en déduit :

Corollaire 5.5.6. *Soit K un corps et soit $a(x) \in K[x] \setminus \{0\}$. Il existe une constante $c \in K^*$, un entier $s \geq 0$, des polynômes irréductibles unitaires <u>distincts</u> $p_1(x), \ldots, p_s(x)$ de $K[x]$ et des entiers positifs $e_1, \ldots, e_s$ tels que*

$$a(x) = c\, p_1(x)^{e_1} \cdots p_s(x)^{e_s}.$$

Cette factorisation est unique à une permutation près des facteurs $p_1(x)^{e_1}, \ldots, p_s(x)^{e_s}$.

Exercices.

5.5.1. Soit a un élément non nul d'un anneau euclidien A. Écrivons $a = up_1^{e_1} \cdots p_s^{e_s}$ où u est une unité de A, où $p_1, \ldots, p_s$ sont des éléments irréductibles de A non associés deux à deux, et où $e_1, \ldots, e_s$ sont des entiers ≥ 0 (on convient que $p^0 = 1$). Montrer que les diviseurs de a sont les produits $vp_1^{d_1} \cdots p_s^{d_s}$ où v est une unité de A et où $d_1, \ldots, d_s$ sont des entiers avec $0 \leq d_i \leq e_i$ pour $i = 1, \ldots, s$.

5.5.2. Soient a et b des éléments non nuls d'un anneau euclidien A et soient $p_1, \ldots, p_s$ un système maximal de diviseurs irréductibles non associés deux à deux de ab.

(i) Montrer qu'on peut écrire

$$a = up_1^{e_1} \cdots p_s^{e_s} \quad \text{et} \quad b = vp_1^{f_1} \cdots p_s^{f_s}$$

avec $u, v \in A^*$ et $e_1, \ldots, e_s, f_1, \ldots, f_s \in \mathbb{N}$.

(ii) Montrer que

$$\operatorname{pgcd}(a, b) \sim p_1^{d_1} \cdots p_s^{d_s}$$

où $d_i = \min(e_i, f_i)$ pour $i = 1, \ldots, s$.

(iii) Montrer que

$$\operatorname{ppcm}(a, b) \sim p_1^{h_1} \cdots p_s^{h_s}$$

où $h_i = \max(e_i, f_i)$ pour $i = 1, \ldots, s$ (voir l'exercice 5.2.4 pour la notion de ppcm).

Indice. Pour (ii), utiliser le résultat de l'exercice 5.5.1.

5.6 Le théorème fondamental de l'algèbre

Fixons un corps K. On note d'abord :

Proposition 5.6.1. *Soient $a(x) \in K[x]$ et $r \in K$. Le polynôme $x - r$ divise $a(x)$ si et seulement si $a(r) = 0$.*

Preuve. En divisant $a(x)$ par $x - r$, on trouve

$$a(x) = q(x)(x - r) + b$$

où $q(x) \in K[x]$ et $b \in K$ car le reste de la division est nul ou de degré $< \deg(x - r) = 1$. En évaluant les deux membres de cette égalité en $x = r$, on obtient $a(r) = b$. La conclusion suit. □

Définition 5.6.2. Soit $a(x) \in K[x]$. On dit qu'un élément r de K est une *racine* de $a(x)$ si $a(r) = 0$.

Le théoréme de d'Alembert-Gauss appelé aussi "théorème fondamental de l'algèbre" est en fait un résultat d'analyse qui s'énonce ainsi :

Théorème 5.6.3. *Tout polynôme $a(x) \in \mathbb{C}[x]$ de degré ≥ 1 admet au moins une racine dans $\mathbb{C}$.*

Nous allons accepter ce résultat sans preuve. En vertu de la proposition 5.6.1, il implique que les polynômes irréductibles unitaires de $\mathbb{C}[x]$ sont de la forme $x - r$ avec $r \in \mathbb{C}$. En vertu du corollaire 5.5.6, on conclut :

Théorème 5.6.4. *Soit $a(x) \in \mathbb{C}[x] \setminus \{0\}$, soient $r_1, \ldots, r_s$ les racines distinctes de $a(x)$ dans $\mathbb{C}$, et soit c le coefficient dominant de $a(x)$. Il existe un et un seul choix d'entiers $e_1, \ldots, e_s \geq 1$ tel que*

$$a(x) = c(x - r_1)^{e_1} \cdots (x - r_s)^{e_s}$$

L'entier e_i s'appelle la multiplicité *de r_i comme racine de $a(x)$.*

Le théorème 5.6.3 nous éclaire aussi sur les polynômes irréductibles de $\mathbb{R}[x]$.

Théorème 5.6.5. *Les polynômes irréductibles unitaires de $\mathbb{R}[x]$ sont de la forme $x - r$ avec $r \in \mathbb{R}$ ou de la forme $x^2 + b\,x + c$ avec $b, c \in \mathbb{R}$ et $b^2 - 4c < 0$.*

Preuve. Soit $p(x)$ un polynôme irréductible unitaire de $\mathbb{R}[x]$. En vertu du théorème 5.6.3, il admet au moins une racine r dans $\mathbb{C}$. Si $r \in \mathbb{R}$, alors $x - r$ divise $p(x)$ dans $\mathbb{R}[x]$ et par suite $p(x) = x - r$. Si $r \notin \mathbb{R}$, alors le conjugué complexe $\bar{r}$ de r est distinct de r et, comme

$$p(\bar{r}) = \overline{p(r)} = 0,$$

c'est aussi une racine de $p(x)$. Par suite, $p(x)$ est divisible par $(x - r)(x - \bar{r})$ dans $\mathbb{C}[x]$. Comme ce produit s'écrit $x^2 + b\,x + c$ avec $b = -(r + \bar{r}) \in \mathbb{R}$ et $c = r\bar{r} \in \mathbb{R}$, le quotient de $p(x)$ par $x^2 + b\,x + c$ est un polynôme de $\mathbb{R}[x]$ et par suite $p(x) = x^2 + b\,x + c$. Puisque les racines de $p(x)$ dans $\mathbb{C}$ ne sont pas réelles, on doit avoir $b^2 - 4c < 0$. □

Exercices.

5.6.1. Soit K un corps et soit $p(x)$ un polynôme non nul de $K[x]$.

(i) Supposons que $\deg(p(x)) = 1$. Montrer que $p(x)$ est irréductible.

(ii) Supposons que $2 \leq \deg(p(x)) \leq 3$. Montrer que $p(x)$ est irréductible si et seulement si il n'admet pas de racine dans K.

(iii) En général, montrer que $p(x)$ est irréductible si et seulement si il n'admet pas de facteur irréductible de degré $\leq \deg(p(x))/2$ dans $K[x]$.

5.6.2. Soit $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$ le corps fini à 3 éléments.

(i) Déterminer tous les polynômes unitaires irréductibles de $\mathbb{F}_3[x]$ de degré au plus 2.
(ii) Montrer que $x^4 + x^3 + \bar{2}$ est un polynôme irréductible de $\mathbb{F}_3[x]$.
(iii) Factoriser $x^5 + \bar{2}x^4 + \bar{2}x^3 + x^2 + x + \bar{2}$ en produit de polynômes irréductibles de $\mathbb{F}_3[x]$.

5.6.3. Déterminer tous les polynômes unitaires irréductibles de $\mathbb{F}_2[x]$ de degré au plus 4 où $\mathbb{F}_2 = \{\bar{0}, \bar{1}\}$ désigne le corps fini à 2 éléments.

5.6.4. Factoriser $x^4 - 2$ en produit de polynômes irréductibles (i) dans $\mathbb{R}[x]$, (ii) dans $\mathbb{C}[x]$.

5.6.5. Soit V un espace vectoriel de dimension finie $n \geq 1$ sur $\mathbb{C}$ et soit $T\colon V \to V$ un opérateur linéaire. Montrer que T possède au moins une valeur propre.

5.6.6. Soit V un espace vectoriel de dimension finie $n \geq 1$ sur un corps K, soit $T\colon V \to V$ un opérateur linéaire sur V, et soient $\lambda_1, \ldots, \lambda_s$ les valeurs propres distinctes de T dans K. Montrer que T est diagonalisable si et seulement si

$$\mathrm{car}_T(x) = (x - \lambda_1)^{e_1} \cdots (x - \lambda_s)^{e_s}$$

où $e_i = \dim_K E_{\lambda_i}(T)$ pour $i = 1, \ldots, s$.

Indice. Utiliser le théorème 3.2.3.

5.6.7. Soit $A \in \mathrm{Mat}_{n\times n}(K)$ où n est un entier positif et K un corps. Énoncer et démontrer une condition nécessaire et suffisante, analogue à celle de l'exercice 5.6.6, pour que A soit diagonalisable sur K.

Chapitre 6

Modules

La notion de module est une extension naturelle de celle d'espace vectoriel. La différence est que les scalaires ne sont plus pris dans un corps mais plutôt dans un anneau. Cela permet d'englober davantage d'objets. Ainsi, on verra qu'un groupe abélien est un module sur l'anneau $\mathbb{Z}$ des entiers. On verra aussi que si V est un espace vectoriel sur un corps K muni d'un opérateur linéaire, alors V est naturellement un module sur l'anneau $K[x]$ des polynômes en une variable sur K. Dans ce chapitre, on introduit de manière générale la théorie des modules sur un anneau commutatif avec les notions de sous-module, de module libre, de somme directe, d'homomorphisme et d'isomorphisme de modules. On réinterprête le sens de ces constructions dans plusieurs contextes, le plus important pour nous étant le cas d'un espace vectoriel sur un corps K muni d'un opérateur linéaire.

6.1 La notion de module

Définition 6.1.1. Un *module* sur un anneau commutatif A est un ensemble M muni d'une addition et d'une multiplication par les éléments de A :

$$\begin{array}{rcl} M \times M & \longrightarrow & M \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \mathbf{u} + \mathbf{v} \end{array} \qquad \text{et} \qquad \begin{array}{rcl} A \times M & \longrightarrow & M \\ (a, \mathbf{u}) & \longmapsto & a\,\mathbf{u} \end{array}$$

qui satisfont les axiomes suivants :

$$\left.\begin{array}{ll} \textbf{Mod1}. & \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \\ \textbf{Mod2}. & \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \end{array}\right\} \text{pour tout choix de } \mathbf{u}, \mathbf{v}, \mathbf{w} \in M.$$

Mod3. Il existe $\mathbf{0} \in M$ tel que $\mathbf{u} + \mathbf{0} = \mathbf{u}$ pour tout $\mathbf{u} \in M$.
Mod4. Pour tout $\mathbf{u} \in M$, il existe $-\mathbf{u} \in M$ tel que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

$$\left.\begin{array}{ll} \textbf{Mod5}. & 1\mathbf{u} = \mathbf{u} \\ \textbf{Mod6}. & a(b\mathbf{u}) = (ab)\mathbf{u} \\ \textbf{Mod7}. & (a+b)\mathbf{u} = a\,\mathbf{u} + b\,\mathbf{u} \\ \textbf{Mod8}. & a(\mathbf{u} + \mathbf{v}) = a\,\mathbf{u} + a\,\mathbf{v} \end{array}\right\} \text{pour tout choix de } a, b \in A \text{ et } \mathbf{u}, \mathbf{v} \in M.$$

Dans la suite, on parlera indifféremment de *module sur A* ou de *A-module* pour désigner un tel objet.

Si on compare la définition de module sur un anneau commutatif A à celle d'un espace vectoriel sur un corps K (voir la définition 1.1.1), on reconnaît les mêmes axiomes. Comme un corps est un type particulier d'anneau commutatif, on en déduit :

Exemple 6.1.2. Un module sur un corps K est simplement un espace vectoriel sur K.

Les axiomes **Mod1** à **Mod4** signifient qu'un module M est un groupe abélien sous l'addition (voir l'appendice A). On en déduit que l'ordre dans lequel on additionne des éléments $\mathbf{u}_1, \ldots, \mathbf{u}_n$ de M n'affecte par leur somme notée

$$\mathbf{u}_1 + \cdots + \mathbf{u}_n \quad \text{ou} \quad \sum_{i=1}^{n} \mathbf{u}_i.$$

L'élément $\mathbf{0}$ de M caractérisé par l'axiome **Mod3** est appelé le *zéro* de M (c'est l'élément neutre pour l'addition). Enfin, pour chaque $\mathbf{u} \in M$, l'élément $-\mathbf{u}$ de M caractérisé par l'axiome **Mod 4** est appelé l'*inverse additif* de $\mathbf{u}$. On définit la soustraction dans M par

$$\mathbf{u} - \mathbf{v} := \mathbf{u} + (-\mathbf{v}).$$

On vérifie aussi, par récurrence sur n, que les axiomes de distributivité **Mod7** et **Mod8** impliquent en général

$$\left(\sum_{i=1}^{n} a_i\right)\mathbf{u} = \sum_{i=1}^{n} a_i\mathbf{u} \quad \text{et} \quad a\sum_{i=1}^{n} \mathbf{u}_i = \sum_{i=1}^{n} a\,\mathbf{u}_i \tag{6.1}$$

quels que soient $a, a_1, \ldots, a_n \in A$ et $\mathbf{u}, \mathbf{u}_1, \ldots, \mathbf{u}_n \in M$.

On montre encore :

Lemme 6.1.3. *Soit M un module sur un anneau commutatif A. Pour tout $a \in A$ et tout $\mathbf{u} \in M$, on a*

(i) $0\mathbf{u} = \mathbf{0}$ *et* $a\mathbf{0} = \mathbf{0}$

(ii) $(-a)\mathbf{u} = a(-\mathbf{u}) = -(a\,\mathbf{u})$.

Preuve. Dans A, on a $0 + 0 = 0$. Alors l'axiome **Mod7** livre

$$0\mathbf{u} = (0+0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u},$$

donc $0\mathbf{u} = \mathbf{0}$. Plus généralement, puisque $a + (-a) = 0$, le même axiome donne

$$0\mathbf{u} = (a + (-a))\mathbf{u} = a\,\mathbf{u} + (-a)\mathbf{u}.$$

Comme $0\mathbf{u} = \mathbf{0}$, cela signifie que $(-a)\mathbf{u} = -(a\,\mathbf{u})$.

Les formules $a\mathbf{0} = \mathbf{0}$ et $a(-\mathbf{u}) = -(a\,\mathbf{u})$ se démontrent de manière semblable en appliquant plutôt l'axiome **Mod8**. □

Supposons que M soit un $\mathbb{Z}$-module, c'est-à-dire un module sur $\mathbb{Z}$. Alors, M est un groupe abélien sous l'addition et, pour tout entier $n \geq 1$ et pour tout $\mathbf{u} \in M$, les formules de distributivité générale (6.1) livrent

$$n\,\mathbf{u} = (\underbrace{1 + \cdots + 1}_{n \text{ fois}})\mathbf{u} = \underbrace{1\mathbf{u} + \cdots + 1\mathbf{u}}_{n \text{ fois}} = \underbrace{\mathbf{u} + \cdots + \mathbf{u}}_{n \text{ fois}}.$$

Gràce au lemme 6.1.3, on trouve aussi

$$(-n)\mathbf{u} = n(-\mathbf{u}) = \underbrace{(-\mathbf{u}) + \cdots + (-\mathbf{u})}_{n \text{ fois}} \quad \text{et} \quad 0\mathbf{u} = \mathbf{0}.$$

Autrement dit, la multiplication par les entiers dans M est entièrement déterminée par la loi d'addition dans M et correspond à la notion naturelle de produit par un entier dans un groupe abélien. Réciproquement, si M est un groupe abélien, la règle des "exposants" (Proposition A.2.11 de l'appendice A) signifie que le produit par un entier dans M satisfait aux axiomes **Mod6** à **Mod8**. Comme il satisfait clairement l'axiome **Mod5**, on en déduit :

Proposition 6.1.4. *Un $\mathbb{Z}$-module est simplement un groupe abélien muni de la multiplication naturelle par les entiers.*

La proposition suivante fournit un troisième exemple de module.

Proposition 6.1.5. *Soit V un espace vectoriel sur un corps K, et soit $T \in \mathrm{End}_K(V)$ un opérateur linéaire sur V. Alors V est un $K[x]$-module pour l'addition de V et la multiplication par les polynômes donnée par*

$$p(x)\mathbf{v} := p(T)(\mathbf{v})$$

pour tout $p(x) \in K[x]$ et tout $\mathbf{v} \in V$.

Preuve. Comme V est un groupe abélien sous l'addition, il suffit de vérifier les axiomes **Mod5** à **Mod8**. Soient $p(x), q(x) \in K[x]$ et $\mathbf{u}, \mathbf{v} \in V$. On doit montrer :

1) $1\mathbf{v} = \mathbf{v}$,
2) $p(x)(q(x)\mathbf{v}) = (p(x)q(x))\mathbf{v}$,
3) $(p(x) + q(x))\mathbf{v} = p(x)\mathbf{v} + q(x)\mathbf{v}$,
4) $p(x)(\mathbf{u} + \mathbf{v}) = p(x)\mathbf{u} + p(x)\mathbf{v}$.

En vertu de la définition de produit par un polynôme, cela revient à montrer

1') $I_V(\mathbf{v}) = \mathbf{v}$,
2') $p(T)(q(T)\mathbf{v}) = (p(T) \circ q(T))\mathbf{v}$,
3') $(p(T) + q(T))\mathbf{v} = p(T)\mathbf{v} + q(T)\mathbf{v}$,
4') $p(T)(\mathbf{u} + \mathbf{v}) = p(T)\mathbf{u} + p(T)\mathbf{v}$,

car l'application

$$\begin{array}{ccc} K[x] & \longrightarrow & \mathrm{End}_K(V)\,, \\ r(x) & \longmapsto & r(T) \end{array}$$

d'évaluation en T, étant un homorphisme d'anneaux, elle applique 1 sur I_V, $p(x)+q(x)$ sur $p(T)+q(T)$ et $p(x)q(x)$ sur $p(T)\circ q(T)$. Les égalités $1'), 2')$ et $3')$ découlent des définitions de I_V, $p(T)\circ q(T)$ et $p(T)+q(T)$ respectivement. Enfin $4')$ découlent du fait que $p(T)$ est une application linéaire. □

Dans le contexte de la proposition 6.1.5, on note aussi que, pour un polynôme constant $p(x)=a$ avec $a\in K$, on a

$$p(x)\mathbf{v} = (a\,I)(\mathbf{v}) = a\,\mathbf{v}$$

quel que soit $\mathbf{v}\in V$. Donc la notation $a\,\mathbf{v}$ possède le même sens que l'on considère a comme un polynôme de $K[x]$ ou comme un scalaire du corps K. On peut donc énoncer le complément suivant à la proposition 6.1.5.

Proposition 6.1.6. *Dans les notations de la proposition 6.1.5, la multiplication externe par les éléments de $K[x]$ sur V étend la multiplication scalaire par les éléments de K.*

Exemple 6.1.7. Soit V un espace vectoriel de dimension 2 sur $\mathbb{Q}$ muni d'une base $\mathcal{B}=\{\mathbf{v}_1,\mathbf{v}_2\}$ et soit $T\colon V\to V$ l'opérateur linéaire déterminé par les conditions

$$T(\mathbf{v}_1)=\mathbf{v}_1-\mathbf{v}_2,\quad T(\mathbf{v}_2)=\mathbf{v}_1.$$

On munit V de la structure de $\mathbb{Q}[x]$-module associée à cet endomorphisme T. Calculons

$$(2x-1)\mathbf{v}_1+(x^2+3x)\mathbf{v}_2.$$

Solution: Par définition, c'est

$$\begin{aligned}(2\,T-I)(\mathbf{v}_1)+(T^2+3\,T)(\mathbf{v}_2) &= 2\,T(\mathbf{v}_1)-\mathbf{v}_1+T^2(\mathbf{v}_2)+3\,T(\mathbf{v}_2)\\ &= 3\,T(\mathbf{v}_1)-\mathbf{v}_1+3\,T(\mathbf{v}_2)\quad(\text{car } T^2(\mathbf{v}_2)=T(T(\mathbf{v}_2))=T(\mathbf{v}_1))\\ &= 3(\mathbf{v}_1-\mathbf{v}_2)-\mathbf{v}_1+3\mathbf{v}_1\\ &= 5\mathbf{v}_1-3\mathbf{v}_2.\end{aligned}$$

Fixons un anneau commutatif A quelconque. On conclut avec trois exemples généraux de A-module. Les détails des preuves sont laissés en exercice.

Exemple 6.1.8. L'anneau A est un A-module pour sa loi d'addition et la multiplication externe

$$\begin{array}{ccc} A\times A & \longrightarrow & A \\ (a,\mathbf{u}) & \longmapsto & a\,\mathbf{u} \end{array}$$

donnée par le produit dans A.

Exemple 6.1.9. Soit $n \in \mathbb{N}^*$. L'ensemble

$$A^n = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \; ; \, a_1, a_2, \ldots, a_n \in A \right\}$$

des n-uplets d'éléments de A est un module sur A pour les opérations

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix} \quad \text{et} \quad c \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} c\,a_1 \\ c\,a_2 \\ \vdots \\ c\,a_n \end{pmatrix}.$$

Pour $n = 1$, on retrouve la structure naturelle de A-module sur $A^1 = A$ décrite par l'exemple 6.1.8.

Exemple 6.1.10. Soient $M_1, \ldots, M_n$ des A-modules. Leur produit cartésien

$$M_1 \times \cdots \times M_n = \{(\mathbf{u}_1, \ldots, \mathbf{u}_n) \, ; \, \mathbf{u}_1 \in M_1, \ldots, \mathbf{u}_n \in M_n\}$$

est un A-module pour les opérations

$$\begin{aligned} (\mathbf{u}_1, \ldots, \mathbf{u}_n) + (\mathbf{v}_1, \ldots, \mathbf{v}_n) &= (\mathbf{u}_1 + \mathbf{v}_1, \ldots, \mathbf{u}_n + \mathbf{v}_n) \\ c\,(\mathbf{u}_1, \ldots, \mathbf{u}_n) &= (c\,\mathbf{u}_1, \ldots, c\,\mathbf{u}_n). \end{aligned}$$

Remarque. Si on applique la construction de l'exemple 6.1.10 avec $M_1 = \cdots = M_n = A$ pour la structure naturelle de A-module sur A donnée par l'exemple 6.1.8, on obtient une structure de A-module sur $A \times \cdots \times A = A^n$ qui est essentiellement celle de l'exemple 6.1.9 excepté que les éléments de A^n sont présentés sous forme de lignes. Pour les applications ultérieures, il est cependant plus commode d'écrire les éléments de A^n en colonnes.

Exercices.

6.1.1. Compléter la preuve du lemme 6.1.3 en montrant que $a\,\mathbf{0} = \mathbf{0}$ et $a(-\mathbf{u}) = -(a\,\mathbf{u})$.

6.1.2. Soit V un espace vectoriel de dimension 3 sur $\mathbb{F}_5$ muni d'une base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ et soit $T\colon V \to V$ l'opérateur linéaire déterminé par les conditions

$$T(\mathbf{v}_1) = \mathbf{v}_1 + \bar{2}\mathbf{v}_2 - \mathbf{v}_3, \quad T(\mathbf{v}_2) = \mathbf{v}_2 + \bar{4}\mathbf{v}_3, \quad T(\mathbf{v}_3) = \mathbf{v}_3.$$

Pour la structure correspondante de $\mathbb{F}_5[x]$-module sur V, calculer

(i) $x\,\mathbf{v}_1$,

(ii) $x^2\,\mathbf{v}_1 - (x+\bar{1})\mathbf{v}_2$,
(iii) $(\bar{2}x^2 - \bar{3})\mathbf{v}_1 + (\bar{3}x)\mathbf{v}_2 + (x^2+\bar{2})\mathbf{v}_3$.

6.1.3. Soit $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'opérateur linéaire dont la matrice en base standard $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$ est $[T]_\mathcal{E} = \begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix}$. Pour la structure de $\mathbb{R}[x]$-module associée sur $\mathbb{R}^2$, calculer
(i) $x\,\mathbf{e}_1$,
(ii) $x^2\,\mathbf{e}_1 + \mathbf{e}_2$,
(iii) $(x^2+x+1)\mathbf{e}_1 + (2x-1)\mathbf{e}_2$.

6.1.4. Soit $n \in \mathbb{N}^*$ et soit A un anneau commutatif. Montrer que A^n est un A-module pour les opérations définies dans l'exemple 6.1.9.

6.1.5. Soient $M_1, \ldots, M_n$ des modules sur un anneau commutatif A. Montrer que leur produit cartésien $M_1 \times \cdots \times M_n$ est un A-module pour les opérations définies dans l'exemple 6.1.10.

6.1.6. Soit M un module sur un anneau commutatif A et soit X un ensemble. Pour tout $a \in A$ et toute paire de fonctions $f\colon X \to M$ et $g\colon X \to M$, on définit $f+g\colon X \to M$ et $af\colon X \to M$ en posant

$$(f+g)(x) = f(x) + g(x) \quad \text{et} \quad (af)(x) = af(x)$$

pour tout $x \in X$. Montrer que l'ensemble $\mathcal{F}(X, M)$ des fonctions de X dans M est un A-module pour ces opérations.

6.2 Sous-modules

Fixons un module M quelconque sur un anneau commutatif A quelconque.

Définition 6.2.1. Un *sous-A-module* de M est un sous-ensemble N de M qui remplit les conditions :

SM1. $\mathbf{0} \in N$.
SM2. Si $\mathbf{u}, \mathbf{v} \in N$, alors $\mathbf{u} + \mathbf{v} \in N$.
SM3. Si $a \in A$ et $\mathbf{u} \in N$, alors $a\,\mathbf{u} \in N$.

On note que la condition **SM3** implique $-\mathbf{u} = (-1)\mathbf{u} \in N$ pour tout $\mathbf{u} \in N$ ce qui, joint aux conditions **SM1** et **SM2**, signifie qu'un sous-A-module de M est en particulier un sous-groupe de M.

Les conditions **SM2** et **SM3** demandent qu'un sous-A-module N de M soit stable sous l'addition dans M et la multiplication par les éléments de A dans M. Par restriction, ces opérations fournissent donc une addition sur N et une multiplication par les éléments de A sur N. On laisse en exercice le soin de vérifier que ces opérations satisfont les 8 axiomes requis pour faire de N un A-module.

Proposition 6.2.2. *Un sous-A-module N de M est lui-même un A-module pour l'addition et la multiplication par les éléments de A restreintes de M à N.*

Exemple 6.2.3. Soit V un espace vectoriel sur un corps K. Un sous-K-module de V est simplement un sous-K-espace vectoriel de V.

Exemple 6.2.4. Soit M un groupe abélien. Un sous-$\mathbb{Z}$-module de M est simplement un sous-groupe de M.

Exemple 6.2.5. Soit A un anneau commutatif considéré comme module sur lui-même (voir l'exemple 6.1.8). Un sous-A-module de A est simplement un idéal de A.

Dans le cas d'un espace vectoriel V muni d'un endomorphisme, le concept de sous-module se traduit ainsi :

Proposition 6.2.6. *Soit V un espace vectoriel sur un corps K et soit $T \in \mathrm{End}_K(V)$. Pour la structure correspondante de $K[x]$-module sur V, un sous-$K[x]$-module de V est simplement un sous-espace vectoriel T-*invariant *de V, c'est-à-dire un sous-espace vectoriel U de V tel que $T(\mathbf{u}) \in U$ pour tout $\mathbf{u} \in U$.*

Preuve. Comme la multiplication par les éléments de $K[x]$ sur V étend la muliplication scalaire par les éléments de K (voir la proposition 6.1.6), un sous-$K[x]$-module U de V est nécessairement un sous-espace vectoriel de V. Comme il satisfait aussi $x\mathbf{u} \in U$ pour tout $\mathbf{u} \in U$, et que $x\,\mathbf{u} = T(\mathbf{u})$ (par définition), on obtient aussi que U doit être T-invariant.

Réciproquement, si U est un sous-espace vectoriel T-invariant de U, on a $T(\mathbf{u}) \in U$ pour tout $\mathbf{u} \in U$. Par récurrence sur n, on en déduit que $T^n(\mathbf{u}) \in U$ pour tout entier $n \geq 1$ et tout $\mathbf{u} \in U$. Par suite, pour tout polynôme $p(x) = a_0 + a_1x + \cdots + a_nx^n \in K[x]$ et tout $\mathbf{u} \in U$, on trouve

$$p(x)\mathbf{u} = (a_0I + a_1T + \cdots + a_nT^n)(\mathbf{u}) = a_0\mathbf{u} + a_1T(\mathbf{u}) + \cdots + a_nT^n(\mathbf{u}) \in U.$$

Donc U est un sous-$K[x]$-module de V : on vient de montrer qu'il remplit la condition **SM3** et les autres conditions **SM1** et **SM2** sont satisfaites en vertu du fait que U est un sous-espace de V. □

On conclut cette section avec quelques constructions générales valables pour un A-module quelconque.

Proposition 6.2.7. *Soient $\mathbf{u}_1, \ldots, \mathbf{u}_s$ des éléments d'un A-module M. L'ensemble*

$$\{a_1\,\mathbf{u}_1 + \cdots + a_s\,\mathbf{u}_s\,;\, a_1, \ldots, a_s \in A\}$$

est un sous-A-module de M qui contient $\mathbf{u}_1, \ldots, \mathbf{u}_s$ et il est le plus petit sous-A-module de M avec cette propriété.

Ce sous-A-module est appelé le sous-A-module de M *engendré* par $\mathbf{u}_1, \ldots, \mathbf{u}_s$. Il noté

$$\langle \mathbf{u}_1, \ldots, \mathbf{u}_s \rangle_A \quad \text{ou} \quad A\,\mathbf{u}_1 + \cdots + A\,\mathbf{u}_s.$$

La preuve de ce résultat est laissée en exercice.

Définition 6.2.8. On dit qu'un A-module M (ou un sous-A-module de M) est *cyclique* s'il est engendré par un seul élément. On dit qu'il est de *type fini* s'il est engendré par un nombre fini d'éléments de A.

Exemple 6.2.9. Si $\mathbf{u}_1, \ldots, \mathbf{u}_s$ sont des éléments d'un groupe abélien M, alors $\langle \mathbf{u}_1, \ldots, \mathbf{u}_s \rangle_{\mathbb{Z}}$ est simplement le sous-groupe de M engendré par $\mathbf{u}_1, \ldots, \mathbf{u}_s$. En particulier, M est cyclique en tant que groupe abélien si et seulement si il est cyclique en tant que $\mathbb{Z}$-module.

Exemple 6.2.10. Soit V un espace vectoriel de dimension finie sur un corps K, soit $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ un système de générateurs de V, et soit $T \in \mathrm{End}_K(V)$. Comme la structure de $K[x]$-module sur V associée à T étend celle de K-module, on a

$$V = \langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle_K \subseteq \langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle_{K[x]},$$

donc $V = \langle \mathbf{v}_1, \ldots, \mathbf{v}_s \rangle_{K[x]}$ est aussi engendré par $\mathbf{v}_1, \ldots, \mathbf{v}_s$ en tant que $K[x]$-module. En particulier, V est un $K[x]$-module de type fini.

Proposition 6.2.11. *Soient $N_1, \ldots, N_s$ des sous-modules d'un A-module M. Leur intersection $N_1 \cap \cdots \cap N_s$ et leur* somme

$$N_1 + \cdots + N_s = \{\mathbf{u}_1 + \cdots + \mathbf{u}_s \,;\, \mathbf{u}_1 \in N_1, \ldots, \mathbf{u}_s \in N_s\}$$

sont des sous-A-modules de M.

Preuve pour la somme. On note d'abord que, puisque $\mathbf{0} \in N_i$ pour $i = 1, \ldots, s$, on a

$$\mathbf{0} = \mathbf{0} + \cdots + \mathbf{0} \in N_1 + \cdots + N_s.$$

Soient $\mathbf{u}, \mathbf{u}'$ des éléments quelconques de $N_1 + \cdots + N_s$, et soit $a \in A$. On peut écrire

$$\mathbf{u} = \mathbf{u}_1 + \cdots + \mathbf{u}_s \quad \text{et} \quad \mathbf{u}' = \mathbf{u}'_1 + \cdots + \mathbf{u}'_s$$

avec $\mathbf{u}_1, \mathbf{u}'_1 \in N_1, \ldots, \mathbf{u}_s, \mathbf{u}'_s \in N_s$. Par suite, on obtient

$$\begin{aligned} \mathbf{u} + \mathbf{u}' &= (\mathbf{u}_1 + \mathbf{u}'_1) + \cdots + (\mathbf{u}_s + \mathbf{u}'_s) \in N_1 + \cdots + N_s \\ a\,\mathbf{u} &= (a\,\mathbf{u}_1) + \cdots + (a\,\mathbf{u}_s) \in N_1 + \cdots + N_s. \end{aligned}$$

Cela montre que $N_1 + \cdots + N_s$ est un sous-A-module de M. □

Notons en terminant cette section que la notation $A\,\mathbf{u}_1 + \cdots + A\,\mathbf{u}_s$ pour le sous-A-module $\langle \mathbf{u}_1, \ldots, \mathbf{u}_s \rangle_A$ engendré par des éléments $\mathbf{u}_1, \ldots, \mathbf{u}_s$ de M est consistante avec la notion de somme de sous-modules car, pour tout élément $\mathbf{u}$ de M, on a

$$A\,\mathbf{u} = \langle \mathbf{u} \rangle_A = \{a\,\mathbf{u} \,;\, a \in A\}.$$

Lorsque $A = K$ est un corps, on reconnaît les notions usuelles d'intersection et de somme de sous-espaces d'un espace vectoriel sur K. Lorsque $A = \mathbb{Z}$, on retrouve plutôt les notions d'intersection et de somme de sous-groupes d'un groupe abélien.

Exercices.

6.2.1. Démontrer la proposition 6.2.2.

6.2.2. Soit M un groupe abélien. Expliquer pourquoi les sous-$\mathbb{Z}$-modules de M sont simplement les sous-groupes de M.

Indice. Montrer que tout sous-$\mathbb{Z}$-module de M est un sous-groupe de M et, réciproquement, que tout sous-groupe de M est un sous-$\mathbb{Z}$-modules de M.

6.2.3. Démontrer la proposition 6.2.7.

6.2.4. Soit M un module sur un anneau commutatif A, et soient $u_1, \dots, u_s, v_1, \dots, v_t \in M$. Posons $L = \langle u_1, \dots, u_s \rangle_A$ et $N = \langle v_1, \dots, v_t \rangle_A$. Montrer que $L+N = \langle u_1, \dots, u_s, v_1, \dots, v_t \rangle_A$.

6.2.5. Soit M un module sur un anneau commutatif A. On dit qu'un élément $\mathbf{u}$ de M est *de torsion* s'il existe un élément non nul a de A tel que $a\,\mathbf{u} = 0$. On dit que M est un *A-module de torsion* si tous ses éléments sont de torsion. Supposons que A soit un anneau intègre.

(i) Montrer que l'ensemble M_{tor} des éléments de torsion de M est un sous-A-module de M.

(ii) Montrer que, si M est de type fini et si tous ses éléments sont de torsion, alors il existe un élément non nul a de A tel que $a\,\mathbf{u} = 0$ pour tout $\mathbf{u} \in M$.

6.3 Modules libres

Définition 6.3.1. Soient $\mathbf{u}_1, \dots, \mathbf{u}_n$ des éléments d'un A-module M.

(i) On dit que $\mathbf{u}_1, \dots, \mathbf{u}_n$ sont *linéairement indépendants* sur A si le seul choix d'éléments $a_1, \dots, a_n$ de A tels que

$$a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0}$$

est $a_1 = \cdots = a_n = 0$.

(ii) On dit que $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ est une *base* de M si $\mathbf{u}_1, \dots, \mathbf{u}_n$ sont linéairement indépendants sur A et engendrent M en tant que A-module.

(iii) On dit que M est un A-module *libre* s'il admet une base.

Par convention le module $\{\mathbf{0}\}$ réduit à un seul élément est un A-module libre qui admet pour base l'ensemble vide.

A la différence des espaces vectoriels de type fini sur un corps K qui admettent toujours une base, les A-modules de type fini n'admettent pas de base en général (voir les exercices).

Exemple 6.3.2. Soit $n \in \mathbb{N}^*$. Le A-module A^n introduit à l'exemple 6.1.9 admet pour base

$$\left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}.$$

En effet, on

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + a_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{6.2}$$

pour tout choix de $a_1, \ldots, a_n \in A$. Donc $\mathbf{e}_1, \ldots, \mathbf{e}_n$ engendrent A^n en tant que A-module. Enfin, si $a_1, \ldots, a_n \in A$ satisfont $a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n = \mathbf{0}$, alors l'égalité (6.2) livre

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

et par suite $a_1 = a_2 = \cdots = a_n = 0$. Donc $\mathbf{e}_1, \ldots, \mathbf{e}_n$ sont aussi linéairement indépendants sur A.

Exercices.

6.3.1. Montrer qu'un groupe abélien fini M admet une base en tant que $\mathbb{Z}$-module si et seulement si $M = \{0\}$ (auquel cas l'ensemble vide est, par convention, une base de M).

6.3.2. En général, montrer que, si M est un module fini sur un anneau commutatif infini A, alors M est un A-module libre si et seulement si $M = \{0\}$.

6.3.3. Soit V un espace vectoriel de dimension finie sur un corps K et soit $T \in \mathrm{End}_K(V)$. Montrer que, pour la structure correspondante de $K[x]$-module, V admet une base si et seulement si $V = \{0\}$.

6.3.4. Soient m et n des entiers positifs et soit A un anneau commutatif. Montrer que $\mathrm{Mat}_{m\times n}(A)$ est un A-module libre pour les opérations usuelles d'addition et de multiplication par les éléments de A.

6.4 Somme directe

Définition 6.4.1. Soient $M_1, \ldots, M_s$ des sous-modules d'un A-module M. On dit que leur somme est *directe* si le seul choix de $\mathbf{u}_1 \in M_1, \ldots, \mathbf{u}_s \in M_s$ tels que

$$\mathbf{u}_1 + \cdots + \mathbf{u}_s = 0$$

est $\mathbf{u}_1 = \cdots = \mathbf{u}_s = 0$. Lorsque cette condition est remplie, on note la somme $M_1 + \cdots + M_s$ sous la forme $M_1 \oplus \cdots \oplus M_s$.

On dit que M est la *somme directe* de $M_1, \ldots, M_s$ si $M = M_1 \oplus \cdots \oplus M_s$.

Les propositions suivantes qui généralisent les résultats analogues de la section 1.4 sont laissées en exercice. La première justifie l'importance de la notion de somme directe.

Proposition 6.4.2. *Soient $M_1, \ldots, M_s$ des sous-modules d'un A-module M. Alors, on a $M = M_1 \oplus \cdots \oplus M_s$ si et seulement si, pour tout $\mathbf{u} \in M$, il existe un et un seul choix de $\mathbf{u}_1 \in M_1, \ldots, \mathbf{u}_s \in M_s$ tels que $\mathbf{u} = \mathbf{u}_1 + \cdots + \mathbf{u}_s$.*

Proposition 6.4.3. *Soient $M_1, \ldots, M_s$ des sous-modules d'un A-module M. Les conditions suivantes sont équivalentes :*

1) la somme de $M_1, \ldots, M_s$ est directe,

2) $M_i \cap (M_1 + \cdots + \widehat{M_i} + \cdots + M_s) = \{\mathbf{0}\}$ pour $i = 1, \ldots, s$,

3) $M_i \cap (M_{i+1} + \cdots + M_s) = \{\mathbf{0}\}$ pour $i = 1, \ldots, s-1$.

Exemple 6.4.4. Soit V un espace vectoriel sur un corps K et soit $T \in \mathrm{End}_K(V)$. Pour la structure correspondante de $K[x]$-module sur V, on sait que les sous-modules de V sont simplement les sous-espaces vectoriels T-invariants de V. Comme la notion de somme directe ne fait intervenir que l'addition, dire que V, en tant que $K[x]$-module, est somme directe de sous-$K[x]$-modules $V_1, \ldots, V_s$ équivaut simplement à dire que V, en tant qu'espace vectoriel sur K, est somme directe de sous-espaces vectoriels T-invariants $V_1, \ldots, V_s$.

Exercices.

6.4.1. Démontrer la proposition 6.4.3.

6.4.2. Soit M un module sur un anneau commutatif A et soient $\mathbf{u}_1, \ldots, \mathbf{u}_n \in M$. Montrer que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ est une base de M si et seulement si on a à la fois $M = A\mathbf{u}_1 \oplus \cdots \oplus A\mathbf{u}_n$ et aucun des éléments $\mathbf{u}_1, \ldots, \mathbf{u}_n$ n'est de torsion (voir l'exercice 6.2.5 pour la définition de ce terme).

6.5 Homomorphismes de modules

Fixons un anneau commutatif A.

Définition 6.5.1. Soient M et N des A-modules. Un *homomorphisme* de A-modules de M dans N est une fonction $\varphi\colon M \to N$ qui satisfait

HM1.	$\varphi(\mathbf{u}_1 + \mathbf{u}_2) = \varphi(\mathbf{u}_1) + \varphi(\mathbf{u}_2)$	pour tout choix de $\mathbf{u}_1, \mathbf{u}_2 \in M$,
HM2.	$\varphi(a\,\mathbf{u}) = a\,\varphi(\mathbf{u})$	pour tout $a \in A$ et tout $\mathbf{u} \in M$.

La condition **HM1** signifie qu'un homomorphisme de A-modules est en particulier un homomorphisme de groupes sous l'addition.

Exemple 6.5.2. Si $A = K$ est un corps, des K-modules M et N sont simplement des espaces vectoriels sur K et un homomorphisme de K-modules $\varphi\colon M \to N$ n'est rien d'autre qu'une application linéaire de M dans N.

Exemple 6.5.3. On rappelle que si $A = \mathbb{Z}$, des $\mathbb{Z}$-modules M et N sont simplement des groupes abéliens équipés de la multiplication naturelle par les entiers. Si $\varphi\colon M \to N$ est un homomorphisme de $\mathbb{Z}$-modules, la condition **HM1** montre que c'est homomorphisme de groupes. Réciproquement si $\varphi\colon M \to N$ est un homomorphisme de groupes abéliens, on a

$$\varphi(n\,\mathbf{u}) = n\,\varphi(\mathbf{u})$$

pour tout $n \in \mathbb{Z}$ et tout $\mathbf{u} \in M$ (exercice), donc φ remplit les conditions **HM1** et **HM2** de la définition 6.5.1. Donc, *un homomorphisme de $\mathbb{Z}$-modules est simplement un homomorphisme de groupes abéliens.*

Exemple 6.5.4. Soit V un espace vectoriel sur un corps K et soit $T \in \mathrm{End}_K(V)$. Pour la structure correspondante de $K[x]$-module sur V, un homomorphisme de $K[x]$-modules $S\colon V \to V$ est simplement une application linéaire telle que $S \circ T = T \circ S$ (exercice).

Exemple 6.5.5. Soient $m, n \in \mathbb{N}^*$, et soit $P \in \mathrm{Mat}_{m\times n}(A)$. Pour tout choix de $\mathbf{u}, \mathbf{u}' \in A^n$ et de $a \in A$, on a :

$$P(\mathbf{u} + \mathbf{u}') = P\mathbf{u} + P\mathbf{u}' \quad \text{et} \quad P(a\,\mathbf{u}) = a\,P\mathbf{u}.$$

Donc la fonction

$$\begin{array}{rcl} A^n & \longrightarrow & A^m \\ \mathbf{u} & \longmapsto & P\mathbf{u} \end{array}$$

est un homomorphisme de A-modules.

Le résultat suivant montre qu'on obtient ainsi tous les homomorphismes de A-modules de A^n dans A^m.

Proposition 6.5.6. *Soient $m, n \in \mathbb{N}^*$ et soit $\varphi\colon A^n \to A^m$ un homomorphisme de A-modules. Il existe une et une seule matrice $P \in \mathrm{Mat}_{m\times n}(A)$ telle que*

$$\varphi(\mathbf{u}) = P\,\mathbf{u} \quad \textit{pour tout } \mathbf{u} \in A^n. \tag{6.3}$$

Preuve. Soit $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base "canonique" de A^n définie à l'exemple 6.3.2 et soit P la matrice $m \times n$ dont les colonnes sont $C_1 := \varphi(\mathbf{e}_1)$, ..., $C_n := \varphi(\mathbf{e}_n)$. Pour tout n-uplet $\mathbf{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in A^n$, on trouve :

$$\varphi(\mathbf{u}) = \varphi(a_1\mathbf{e}_1 + \cdots + a_n\mathbf{e}_n) = a_1\,C_1 + \cdots + a_n\,C_n = \begin{pmatrix} C_1 \cdots C_n \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = P\,\mathbf{u}\,.$$

Donc P satisfait la condition (6.3). C'est la seule matrice qui remplisse cette condition car si une matrice $P' \in \mathrm{Mat}_{m\times n}(A)$ satisfait aussi $\varphi(\mathbf{u}) = P'\mathbf{u}$ pour tout $\mathbf{u} \in A^n$, alors sa j-ième colonne est $P'\mathbf{e}_j = \varphi(\mathbf{e}_j) = C_j$ pour tout $j = 1, \ldots, n$. □

On reprend la théorie générale avec le résultat suivant.

Proposition 6.5.7. *Soient $\varphi\colon M \to N$ et $\psi\colon N \to P$ des homomorphismes de A-modules :*

(i) La composée $\psi \circ \varphi\colon M \to P$ est aussi un homomorphisme de A-modules.

(ii) Si $\varphi\colon M \to N$ est bijective, son inverse $\varphi^{-1}\colon N \to M$ est un homomorphisme de A-modules.

La partie (i) est laissé en exercice.

Preuve de (ii). Supposons que φ soit bijective. Soient $\mathbf{v}, \mathbf{v}' \in N$ et $a \in A$. On pose

$$\mathbf{u} := \varphi^{-1}(\mathbf{v}) \quad \text{et} \quad \mathbf{u}' := \varphi^{-1}(\mathbf{v}').$$

Comme $\mathbf{u}, \mathbf{u}' \in M$ et que φ est un homomorphisme de A-modules, on trouve

$$\varphi(\mathbf{u} + \mathbf{u}') = \varphi(\mathbf{u}) + \varphi(\mathbf{u}') = \mathbf{v} + \mathbf{v}' \quad \text{et} \quad \varphi(a\,\mathbf{u}) = a\,\varphi(\mathbf{u}) = a\,\mathbf{v}.$$

Par conséquent,

$$\varphi^{-1}(\mathbf{v} + \mathbf{v}') = \mathbf{u} + \mathbf{u}' = \varphi^{-1}(\mathbf{v}) + \varphi^{-1}(\mathbf{v}') \quad \text{et} \quad \varphi^{-1}(a\,\mathbf{v}) = a\,\mathbf{u} = a\,\varphi^{-1}(\mathbf{v}).$$

Cela montre bien que $\varphi^{-1}\colon N \to M$ est un homomorphisme de A-modules. □

Définition 6.5.8. Un homomorphisme de A-modules qui est bijectif est appelé un *isomorphisme* de A-modules. On dit qu'un A-module M est *isomorphe* à un A-module N s'il existe un isomorphisme de A-modules $\varphi\colon M \to N$. On écrit alors $M \simeq N$.

En vertu de ces définitions, la proposition 6.5.7 admet pour conséquence immédiate :

Corollaire 6.5.9. *Soient $\varphi\colon M \to N$ et $\psi\colon N \to P$ des isomorphismes de A-modules. Alors $\psi \circ \varphi\colon M \to P$ et $\varphi^{-1}\colon N \to M$ sont aussi des isomorphismes de A-modules.*

On en déduit encore que l'isomorphisme est une relation d'équivalence sur la classe des A-modules. De plus, l'ensemble des isomorphismes de A-modules d'un A-module M dans lui-même, appelés *automorphismes* de M, forment un sous-groupe du groupe des bijections de M dans lui-même. On l'appelle le *groupe des automorphismes* de M.

Définition 6.5.10. Soit $\varphi\colon M \to N$ un homomorphisme de A-modules. Le *noyau* de φ est

$$\ker(\varphi) := \{\mathbf{u} \in M\,;\, \varphi(\mathbf{u}) = \mathbf{0}\}$$

et l'*image* de φ est

$$\mathrm{Im}(\varphi) := \{\varphi(\mathbf{u})\,;\, \mathbf{u} \in M\}.$$

Le noyau et l'image d'un homomorphisme de A-modules $\varphi\colon M \to N$ sont donc simplement le noyau et l'image de φ considéré comme homomorphisme de groupes abéliens. Le résultat suivant est laissé en exercice.

Proposition 6.5.11. *Soit $\varphi\colon M \to N$ un homomorphisme de A-modules :*

(i) $\ker(\varphi)$ *est un sous-A-module de M.*

(ii) $\mathrm{Im}(\varphi)$ *est un sous-A-module de N.*

(iii) L'homomorphisme φ est injectif si et seulement si $\ker(\varphi) = \{\mathbf{0}\}$. *Il est surjectif si et seulement si* $\mathrm{Im}(\varphi) = N$.

Exemple 6.5.12. Soit M un A-module et soit $a \in A$. L'application

$$\begin{array}{rccc} \varphi: & M & \to & M \\ & \mathbf{u} & \to & a\mathbf{u} \end{array}$$

est un homomorphisme de A-modules (exercice). Son noyau est

$$M_a := \{\mathbf{u} \in M\,;\, a\mathbf{u} = \mathbf{0}\}$$

et son image est

$$aM := \{a\mathbf{u}\,;\, \mathbf{u} \in M\}.$$

Donc, en vertu de la proposition 6.5.11, M_a et aM sont des sous-A-modules de M.

On conclut ce chapitre par le résultat suivant.

Proposition 6.5.13. *Soit $\varphi\colon M \to N$ un homomorphisme de A-modules. Supposons que M soit libre avec base $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. Alors φ est un isomorphisme si et seulement si $\{\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)\}$ est une base de N.*

Preuve. On va montrer plus précisément que

i) φ est injectif si et seulement si $\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)$ sont linéairement indépendants sur A,

ii) φ est surjectif si et seulement si $\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)$ engendrent N.

Par conséquent, φ est bijectif si et seulement si $\{\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)\}$ est une base de N.

Pour démontrer (i), on utilise le fait que φ est injectif si et seulement si $\ker(\varphi) = \{\mathbf{0}\}$ (proposition 6.5.11(iii)). On note que, pour tout choix de $a_1, \ldots, a_n \in A$, on a

$$\begin{aligned} a_1\varphi(\mathbf{u}_1) + \cdots + a_n\varphi(\mathbf{u}_n) = 0 &\iff \varphi(a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n) = 0 \\ &\iff a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n \in \ker(\varphi). \end{aligned} \tag{6.4}$$

Si $\ker(\varphi) = \{\mathbf{0}\}$, la dernière condition devient $a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0}$ qui implique $a_1 = \cdots = a_n = 0$ puisque $\mathbf{u}_1, \ldots, \mathbf{u}_n$ sont linéairement indépendants sur A. Dans ce cas, on conclut que $\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)$ sont linéairement indépendants sur A. Réciproquement, si $\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)$ sont linéairement indépendants, les équivalences (6.4) nous apprennnent que le seul choix de $a_1, \ldots, a_n \in A$ tel que $a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n \in \ker(\varphi)$ est $a_1 = \cdots = a_n = 0$. Comme $\mathbf{u}_1, \ldots, \mathbf{u}_n$ engendrent M, cela implique que $\ker(\varphi) = \{\mathbf{0}\}$. Ainsi $\ker(\varphi) = \{\mathbf{0}\}$ si et seulement si $\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)$ sont linéairement indépendants. Pour démontrer (ii), on utilise le fait que φ est surjectif si et seulement si $\mathrm{Im}(\varphi) = N$. Comme $M = \langle \mathbf{u}_1, \ldots, \mathbf{u}_n \rangle_A$, on a

$$\begin{aligned} \mathrm{Im}(\varphi) &= \{\varphi(a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n)\,;\, a_1, \ldots, a_n \in A\} \\ &= \{a_1\varphi(\mathbf{u}_1) + \cdots + a_n\varphi(\mathbf{u}_n)\,;\, a_1, \ldots, a_n \in A\} \\ &= \langle \varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n) \rangle_A. \end{aligned}$$

Donc $\mathrm{Im}(\varphi) = N$ si et seulement si $\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_n)$ engendrent N. □

Exercices.

6.5.1. Démontrer la proposition 6.5.7 (i).

6.5.2. Démontrer la proposition 6.5.11.

6.5.3. Démontrer que l'application φ de l'exemple 6.5.12 est un homomorphisme de A-modules.

6.5.4. Soit $\varphi\colon M \to M'$ un homomorphisme de A-modules et soit N un sous-A-module de M. On définit l'*image* de N sous φ par

$$\varphi(N) := \{\varphi(\mathbf{u})\,;\, \mathbf{u} \in M\}.$$

Montrer que

(i) $\varphi(N)$ est un sous-A-module de M' ;

(ii) si $N = \langle \mathbf{u}_1, \ldots, \mathbf{u}_m\rangle_A$, alors $\varphi(N) = \langle \varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_m)\rangle_A$;

(iii) si N admet une base $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ et si φ est injective, alors $\{\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_m)\}$ est une base de $\varphi(N)$.

6.5.5. Soit $\varphi\colon M \to M'$ un homomorphisme de A-modules et soit N' un sous-A-module de M'. On définit l'*image réciproque* de N' sous φ par

$$\varphi^{-1}(N) := \{\mathbf{u} \in M\,;\, \varphi(\mathbf{u}) \in N'\}$$

(cette notation ne suppose pas que φ^{-1} existe). Montrer que $\varphi^{-1}(N)$ est un sous-A-module de M.

6.5.6. Soit M un A-module libre avec base $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, soit N un A-module quelconque, et soient $\mathbf{v}_1, \ldots, \mathbf{v}_n$ des éléments quelconques de N. Montrer qu'il existe un et un seul homomorphisme de A-modules $\varphi\colon M \to N$ tel que $\varphi(\mathbf{u}_i) = \mathbf{v}_i$ pour $i = 1, \ldots, n$.

6.5.7. Soit N un A-module de type fini engendré par n éléments. Montrer qu'il existe un homomorphisme de A-modules surjectif $\varphi\colon A^n \to N$.

Indice. Utiliser le résultat de l'exercice 6.5.6.

Chapitre 7

Structure des modules de type fini sur un anneau euclidien

Dans ce chapitre, on définit les notions d'annulateur d'un module et d'annulateur d'un élément d'un module. On étudie en détail le sens de ces notions pour un groupe abélien fini considéré comme $\mathbb{Z}$-module puis pour un espace vectoriel de dimension finie sur un corps K muni d'un endomorphisme. On cite ensuite sans démonstration le théorème de structure pour les modules de type fini sur un anneau euclidien et on analyse ses conséquences dans les mêmes deux situations. On montre enfin un raffinement de ce résultat et on conclut avec la forme canonique de Jordan pour un opérateur linéaire sur un espace vectoriel complexe de dimension finie.

7.1 Annulateurs

Définition 7.1.1. Soit M un module sur un anneau commutatif A. L'*annulateur* d'un élément $\mathbf{u}$ de M est

$$\mathrm{Ann}(\mathbf{u}) := \{a \in A\,;\, a\mathbf{u} = \mathbf{0}\}.$$

L'*annulateur* de M est

$$\mathrm{Ann}(M) := \{a \in A\,;\, a\mathbf{u} = \mathbf{0} \text{ pour tout } \mathbf{u} \in M\}.$$

Le résultat suivant est laissé en exercice :

Proposition 7.1.2. *Soient A et M comme dans la définition ci-dessus. Les annulateurs de M et des éléments de M sont des idéaux de A. De plus, si $M = \langle \mathbf{u}_1, \ldots, \mathbf{u}_s \rangle_A$, alors*

$$\mathrm{Ann}(M) = \mathrm{Ann}(\mathbf{u}_1) \cap \cdots \cap \mathrm{Ann}(\mathbf{u}_s).$$

Rappelons qu'un A-module est dit *cyclique* s'il est engendré par un seul élément (voir la section 6.2). Pour un tel module, la dernière assertion de la proposition 7.1.2 livre :

Corollaire 7.1.3. *Soit M un module cyclique sur un anneau commutatif A. On a* $\mathrm{Ann}(M) = \mathrm{Ann}(\mathbf{u})$ *pour tout générateur* $\mathbf{u}$ *de M.*

Dans le cas d'un groupe abélien G noté additivement et considéré comme $\mathbb{Z}$-module, ces notions se traduisent ainsi :

1° L'annulateur d'un élément g de G est

$$\mathrm{Ann}(g) = \{m \in \mathbb{Z}\,;\, mg = 0\}.$$

C'est un idéal de $\mathbb{Z}$. Par définition, si $\mathrm{Ann}(g) = \{0\}$, on dit que g est d'*ordre infini*. Sinon, on dit que g est d'*ordre fini* et son *ordre*, noté $\mathrm{o}(g)$, est le plus petit élément positif de $\mathrm{Ann}(g)$, donc c'est aussi le générateur de $\mathrm{Ann}(g)$.

2° L'annulateur de G est

$$\mathrm{Ann}(G) = \{m \in \mathbb{Z}\,;\, mg = 0 \text{ pour tout } g \in G\}.$$

C'est aussi un idéal de $\mathbb{Z}$. Si $\mathrm{Ann}(G) \neq \{0\}$, alors le plus petit élément positif de $\mathrm{Ann}(G)$, c'est-à-dire son générateur positif, est appelé l'*exposant* de G.

Si G est un groupe fini, alors l'*ordre* $|G|$ du groupe (sa cardinalité) est un élément de $\mathrm{Ann}(G)$. Cela découle du théorème de Lagrange qui dit que, dans un groupe fini (commutatif ou non), l'ordre de tout élément divise l'ordre du groupe. Dans ce cas, on a $\mathrm{Ann}(G) \neq \{0\}$ et l'exposant de G est un diviseur de $|G|$. Plus précisément, on montre :

Proposition 7.1.4. *Soit G un groupe abélien fini considéré comme $\mathbb{Z}$-module. L'exposant de G est le ppcm des ordres des éléments de G et c'est un diviseur de l'ordre $|G|$ de G. Si G est cyclique engendré par un élément g, l'exposant de G, l'ordre de G et l'ordre de g sont égaux.*

Rappelons qu'un *plus petit commun multiple* (ppcm) d'une famille finie d'entiers non nuls $a_1, \ldots, a_s$ est tout entier m non nul divisible par chacun de $a_1, \ldots, a_s$ et qui divise tout autre entier avec cette propriété (voir les exercices 5.2.4 et 5.4.6 pour l'existence du ppcm).

Preuve. Soit m le ppcm des ordres des éléments de G. Pour tout $g \in G$, on a $mg = 0$ car $\mathrm{o}(g) \,|\, m$, donc $m \in \mathrm{Ann}(G)$. Par ailleurs, si $n \in \mathrm{Ann}(G)$, alors $ng = O$ pour tout $g \in G$, donc $\mathrm{o}(g) \,|\, n$ pour tout $g \in G$ et par suite $m \,|\, n$. Cela montre que $\mathrm{Ann}(G) = m\mathbb{Z}$. Donc l'exposant de G est m. Le commentaire qui précède la proposition montre que c'est un diviseur de $|G|$.

Si $G = \mathbb{Z}g$ est cyclique engendré par g, on sait que $|G| = \mathrm{o}(g)$, et le corollaire 7.1.3 donne $\mathrm{Ann}(G) = \mathrm{Ann}(g)$. Donc l'exposant de G est $\mathrm{o}(g) = |G|$. □

Exemple 7.1.5. Soient $C_1 = \mathbb{Z}g_1$ et $C_2 = \mathbb{Z}g_2$ des groupes cycliques engendrés respectivement par un élément g_1 d'ordre 6 et un élément g_2 d'ordre 8. On forme leur produit cartésien

$$C_1 \times C_2 = \{(k\,g_1,\, \ell\,g_2)\,;\, k, \ell \in \mathbb{Z}\} = \mathbb{Z}(g_1, 0) + \mathbb{Z}(0, g_2).$$

1° Cherchons l'ordre de $(2g_1, 3g_2)$.

Pour un entier $m \in \mathbb{Z}$, on a

$$\begin{aligned} m(2g_1, 3g_2) = (0,0) &\iff 2mg_1 = 0 \quad \text{et} \quad 3mg_2 = 0 \\ &\iff 6 \,|\, 2m \quad \text{et} \quad 8 \,|\, 3m \\ &\iff 3 \,|\, m \quad \text{et} \quad 8 \,|\, m \\ &\iff 24 \,|\, m\,. \end{aligned}$$

Donc l'ordre de $(2g_1, 3g_2)$ est 24.

2º Cherchons l'exposant de $C_1 \times C_2$.

Pour un entier $m \in \mathbb{Z}$, on a

$$\begin{aligned} m \in \mathrm{Ann}(C_1 \times C_2) &\iff m(kg_1, \ell g_2) = (0,0) \quad \forall k, \ell \in \mathbb{Z} \\ &\iff mk\, g_1 = 0 \quad \text{et} \quad m\ell\, g_2 = 0 \quad \forall k, \ell \in \mathbb{Z} \\ &\iff m\, g_1 = 0 \quad \text{et} \quad m\, g_2 = 0 \\ &\iff 6 \,|\, m \quad \text{et} \quad 8 \,|\, m \\ &\iff 24 \,|\, m\,. \end{aligned}$$

Donc l'exposant de $C_1 \times C_2$ est 24.

Comme $|C_1 \times C_2| = |C_1|\ |C_2| = 6 \cdot 8 = 48$ est un multiple strict de l'exposant de $C_1 \times C_2$, le groupe $C_1 \times C_2$ n'est pas cyclique (voir la dernière partie de la proposition 7.1.4).

Dans le cas d'un espace vectoriel muni d'un endomorphisme, les notions d'annulateurs se traduisent ainsi :

Proposition 7.1.6. *Soit V un espace vectoriel sur un corps K et soit $T \in \mathrm{End}_K(V)$. Pour la structure correspondante de $K[x]$-module sur V, l'annulateur d'un élément $\mathbf{v}$ de V est*

$$\mathrm{Ann}(\mathbf{v}) = \{p(x) \in K[x]\,;\, p(T)(\mathbf{v}) = \mathbf{0}\},$$

tandis que l'annulateur de V est

$$\mathrm{Ann}(V) = \{p(x) \in K[x]\,;\, p(T) = 0\}.$$

Ce sont des idéaux de $K[x]$.

Preuve. La multiplication par un polynôme $p(x) \in K[x]$ est définie par $p(x)\mathbf{v} = p(T)\mathbf{v}$ pour tout $\mathbf{v} \in V$. Donc, pour $\mathbf{v} \in V$ fixé, on a :

$$p(x) \in \mathrm{Ann}(\mathbf{v}) \quad \iff \quad p(T)(\mathbf{v}) = \mathbf{0}$$

et en conséquence

$$p(x) \in \mathrm{Ann}(V) \quad \iff \quad p(T)(\mathbf{v}) = \mathbf{0} \quad \forall \mathbf{v} \in V \quad \iff \quad p(T) = 0.$$

□

On sait que tout idéal *non nul* de $K[x]$ possède un unique générateur unitaire, à savoir le polynôme unitaire de plus petit degré qui appartienne à cet idéal. Dans le contexte de la proposition 7.1.6, le résultat suivant fournit le moyen de calculer le générateur unitaire de $\mathrm{Ann}(\mathbf{v})$ lorsque V est de dimension finie.

Proposition 7.1.7. *Avec les notations de la proposition 7.1.6 supposons que V soit de dimension finie. Soit $\mathbf{v} \in V$ avec $\mathbf{v} \neq \mathbf{0}$. Alors :*

(i) il existe un plus grand entier positif m tel que $\mathbf{v}, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v})$ soient linéairement indépendants sur K ;

(ii) pour cet entier m, il existe un et seul choix d'éléments $a_0, a_1, \ldots, a_{m-1}$ de K tels que

$$T^m(\mathbf{v}) = a_0\mathbf{v} + a_1T(\mathbf{v}) + \cdots + a_{m-1}T^{m-1}(\mathbf{v});$$

(iii) le polynôme $p(x) = x^m - a_{m-1}x^{m-1} - \cdots - a_1x - a_0$ est le générateur unitaire de $\mathrm{Ann}(\mathbf{v})$;

(iv) le sous-$K[x]$-module U de V engendré par $\mathbf{v}$ est un sous-espace vectoriel T-invariant de V qui admet pour base $\mathcal{C} = \{\mathbf{v}, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v})\}$;

(v) on a $[T|_U]_{\mathcal{C}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{m-1} \end{pmatrix}$.

Avant de passer à la preuve de ce résultat, on donne un nom à la matrice qui apparaît dans la dernière assertion de la proposition.

Définition 7.1.8. Soit $m \in \mathbb{N}^*$ et soient $a_0, \ldots, a_{m-1}$ des éléments d'un corps K. La matrice

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{m-1} \end{pmatrix}$$

s'appelle la *matrice compagnon* du polynôme unitaire $x^m - a_{m-1}x^{m-1} - \cdots - a_1x - a_0 \in K[x]$.

Preuve de la proposition 7.1.7. Soit $n = \dim_K(V)$. On observe d'abord que, pour tout entier $m \geq 1$, les vecteurs $\mathbf{v}, T(\mathbf{v}), \ldots, T^m(\mathbf{v})$ sont linéairement dépendants si et seulement si l'annulateur $\mathrm{Ann}(\mathbf{v})$ de $\mathbf{v}$ contient un polynôme non nul de degré au plus m. Comme $\dim_K(V) = n$, on sait que $\mathbf{v}, T(\mathbf{v}), \ldots, T^n(\mathbf{v})$ sont linéairement dépendants et par suite $\mathrm{Ann}(\mathbf{v})$ contient un polynôme non nul de degré $\leq n$. En particulier $\mathrm{Ann}(\mathbf{v})$ est un idéal non nul de $K[x]$, et en tant que tel il possède un unique générateur unitaire $p(x)$, à savoir son élément unitaire de plus petit degré. Soit m le degré de $p(x)$. Comme $\mathbf{v} \neq \mathbf{0}$, on a $m \geq 1$. Écrivons

$$p(x) = x^m - a_{m-1}x^{m-1} - \cdots - a_1x - a_0.$$

Alors l'entier m est le plus grand entier tel que $\mathbf{v}, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v})$ soient linéairement indépendants et les éléments $a_0, a_1, \ldots, a_{m-1}$ de A sont les seuls pour lesquels

$$T^m(\mathbf{v}) - a_{m-1}T^{m-1}(\mathbf{v}) - \cdots - a_1 T(\mathbf{v}) - a_0\mathbf{v} = \mathbf{0},$$

c'est-à-dire tels que

$$T^m(\mathbf{v}) = a_0\mathbf{v} + a_1 T(\mathbf{v}) + \cdots + a_{m-1}T^{m-1}(\mathbf{v}). \tag{7.1}$$

Cela démontre (i), (ii) et (iii) tout à la fois. Par ailleurs, on a clairement

$$T(T^i(\mathbf{v})) \in \langle \mathbf{v}, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v})\rangle_K$$

pour $i = 0, \ldots, m-2$ et la formule (7.1) montre que cela est encore vrai pour $i = m-1$. Alors, en vertu de la proposition 2.6.4, le sous-espace $\langle \mathbf{v}, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v})\rangle_K$ est T-invariant. Par suite, c'est un sous-$K[x]$-module de V (voir la proposition 6.2.6). Comme il contient $\mathbf{v}$ et qu'il est contenu dans $U := K[x]\mathbf{v}$, on en déduit que ce sous-espace est U. Enfin, puisque $\mathbf{v}, T(\mathbf{v}), \ldots, T^{m-1}(\mathbf{v})$ sont linéairement indépendants, on conclut que ces m vecteurs forment une base de U. Cela démontre (iv). La dernière assertion (v) découle directement de (iv) en utilisant la relation (7.1). □

Exemple 7.1.9. Soit V un espace vectoriel de dimension 3 sur $\mathbb{Q}$, soit $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ une base de V, et soit $T \in \mathrm{End}_{\mathbb{Q}}(V)$. Supposons que

$$[T]_{\mathcal{B}} = \begin{pmatrix} 2 & -2 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix},$$

et cherchons le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_1)$.

On trouve

$$[T(\mathbf{v}_1)]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{v}_1]_{\mathcal{B}} = \begin{pmatrix} 2 & -2 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix},$$

$$[T^2(\mathbf{v}_1)]_{\mathcal{B}} = [T]_{\mathcal{B}}[T(\mathbf{v}_1)]_{\mathcal{B}} = \begin{pmatrix} 2 & -2 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix},$$

$$[T^3(\mathbf{v}_1)]_{\mathcal{B}} = [T]_{\mathcal{B}}[T^2(\mathbf{v}_1)]_{\mathcal{B}} = \begin{pmatrix} 2 & -2 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix},$$

donc

$$T(\mathbf{v}_1) = 2\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3, \quad T^2(\mathbf{v}_1) = \mathbf{v}_1 + 3\mathbf{v}_3, \quad T^3(\mathbf{v}_1) = -\mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3.$$

Comme

$$\det\Big([\mathbf{v}_1]_{\mathcal{B}}\,[T(\mathbf{v}_1)]_{\mathcal{B}}\,[T^2(\mathbf{v}_1)]_{\mathcal{B}}\Big) = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 3 \end{vmatrix} = 3 \neq 0,$$

les vecteurs $\mathbf{v}_1, T(\mathbf{v}_1), T^2(\mathbf{v}_1)$ sont linéairement indépendants. Comme $\dim_{\mathbb{Q}}(V) = 3$, le plus grand entier m tel que $\mathbf{v}_1, T(\mathbf{v}_1), \ldots, T^{m-1}(\mathbf{v}_1)$ soient linéairement indépendants est nécessairement $m = 3$. En vertu de la proposition 7.1.7, cela signifie que le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_1)$ est de degré 3. On trouve

$$\begin{aligned}
T^3(\mathbf{v}_1) = a_0\mathbf{v}_1 + a_1T(\mathbf{v}_1) + a_2T^2(\mathbf{v}_1) &\iff [T^3(\mathbf{v}_1)]_{\mathcal{B}} = a_0[\mathbf{v}_1]_{\mathcal{B}} + a_1[T(\mathbf{v}_1)]_{\mathcal{B}} + a_2[T^2(\mathbf{v}_1)]_{\mathcal{B}} \\
&\iff \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} = a_0 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_1 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \\
&\iff \begin{cases} a_0 + 2a_1 + a_2 = -1 \\ a_1 = -2 \\ a_1 + 3a_2 = 1 \end{cases} \\
&\iff a_0 = 2, \quad a_1 = -2 \quad \text{et} \quad a_2 = 1.
\end{aligned}$$

Ainsi on a

$$T^3(\mathbf{v}_1) = 2\mathbf{v}_1 - 2T(\mathbf{v}_1) + T^2(\mathbf{v}_1).$$

En vertu de la proposition 7.1.7, on conclut que le polynôme

$$p(x) = x^3 - x^2 + 2x - 2$$

est le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_1)$.

Comme $\dim_{\mathbb{Q}}(V) = 3$ et que $\mathbf{v}_1, T(\mathbf{v}_1), T^2(\mathbf{v}_1)$ sont linéairement indépendants sur $\mathbb{Q}$, l'ensemble ordonné $\mathcal{C} = \{\mathbf{v}_1, T(\mathbf{v}_1), T^2(\mathbf{v}_1)\}$ forme une base de V. Donc,

$$V = \langle \mathbf{v}_1, T(\mathbf{v}_1), T^2(\mathbf{v}_1) \rangle_{\mathbb{Q}} = \langle \mathbf{v}_1 \rangle_{\mathbb{Q}[x]}$$

est un $\mathbb{Q}[x]$-module cyclique engendré par $\mathbf{v}_1$ et la dernière assertion de la proposition 7.1.7 donne

$$[\,T\,]_{\mathcal{C}} = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}.$$

De la proposition 7.1.7, on déduit :

Proposition 7.1.10. *Soit V un espace vectoriel de dimension finie sur un corps K et soit $T \in \mathrm{End}_K(V)$. Pour la structure de $K[x]$-module correspondante sur V, l'idéal $\mathrm{Ann}(V)$ n'est pas nul.*

Preuve. Soit $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V. On a

$$V = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_K = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_{K[x]}.$$

Alors la proposition 7.1.2 donne

$$\mathrm{Ann}(V) = \mathrm{Ann}(\mathbf{v}_1) \cap \cdots \cap \mathrm{Ann}(\mathbf{v}_n).$$

Pour chaque $i = 1, \ldots, n$, la proposition 7.1.7 montre que $\mathrm{Ann}(\mathbf{v}_i)$ est engendré par un polynôme unitaire $p_i(x) \in K[x]$ de degré $\leq n$. Le produit $p_1(x) \cdots p_n(x)$ annule chacun des $\mathbf{v}_i$, donc il annule V tout entier. Cela montre que $\mathrm{Ann}(V)$ contient un polynôme non nul de degré $\leq n^2$. □

On conclut cette section avec la définition suivante.

Définition 7.1.11. Soit V un espace vectoriel sur un corps K et soit $T \in \mathrm{End}_K(V)$. Supposons que $\mathrm{Ann}(V) \neq \{0\}$ pour la structure correspondante de $K[x]$-module sur V. Alors le générateur unitaire de $\mathrm{Ann}(V)$ est appelé le *polynôme minimal* de T et noté $\min_T(x)$. En vertu de la proposition 7.1.6, c'est le polynôme unitaire $p(x) \in K[x]$ de plus petit degré tel que $p(T) = 0$.

Exemple 7.1.12. Dans l'exemple 7.1.9, V est un $\mathbb{Q}[x]$-module cyclique engendré par $\mathbf{v}_1$. En vertu du corollaire 7.1.3, on a donc

$$\mathrm{Ann}(V) = \mathrm{Ann}(\mathbf{v}_1),$$

et par suite le polynôme minimal de T est le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_1)$, c'est-à-dire que $\min_T(x) = x^3 - x^2 + 2x - 2$.

Exercices.

7.1.1. Démontrer la proposition 7.1.2 et son corollaire 7.1.3.

7.1.2. Soient $C_1 = \mathbb{Z}g_1$, $C_2 = \mathbb{Z}g_2$ et $C_3 = \mathbb{Z}g_3$ des groupes abéliens cycliques d'ordres respectifs 6, 10 et 15.

(i) Déterminer l'ordre de $(3g_1, 5g_2) \in C_1 \times C_2$.

(ii) Déterminer l'ordre de $(g_1, g_2, g_3) \in C_1 \times C_2 \times C_3$.

(iii) Déterminer l'exposant de $C_1 \times C_2$.

(iv) Déterminer l'exposant de $C_1 \times C_2 \times C_3$.

7.1.3. Soit V un espace vectoriel de dimension 3 sur $\mathbb{R}$, soit $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ une base de V, et soit $T \in \mathrm{End}_{\mathbb{R}}(V)$. Supposons que

$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 2 \\ 1 & -1 & -2 \end{pmatrix}.$$

Pour la structure correspondante de $\mathbb{R}[x]$-module sur V :

(i) déterminer le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_1)$;

(ii) déterminer le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_2)$ et en déduire celui de $\mathrm{Ann}(V)$.

7.1.4. Soit $T\colon \mathbb{Q}^3 \to \mathbb{Q}^3$ l'application linéaire dont la matrice relative à la base standard $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de $\mathbb{Q}^3$ est

$$[T]_{\mathcal{E}} = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 2 & 2 \\ -2 & 2 & 3 \end{pmatrix}.$$

Pour la structure correspondante de $\mathbb{Q}[x]$-module sur $\mathbb{Q}^3$:

(i) déterminer le générateur unitaire de $\mathrm{Ann}(\mathbf{e}_1 + \mathbf{e}_3)$;

(ii) montrer que $\mathbb{Q}^3$ est T-cyclique et en déduire $\min_T(x)$.

7.1.5. Soit V un espace vectoriel de dimension n sur un corps K et soit $T \in \mathrm{End}_K(V)$.

(i) Montrer que $I, T, \ldots, T^{n^2}$ sont linéairement dépendants sur K.

(ii) En déduire qu'il existe un polynôme non nul $p(x) \in K[x]$ de degré au plus n^2 tel que $p(T) = 0$.

(iii) Énoncer et démontrer le résultat correspondant pour une matrice $A \in \mathrm{Mat}_{n\times n}(K)$.

7.1.6. Soit V un espace vectoriel de dimension n sur un corps K et soit $T \in \mathrm{End}_K(V)$. Supposons que V soit T-cyclique. Montrer qu'il existe un polynôme non nul $p(x) \in K[x]$ de degré au plus n tel que $p(T) = 0$.

7.2 Modules sur un anneau euclidien

On énonce d'abord le théorème de structure pour un module de type fini sur un anneau euclidien, puis on examine ses conséquences pour les groupes abéliens et les espaces vectoriels munis d'un endomorphisme. La preuve du théorème sera donnée au chapitre 8.

Théorème 7.2.1. *Soit M un module de type fini sur un anneau euclidien A. Alors M est une somme directe de sous-modules cycliques*

$$M = A\mathbf{u}_1 \oplus A\mathbf{u}_2 \oplus \cdots \oplus A\mathbf{u}_s$$

avec $A \supsetneq \mathrm{Ann}(\mathbf{u}_1) \supseteq \mathrm{Ann}(\mathbf{u}_2) \supseteq \cdots \supseteq \mathrm{Ann}(\mathbf{u}_s)$.

Puisque A est euclidien, chacun des idéaux $\mathrm{Ann}(\mathbf{u}_i)$ est engendré par un élément d_i de A. La condition $\mathrm{Ann}(\mathbf{u}_i) \supseteq \mathrm{Ann}(\mathbf{u}_{i+1})$ signifie dans ce cas que $d_i \mid d_{i+1}$ si $d_i \neq 0$ tandis qu'elle implique $d_{i+1} = 0$ si $d_i = 0$. Il existe donc un entier t tel que $d_1, \ldots, d_t$ ne soient pas nuls avec $d_1 \mid d_2 \mid \ldots \mid d_t$ et $d_{t+1} = \cdots = d_s = 0$. La condition $A \neq \mathrm{Ann}(\mathbf{u}_1)$ signifie en outre que d_1 n'est pas une unité. Elle est facile à remplir car si $\mathrm{Ann}(\mathbf{u}_1) = A$, alors $\mathbf{u}_1 = 0$ et par suite on peut omettre le facteur $A\mathbf{u}_1$ de la somme directe.

Bien que la décomposition de M spécifiée par le théorème 7.2.1 ne soit en général pas unique, on peut montrer que l'entier s et les idéaux $\mathrm{Ann}(\mathbf{u}_1), \ldots, \mathrm{Ann}(\mathbf{u}_s)$ ne dépendent que de M et non du choix des $\mathbf{u}_i$. Par exemple, on a la caractérisation suivante de $\mathrm{Ann}(\mathbf{u}_s)$, dont la preuve est laissée en exercice :

Corollaire 7.2.2. *Avec les notations du théorème 7.2.1, on a* $\mathrm{Ann}(M) = \mathrm{Ann}(\mathbf{u}_s)$.

Dans le cas où $A = \mathbb{Z}$, un A-module est simplement un groupe abélien. Pour un groupe abélien fini, le théorème 7.2.1 et son corollaire livrent :

Théorème 7.2.3. *Soit G un groupe abélien fini. Alors G est une somme directe de sous-groupes cycliques*

$$G = \mathbb{Z}g_1 \oplus \mathbb{Z}g_2 \oplus \cdots \oplus \mathbb{Z}g_s$$

avec $o(g_1) \neq 1$ et $o(g_1) \,|\, o(g_2) \,|\, \ldots \,|\, o(g_s)$. De plus l'exposant de G est $o(g_s)$.

Pour la situation qui nous intéresse plus particulièrement dans ce cours, on obtient :

Théorème 7.2.4. *Soit V un espace vectoriel de dimension finie sur un corps K et soit $T \in \mathrm{End}_K(V)$. Pour la structure correspondante de $K[x]$-module sur V, on peut écrire*

$$V \;=\; K[x]\mathbf{v}_1 \oplus K[x]\mathbf{v}_2 \oplus \cdots \oplus K[x]\mathbf{v}_s \tag{7.2}$$

de telle sorte qu'en désignant par $d_i(x)$ le générateur unitaire de $\mathrm{Ann}(\mathbf{v}_i)$ pour $i = 1, \ldots, s$, on ait $\deg(d_1(x)) \neq 0$ et

$$d_1(x) \,|\, d_2(x) \,|\, \ldots \,|\, d_s(x).$$

Le polynôme minimal de T est $d_s(x)$.

En combinant ce résultat avec la proposition 7.1.7, on obtient :

Théorème 7.2.5. *Avec les notations du théorème 7.2.4, posons*

$$V_i = K[x]\mathbf{v}_i \quad \textit{et} \quad n_i = \deg(d_i(x))$$

pour $i = 1, \ldots, s$. Alors, l'ensemble ordonné

$$\mathcal{C} = \{\mathbf{v}_1, T(\mathbf{v}_1), \ldots, T^{n_1-1}(\mathbf{v}_1), \ldots\ldots, \mathbf{v}_s, T(\mathbf{v}_s), \ldots, T^{n_s-1}(\mathbf{v}_s)\}$$

est une base de V relative à laquelle la matrice de T est diagonale par blocs avec s blocs sur sa diagonale :

$$[T]_{\mathcal{C}} \;=\; \begin{pmatrix} D_1 & & & 0 \\ & D_2 & & \\ & & \ddots & \\ 0 & & & D_s \end{pmatrix}, \tag{7.3}$$

le i-ième bloc D_i de la diagonale étant la matrice compagnon du polynôme $d_i(x)$.

Preuve. En vertu de la proposition 7.1.7, chaque V_i est un sous-espace T-cyclique de V qui admet pour base

$$\mathcal{C}_i = \{\mathbf{v}_i, T(\mathbf{v}_i), \ldots, T^{n_i-1}(\mathbf{v}_i)\}$$

et on a

$$[T|_{V_i}]_{\mathcal{C}_i} = D_i.$$

Comme $V = V_1 \oplus \cdots \oplus V_s$, le théorème 2.6.6 montre que $\mathcal{C} = \mathcal{C}_1 \amalg \cdots \amalg \mathcal{C}_s$ est une base de V telle que

$$[T]_{\mathcal{C}} = \begin{pmatrix} [T|_{V_1}]_{\mathcal{C}_1} & & 0 \\ & \ddots & \\ 0 & & [T|_{V_s}]_{\mathcal{C}_s} \end{pmatrix} = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_s \end{pmatrix}.$$

□

Le polynômes $d_1(x), \ldots, d_s(x)$ qui apparaissent dans le théorème 7.2.4 s'appellent les *diviseurs élémentaires* de T et la matrice correspondante (7.3) fournie par le théorème 7.2.5 s'appelle la *forme rationnelle* de T. Il découle du théorème 7.2.5 que la somme des degrés des diviseurs élémentaires de T est égale à la dimension de V.

Exemple 7.2.6. Soit V un espace vectoriel sur $\mathbb{R}$ de dimension 5. Supposons que les diviseurs élémentaires de T soient

$$d_1(x) = x - 1, \quad d_2(x) = d_3(x) = x^2 - 2x + 1.$$

Alors il existe une base $\mathcal{C}$ de V telle que

$$[T]_{\mathcal{C}} = \begin{pmatrix} \boxed{1} & & & 0 \\ & \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & & \\ 0 & & \boxed{\begin{matrix} 0 & -1 \\ 1 & 2 \end{matrix}} & \end{pmatrix}$$

car la matrice compagnon de $x - 1$ est $\begin{pmatrix} 1 \end{pmatrix}$ et celle de $x^2 - 2x + 1$ est $\begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$.

La preuve du théorème 7.2.5 n'utilise pas les relations de divisibilité qui lient les polynômes $d_1(x), \ldots, d_s(x)$. Elle s'adapte en fait à toute décomposition de la forme (7.2).

Exemple 7.2.7. Soit V un espace vectoriel sur $\mathbb{C}$ de dimension 5 et soit $T\colon V \to V$ un opérateur linéaire sur $\mathbb{C}$. Supposons que, pour la structure correspondante de $\mathbb{C}[x]$-module sur V, on ait

$$V = \mathbb{C}[x]\mathbf{v}_1 \oplus \mathbb{C}[x]\mathbf{v}_2 \oplus \mathbb{C}[x]\mathbf{v}_3$$

et que les générateurs unitaires des annulateurs de $\mathbf{v}_1, \mathbf{v}_2$ et $\mathbf{v}_3$ soient respectivement

$$x^2 - i\,x + 1 - i, \quad x - 1 + 2i, \quad x^2 + (2+i)x - 3.$$

Alors $\mathcal{C} = \{\mathbf{v}_1, T(\mathbf{v}_1), \mathbf{v}_2, \mathbf{v}_3, T(\mathbf{v}_3)\}$ est une base de V et

$$[T]_{\mathcal{C}} = \begin{pmatrix} \boxed{\begin{matrix} 0 & -1+i \\ 1 & i \end{matrix}} & & & 0 \\ & \boxed{1-2i} & & \\ 0 & & \boxed{\begin{matrix} 0 & 3 \\ 1 & -2-i \end{matrix}} & \end{pmatrix}.$$

Exercices.

7.2.1. Démontrer le corollaire 7.2.2 sur la base du théorème 7.2.1.

7.2.2. Démontrer le théorème 7.2.3 sur la base du théorème 7.2.1 et de son corollaire 7.2.2.

7.2.3. Démontrer le théorème 7.2.4 sur la base du théorème 7.2.1 et de son corollaire 7.2.2.

7.2.4. Soit V un espace vectoriel réel de dimension finie muni d'un endomorphisme $T \in \mathrm{End}_{\mathbb{R}}(V)$. Supposons que, pour la structure correspondante de $\mathbb{R}[x]$-module sur V, on ait

$$V = \mathbb{R}[x]\mathbf{v}_1 \oplus \mathbb{R}[x]\mathbf{v}_2 \oplus \mathbb{R}[x]\mathbf{v}_3,$$

où $\mathbf{v}_1$, $\mathbf{v}_2$ et $\mathbf{v}_3$ sont des éléments de V dont les annulateurs respectifs sont engendrés par $x^2 - x + 1$, x^3 et $x - 2$. Déterminer la dimension de V, une base $\mathcal{C}$ de V associée à cette décomposition et, pour cette base, déterminer $[T]_{\mathcal{C}}$.

7.2.5. Soit $A \in \mathrm{Mat}_{n\times n}(K)$ où K est un corps. Montrer qu'il existe des polynômes unitaires non constants $d_1(x), \ldots, d_s(x) \in K[x]$ avec $d_1(x) \mid \cdots \mid d_s(x)$ et une matrice inversible $P \in \mathrm{Mat}_{n\times n}(K)$ telle que PAP^{-1} soit diagonale par blocs, avec les matrices compagnon de $d_1(x), \ldots, d_s(x)$ sur la diagonale.

7.3 Décomposition primaire

Le théorème 7.2.1 nous apprend que tout module de type fini sur un anneau euclidien est une somme directe de sous-modules cycliques. Pour obtenir une décomposition plus fine, il suffit de décomposer les modules cycliques. C'est l'objet de la décomposition primaire. Comme elle s'applique à tout module d'annulateur non nul, on considère le cas général qui n'est pas plus compliqué.

Avant de citer le résultat principal, rappelons que, si M est un module sur un anneau commutatif A, alors, pour chaque $a \in A$, on définit des sous-A-modules de M en posant

$$M_a := \{\mathbf{u} \in M; a\mathbf{u} = \mathbf{0}\} \quad \text{et} \quad aM = \{a\mathbf{u}; \mathbf{u} \in M\}$$

(voir l'exemple 6.5.12).

Théorème 7.3.1. *Soit M un module sur un anneau euclidien A. Supposons que $M \neq \{\mathbf{0}\}$ et que $\mathrm{Ann}(M)$ contienne un élément non nul a. Écrivons*

$$a = \epsilon q_1 \ldots q_s$$

où ϵ est une unité de A, s un entier ≥ 0, et $q_1, \ldots, q_s$ des éléments non nuls de A premiers deux à deux. Alors l'entier s est ≥ 1 et on a :

1) $M = M_{q_1} \oplus \cdots \oplus M_{q_s}$,

2) $M_{q_i} = q_1 \cdots \widehat{q_i} \cdots q_s\, M$ *pour* $i = 1, \ldots, s$.
De plus, si $\mathrm{Ann}(M) = (a)$, *alors* $\mathrm{Ann}(M_{q_i}) = (q_i)$ *pour* $i = 1, \ldots, s$.

Dans la formule 2), le chapeau sur q_i signifie que q_i est omis du produit. Si $s = 1$, ce produit est vide et on convient que $q_1 \cdots \widehat{q_i} \cdots q_s = 1$.

Preuve. Comme $\mathrm{Ann}(M)$ est un idéal de A, il contient

$$\epsilon^{-1} a = q_1 \cdots q_s.$$

Par ailleurs, puisque $M \neq \{0\}$, il ne contient pas 1. Donc, on doit avoir $s \geq 1$.

Si $s = 1$, on trouve que $M = M_{q_1} = \widehat{q_1} M$, et les formules 1) et 2) sont vérifiées. Supposons désormais $s \geq 2$ et définissons

$$a_i := q_1 \cdots \widehat{q_i} \cdots q_s \quad \text{pour } i = 1, \ldots, s.$$

Comme $q_1, \ldots, q_s$ sont premiers deux à deux, les produits $a_1, \ldots, a_s$ sont premiers dans leur ensemble et par suite, il existe $r_1, \ldots, r_s \in A$ tels que

$$r_1 a_1 + \cdots + r_s a_s = 1$$

(voir l'exercice 7.3.1). Alors, pour tout $u \in M$, on a :

$$u = (r_1 a_1 + \cdots + r_s a_s) u = r_1 a_1 u + \cdots + r_s a_s u. \tag{7.4}$$

Comme $r_i a_i u = a_i (r_i u) \in a_i M$ pour $i = 1, \ldots, s$, cela montre que

$$M = a_1 M + \cdots + a_s M. \tag{7.5}$$

Si $u \in M_{q_i}$, on a $a_j u = 0$ pour tout indice j distinct de i car $q_i \,|\, a_j$. Donc la formule (7.4) livre :

$$u = r_i a_i u \quad \text{pour tout } u \in M_{q_i}. \tag{7.6}$$

Cela montre que $M_{q_i} \subseteq a_i M$. Comme

$$q_i a_i M = \epsilon^{-1} b M = \{0\},$$

on a aussi $a_i M \subseteq M_{q_i}$, donc $a_i M = M_{q_i}$. Par suite l'égalité (7.5) devient

$$M = M_{q_1} + \cdots + M_{q_s}. \tag{7.7}$$

Pour montrer que cette somme est directe, choisissons $u_1 \in M_{q_1}, \ldots, u_s \in M_{q_s}$ tels que

$$u_1 + \cdots + u_s = 0.$$

En multipliant les deux membres de cette égalité par $r_i a_i$ et en notant que $r_i a_i u_j = 0$ pour $j \neq i$ (car $q_j \,|\, a_i$ si $j \neq i$), on obtient

$$r_i a_i u_i = 0.$$

Comme (7.6) donne $u_i = r_i a_i u_i$, on conclut que $u_i = 0$ pour $i = 1, \ldots, s$. En vertu de la définition 6.4.1, cela montre que la somme (7.7) est directe.

Enfin, supposons que $\mathrm{Ann}(M) = (a)$. Comme $M_{q_i} = a_i M$, on trouve

$$\begin{aligned} \mathrm{Ann}(M_{q_i}) &= \mathrm{Ann}(a_i M) \\ &= \{r \in A\,;\, r\, a_i M = \{0\}\} \\ &= \{r \in A\,;\, r\, a_i \in \mathrm{Ann}(M)\} \\ &= \{r \in A\,;\, a \,|\, r\, a_i\} \\ &= (q_i). \end{aligned}$$

□

Pour un module cyclique, le théorème 7.3.1 admet la conséquence suivante :

Corollaire 7.3.2. *Soit A un anneau euclidien et soit $M = A\,\mathbf{u}$ un A-module cyclique. Supposons que $\mathbf{u} \neq \mathbf{0}$ et que $\mathrm{Ann}(\mathbf{u}) \neq \{\mathbf{0}\}$. Choisissons un générateur a de $\mathrm{Ann}(\mathbf{u})$ et factorisons-le sous la forme*

$$a = \epsilon p_1^{e_1} \cdots p_s^{e_s}$$

où ϵ est une unité de A, s un entier ≥ 1, $p_1, \ldots, p_s$ des éléments irréductibles de A non associés deux à deux, et $e_1, \ldots, e_s$ des entiers positifs. Alors, pour $i = 1, \ldots, s$, le produit

$$\mathbf{u}_i = p_1^{e_1} \cdots \widehat{p_i^{e_i}} \cdots p_s^{e_s} \mathbf{u} \in M$$

satisfait $\mathrm{Ann}(\mathbf{u}_i) = (p_i^{e_i})$, et on a :

$$M = A\mathbf{u}_1 \oplus \cdots \oplus A\mathbf{u}_s. \tag{7.8}$$

Preuve. Comme $\mathrm{Ann}(M) = \mathrm{Ann}(\mathbf{u}) = (a)$, le théorème 7.3.1 s'applique avec $q_1 = p_1^{e_1}, \ldots, q_s = p_s^{e_s}$. Puisque

$$q_1 \cdots \widehat{q_i} \cdots q_s\, M = A\mathbf{u}_i \quad (1 \leq i \leq s),$$

il fournit la décomposition (7.8) et livre

$$\mathrm{Ann}(\mathbf{u}_i) = \mathrm{Ann}(A\mathbf{u}_i) = (q_i) = (p_i^{e_i}) \quad (1 \leq i \leq s).$$

□

En combinant ce résultat au théorème 7.2.1, on obtient :

Théorème 7.3.3. *Soit M un module de type fini sur un anneau euclidien. Supposons que $M \neq \{\mathbf{0}\}$ et que $\mathrm{Ann}(M) \neq \{0\}$. Alors M est une somme directe de sous-module cycliques dont l'annulateur est engendré par une puissance d'un élément irréductible de A.*

Preuve. Le corollaire 7.3.2 montre que ce résultat est vrai pour un module cyclique. Le cas général s'ensuit puisque, selon le théorème 7.2.1, le module M est une somme directe de sous-modules cycliques, et que l'annulateur de chacun de ces sous-modules n'est pas nul puisqu'il contient $\mathrm{Ann}(M)$. □

Exemple 7.3.4. Reprenons les notations de l'exemple 7.1.9. L'espace vectoriel V est un $\mathbb{Q}[x]$-module cyclique engendré par le vecteur $\mathbf{v}_1$ et l'idéal $\mathrm{Ann}(\mathbf{v}_1)$ est engendré par

$$p(x) = x^3 - x^2 + 2x - 2 = (x-1)(x^2+2).$$

Comme x^2+2 n'admet pas de racine réelle, il n'admet pas de racine dans $\mathbb{Q}$, donc c'est un polynôme irréductible de $\mathbb{Q}[x]$. Le polynôme $x-1$ est aussi irréductible, car son degré est 1. Alors le corollaire 7.3.2 donne

$$V = \mathbb{Q}[x]\mathbf{u}_1 \oplus \mathbb{Q}[x]\mathbf{u}_2$$

où

$$\begin{array}{rclcrcl} \mathbf{u}_1 & = & (x^2+2)\mathbf{v}_1 & = & T^2(\mathbf{v}_1) + 2\mathbf{v}_1 & = & 3\mathbf{v}_1 + 3\mathbf{v}_3, \\ \mathbf{u}_2 & = & (x-1)\mathbf{v}_1 & = & T(\mathbf{v}_1) - \mathbf{v}_1 & = & \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 \end{array}$$

satisfont

$$\mathrm{Ann}(\mathbf{u}_1) = (x-1) \quad \text{et} \quad \mathrm{Ann}(\mathbf{u}_2) = (x^2+2).$$

En raisonnant comme dans la preuve du théorème 7.2.5, on en déduit que

$$\mathcal{C} = \{\mathbf{u}_1, \mathbf{u}_2, T(\mathbf{u}_2)\} = \{3\mathbf{v}_1 + 3\mathbf{v}_3, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3, -\mathbf{v}_1 - \mathbf{v}_2 + 2\mathbf{v}_3)\}$$

est une base de V telle que

$$[T]_{\mathcal{C}} = \begin{pmatrix} \boxed{1} & 0 \\ 0 & \boxed{\begin{matrix} 0 & -2 \\ 1 & 0 \end{matrix}} \end{pmatrix}$$

soit diagonale par blocs avec les matrices compagnons de $x-1$ et de x^2+2 sur la diagonale.

On laisse au lecteur le soin d'analyser la signification du théorème 7.3.3 pour les groupes abéliens. Pour la situation qui nous intéresse plus particulièrement, on trouve :

Théorème 7.3.5. *Soit V un espace vectoriel de dimension finie sur un coprs K et soit $T\colon V \to V$ un opérateur linéaire sur V. Il existe une base de V relative à laquelle la matrice de T soit diagonale par blocs avec sur la diagonale les matrices compagnons de puissances de polynômes unitaires irréductibles de $K[x]$.*

Cette matrice s'appelle la *forme rationnelle primaire* de T. On peut montrer qu'elle est unique à une permutation près des blocs sur la diagonale.

On laisse la démonstration de ce résultat en exercice et on se contente de donner un exemple.

Exemple 7.3.6. Soit V un espace vectoriel de dimension 6 sur $\mathbb{R}$ et soit $T\colon V \to V$ un opérateur linéaire. Supposons que, pour la structure correspondante de $\mathbb{R}[x]$-module sur V, on ait

$$V = \mathbb{R}[x]\mathbf{v}_1 \oplus \mathbb{R}[x]\mathbf{v}_2$$

avec $\mathrm{Ann}(\mathbf{v}_1) = (x^3 - x^2)$ et $\mathrm{Ann}(\mathbf{v}_2) = (x^3 - 1)$.

1° On a $x^3 - x^2 = x^2(x-1)$ où x et $x-1$ sont irréductibles. Par suite,

$$\mathbb{R}[x]\mathbf{v}_1 = \mathbb{R}[x]\mathbf{v}_1' \oplus \mathbb{R}[x]\mathbf{v}_1''$$

où $\mathbf{v}_1' = (x-1)\mathbf{v}_1 = T(\mathbf{v}_1) - \mathbf{v}_1$ et $\mathbf{v}_1'' = x^2\mathbf{v}_1 = T^2(\mathbf{v}_1)$ satisfont

$$\mathrm{Ann}(\mathbf{v}_1') = (x^2) \quad \text{et} \quad \mathrm{Ann}(\mathbf{v}_1'') = (x-1).$$

2° De même, $x^3 - 1 = (x-1)(x^2+x+1)$ est la factorisation de x^3-1 en produit de puissances de polynômes irréductibles de $\mathbb{R}[x]$. Par suite,

$$\mathbb{R}[x]\mathbf{v}_2 = \mathbb{R}[x]\mathbf{v}_2' \oplus \mathbb{R}[x]\mathbf{v}_2''$$

où $\mathbf{v}_2' = (x^2+x+1)\mathbf{v}_2 = T^2(\mathbf{v}_2) + T(\mathbf{v}_2) + \mathbf{v}_2$ et $\mathbf{v}_2'' = (x-1)\mathbf{v}_2 = T(\mathbf{v}_2) - \mathbf{v}_2$ satisfont

$$\mathrm{Ann}(\mathbf{v}_2') = (x-1) \quad \text{et} \quad \mathrm{Ann}(\mathbf{v}_2'') = (x^2+x+1).$$

3° On conclut que

$$V = \mathbb{R}[x]\mathbf{v}_1' \oplus \mathbb{R}[x]\mathbf{v}_1'' \oplus \mathbb{R}[x]\mathbf{v}_2' \oplus \mathbb{R}[x]\mathbf{v}_2''$$

et que

$$\mathcal{C} = \{\mathbf{v}_1', T(\mathbf{v}_1'), \mathbf{v}_1'', \mathbf{v}_2', \mathbf{v}_2'', T(\mathbf{v}_2'')\}$$

est une base de V pour laquelle

$$[T]_{\mathcal{C}} = \begin{pmatrix} \boxed{\begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}} & & & 0 \\ & \boxed{1} & & \\ & & \boxed{1} & \\ 0 & & & \boxed{\begin{matrix} 0 & -1 \\ 1 & -1 \end{matrix}} \end{pmatrix},$$

avec sur la diagonale les matrices compagnons de x^2, $x-1$, $x-1$ et x^2+x+1.

Exercices.

7.3.1. Soit A un anneau euclidien, soient $p_1, \ldots, p_s$ des éléments irréductibles non associés de A et soient $e_1, \ldots, e_s$ des entiers positifs. Posons $q_i = p_i^{e_i}$ et $a_i = q_1 \cdots \widehat{q_i} \cdots q_s$ pour $i = 1, \ldots, s$. Montrer qu'il existe $r_1, \ldots, r_s \in A$ tels que $1 = r_1a_1 + \cdots + r_sa_s$.

7.3.2. Soit M un module sur un anneau euclidien A. Supposons que $\mathrm{Ann}(M) = (a)$ avec $a \neq 0$, et soit $c \in A$ avec $c \neq 0$. Montrer que $M_c = M_d$ où $d = \gcd(a, c)$.

7.3.3. Soit V un espace vectoriel sur un corps K et soit $T \in \mathrm{End}_K(V)$. On considère V comme un $k[x]$-module pour ce choix de T. Montrer que, dans ce contexte, les notions de sous-modules M_a et aM introduites dans l'exemple 6.5.12 et rappelées avant le théorème 7.3.1 se lisent :

$$V_{p(x)} = \ker(p(T)) \quad \text{et} \quad p(x)V = \mathrm{Im}(p(T)).$$

pour tout $p(x) \in k[x]$. En déduire que $\ker(p(T))$ et $\mathrm{Im}(p(T))$ sont des sous-espaces T-invariants de V.

7.3.4. Soit V un espace vectoriel de dimension finie sur un corps K et soit $T \in \mathrm{End}_K(V)$. Supposons que $p(x) \in k[x]$ soit un polynôme unitaire tel que $p(T) = 0$, et écrivons $p(x) = p_1(x)^{e_1} \cdots p_s(x)^{e_s}$ où $p_1(x), \ldots, p_s(x)$ sont des polynômes irréductibles unitaires distincts de $k[x]$ et $e_1, \ldots, e_s$ des entiers positifs. En combinant l'exercise 7.3.3 et le théorème 7.3.1, montrer que

(i) $V = (\ker p_1(T)^{e_1}) \oplus \cdots \oplus (\ker p_s(T)^{e_s})$,

(ii) $\ker p_i(T)^{e_i} = \mathrm{Im}\Big(p_1(T)^{e_1} \ldots \widehat{p_i(T)^{e_i}} \ldots p_s(T)^{e_s}\Big)$ pour $i = 1, \ldots s$.

7.3.5. Soit V un espace vectoriel de dimension finie n sur un corps K, et soit $T \in \mathrm{End}_K(V)$. Montrer que T est diagonalisable si et seulement si son polynôme minimal $\min_T(x)$ admet une factorisation de la forme

$$\min_T(x) = (x - \alpha_1) \cdots (x - \alpha_s)$$

où $\alpha_1, \ldots, \alpha_s$ sont des éléments distincts de K.

7.3.6. Soit $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'application linéaire dont la matrice relative à la base canonique $\mathcal{C}$ de $\mathbb{R}^3$ est

$$[T]_{\mathcal{C}} = \begin{pmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ 2 & 4 & -1 \end{pmatrix}.$$

(i) Montrer que $p(T) = 0$ où $p(x) = (x-1)(x-2) \in \mathbb{R}[x]$.

(ii) Pourquoi peut-on affirmer que $\mathbb{R}^3 = \ker(T - I) \oplus \ker(T - 2I)$?

(iii) Déterminer une base $\mathcal{B}_1$ pour $\ker(T - I)$ et une base $\mathcal{B}_2$ pour $\ker(T - 2I)$.

(iv) Donner la matrice de T par rapport à la base $\mathcal{B} = \mathcal{B}_1 \coprod \mathcal{B}_2$ de $\mathbb{R}^3$.

7.3.7. Soit V un espace vectoriel sur $\mathbb{Q}$ avec pour base $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, et soit $T\colon V \to V$ l'application linéaire dont la matrice relative à la base $\mathcal{A}$ est

$$[T]_{\mathcal{A}} = \begin{pmatrix} 5 & 1 & 5 \\ -3 & -2 & -2 \\ -3 & -1 & -3 \end{pmatrix}.$$

(i) Montrer que $p(T) = 0$ où $p(x) = (x+1)^2(x-2) \in \mathbb{Q}[x]$.

(ii) Pourquoi peut-on affirmer que $V = \ker(T + I)^2 \oplus \ker(T - 2I)$?

(iii) Déterminer une base $\mathcal{B}_1$ pour $\ker(T+I)^2$ et une base $\mathcal{B}_2$ pour $\ker(T-2I)$.

(iv) Donner la matrice de T par rapport à la base $\mathcal{B} = \mathcal{B}_1 \coprod \mathcal{B}_2$ de V.

7.3.8. Soit V un espace vectoriel de dimension finie sur $\mathbb{C}$ et soit $T\colon V \to V$ une application $\mathbb{C}$-linéaire. Supposons que le polynôme $p(x) = x^4 - 1$ satisfasse $p(T) = 0$. Factoriser $p(x)$ en produit de puissances de polynômes irréductibles dans $\mathbb{C}[x]$ et donner une décomposition correspondante de V comme somme directe de sous-espaces T-invariants.

7.3.9. Soit V un espace vectoriel de dimension finie sur $\mathbb{R}$ et soit $T\colon V \to V$ une application $\mathbb{R}$-linéaire. Supposons que le polynôme $p(x) = x^3 + 2x^2 + 2x + 4$ satisfasse $p(T) = 0$. Factoriser $p(x)$ en produit de puissances de polynômes irréductibles dans $\mathbb{R}[x]$ et donner une décomposition correspondante de V comme somme directe de sous-espaces T-invariants.

7.3.10 (Application aux équations différentielles). Soit D l'opérateur de dérivation sur l'espace vectoriel $\mathcal{C}_\infty(\mathbb{R})$ des fonctions indéfiniment différentiables de $\mathbb{R}$ dans $\mathbb{R}$. Le théorème de Rolle nous apprend que si une fonction $f\colon \mathbb{R} \to \mathbb{R}$ est différentiable en chaque point $x \in \mathbb{R}$ avec $f'(x) = 0$, alors f est une constante.

(i) Soit $\lambda \in \mathbb{R}$. Supposons qu'une fonction satisfasse $f'(x) = \lambda f(x)$ pour tout $x \in \mathbb{R}$. En appliquant le théorème de Rolle à la fonction $f(x)e^{-\lambda x}$, montrer qu'il existe une constante $C \in \mathbb{R}$ telle que $f(x) = Ce^{\lambda x}$ pour tout $x \in \mathbb{R}$. En déduire que $\ker(D - \lambda I) = \langle e^{\lambda x} \rangle_{\mathbb{R}}$.

(ii) Soient $\lambda_1, \ldots, \lambda_s$ des nombres réels distincts, et soit $p(x) = (x - \lambda_1) \cdots (x - \lambda_s) \in \mathbb{R}[x]$. En combinant le résultat de (i) avec le théorème 7.3.1, montrer que $\ker p(D)$ est un sous-espace de $\mathcal{C}_\infty(\mathbb{R})$ de dimension s qui admet pour base $\{e^{\lambda_1 x}, \ldots, e^{\lambda_s x}\}$.

(iii) Déterminer la forme générale des fonctions $f \in \mathcal{C}_\infty(\mathbb{R})$ telles que $f''' - 2f'' - f' + 2f = 0$.

7.4 Forme canonique de Jordan

Dans cette section, on fixe un espace vectoriel V de dimension finie sur $\mathbb{C}$ et un opérateur linéaire $T\colon V \to V$. On munit aussi V de la structure correspondante de $\mathbb{C}[x]$-module.

Comme les polynômes irréductibles unitaires de $\mathbb{C}[x]$ sont de la forme $x - \lambda$ avec $\lambda \in \mathbb{C}$, le théorème 7.3.3 donne

$$V = \mathbb{C}[x]\mathbf{v}_1 \oplus \cdots \oplus \mathbb{C}[x]\mathbf{v}_s$$

pour des éléments $\mathbf{v}_1, \ldots, \mathbf{v}_s$ de V dont les annulateurs sont de la forme

$$\mathrm{Ann}(\mathbf{v}_1) = \big((x - \lambda_1)^{e_1}\big),\ \ldots,\ \mathrm{Ann}(\mathbf{v}_s) = \big((x - \lambda_s)^{e_s}\big)$$

avec $\lambda_1, \ldots, \lambda_s \in \mathbb{C}$ et $e_1, \ldots, e_s \in \mathbb{N}^*$.

Posons $U_i = \mathbb{C}[x]\mathbf{v}_i$ pour $i = 1, \ldots, s$. Pour chaque i, la proposition 7.1.7 montre que $\mathcal{C}_i := \{\mathbf{v}_i, T(\mathbf{v}_i), \ldots, T^{e_i - 1}(\mathbf{v}_i)\}$ est une base de U_i et que $[T|_{U_i}]_{\mathcal{C}_i}$ est la matrice compagnon de $(x - \lambda_i)^{e_i}$. Alors le théorème 2.6.6, nous apprend que $\mathcal{C} = \mathcal{C}_1 \amalg \cdots \amalg \mathcal{C}_s$ est une base de

V relative à laquelle la matrice de $[T]_{\mathcal{C}}$ de T est diagonale par blocs avec, sur la diagonale, les matrices compagnons des polynômes $(x-\lambda_1)^{e_1}, \ldots, (x-\lambda_s)^{e_s}$. C'est la forme rationnelle primaire de T. Le lemme suivant propose un choix de base différent pour chacun des sous-espaces U_i.

Lemme 7.4.1. *Soit $\mathbf{v} \in V$ et soit $U := \mathbb{C}[x]\mathbf{v}$. Supposons que $\mathrm{Ann}(\mathbf{v}) = ((x-\lambda)^e)$ avec $\lambda \in \mathbb{C}$ et $e \in \mathbb{N}^*$. Alors,*

$$\mathcal{B} = \{(T-\lambda I)^{e-1}(\mathbf{v}), \ldots, (T-\lambda I)(\mathbf{v}), \mathbf{v}\}$$

est une base de U pour laquelle

$$[T|_U]_{\mathcal{B}} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \tag{7.9}$$

Preuve. Puisque $\mathrm{Ann}(\mathbf{v}) = ((x-\lambda)^e)$, la proposition 7.1.7 montre que

$$\mathcal{C} := \{\mathbf{v}, T(\mathbf{v}), \ldots, T^{e-1}(\mathbf{v})\}$$

est une base de U et, par conséquent, $\dim_{\mathbb{C}}(U) = e$. Posons

$$\mathbf{u}_i = (T-\lambda I)^{e-i}(\mathbf{v}) \quad \text{pour } i = 1, \ldots, e,$$

de sorte que $\mathcal{B}$ s'écrive

$$\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_e\}$$

(on convient que $\mathbf{u}_e = (T-\lambda I)^0(\mathbf{v}) = \mathbf{v}$). On trouve

$$T(\mathbf{u}_1) = (T-\lambda I)(\mathbf{u}_1) + \lambda \mathbf{u}_1 = (T-\lambda I)^e(\mathbf{v}) + \lambda \mathbf{u}_1 = \lambda \mathbf{u}_1 \tag{7.10}$$

car $(T-\lambda I)^e(\mathbf{v}) = (x-\lambda)^e \mathbf{v} = \mathbf{0}$. Pour $i = 2, \ldots, e$, on trouve aussi

$$T(\mathbf{u}_i) = (T-\lambda I)(\mathbf{u}_i) + \lambda \mathbf{u}_i = \mathbf{u}_{i-1} + \lambda \mathbf{u}_i\,. \tag{7.11}$$

Les relations (7.10) et (7.11) montrent en particulier que

$$T(\mathbf{u}_i) \in \langle \mathbf{u}_1, \ldots, \mathbf{u}_e \rangle_{\mathbb{C}} \quad \text{pour } i = 1, \ldots, e.$$

Donc $\langle \mathbf{u}_1, \ldots, \mathbf{u}_e \rangle_{\mathbb{C}}$ est un sous-espace T-invariant de U. Comme il contient $\mathbf{v} = \mathbf{u}_e$, il contient tous les éléments de $\mathcal{C}$ et par suite

$$\langle \mathbf{u}_1, \ldots, \mathbf{u}_e \rangle_{\mathbb{C}} = U.$$

Comme $\dim_{\mathbb{C}}(U) = e = |\mathcal{B}|$, on conclut que $\mathcal{B}$ est une base de U. De (7.10) et (7.11), on déduit

$$[T(\mathbf{u}_1)]_{\mathcal{B}} = [\lambda \mathbf{u}_1]_{\mathcal{B}} = \begin{pmatrix} \lambda \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{et} \quad [T(\mathbf{u}_i)]_{\mathcal{B}} = [\mathbf{u}_{i-1} + \lambda \mathbf{u}_i]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \lambda \\ \vdots \\ 0 \end{pmatrix} \leftarrow i \quad \text{pour } i = 2, \ldots, e,$$

d'où la forme (7.9) pour la matrice $[T|_U]_{\mathcal{B}}$. □

Définition 7.4.2. Pour chaque $\lambda \in \mathbb{C}$ et chaque $e \in \mathbb{N}^*$, on désigne par $J_\lambda(e)$ la matrice carrée de taille e avec les coefficients de la diagonale égaux à λ, les coefficients juste au-dessus de la diagonale égaux à 1, et tous les autres coefficients égaux à 0 :

$$J_\lambda(e) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ 0 & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathrm{Mat}_{e\times e}(\mathbb{C}).$$

On l'appelle le *bloc de Jordan* de format $e \times e$ pour λ.

Grâce au lemme 7.4.1 et à la discussion qui précède, on conclut :

Théorème 7.4.3. *Il existe une base $\mathcal{B}$ de V telle que $[T]_\mathcal{B}$ soit diagonale par blocs avec des blocs de Jordan sur la diagonale.*

Preuve. Le lemme 7.4.1 montre que, pour $i = 1, \ldots, s$, le sous-espace U_i admet pour base

$$\mathcal{B}_i = \{(T - \lambda_i I)^{e_i - 1}(\mathbf{v}_i), \ldots, (T - \lambda_i I)(\mathbf{v}_i), \mathbf{v}_i\}$$

et que

$$[T|_{U_i}]_{\mathcal{B}_i} = J_{\lambda_i}(e_i).$$

Alors, suivant le théorème 2.6.6, l'ensemble ordonné $\mathcal{B} := \mathcal{B}_1 \amalg \cdots \amalg \mathcal{B}_s$ est une base de V pour laquelle

$$[T]_\mathcal{B} = \begin{pmatrix} J_{\lambda_1}(e_1) & & & 0 \\ & J_{\lambda_2}(e_2) & & \\ & & \ddots & \\ 0 & & & J_{\lambda_s}(e_s) \end{pmatrix}. \tag{7.12}$$

□

On dit que (7.12) est une *forme canonique de Jordan* de T. Le résultat suivant montre que celle-ci est essentiellement unique, et donne le moyen de la calculer :

Théorème 7.4.4. *Les formes canoniques de Jordan de T s'obtiennent l'une de l'autre par une permutation des blocs de Jordan sur la diagonale. Les blocs en question sont de la forme $J_\lambda(e)$ où λ est une valeur propre de T et où e est un entier ≥ 1. Pour chaque $\lambda \in \mathbb{C}$ et chaque entier $k \geq 1$, on a*

$$\dim_\mathbb{C} \ker(T - \lambda I)^k = \sum_{e \geq 1} \min(k, e)\, n_\lambda(e) \tag{7.13}$$

où $n_\lambda(e)$ désigne le nombre de blocs de Jordan égaux à $J_\lambda(e)$ sur la diagonale.

Avant de passer à la démonstration de ce théorème, on explique d'abord comment la formule (7.13) permet de calculer les entiers $n_\lambda(e)$ et donc de déterminer la forme canonique de Jordan de T. Pour cela, on fixe $\lambda \in \mathbb{C}$. On note qu'il existe un entier $r \geq 1$ tel que $n_\lambda(e) = 0$ pour tout $e > r$ et on pose

$$m_k = \dim_{\mathbb{C}} \ker(T - \lambda I)^k$$

pour $k = 1, \ldots, r$. Alors la formule (7.13) donne

$$\begin{aligned}
m_1 &= n_\lambda(1) + \ n_\lambda(2) + \ n_\lambda(3) + \cdots + \ n_\lambda(r), \\
m_2 &= n_\lambda(1) + 2n_\lambda(2) + 2n_\lambda(3) + \cdots + 2n_\lambda(r), \\
m_3 &= n_\lambda(1) + 2n_\lambda(2) + 3n_\lambda(3) + \cdots + 3n_\lambda(r), \\
&\cdots \\
m_r &= n_\lambda(1) + 2n_\lambda(2) + 3n_\lambda(3) + \cdots + rn_\lambda(r).
\end{aligned} \tag{7.14}$$

On en déduit

$$\begin{aligned}
m_1 &= n_\lambda(1) + n_\lambda(2) + n_\lambda(3) + \cdots + n_\lambda(r), \\
m_2 - m_1 &= n_\lambda(2) + n_\lambda(3) + \cdots + n_\lambda(r), \\
m_3 - m_2 &= n_\lambda(3) + \cdots + n_\lambda(r), \\
&\cdots \\
m_r - m_{r-1} &= n_\lambda(r).
\end{aligned}$$

Donc, on peut retrouver $n_\lambda(1), \ldots, n_\lambda(r)$ à partir de la donnée de $m_1, \ldots, m_r$. Explicitement, cela donne

$$n_\lambda(e) = \begin{cases} 2m_1 - m_2 & \text{si } e = 1, \\ 2m_e - m_{e-1} - m_{e+1} & \text{si } 1 < e < r, \\ m_r - m_{r-1} & \text{si } e = r. \end{cases} \tag{7.15}$$

Preuve du théorème 7.4.4. Supposons que $[T]_{\mathcal{B}}$ soit sous la forme de Jordan (7.12) pour une base $\mathcal{B}$ de V.

L'exercice 7.4.1 à la fin de cette section montre que, pour toute permutation $J'_1, \ldots, J'_s$ des blocs $J_{\lambda_1}(e_1), \ldots, J_{\lambda_s}(e_s)$, il existe une base $\mathcal{B}'$ de V telle que

$$[T]_{\mathcal{B}'} = \begin{pmatrix} J'_1 & & 0 \\ & \ddots & \\ 0 & & J'_s \end{pmatrix}.$$

Pour montrer la première affirmation du théorème, il suffit donc de vérifier la formule (7.13) car, comme le commentaire qui suit le théorème le montre, cette formule permet de calculer tous les entiers $n_\lambda(e)$ et par suite détermine la forme de Jordan de T à une permutation près des blocs sur la diagonale.

Fixons $\lambda \in \mathbb{C}$ et $k \in \mathbb{N}^*$, et posons $n = \dim_{\mathbb{C}}(V)$. On a

$$\begin{aligned}
\dim_{\mathbb{C}} \ker(T-\lambda I)^k &= n - \operatorname{rang}[(T-\lambda I)^k]_{\mathcal{B}} \\
&= n - \operatorname{rang}([T]_{\mathcal{B}} - \lambda I)^k \\
&= n - \operatorname{rang} \begin{pmatrix} J_{\lambda_1}(e_1) - \lambda I & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_s}(e_s) - \lambda I \end{pmatrix}^k \\
&= n - \operatorname{rang} \begin{pmatrix} (J_{\lambda_1}(e_1) - \lambda I)^k & & 0 \\ & \ddots & \\ 0 & & (J_{\lambda_s}(e_s) - \lambda I)^k \end{pmatrix} \\
&= n - \big(\operatorname{rang}(J_{\lambda_1}(e_1) - \lambda I)^k + \cdots + \operatorname{rang}(J_{\lambda_s}(e_s) - \lambda I)^k\big).
\end{aligned}$$

Comme par ailleurs, on a $n = e_1 + \cdots + e_s$, on conclut que

$$\dim_{\mathbb{C}} \ker(T-\lambda I)^k = \sum_{i=1}^{s} \big(e_i - \operatorname{rang}(J_{\lambda_i}(e_i) - \lambda I)^k\big).$$

Pour $i = 1, \ldots, s$, on trouve

$$J_{\lambda_i}(e_i) - \lambda I = \begin{pmatrix} \lambda_i - \lambda & 1 & & 0 \\ & \lambda_i - \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i - \lambda \end{pmatrix} = J_{\lambda_i - \lambda}(e_i)$$

donc la formule devient

$$\dim_{\mathbb{C}} \ker(T-\lambda I)^k = \sum_{i=1}^{s} \big(e_i - \operatorname{rang}(J_{\lambda_i - \lambda}(e_i))^k\big). \tag{7.16}$$

Si $\lambda_i \neq \lambda$, on trouve

$$\det(J_{\lambda_i - \lambda}(e_i)) = (\lambda_i - \lambda)^{e_i} \neq 0,$$

donc $J_{\lambda_i - \lambda}(e_i)$ est inversible. Alors $J_{\lambda_i - \lambda}(e_i)^k$ est aussi inversible pour chaque entier $k \geq 1$. Par suite, on a

$$\operatorname{rang}(J_{\lambda_i - \lambda}(e_i)^k) = e_i \quad \text{ si } \lambda_i \neq \lambda. \tag{7.17}$$

Si $\lambda_i = \lambda$, on a

$$J_{\lambda_i - \lambda}(e_i) = J_0(e_i) = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

et on vérifie sans peine par récurrence sur k, que

$$J_0(e_i)^k = \left(\begin{array}{c|c} 0 & I_{e_i-k} \\ \hline 0_k & 0 \end{array}\right) \qquad \text{pour } k = 1, \ldots, e_i - 1 \tag{7.18}$$

et que $J_0(e_i)^k = 0$ si $k \geq e_i$ (dans la formule (7.18), le symbole I_{e_i-k} représente la matrice identité de format $(e_i - k) \times (e_i - k)$, le symbole 0_k représente la matrice nulle de format $k \times k$, et les symboles 0 sur la diagonale sont des matrices nulles de formats mixtes). On en déduit que

$$\operatorname{rang}(J_0(e_i)^k) = \begin{cases} e_i - k & \text{si } 1 \leq k < e_i, \\ 0 & \text{si } k \geq e_i. \end{cases} \tag{7.19}$$

Grâce à (7.17) et (7.19), la formule (7.16) devient

$$\dim_{\mathbb{C}} \ker(T - \lambda I)^k = \sum_{\{i\,;\,\lambda_i = \lambda\}} \min(k, e_i) = \sum_{e \geq 1} \min(k, e) n_\lambda(e)$$

où, pour chaque entier $e \geq 1$, l'entier $n_\lambda(e)$ représente le nombre d'indices $i \in \{1, \ldots, s\}$ tels que $(\lambda_i, e_i) = (\lambda, e)$, donc le nombre de blocs égaux à $J_\lambda(e)$ sur la diagonale de (7.12). Cela démontre la formule (7.13) et par la même occasion complète la preuve de la première assertion du théorème. Les exercices à la fin de cette section proposent une autre preuve de (7.13).

Enfin, on trouve que

$$\begin{aligned} \operatorname{car}_T(x) = \det(xI - [\,T\,]_{\mathcal{B}}) &= \begin{vmatrix} xI - J_{\lambda_1}(e_1) & & 0 \\ & \ddots & \\ 0 & & xI - J_{\lambda_s}(e_s) \end{vmatrix} \\ &= \det(xI - J_{\lambda_1}(e_1)) \cdots \det(xI - J_{\lambda_s}(e_s)) \\ &= (x - \lambda_1)^{e_1} \cdots (x - \lambda_s)^{e_s}. \end{aligned}$$

Par suite, les nombres λ pour lesquels la forme de Jordan (7.12) de T contient un bloc $J_\lambda(e)$ avec $e \geq 1$ sont les valeurs propres de T. Cela démontre la seconde assertion du théorème et conclut la démonstration. □

Pour déterminer les entiers $n_\lambda(e)$ grâce aux formules (7.14), il faut connaître, pour chaque valeur propre λ de T, le plus grand entier r tel que $n_\lambda(r) \geq 1$. Le corollaire suivant fournit le moyen de calculer cet entier.

Corollaire 7.4.5. *Soit λ une valeur propre de T. Posons $m_k = \dim_{\mathbb{C}} \ker(T - \lambda I)^k$ pour tout $k \geq 1$. Alors on a*

$$m_1 < m_2 < \cdots < m_r = m_{r+1} = m_{r+2} = \cdots$$

où r désigne le plus grand entier pour lequel $n_\lambda(r) \geq 1$.

Preuve. Soit r le plus grand entier pour lequel $n_\lambda(r) \geq 1$. Les égalités (7.14) montrent que $m_k - m_{k-1} \geq n_\lambda(r) \geq 1$ pour $k = 2, \ldots, r$ et par suite $m_1 < m_2 < \cdots < m_r$. Pour $k \geq r$, la formule (7.13) donne $m_k = m_r$. □

Exemple 7.4.6. Soit $T\colon \mathbb{C}^{10} \to \mathbb{C}^{10}$ une application linéaire. Supposons que ses valeurs propres soient $1+i$ et 2, et que les entiers

$$m_k = \dim_{\mathbb{C}} \ker(T - (1+i)I)^k \quad \text{et} \quad m'_k = \dim_{\mathbb{C}} \ker(T - 2I)^k$$

satisfassent

$$m_1 = 3, \quad m_2 = 5, \quad m_3 = 6, \quad m_4 = 6, \quad m'_1 = 2, \quad m'_2 = 4 \quad \text{et} \quad m'_3 = 4.$$

Comme $m_1 < m_2 < m_3 = m_4$, le corollaire 7.4.5 montre que le plus grand bloc de Jordan de T pour la valeur propre $\lambda = 1+i$ est de format 3×3. De même, comme $m'_1 < m'_2 = m'_3$, le plus grand bloc de Jordan de T pour la valeur propre $\lambda = 2$ est de format 2×2.

Les formules (7.14) pour $\lambda = 1+i$ se lisent

$$\begin{aligned} m_1 = 3 &= n_{1+i}(1) + \ n_{1+i}(2) + \ n_{1+i}(3), \\ m_2 = 5 &= n_{1+i}(1) + 2n_{1+i}(2) + 2n_{1+i}(3), \\ m_3 = 6 &= n_{1+i}(1) + 2n_{1+i}(2) + 3n_{1+i}(3), \end{aligned}$$

donc $n_{1+i}(1) = n_{1+i}(2) = n_{1+i}(3) = 1$. Pour $\lambda = 2$, elles s'écrivent

$$\begin{aligned} m'_1 = 2 &= n_2(1) + \ n_2(2), \\ m'_2 = 4 &= n_2(1) + 2n_2(2), \end{aligned}$$

donc $n_2(1) = 0$ et $n_2(2) = 2$. On conclut que la forme canonique de Jordan de T est la matrice diagonale par blocs avec, sur la diagonale, les blocs

$$(1+i), \begin{pmatrix} 1+i & 1 \\ 0 & 1+i \end{pmatrix}, \begin{pmatrix} 1+i & 1 & 0 \\ 0 & 1+i & 1 \\ 0 & 0 & 1+i \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

Exemple 7.4.7. Soit $T\colon \mathbb{C}^4 \to \mathbb{C}^4$ l'application linéaire dont la matrice relative à la base canonique $\mathcal{E}$ de $\mathbb{C}^4$ est

$$[T]_{\mathcal{E}} = \begin{pmatrix} 1 & 0 & -1 & 4 \\ 0 & 1 & -1 & 5 \\ 1 & 0 & -1 & 8 \\ 1 & -1 & 0 & 3 \end{pmatrix}.$$

On cherche à déterminer la forme canonique de Jordan de T.

Solution: Les calculs donnent

$$\mathrm{car}_T(x) = \det(xI - [T]_{\mathcal{E}}) = x^2(x-2)^2.$$

Donc les valeurs propres de T sont 0 et 2. On pose

$$m_k = \dim_{\mathbb{C}} \ker(T-0I)^k = 4 - \operatorname{rang}[\,T^k\,]_{\mathcal{E}} = 4 - \operatorname{rang}([\,T\,]_{\mathcal{E}})^k$$
$$m'_k = \dim_{\mathbb{C}} \ker(T-2I)^k = 4 - \operatorname{rang}[(T-2I)^k]_{\mathcal{E}} = 4 - \operatorname{rang}([\,T\,]_{\mathcal{E}} - 2I)^k$$

pour tout entier $k \geq 1$.

On trouve, après des suites de transformations élémentaires de lignes, les formes échelonnées réduites suivantes :

$$[\,T\,]_{\mathcal{E}} \sim \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$[\,T\,]^2_{\mathcal{E}} = \begin{pmatrix} 4 & -4 & 0 & 8 \\ 4 & -4 & 0 & 12 \\ 8 & -8 & 0 & 20 \\ 4 & -4 & 0 & 8 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$[\,T\,]^3_{\mathcal{E}} = \begin{pmatrix} 12 & -12 & 0 & 20 \\ 16 & -16 & 0 & 32 \\ 28 & -28 & 0 & 52 \\ 12 & -12 & 0 & 20 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$[\,T\,]_{\mathcal{E}} - 2I = \begin{pmatrix} -1 & 0 & -1 & 4 \\ 0 & -1 & -1 & 5 \\ 1 & 0 & -3 & 8 \\ 1 & -1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$([\,T\,]_{\mathcal{E}} - 2I)^2 = \begin{pmatrix} 4 & -4 & 4 & -8 \\ 4 & -4 & 4 & -8 \\ 4 & -8 & 8 & -12 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$([\,T\,]_{\mathcal{E}} - 2I)^3 = \begin{pmatrix} -8 & -12 & -12 & 20 \\ -8 & -12 & -12 & 20 \\ -8 & -20 & -20 & 28 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

donc $m_1 = 1$, $m_2 = m_3 = 2$, $m'_1 = 1$ et $m'_2 = m'_3 = 2$. Ainsi le plus grand bloc de Jordan de T pour la valeur propre 0 est de format 2×2 et celui pour la valeur propre 2 est aussi de format 2×2. Comme la somme des tailles des blocs doit être égale à 4, on conclut que la

forme canonique de Jordan de T est

$$\begin{pmatrix} J_0(2) & 0 \\ 0 & J_2(2) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Par contre, cela ne dit pas comment déterminer une base $\mathcal{B}$ de $\mathbb{C}^4$ telle que $[T]_{\mathcal{B}}$ ait cette forme. Les considérations du chapitre suivant donnent en principe le moyen de le faire.

Le pendant matriciel du théorème 7.4.3 est l'énoncé suivant :

Théorème 7.4.8. *Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. Il existe une matrice inversible $P \in GL_n(\mathbb{C})$ telle que $P^{-1}AP$ soit diagonale par blocs avec des blocs de Jordan sur sa diagonale.*

Un tel produit $P^{-1}AP = \begin{pmatrix} J_{\lambda_1}(e_1) & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_s}(e_s) \end{pmatrix}$ s'appelle une *forme canonique de Jordan de A.*

Le pendant matriciel du théorème 7.4.4 s'énonce ainsi :

Théorème 7.4.9. *Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. Les formes canoniques de Jordan de A se déduisent l'une de l'autre par une permutation des blocs de Jordan sur la diagonale. Les blocs en question sont de la forme $J_\lambda(e)$ où λ est une valeur propre de A et où e est un entier ≥ 1. Pour chaque $\lambda \in \mathbb{C}$ et chaque entier $k \geq 1$, on a*

$$n - \mathrm{rang}(A - \lambda I)^k = \sum_{e\geq 1} \min(k, e) n_\lambda(e)$$

où $n_\lambda(e)$ désigne le nombre de blocs égaux à $J_\lambda(e)$ sur la diagonale.

Ce résultat fournit le moyen de déterminer si deux matrices à coefficients complexes sont semblables ou non :

Corollaire 7.4.10. *Soient $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. Alors A et B sont semblables si et seulement si elles admettent une forme canonique de Jordan commune.*

La déduction de cet énoncé est laissée en exercice.

Exercices.

7.4.1. Soit V un espace vectoriel sur $\mathbb{C}$ de dimension finie et soit $T\colon V \to V$ un opérateur linéaire. Supposons qu'il existe une base $\mathcal{B}$ de V telle que

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}$$

où $A_i \in \mathrm{Mat}_{e_i \times e_i}(\mathbb{C})$ pour $i = 1, \ldots, s$. Écrivons

$$\mathcal{B} = \mathcal{B}_1 \amalg \cdots \amalg \mathcal{B}_s$$

où $\mathcal{B}_i$ consiste de e_i éléments pour $i = 1, \ldots, s$. Enfin, pour chaque i, désignons par V_i le sous-espace de V engendré par $\mathcal{B}_i$.

(a) Montrer que chaque sous-espace V_i est T-invariant, qu'il admet $\mathcal{B}_i$ pour base et que

$$[T|_{V_i}]_{\mathcal{B}_i} = A_i.$$

(b) Montrer que si $(i_1, \ldots, i_s)$ est une permutation de $(1, \ldots, s)$, alors $\mathcal{B}' = \mathcal{B}_{i_1} \amalg \cdots \amalg \mathcal{B}_{i_s}$ est une base de V telle que

$$[T]_{\mathcal{B}'} = \begin{pmatrix} A_{i_1} & & 0 \\ & \ddots & \\ 0 & & A_{i_s} \end{pmatrix}.$$

7.4.2. Soit V un espace vectoriel sur $\mathbb{C}$ et soit $T\colon V \to V$ un opérateur linéaire. Supposons que

$$V = V_1 \oplus \cdots \oplus V_s$$

où V_i est un sous-espace T-invariant de V pour $i = 1, \ldots, s$.

(a) Montrer que $\ker(T) = \ker(T|_{V_1}) \oplus \cdots \oplus \ker(T|_{V_s})$.

(b) En déduire que, pour tout $\lambda \in \mathbb{C}$ et tout $k \in \mathbb{N}^*$, on a

$$\ker(T - \lambda I_V)^k = \ker(T|_{V_1} - \lambda I_{V_1})^k \oplus \cdots \oplus \ker(T|_{V_s} - \lambda I_{V_s})^k.$$

7.4.3. Soit $T\colon V \to V$ un opérateur linéaire sur un espace vectoriel V de dimension n sur $\mathbb{C}$. Supposons qu'il existe une base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ de V et $\lambda \in \mathbb{C}$ tels que

$$[T]_{\mathcal{B}} = J_\lambda(n).$$

(a) Montrer que, si $\lambda \neq 0$, alors $\dim_{\mathbb{C}}(\ker T^k) = 0$ pour tout entier $k \geq 1$.

(b) Supposons que $\lambda = 0$. Montrer que

$$\ker(T^k) = \langle \mathbf{v}_1, \ldots, \mathbf{v}_k \rangle_{\mathbb{C}}$$

pour $k = 1, \ldots, n$. En déduire que $\dim_{\mathbb{C}} \ker(T^k) = \min(k, n)$ pour tout entier $k \geq 1$.

7.4.4. Donner une démonstration alternative de la formule (7.13) du théorème 7.4.4 en procédant de la manière suivante. On suppose d'abord que $\mathcal{B}$ est une base de V telle que

$$[T]_{\mathcal{B}} = \begin{pmatrix} J_{\lambda_1}(e_1) & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_s}(e_s) \end{pmatrix}.$$

D'après l'exercice 7.4.1, on peut écrire $\mathcal{B} = \mathcal{B}_1 \amalg \cdots \amalg \mathcal{B}_s$ où, pour chaque i, $\mathcal{B}_i$ est une base d'un sous-espace T-invariant V_i de V avec

$$[T|_{V_i}]_{\mathcal{B}_i} = J_{\lambda_i}(e_i)$$

et on a $V = V_1 \oplus \cdots \oplus V_s$.

(a) En appliquant l'exercice 7.4.2, montrer que

$$\dim_{\mathbb{C}} \ker(T - \lambda I)^k = \sum_{i=1}^{s} \dim_{\mathbb{C}} \ker(T|_{V_i} - \lambda I_{V_i})^k$$

pour chaque $\lambda \in \mathbb{C}$ et chaque entier $k \geq 1$.

(b) Montrer que $[T|_{V_i} - \lambda I_{V_i}]_{\mathcal{B}_i} = J_{\lambda_i - \lambda}(e_i)$ et, grâce à l'exercice 7.4.3, en déduire que

$$\dim_{\mathbb{C}} \ker(T|_{V_i} - \lambda I_{V_i})^k = \begin{cases} 0 & \text{si } \lambda \neq \lambda_i\,, \\ \min(e_i, k) & \text{si } \lambda = \lambda_i\,. \end{cases}$$

(c) En déduire la formule (7.13).

7.4.5. Déduire le théorème 7.4.8 du théorème 7.4.3 en appliquant ce dernier à l'opérateur linéaire $T_A \colon \mathbb{C}^n \to \mathbb{C}^n$ associé à A. De la même façon, déduire le théorème 7.4.9 du théorème 7.4.4.

7.4.6. Démontrer le corollaire 7.4.10.

7.4.7. Soit $T \colon \mathbb{C}^7 \to \mathbb{C}^7$ une application linéaire. Supposons que ses valeurs propres soient 0, 1 et $-i$ et que les entiers

$$m_k = \dim_{\mathbb{C}} \ker T^k, \quad m'_k = \dim_{\mathbb{C}} \ker(T - I)^k, \quad m''_k = \dim_{\mathbb{C}} \ker(T + iI)^k$$

satisfassent $m_1 = m_2 = 2$, $m'_1 = 1$, $m'_2 = 2$, $m'_3 = m'_4 = 3$, $m''_1 = 1$ et $m''_2 = m''_3 = 2$. En déduire la forme canonique de Jordan de T.

7.4.8. Déterminer la forme canonique de Jordan de la matrice

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

Chapitre 8

Le théorème des diviseurs élémentaires

Ce chapitre conclut la première partie du cours. On énonce d'abord le théorème des diviseurs élémentaires puis on montre qu'il entraîne le théorème de structure pour les modules de type fini sur un anneau euclidien. Dans la section 8.2, on démontre le théorème de Cayley-Hamilton selon lequel, pour tout opérateur linéaire $T\colon V \to V$ sur un espace vectoriel V de dimension finie, on a $\mathrm{car}_T(T) = 0$. Le reste du chapitre est consacré à la preuve du théorème des diviseurs élémentaires. Dans la section 8.3, on montre que, si A est un anneau euclidien, alors tout sous-module N de A^n possède une base. Dans la section 8.4, on donne un algorithme pour déterminer une telle base connaissant un système de générateurs de N. Dans la section 8.5, on présente un algorithme plus complexe qui permet de transformer toute matrice à coefficients dans A en sa "forme normale de Smith" et on en déduit le théorème des diviseurs élémentaires. Enfin, étant donné un opérateur linéaire $T\colon V \to V$ comme ci-dessus, la section 8.6 fournit une méthode pour déterminer une base de $\mathcal{B}$ de V telle que $[T]_{\mathcal{B}}$ soit sous forme rationnelle.

Dans tout ce chapitre, A désigne un anneau euclidien, et n un entier positif.

8.1 Preuve du théorème de structure

On énonce d'abord :

Théorème 8.1.1 (théorème des diviseurs élémentaires). *Soit N un sous-A-module de A^n. Alors il existe une base $\{C_1, \ldots, C_n\}$ de A^n, un entier r avec $0 \leq r \leq n$ et des éléments non nuls $d_1, \ldots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$ tels que $\{d_1C_1, \ldots, d_rC_r\}$ soit une base de N.*

La preuve de ce résultat est reportée à la section 8.5. Les éléments $d_1, \ldots, d_r$ qui apparaissent dans son énoncé sont appelés les *diviseurs élémentaires* de N. On verra qu'ils sont entièrement déterminés par N au produit près par des unités de A. On se contente ici d'expliquer comment on en déduit le théorème de structure pour les A-modules de type fini.

Preuve du théorème 7.2.1. Soit $M = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_A$ un A-module de type fini. L'application

$$\begin{array}{rccc} \psi : & A^n & \longrightarrow & M \\ & \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} & \longmapsto & a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \end{array} \tag{8.1}$$

est un homorphisme surjectif de A-modules. Posons

$$N = \ker \psi.$$

Alors N est un sous-A-module de A^n. En vertu du théorème des diviseurs élémentaires, il existe une base $\{C_1, \ldots, C_n\}$ de A^n, un entier r avec $0 \leq r \leq n$ et des éléments non nuls $d_1, \ldots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$ tels que $\{d_1C_1, \ldots, d_rC_r\}$ soit une base de N. Posons encore

$$\mathbf{u}_i = \psi(C_i) \quad (i = 1, \ldots, n).$$

Comme $A^n = \langle C_1, \ldots, C_n \rangle_A$, on a

$$M = \mathrm{Im}(\psi) = \langle \mathbf{u}_1, \ldots, \mathbf{u}_n \rangle_A = A\mathbf{u}_1 + \cdots + A\mathbf{u}_n. \tag{8.2}$$

D'autre part, pour $a_1, \ldots, a_n \in A$, on trouve

$$\begin{array}{rcl} a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0} & \iff & a_1\psi(C_1) + \cdots + a_n\psi(C_n) = \mathbf{0} \\ & \iff & \psi(a_1C_1 + \cdots + a_nC_n) = \mathbf{0} \\ & \iff & a_1C_1 + \cdots + a_nC_n \in N. \end{array}$$

Or, $\{d_1C_1, \ldots, d_rC_r\}$ est une base de N. Donc la condition $a_1C_1 + \cdots + a_nC_n \in N$ équivaut à l'existence de $b_1, \ldots, b_r \in A$ tels que

$$a_1C_1 + \cdots + a_nC_n = b_1(d_1C_1) + \cdots + b_r(d_rC_r).$$

Comme $C_1, \ldots, C_n$ sont linéairement indépendants sur A, cela n'est possible que si

$$a_1 = b_1d_1, \ \ldots, \ a_r = b_rd_r \quad \text{et} \quad a_{r+1} = \cdots = a_n = 0.$$

On conclut que

$$\begin{array}{rcl} a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0} & \iff & a_1C_1 + \cdots + a_nC_n \in N \\ & \iff & a_1 \in (d_1), \ \ldots, \ a_r \in (d_r) \quad \text{et} \quad a_{r+1} = \cdots = a_n = 0. \end{array} \tag{8.3}$$

En particulier, cela montre que

$$a_i\mathbf{u}_i = \mathbf{0} \iff \begin{cases} a_i \in (d_i) & \text{si } 1 \leq i \leq r, \\ a_i = 0 & \text{si } r < i \leq n. \end{cases} \tag{8.4}$$

Donc $\mathrm{Ann}(\mathbf{u}_i) = (d_i)$ pour $i = 1, \ldots, r$ et $\mathrm{Ann}(\mathbf{u}_i) = (0)$ pour $i = r+1, \ldots, n$. Comme $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$, on obtient

$$\mathrm{Ann}(\mathbf{u}_1) \supseteq \mathrm{Ann}(\mathbf{u}_2) \supseteq \cdots \supseteq \mathrm{Ann}(\mathbf{u}_n). \tag{8.5}$$

Les formules (8.3) et (8.4) impliquent aussi que

$$a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0} \quad \Longleftrightarrow \quad a_1\mathbf{u}_1 = \cdots = a_n\mathbf{u}_n = \mathbf{0}.$$

En vertu de (8.2), cette propriété signifie que

$$M = A\mathbf{u}_1 \oplus \cdots \oplus A\mathbf{u}_n \tag{8.6}$$

(voir la définition 6.4.1). Cela démontre le théorème 7.2.1 sauf en ce qui concerne la condition $\mathrm{Ann}(\mathbf{u}_1) \neq A$.

Partant de la décomposition (8.6) satisfaisant (8.5), il est facile d'en obtenir une autre avec $\mathrm{Ann}(\mathbf{u}_1) \neq A$. En effet, pour tout $\mathbf{u} \in M$, on a

$$\mathrm{Ann}(\mathbf{u}) = A \quad \Longleftrightarrow \quad \mathbf{u} = \mathbf{0} \quad \Longleftrightarrow \quad A\mathbf{u} = \{\mathbf{0}\}.$$

Donc si $M \neq \{\mathbf{0}\}$ et si k désigne le plus petit entier tel que $\mathbf{u}_k \neq 0$, on obtient

$$A \neq \mathrm{Ann}(\mathbf{u}_k) \supseteq \cdots \supseteq \mathrm{Ann}(\mathbf{u}_n)$$

et

$$M = \{\mathbf{0}\} \oplus \cdots \oplus \{\mathbf{0}\} \oplus A\mathbf{u}_k \oplus \cdots \oplus A\mathbf{u}_n = A\mathbf{u}_k \oplus \cdots \oplus A\mathbf{u}_n$$

comme requis. Si $M = \{\mathbf{0}\}$, on obtient une somme directe vide. □

La section 8.6 fournit un algorithme pour déterminer $\{C_1, \ldots, C_n\}$ et $d_1, \ldots, d_r$ comme dans l'énoncé du théorème 8.1.1, pourvu qu'on connaisse un système de générateurs de N. Dans l'application au théorème 7.2.1, cela requiert un système de générateurs de $\ker(\psi)$ où ψ est donné par (8.1). Le résultat ci-dessous comble ce besoin dans la situation qui nous intéresse plus particulièrement.

Proposition 8.1.2. *Soit V un espace vectoriel de dimension finie $n > 0$ sur un corps K, soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V et soit $T\colon V \to V$ un opérateur linéaire. Pour la structure correspondante de $K[x]$-module sur V, l'application*

$$\begin{array}{rccc} \psi : & K[x]^n & \longrightarrow & V \\ & \begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} & \longmapsto & p_1(x)\mathbf{v}_1 + \cdots + p_n(x)\mathbf{v}_n \end{array}$$

est un homomorphisme surjectif de $K[x]$-modules, et son noyau est le sous-$K[x]$-module de $K[x]^n$ engendré par les colonnes de la matrice $xI - [T]_{\mathcal{B}}$.

Preuve. Écrivons $[T]_{\mathcal{B}} = (a_{ij})$. Soit N le sous-$K[x]$-module de $K[x]^n$ engendré par les colonnes $C_1, \ldots, C_n$ de $xI - [T]_{\mathcal{B}}$. Pour $j = 1, \ldots, n$, on a

$$C_j = \begin{pmatrix} -a_{1,j} \\ \vdots \\ x - a_{jj} \\ \vdots \\ -a_{nj} \end{pmatrix},$$

donc

$$\begin{aligned}\psi(C_j) &= (-a_{1j})\mathbf{v}_1 + \cdots + (x - a_{jj})\mathbf{v}_j + \cdots + (-a_{nj})\mathbf{v}_n \\ &= T(\mathbf{v}_j) - (a_{1j}\mathbf{v}_1 + \cdots + a_{nj}\mathbf{v}_n) \\ &= \mathbf{0}\end{aligned}$$

et par suite $C_j \in \ker(\psi)$. Cela montre que $N \subseteq \ker(\psi)$.

Pour établir l'inclusion réciproque, on montre d'abord que tout n-uplet $(p_1(x), \ldots, p_n(x))^t$ de $K[x]^n$ peut s'écrire comme somme d'un élément de N et d'un élément de K^n.

On démontre cette assertion par récurrence sur le maximum m des degrés des polynômes $p_1(x), \ldots, p_n(x)$ en convenant, pour les besoins de l'argument, que le polynôme 0 est de degré 0. Si $m = 0$, alors $(p_1(x), \ldots, p_n(x))^t$ est un élément de K^n et on a terminé. Supposons $m \geq 1$ et le résultat vrai pour un n-uplet de polynômes de degrés $\leq m-1$. En désignant par b_i le coefficient de x^m dans $p_i(x)$, on trouve :

$$\begin{aligned}\begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} &= \begin{pmatrix} b_1x^m + (\text{ termes de degrés } \leq m-1) \\ \cdots\cdots \\ b_nx^m + (\text{ termes de degrés } \leq m-1) \end{pmatrix} \\ &= \underbrace{b_1x^{m-1}C_1 + \cdots + b_nx^{m-1}C_n}_{\in N} + \begin{pmatrix} p_1^*(x) \\ \vdots \\ p_n^*(x) \end{pmatrix}\end{aligned}$$

pour des polynômes $p^*(x), \ldots, p_n^*(x)$ de $K[x]$ de degrés $\leq m-1$. En appliquant l'hypothèse de récurrence au n-uplet $(p_1^*(x), \ldots, p_n^*(x))^t$, on conclut que $(p_1(x), \ldots, p_n(x))^t$ est la somme d'un élément de N et d'un élément de K^n.

Soit $(p_1(x), \ldots, p_n(x))^t$ un élément quelconque de $\ker(\psi)$. En vertu de l'observation précédente, on peut écrire

$$\begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} = \begin{pmatrix} q_1(x) \\ \vdots \\ q_n(x) \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \tag{8.7}$$

où $(q_1(x), \ldots, q_n(x))^t \in N$ et $(c_1, \ldots, c_n)^t \in K^n$. Comme $N \subseteq \ker(\psi)$, on trouve que

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} - \begin{pmatrix} q_1(x) \\ \vdots \\ q_n(x) \end{pmatrix} \in \ker(\psi),$$

et par suite, en appliquant ψ au vecteur $(c_1, \ldots, c_n)^t$, on obtient

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}.$$

Comme $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sont linéairement indépendants sur K, cela implique que $c_1 = \cdots = c_n = 0$ et l'égalité (8.7) livre

$$\begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} = \begin{pmatrix} q_1(x) \\ \vdots \\ q_n(x) \end{pmatrix} \in N.$$

Le choix de $(p_1(x), \ldots, p_n(x))^t \in \ker(\psi)$ étant arbitraire, cela montre que $\ker(\psi) \subseteq N$, et par suite $\ker(\psi) = N$. □

Exemple 8.1.3. Soit V un espace vectoriel de dimension 2 sur $\mathbb{R}$, soit $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2\}$ une base de V, et soit $T \in \mathrm{End}_{\mathbb{R}}(V)$ avec $[T]_{\mathcal{A}} = \begin{pmatrix} -1 & 2 \\ 4 & 1 \end{pmatrix}$. Pour la structure correspondante de $\mathbb{R}[x]$-module sur V, on considère l'homomorphisme de $\mathbb{R}[x]$-modules $\psi\colon \mathbb{R}[x]^2 \to V$ donné par

$$\psi \begin{pmatrix} p_1(x) \\ p_2(x) \end{pmatrix} = p_1(x)\mathbf{v}_1 + p_2(x)\mathbf{v}_2(x).$$

(i) En vertu de la proposition 8.1.2, $\ker(\psi)$ est le sous-module de $\mathbb{R}[x]^2$ engendré par les colonnes de

$$xI - [T]_{\mathcal{A}} = \begin{pmatrix} x+1 & -2 \\ -4 & x-1 \end{pmatrix},$$

c'est-à-dire que

$$\ker(\psi) = \left\langle \begin{pmatrix} x+1 \\ -4 \end{pmatrix}, \begin{pmatrix} -2 \\ x-1 \end{pmatrix} \right\rangle_{\mathbb{R}[x]}.$$

(ii) On trouve que $(x^2-9)\mathbf{v}_1 = \mathbf{0}$ car

$$(x^2-9)\mathbf{v}_1 = T^2(\mathbf{v}_1) - 9\mathbf{v}_1 = T(-\mathbf{v}_1 + 4\mathbf{v}_2) - 9\mathbf{v}_1 = \mathbf{0}.$$

Ainsi on a $\begin{pmatrix} x^2-9 \\ 0 \end{pmatrix} \in \ker(\psi)$. Pour écrire cet élément de $\ker(\psi)$ comme combinaison linéaire des générateurs de $\ker(\psi)$ trouvés en (i), on procède comme dans la preuve de la proposition 8.1.2. On obtient

$$\begin{pmatrix} x^2-9 \\ 0 \end{pmatrix} = x \begin{pmatrix} x+1 \\ -4 \end{pmatrix} + 0 \begin{pmatrix} -2 \\ x-1 \end{pmatrix} + \begin{pmatrix} -x-9 \\ 4x \end{pmatrix}$$

et

$$\begin{pmatrix} -x-9 \\ 4x \end{pmatrix} = (-1) \begin{pmatrix} x+1 \\ -4 \end{pmatrix} + 4 \begin{pmatrix} -2 \\ x-1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

donc

$$\begin{pmatrix} x^2-9 \\ 0 \end{pmatrix} = (x-1) \begin{pmatrix} x+1 \\ -4 \end{pmatrix} + 4 \begin{pmatrix} -2 \\ x-1 \end{pmatrix}.$$

Exercices.

8.1.1. Soit V un espace vectoriel sur $\mathbb{Q}$ muni d'une base $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2\}$, et soit $T \in \mathrm{End}_{\mathbb{Q}}(V)$ avec $[T]_{\mathcal{A}} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$. On considère V comme un $\mathbb{Q}[x]$-module et on désigne par N le noyau de l'homomorphisme de $\mathbb{Q}[x]$-modules $\psi\colon \mathbb{Q}[x]^2 \longrightarrow V$ donné par

$$\psi \begin{pmatrix} p_1(x) \\ p_2(x) \end{pmatrix} = p_1(x) \cdot \mathbf{v}_1 + p_2(x) \cdot \mathbf{v}_2.$$

(i) Déterminer un système de générateurs de N en tant que $\mathbb{Q}[x]$-module.

(ii) Montrer que $\begin{pmatrix} x^2 - 3 \\ -2x - 5 \end{pmatrix} \in N$ et exprimer ce vecteur colonne comme combinaison linéaire des générateurs de N trouvés en (i).

8.1.2. Soit V un espace vectoriel sur $\mathbb{R}$ muni d'une base $\mathcal{A} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ et soit $T \in \mathrm{End}_{\mathbb{R}}(V)$ avec $[T]_{\mathcal{A}} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. On considère V comme un $\mathbb{R}[x]$-module et on désigne par N le noyau de l'homomorphisme de $\mathbb{R}[x]$-modules $\psi\colon \mathbb{R}[x]^3 \longrightarrow V$ donné par

$$\psi \begin{pmatrix} p_1(x) \\ p_2(x) \\ p_3(x) \end{pmatrix} = p_1(x) \cdot \mathbf{v}_1 + p_2(x) \cdot \mathbf{v}_2 + p_3(x) \cdot \mathbf{v}_3.$$

(i) Déterminer un système de générateurs de N en tant que $\mathbb{R}[x]$-module.

(ii) Montrer que $\begin{pmatrix} x^2 - x - 6 \\ x - 1 \\ x^2 - 2x + 1 \end{pmatrix} \in N$ et exprimer ce vecteur colonne comme combinaison linéaire des générateurs de N trouvés en (i).

8.2 Le théorème de Cayley-Hamilton

On commence par l'observation suivante :

Proposition 8.2.1. *Soit N le sous-A-module de A^n engendré par les colonnes d'une matrice $U \in \mathrm{Mat}_{n\times n}(A)$. Alors N contient* $\det(U)A^n$.

Preuve. On sait que l'adjointe $\mathrm{Adj}(U)$ de U est une matrice de $\mathrm{Mat}_{n\times n}(A)$ qui satisfait

$$U\mathrm{Adj}(U) = \det(U)I \tag{8.8}$$

(voir le théorème B.3.4). Désignons par $C_1, \ldots, C_n$ les colonnes de U et par $\mathbf{e}_1, \ldots, \mathbf{e}_n$ celles de I (ce sont les éléments de base canonique de A^n). Écrivons aussi $\mathrm{Adj}(U) = (a_{ij})$. En comparant les j-ièmes colonnes des matrices dans chaque membre de (8.8), on trouve

$$a_{1j}C_1 + \cdots + a_{nj}C_n = \det(U)\mathbf{e}_j \quad (1 \leq j \leq n).$$

Comme $N = \langle C_1, \ldots, C_n \rangle_A$, on en déduit que

$$\det(U)A^n = \langle \det(U)\mathbf{e}_1, \ldots, \det(U)\mathbf{e}_n \rangle_A \subseteq N.$$

□

On peut maintenant démontrer le résultat fondamental suivant :

Théorème 8.2.2 (Cayley-Hamilton). *Soit V un espace vectoriel de dimension finie sur un corps K et soit $T\colon V \to V$ un opérateur linéaire sur V. Alors on a*

$$\mathrm{car}_T(T) = 0.$$

Preuve. Soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V sur K. D'après la proposition 8.1.2, l'homomorphisme de $K[x]$-modules

$$\begin{array}{rccc} \psi : & K[x]^n & \longrightarrow & V \\ & \begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} & \longmapsto & p_1(x)\mathbf{v}_1 + \cdots + p_n(x)\mathbf{v}_n \end{array}$$

a pour noyau le sous-module de $K[x]^n$ engendré par les colonnes de $xI - [\,T\,]_{\mathcal{B}} \in \mathrm{Mat}_{n\times n}(K[x])$. D'après la proposition précédente (appliquée à $A = K[x]$), cela implique que

$$\ker(\psi) \supseteq \det(xI - [\,T\,]_{\mathcal{B}}) \cdot K[x]^n = \mathrm{car}_T(x) \cdot K[x]^n.$$

En particulier, pour tout $\mathbf{u} \in K[x]^n$, on a $\mathrm{car}_T(x)\mathbf{u} \in \ker(\psi)$ et par suite

$$\mathbf{0} = \psi(\mathrm{car}_T(x)\,\mathbf{u}) = \mathrm{car}_T(x)\,\psi(\mathbf{u}).$$

Comme ψ est un homomorphisme surjectif, cela signifie que $\mathrm{car}_T(x) \in \mathrm{Ann}(V)$, c'est-à-dire que $\mathrm{car}_T(T) = 0$. □

La conséquence suivante est laissée en exercice.

Corollaire 8.2.3. *Pour toute matrice $A \in \mathrm{Mat}_{n\times n}(K)$, on a $\mathrm{car}_A(A) = 0$.*

Exemple 8.2.4. Pour $n = 2$, on peut en donner une preuve directe. Écrivons $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On trouve

$$\mathrm{car}_T(x) = \begin{vmatrix} x-a & -b \\ -c & x-d \end{vmatrix} = x^2 - (a+d)x + (ad - bc)$$

donc

$$\begin{aligned} \mathrm{car}_A(A) &= A^2 - (a+d)A + (ad-bc)I \\ &= \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Exercices.

8.2.1. Démontrer le corollaire 8.2.3.

8.2.2. Soit V un espace vectoriel de dimension finie sur un corps K et soit $T\colon V \to V$ un opérateur linéaire sur V.

(i) Montrer que $\min_T(x)$ divise $\mathrm{car}_T(x)$.

(ii) Si $K = \mathbb{Q}$ et si $\mathrm{car}_T(x) = x^3 - 2x^2 + x$, quelles sont les posibilités pour $\min_T(x)$?

8.2.3 (Preuve alternative du théorème de Cayley-Hamilton). Soit V un espace vectoriel de dimension finie n sur un corps K et soit $T\colon V \to V$ un opérateur linéaire sur V. Le but de cet exercice est de donner une autre démonstration du théorème de Cayley-Hamilton. Pour ce faire, on munit V de la structure correspondante de K[x]-module, et on choisit un élément quelconque $\mathbf{v}$ de V. Soit $p(x)$ le générateur unitaire de $\mathrm{Ann}(\mathbf{v})$, et soit m son degré.

(i) Montrer qu'il existe une base $\mathcal{B}$ de V telle que $[T]_{\mathcal{B}} = \left(\begin{array}{c|c} P & Q \\ \hline 0 & R \end{array}\right)$ où P désigne la matrice compagnon de $p(x)$ et où R est une matrice carrée $(n-m)\times(n-m)$.

(ii) Montrer que $\mathrm{car}_T(x) = p(x)\mathrm{car}_R(x) \in \mathrm{Ann}(\mathbf{v})$.

(iii) Pourquoi peut-on conclure que $\mathrm{car}_T(T) = 0$?

8.3 Les sous-modules de A^n

Le but de cette section est de démontrer le cas particulier suivant du théorème des diviseurs élémentaires comme première étape dans la preuve de ce théorème.

Théorème 8.3.1. *Tout sous-A-module de A^n est libre et possède une base avec au plus n éléments.*

Ce résultat est faux en général pour un anneau commutatif A quelconque (voir l'exercice 8.3.1).

Preuve. On procède par récurrence sur n. Pour $n = 1$, un sous-A-module N de $A^1 = A$ est simplement un idéal de A. Si $N = \{\mathbf{0}\}$, alors par convention N est libre et il admet pour base $\emptyset$. Sinon, on sait que $N = (b) = A\,b$ pour un élément non nul b de A car tout idéal de A est principal (théorème 5.4.3). On note que si $ab = 0$ avec $a \in A$ alors $a = 0$ car A est un anneau intègre. Donc, dans ce cas, $\{b\}$ est une base de N.

Supposons maintenant $n \geq 2$ et que tout sous-A-module de A^{n-1} admette une base avec au plus $n-1$ éléments. Soit N un sous-A-module de A^n. On définit

$$I = \Big\{ a \in A \,;\, \begin{pmatrix} a \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in N \text{ pour certains } a_2, \ldots, a_n \in A \Big\},$$

$$N' = \Big\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in N \,;\, a_1 = 0 \Big\} \quad \text{et} \quad N'' = \Big\{ \begin{pmatrix} a_2 \\ \vdots \\ a_n \end{pmatrix} \in A^{n-1} \,;\, \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in N \Big\}.$$

On vérifie que I est un idéal de A, que N' est un sous-module de N et que N'' est un sous-module de A^{n-1}. De plus, l'application

$$\begin{array}{rccc} \varphi : & N'' & \longrightarrow & N' \\ & \begin{pmatrix} a_2 \\ \vdots \\ a_n \end{pmatrix} & \longmapsto & \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \end{array}$$

est un isomorphisme de A-modules. En vertu de l'hypothèse de récurrence, N'' admet une base avec au plus $n-1$ éléments. L'image de cette base sous φ est donc une base $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ de N' avec $0 \leq m \leq n-1$ (proposition 6.5.13).

Si $I = \{\mathbf{0}\}$, alors on a $N = N'$. Dans ce cas, $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ est une base de N et, comme $0 \leq m \leq n$, on a terminé.

Sinon le théorème 5.4.3 donne $I = (b_1) = Ab_1$ pour un élément non nul b_1 de A. Puisque $b_1 \in I$, il existe $b_2, \ldots, b_n \in A$ tel que

$$\mathbf{u}_{m+1} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in N.$$

On va démontrer que $N = N' \oplus A\mathbf{u}_{m+1}$. Si on accepte ce résultat, alors $\{\mathbf{u}_1, \ldots, \mathbf{u}_{m+1}\}$ est une base de N et, puisque $m+1 \leq n$, le pas de récurrence est complété.

Soit $\mathbf{u} = (c_1, \ldots, c_n)^t \in N$. Par définition de I, on a $c_1 \in I$. Donc il existe $a \in A$ tel que $c_1 = a\, b_1$. On trouve que

$$\mathbf{u} - a\mathbf{u}_{m+1} = \begin{pmatrix} 0 \\ c_2 - a\, b_2 \\ \vdots \\ c_n - a\, b_n \end{pmatrix} \in N',$$

et par suite

$$\mathbf{u} = (\mathbf{u} - a\mathbf{u}_{m+1}) + a\mathbf{u}_{m+1} \in N' + A\mathbf{u}_{m+1}.$$

Le choix de $\mathbf{u}$ étant arbitraire, cela montre que

$$N = N' + A\mathbf{u}_{m+1}.$$

On note aussi que, pour $a \in A$,

$$a\mathbf{u}_{m+1} \in N' \quad \Longleftrightarrow \quad a\, b_1 = 0 \quad \Longleftrightarrow \quad a = 0.$$

Cela montre que

$$N' \cap A\mathbf{u}_{m+1} = \{\mathbf{0}\},$$

donc $N = N' \oplus A\mathbf{u}_{m+1}$ comme annoncé. □

Exemple 8.3.2. Soit $N = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \,;\, 2x + 3y + 6z = 0 \right\}$.

(i) Montrer que N est un sous-$\mathbb{Z}$-module de $\mathbb{Z}^3$.

(ii) Déterminer une base de N comme $\mathbb{Z}$-module.

Solution: Pour (i), il s'agit de montrer que N est un sous-groupe de $\mathbb{Z}^3$. On le laisse en exercice. Pour (ii), on pose

$$I = \{x \in \mathbb{Z} \,;\, 2x + 3y + 6z = 0 \text{ pour certains } y, z \in \mathbb{Z}\Big\},$$

$$N' = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in N \,;\, x = 0 \right\} \quad \text{et} \quad N'' = \left\{ \begin{pmatrix} y \\ z \end{pmatrix} \in \mathbb{Z}^2 \,;\, 3y + 6z = 0 \right\},$$

comme dans la preuve du théorème 8.3.1. On trouve que

$$3y + 6z = 0 \quad \Longleftrightarrow \quad y = -2z,$$

donc

$$N'' = \left\{ \begin{pmatrix} -2z \\ z \end{pmatrix} \,;\, z \in \mathbb{Z} \right\} = \mathbb{Z} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

admet pour base $\left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$ et par suite $\left\{ \mathbf{u}_1 := \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}$ est une base de N'. On trouve aussi que, pour $x \in \mathbb{Z}$, on a

$$\begin{aligned} x \in I \quad &\Longleftrightarrow \quad 2x = -3y - 6z \quad \text{pour certains } y, z \in \mathbb{Z} \\ &\Longleftrightarrow \quad 2x \in (-3, -6) = (3) \\ &\Longleftrightarrow \quad 3 \,|\, 2x \\ &\Longleftrightarrow \quad 3 \,|\, x. \end{aligned}$$

Donc $I = (3)$. De plus 3 est la première composante de

$$\mathbf{u}_2 = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} \in N.$$

En reprenant les arguments de la preuve du théorème 8.3.1, on obtient que

$$N = N' \oplus \mathbb{Z}\mathbf{u}_2 = \mathbb{Z}\mathbf{u}_1 \oplus \mathbb{Z}\mathbf{u}_2.$$

Donc $\left\{ \mathbf{u}_1 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} \right\}$ est une base de N.

Exercices.

8.3.1. Soit A un anneau intègre quelconque et soit I un idéal de A considéré comme sous-A-module de $A^1 = A$. Montrer que I est un A-module libre si et seulement si I est principal.

8.3.2. Soit M un module libre sur un anneau euclidien A. Montrer que tout sous-A-module de M est libre.

8.3.3. Soit $N = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{Z}^3 \,;\, 3x + 10y - 16z = 0 \right\}$.

(i) Montrer que N est un sous-module de $\mathbb{Z}^3$.
(ii) Déterminer une base de N en tant que $\mathbb{Z}$-module.

8.3.4. Soit $N = \left\{ \begin{pmatrix} p_1(x) \\ p_2(x) \\ p_3(x) \end{pmatrix} \in \mathbb{Q}[x]^3 \,;\, (x^2 - x)p_1(x) + x^2 p_2(x) - (x^2 + x)p_3(x) = 0 \right\}$.

(i) Montrer que N est un sous-$\mathbb{Q}[x]$-module de $\mathbb{Q}[x]^3$.
(ii) Déterminer une base de N en tant que $\mathbb{Q}[x]$-module.

8.4 Le module-colonne d'une matrice

Définition 8.4.1. On définit le module-colonne d'une matrice $U \in \mathrm{Mat}_{n\times m}(A)$ comme étant le sous-A-module de A^n engendré par les colonnes de A. On le note $\mathrm{Col}_A(U)$.

Le théorème 8.3.1 montre que tout sous-module de A^n possède une base avec au plus n éléments. Donc le module-colonne d'une matrice $U \in \mathrm{Mat}_{n\times m}(A)$ possède une base. Le but de cette section est de donner un algorithme pour déterminer une telle base. Plus généralement, cet algorithme permet de calculer une base d'un sous-module quelconque N de A^n pourvu qu'on connaisse un système de générateurs $C_1, \ldots, C_m$ de N car N est alors le module-colonne de la matrice $(C_1 \cdots C_m) \in \mathrm{Mat}_{n\times m}(A)$ dont les colonnes sont $C_1, \ldots, C_m$.

On commence par l'observation suivante :

Lemme 8.4.2. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$ et soit U' une matrice obtenue à partir de U en lui appliquant une opération d'un des types suivants :*

(I) permuter deux colonnes,

(II) ajouter à une colonne le produit d'une autre colonne par un élément de A,

(III) multiplier une colonne par un élément de A^*.

Alors U et U' ont le même module-colonne.

Preuve. Par construction les colonnes de U' appartiennent à $\mathrm{Col}_A(U)$ car ce sont des combinaisons linéaires des colonnes de U. Donc on a $\mathrm{Col}_A(U') \subseteq \mathrm{Col}_A(U)$. Par ailleurs, chacune des opérations de type $(I), (II)$ ou (III) est réversible : on peut retrouver U à partir de U' en lui appliquant une opération du même type. Donc on a aussi $\mathrm{Col}_A(U) \subseteq \mathrm{Col}_A(U')$, et par suite $\mathrm{Col}_A(U) = \mathrm{Col}_A(U')$. □

Les opérations $(I), (II), (III)$ du lemme 8.4.2 s'appellent les *opérations élémentaires de colonnes* (sur A). On déduit de ce lemme :

Lemme 8.4.3. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$ et soit une matrice U' obtenue à partir de U en lui appliquant une suite d'opérations élémentaires de colonnes. Alors* $\mathrm{Col}_A(U) = \mathrm{Col}_A(U')$.

Définition 8.4.4. On dit qu'une matrice $U \in \mathrm{Mat}_{n\times m}(A)$ est de *forme échelonnée* si elle est de la forme

$$U = \begin{pmatrix}
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & & \vdots & \vdots & & \vdots \\
0 & \vdots & & & \vdots & \vdots & & \vdots \\
u_{i_1,1} & \vdots & \cdots & \cdots & \vdots & \vdots & & \vdots \\
* & 0 & & & \vdots & \vdots & & \vdots \\
\vdots & u_{i_2,2} & & & \vdots & \vdots & & \vdots \\
\vdots & * & & & 0 & \vdots & & \vdots \\
\vdots & \vdots & & & u_{i_r,r} & \vdots & & \vdots \\
\vdots & \vdots & & & * & \vdots & & \vdots \\
\vdots & \vdots & & & \vdots & \vdots & & \vdots \\
* & * & \cdots & \cdots & * & 0 & & 0
\end{pmatrix} \tag{8.9}$$

avec $1 \leq i_1 < i_2 < \cdots < i_r \leq n$ et $u_{i_1,1} \neq 0, \ldots, u_{i_r,r} \neq 0$.

Lemme 8.4.5. *Supposons que $U \in \mathrm{Mat}_{n\times m}(A)$ soit de forme échelonnée. Alors les colonnes non nulles de U forment une base de $\mathrm{Col}_A(U)$.*

Preuve. Supposons que U soit donnée par (8.9). Alors les r premières colonnes $C_1, \ldots, C_r$ de U engendrent $\mathrm{Col}_A(U)$. Supposons qu'on ait

$$a_1C_1 + \cdots + a_rC_r = \mathbf{0}$$

pour des éléments $a_1, \ldots, a_r$ de A. Si $a_1, \ldots, a_r$ ne sont pas tous nuls, il existe un plus petit entier k tel que $a_k \neq 0$. Alors on obtient

$$a_kC_k + \cdots + a_rC_r = \mathbf{0}.$$

En prenant la i_k-ième composante de chaque membre de cette égalité, on trouve

$$a_k u_{i_k,k} = 0.$$

Comme $a_k \neq 0$ et $u_{i_k,k} \neq 0$, c'est impossible (car A est un anneau intègre). Cette contradiction montre qu'on doit avoir $a_1 = \cdots = a_r = 0$. Donc les colonnes $C_1, \ldots, C_r$ sont linéairement indépendantes et par suite elles forment une base de $\mathrm{Col}_A(U)$. □

Les trois lemmes précédents utilisent seulement que le fait que A est un anneau intègre. Le résultat suivant utilise de manière cruciale l'hypothèse que A est un anneau euclidien.

Théorème 8.4.6. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$. Il est possible d'amener U sous la forme échelonnée par une suite d'opérations élémentaires de colonnes.*

Preuve. Soit $\varphi\colon A \setminus \{0\} \to \mathbb{N}$ le stathme euclidien associé à A (voir définition 5.4.1). On considère l'algorithme suivant :

Algorithme 8.4.7.

1. *Si $U = 0$, la matrice U est déjà sous forme échelonnée et on a terminé.*

2. *Supposons que $U \neq 0$. On localise la première ligne non nulle de U.*

$$\left(\begin{array}{c} 0 \\ \hline \boxed{\neq 0} \\ * \end{array}\right) \leftarrow$$

2.1. *On choisit sur cette ligne un élément $a \neq 0$ pour lequel $\varphi(a)$ est minimal et on permute les colonnes pour amener ce dernier dans la première colonne.*

$$\left(\begin{array}{c|ccc} 0 & & & \\ \hline a \neq 0 & a_2 & \cdots & a_m \\ * & & & \end{array}\right)$$

2.2. *On divise chacun des autres éléments $a_2, \ldots, a_m$ de cette ligne par a en écrivant*

$$a_j = q_j a + r_j$$

avec $q_j, r_j \in A$ satisfaisant $r_j = 0$ ou $\varphi(r_j) < \varphi(a)$, puis, pour $j = 2, \ldots, m$, on soustrait de la j-ième colonne le produit de la première colonne par q_j.

$$\left(\begin{array}{ccccc} & C_2 - q_2C_1 & & \cdots & C_m - q_mC_1 \\ & \downarrow & & & \downarrow \\ & 0 & & & \\ \hline a & r_2 & & \cdots & r_m \\ & * & & & \end{array}\right)$$

2.3. *Si $r_2, \ldots, r_m$ ne sont pas tous nuls, on retourne à l'étape 2.1.*

S'ils sont tous nuls, la nouvelle matrice est de la forme indiquée à droite et on retourne à l'étape 1 en se restreignant à la sous-matrice formée par les $m-1$ dernières colonnes, c'est-à-dire en faisant abstraction de la première colonne.

$$\left(\begin{array}{c|cccc} 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \\ \hline a & 0 & \cdots & \cdots & 0 \\ * & & & * & \end{array}\right)$$

Cet algorithme termine après un nombre fini d'étapes car chaque itération des pas 2.1 à 2.3 sur une ligne donnée remplace le premier élément $a \neq 0$ de cette ligne par un élément $a' \neq 0$ de A avec $\varphi(a') < \varphi(a)$, et cela ne peut pas être répété indéfiniment puisque φ prend ses valeurs dans $\mathbb{N} = \{0, 1, 2, \ldots\}$. A la fin, cet algorithme produit une matrice de forme échelonnée comme requis. □

En combinant ce théorème aux lemmes précédents, on conclut :

Corollaire 8.4.8. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$. Par une suite d'opérations élémentaires de colonnes, on peut amener U sous la forme d'une matrice échelonnée U'. Alors les colonnes non nulles de U' forment une base de $\mathrm{Col}_A(U)$.*

Preuve. La première assertion découle du théorème 8.4.6. En vertu du lemme 8.4.3, on a $\mathrm{Col}_A(U) = \mathrm{Col}_A(U')$. Finalement le lemme 8.4.5 montre que les colonnes non nulles de U' forment une base de $\mathrm{Col}_A(U')$, donc celles-ci forment aussi une base de $\mathrm{Col}_A(U)$. □

Exemple 8.4.9. Donner une base de $\mathrm{Col}_{\mathbb{Z}}(U)$ où $U = \begin{pmatrix} 3 & 7 & 4 \\ -4 & -2 & 2 \\ 1 & -5 & -6 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{Z})$.

Solution: On trouve

$$U \sim \begin{array}{c} \\ \\ \end{array}\!\!\begin{array}{c} \\ C_2 - 2C_1 \quad C_3 - C_1 \end{array}\!\!\begin{pmatrix} 3 & 1 & 1 \\ -4 & 6 & 6 \\ 1 & -7 & -7 \end{pmatrix} \sim \begin{array}{c} \\ C_2 \quad C_1 \end{array}\!\!\begin{pmatrix} 1 & 3 & 1 \\ 6 & -4 & 6 \\ -7 & 1 & -7 \end{pmatrix}$$

$$\sim \begin{array}{c} \\ C_2 - 3C_1 \quad C_3 - C_1 \end{array}\!\!\begin{pmatrix} 1 & 0 & 0 \\ 6 & -22 & 0 \\ -7 & 22 & 0 \end{pmatrix} \sim \begin{array}{c} \\ -C_2 \end{array}\!\!\begin{pmatrix} 1 & 0 & 0 \\ 6 & 22 & 0 \\ -7 & -22 & 0 \end{pmatrix}.$$

Donc $\left\{ \begin{pmatrix} 1 \\ 6 \\ -7 \end{pmatrix}, \begin{pmatrix} 0 \\ 22 \\ -22 \end{pmatrix} \right\}$ est une base de $\mathrm{Col}_{\mathbb{Z}}(U)$.

Exemple 8.4.10. Déterminer une base de

$$N := \left\langle \begin{pmatrix} x^2-1 \\ x+1 \end{pmatrix}, \begin{pmatrix} x^2-x \\ x \end{pmatrix}, \begin{pmatrix} x^3-7 \\ x-1 \end{pmatrix} \right\rangle_{\mathbb{Q}[x]} \subseteq \mathbb{Q}[x]^2.$$

Solution: On a $N = \mathrm{Col}_{\mathbb{Q}[x]}(U)$ où

$$\begin{aligned} U &= \begin{pmatrix} x^2-1 & x^2-x & x^3-7 \\ x+1 & x & x-1 \end{pmatrix} \sim \begin{matrix} \\ \end{matrix} \overset{\quad\quad C_2-C_1 \quad C_3-xC_1}{\begin{pmatrix} x^2-1 & -x+1 & x-7 \\ x+1 & -1 & -x^2-1 \end{pmatrix}} \\ &\sim \overset{C_2 \quad\quad C_1}{\begin{pmatrix} -x+1 & x^2-1 & x-7 \\ -1 & x+1 & -x^2-1 \end{pmatrix}} \sim \overset{\quad -C_1}{\begin{pmatrix} x-1 & x^2-1 & x-7 \\ 1 & x+1 & -x^2-1 \end{pmatrix}} \\ &\sim \overset{\quad C_2-(x+1)C_1 \quad C_3-C_1}{\begin{pmatrix} x-1 & 0 & -6 \\ 1 & 0 & -x^2-2 \end{pmatrix}} \sim \overset{\quad -C_3 \quad C_1 \quad C_2}{\begin{pmatrix} 6 & x-1 & 0 \\ x^2+2 & 1 & 0 \end{pmatrix}} \\ &\sim \overset{\quad 6C_2}{\begin{pmatrix} 6 & 6(x-1) & 0 \\ x^2+2 & 6 & 0 \end{pmatrix}} \quad \text{car } 6 \in \mathbb{Q}[x]^* = \mathbb{Q}^*, \\ &\sim \overset{\quad C_2-(x-1)C_1}{\begin{pmatrix} 6 & 0 & 0 \\ x^2+2 & -x^3+x^2-2x+8 & 0 \end{pmatrix}} \end{aligned}$$

Donc $\left\{ \begin{pmatrix} 6 \\ x^2+2 \end{pmatrix}, \begin{pmatrix} 0 \\ -x^3+x^2-2x+8 \end{pmatrix} \right\}$ est une base de N sur $\mathbb{Q}[x]$.

Exercices.

8.4.1. Déterminer une base du sous-groupe de $\mathbb{Z}^3$ engendré par $\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$ et $\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$.

8.4.2. Déterminer une base du sous-$\mathbb{R}[x]$-module de $\mathbb{R}[x]^2$ engendré par les couples $\begin{pmatrix} x^2-1 \\ x \end{pmatrix}$, $\begin{pmatrix} x^2-2x+1 \\ x-2 \end{pmatrix}$ et $\begin{pmatrix} x^3-1 \\ x^2-1 \end{pmatrix}$.

8.4.3. Soit A un anneau euclidien, et soit $U \in \mathrm{Mat}_{n\times m}(A)$. Montrer que l'idéal I de A engendré par les éléments de la premiére ligne de U n'est pas affecté par les opérations élémentaires de colonnes. En déduire que si une suite d'opérations élémentaires de colonnes transforme U en une matrice dont la première ligne est $(a, 0, \ldots, 0)$, alors a est un générateur de I.

8.5 Forme normale de Smith

Pour définir la notion d'opération élémentaire de lignes sur une matrice, on remplace partout le mot "colonne" par "ligne" dans les énoncés (I) à (III) du lemme 8.4.2 :

Définition 8.5.1. Une *opération élémentaire de lignes* sur une matrice de $\mathrm{Mat}_{n\times m}(A)$ est toute opération d'un des types suivants :

(I) permuter deux lignes,

(II) ajouter à une ligne le produit d'une autre ligne par un élément de A,

(III) multiplier une ligne par un élément de A^*.

Pour étudier l'impact d'une opération élémentaire de lignes sur le module-colonne d'une matrice, on fait appel au résultat suivant dont la preuve est laissée en exercice (voir l'exercice 6.5.4).

Lemme 8.5.2. *Soit $\varphi\colon M \to M'$ un homomorphisme de A-modules et soit N un sous-A-module de M. Alors*

$$\varphi(N) := \{\varphi(\mathbf{u})\,;\, \mathbf{u} \in M\}$$

est un sous-A-module de M'. De plus :

(i) si $N = \langle \mathbf{u}_1, \ldots, \mathbf{u}_m\rangle_A$, alors $\varphi(N) = \langle \varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_m)\rangle_A$;

(ii) si N admet une base $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ et si φ est injective, alors $\{\varphi(\mathbf{u}_1), \ldots, \varphi(\mathbf{u}_m)\}$ est une base de $\varphi(N)$.

On peut maintenant montrer :

Lemme 8.5.3. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$ et soit U' une matrice obtenue à partir de U en lui appliquant une opération élémentaire de lignes. Alors il existe un automorphisme φ de A^n tel que*

$$\mathrm{Col}_A(U') = \varphi(\mathrm{Col}_A(U)).$$

On rappelle qu'un *automorphisme* d'un A-module est un isomorphisme de ce module dans lui-même.

Preuve. Soient $C_1, \ldots, C_m$ les colonnes de U et soient $C'_1, \ldots, C'_m$ celles de U'. On distingue trois cas.

Si U' est obtenue à partir de U en lui appliquant une opération du type (I), on supposera simplement qu'on a permuté les lignes 1 et 2. Le cas général est semblable mais plus compliqué à écrire. Dans ce cas, l'application $\varphi\colon A^n \to A^n$ donnée par

$$\varphi\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} u_2 \\ u_1 \\ u_3 \\ \vdots \\ u_n \end{pmatrix}$$

est un automorphisme de A^n tel que $\varphi(C_j) = C'_j$ pour $j = 1, \ldots, m$. Comme $\mathrm{Col}_A(U) = \langle C_1, \ldots, C_m\rangle_A$ et que $\mathrm{Col}_A(U') = \langle C'_1, \ldots, C'_m\rangle_A$, le lemme 8.5.2 livre alors $\mathrm{Col}_A(U') = \varphi(\mathrm{Col}_A(U'))$.

Si U' est obtenue par une opération du type (II), on supposera qu'on ajouté à la ligne 1 le produit de la ligne 2 par un élément a de A. Alors on a $C'_j = \varphi(C_j)$ pour $j = 1, \ldots, m$ où $\varphi\colon A^n \to A^n$ est l'application donnée par

$$\varphi\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} u_1 + a\,u_2 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

On vérifie que φ est un automorphisme de A^n et la conclusion découle alors du lemme 8.5.2.

Enfin si U' est obtenue par une opération de type (III), on supposera qu'on a multiplié la ligne 1 de U par une unité $a \in A^*$. Alors $C'_j = \varphi(C_j)$ pour $j = 1, \ldots, m$ où $\varphi\colon A^n \to A^n$ est l'automorphisme de A^n donné par

$$\varphi\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} a\,u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},$$

et le lemme 8.5.2 donne $\mathrm{Col}_A(U') = \varphi(\mathrm{Col}_A(U))$. □

En combinant les lemmes 8.4.2 et 8.5.3, on obtient :

Théorème 8.5.4. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$ et soit U' une matrice obtenue à partir de U en lui appliquant une suite d'opérations élémentaires de lignes et de colonnes. Alors il existe un automorphisme $\varphi\colon A^n \to A^n$ tel que*

$$\mathrm{Col}_A(U') = \varphi(\mathrm{Col}_A(U)).$$

Preuve. Par hypothèse, il existe une suite de matrices

$$U_1 = U,\ U_2,\ U_3,\ \ldots,\ U_k = U'$$

telles que, pour $i = 1, \dots, k-1$, la matrice U_{i+1} soit obtenue en appliquant à U_i une opération élémentaire de lignes ou de colonnes. S'il s'agit d'une opération de lignes, le lemme 8.5.3 donne

$$\mathrm{Col}_A(U_{i+1}) = \varphi_i(\mathrm{Col}_A(U_i))$$

pour un automorphisme φ_i de A^n. S'il s'agit d'une opération de colonnes, le lemme 8.4.2 donne

$$\mathrm{Col}_A(U_{i+1}) = \mathrm{Col}_A(U_i) = \varphi_i(\mathrm{Col}_A(U_i))$$

où $\varphi_i = I_{A^n}$ est l'application identité de A^n. On en déduit que

$$\mathrm{Col}_A(U') = \mathrm{Col}_A(U_k) = \varphi_{k-1}(\dots\varphi_2(\varphi_1(\mathrm{Col}_A(U_1)))\dots) = \varphi(\mathrm{Col}_A(U))$$

où $\varphi = \varphi_{k-1} \circ \cdots \circ \varphi_2 \circ \varphi_1$ est un automorphisme de A^n. □

Le théorème 8.5.4 est valable sur un anneau commutatif A quelconque. En utilisant l'hypothèse que A est euclidien, on montre :

Théorème 8.5.5. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$. En appliquant à U une suite d'opérations élémentaires de lignes et de colonnes, on peut l'amener sous la forme*

$$\left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right) \tag{8.10}$$

pour un entier $r \geq 0$ et des éléments non nuls $d_1, \dots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \dots \,|\, d_r$.

On dit que (8.1) est la *forme normale de Smith* de U. On peut montrer que les éléments $d_1, \dots, d_r$ sont uniquement déterminés par U au produit près par une unité de A (voir l'exercice 8.5.4). On les appelle les *diviseurs élémentaires* de U.

Preuve. Si $U = 0$, la matrice U est déjà sous forme normale de Smith avec $r = 0$. On suppose donc que $U \neq 0$. Alors il suffit de montrer que U peut être transformée en une matrice du type

$$\left(\begin{array}{c|ccc} d & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U^* & \\ 0 & & & \end{array}\right) \tag{8.11}$$

où d est un élément non nul de A qui divise tous les coefficients de U^*.

Puisque $U \neq 0$, il existe un élément non nul a de U pour lequel $\varphi(a)$ est minimal. En permutant les lignes et les colonnes de U, on peut amener cet élément en position (1,1) :

$$\begin{pmatrix} & \vdots & \\ \cdots & a & \cdots \\ & \vdots & \end{pmatrix} \sim \begin{pmatrix} \cdots & a & \cdots \\ & \vdots & \\ & \vdots & \end{pmatrix} \sim \begin{pmatrix} a & \cdots & \cdots \\ \vdots & & \\ \vdots & & \end{pmatrix}$$

Étape 1. Soient $a_2, \dots, a_m$ les autres éléments de la première ligne. Pour $j = 2, \dots, m$, on écrit $a_j = q_j a + r_j$ avec $q_j, r_j \in A$ qui satisfont $r_j = 0$ ou $\varphi(r_j) < \varphi(a)$, puis on soustrait de la colonne j la colonne 1 multipliée par q_j. Cela donne

$$\begin{array}{c} \\ \left(\begin{array}{c|ccc} a & r_2 & \cdots & r_m \\ \hline * & & & \\ \vdots & & * & \\ * & & & \end{array}\right) \end{array} \begin{array}{c} C_2 - q_2 C_1 \quad \cdots \quad C_m - q_m C_1 \\ \\ \end{array}$$

Si $r_1 = \cdots = r_m = 0$ (cela arrive lorsque a divise tous les éléments de la première ligne), cette matrice est de la forme

$$\left(\begin{array}{c|ccc} a & 0 & \cdots & 0 \\ \hline * & & & \\ \vdots & & * & \\ * & & & \end{array}\right).$$

Sinon, on choisit un indice j avec $r_j \neq 0$ pour lequel $\varphi(r_j)$ est minimal et on permute les colonnes 1 et j pour amener cet élément r_j en position $(1,1)$. Cela fournit une matrice de la forme

$$\left(\begin{array}{c|ccc} a' & * & \cdots & * \\ \hline * & & & \\ \vdots & & * & \\ * & & & \end{array}\right)$$

avec $a' \neq 0$ et $\varphi(a') < \varphi(a)$. On répète ensuite le procédé. Comme toute suite d'éléments non nuls $a, a', a'', \dots$ de A avec $\varphi(a) > \varphi(a') > \varphi(a'') > \cdots$ ne peut se poursuivre indéfiniment, on atteint après un nombre fini d'itérations une matrice de la forme

$$\left(\begin{array}{c|ccc} b & 0 & \cdots & 0 \\ \hline * & & & \\ \vdots & & * & \\ * & & & \end{array}\right) \tag{8.12}$$

avec $b \neq 0$ et $\varphi(b) \leq \varphi(a)$. De plus on a $\varphi(b) < \varphi(a)$ si au départ a ne divise pas tous les éléments de la ligne 1.

Étape 2. Soient $b_2, \dots, b_m$ les autres éléments de la première colonne de la matrice (8.12). Pour $i = 2, \dots, n$, on écrit $b_i = q_i' b + r_i'$ avec $q_i', r_i' \in A$ satisfaisant $r_i' = 0$ ou $\varphi(r_i') < \varphi(b)$, puis on soustrait de la ligne i la ligne 1 multipliée par q_i'. Cela donne

$$\begin{array}{c} \\ \left(\begin{array}{c|ccc} b & 0 & \cdots & 0 \\ \hline r_2' & & & \\ \vdots & & * & \\ r_n' & & & \end{array}\right) \end{array} \begin{array}{c} \\ L_2 - q_2' L_1 \\ \vdots \\ L_n - q_n' L_1 \end{array}$$

Si $r'_2 = \cdots = r'_n = 0$ (cela arrive lorsque b divise tous les éléments de la première colonne), cette matrice est de la forme

$$\left(\begin{array}{c|ccc} b & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & * & \\ 0 & & & \end{array}\right).$$

Sinon on choisit un indice i avec $r'_i \neq 0$ pour lequel $\varphi(r'_i)$ est minimal puis on permute les lignes 1 et i pour amener cet élément r'_i en position $(1,1)$. Cela donne une matrice de la forme

$$\left(\begin{array}{c|ccc} b' & * & \cdots & * \\ \hline * & & & \\ \vdots & & * & \\ * & & & \end{array}\right).$$

avec $b' \neq 0$ et $\varphi(b') < \varphi(b)$. On retourne ensuite à l'étape 1. Comme toute suite d'éléments non nuls $b, b', b'', \ldots$ de A avec $\varphi(b) > \varphi(b') > \varphi(b'') > \cdots$ est nécessairement finie, on obtient, après un nombre fini d'étapes une matrice de la forme

$$\left(\begin{array}{c|ccc} c & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{array}\right) \tag{8.13}$$

avec $c \neq 0$ et $\varphi(c) \leq \varphi(a)$. De plus, on a $\varphi(c) < \varphi(a)$ si initialement a ne divise pas tous les éléments de la première ligne et de la première colonne.

Étape 3. Si c divise tous les coefficients de la sous-matrice U' de (8.13), on a terminé. Sinon, on ajoute à la ligne 1 une ligne qui contient un élément de U' non divisible par c. On retourne ensuite à l'étape 1 avec la matrice résultante

$$\left(\begin{array}{c|ccc} c & c_2 & \cdots & c_m \\ \hline 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{array}\right) \quad L_1 + L_i \quad \text{où} \quad c \nmid L_i.$$

Puisque c ne divise pas tous les éléments de la ligne 1 et de la colonne 1, cela conduit à une nouvelle matrice

$$\left(\begin{array}{c|ccc} c' & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & U'' & \\ 0 & & & \end{array}\right)$$

avec $c' \neq 0$ et $\varphi(c') < \varphi(c)$. Puisque toute suite d'éléments non nuls $c, c', c'', \ldots$ de A avec $\varphi(c) > \varphi(c') > \varphi(c'') > \cdots$ est nécessairement finie, on atteint après un nombre fini d'itérations, une matrice de la forme requise (8.10) où $d \neq 0$ divise tous les coefficients de U^* □

Exemple 8.5.6. Déterminer la forme normale de Smith de $U = \begin{pmatrix} 4 & 8 & 4 \\ 4 & 10 & 8 \end{pmatrix} \in \mathrm{Mat}_{2\times 3}(\mathbb{Z})$.

Solution: Le calcul donne :

$$U \sim \overset{\quad C_2-2C_1 \quad C_3-C_1}{\begin{pmatrix} 4 & 0 & 0 \\ 4 & 2 & 4 \end{pmatrix}} \sim \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 4 \end{pmatrix} \begin{matrix} \\ L_2-L_1 \end{matrix} \sim \begin{pmatrix} 4 & 2 & 4 \\ 0 & 2 & 4 \end{pmatrix} \begin{matrix} \\ L_1+L_2 \end{matrix}$$

$$\begin{matrix} \\ \uparrow \\ \text{non divisible} \\ \text{par } 4 \end{matrix}$$

$$\sim \overset{\quad C_2 \quad C_1}{\begin{pmatrix} 2 & 4 & 4 \\ 2 & 0 & 4 \end{pmatrix}} \sim \overset{\quad C_2-2C_1 \quad C_3-2C_1}{\begin{pmatrix} 2 & 0 & 0 \\ 2 & -4 & 0 \end{pmatrix}} \sim \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \end{pmatrix} \begin{matrix} \\ L_2-L_1 \end{matrix}$$

$$\sim \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \end{pmatrix} \begin{matrix} \\ -L_2 \end{matrix}$$

Exemple 8.5.7. Déterminer les diviseurs élémentaires de

$$U = \begin{pmatrix} x & 1 & 1 \\ 3 & x-2 & -3 \\ -4 & 4 & x+5 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{Q}[x]).$$

Solution: On trouve

$$U \sim \overset{\quad C_2 \quad C_1}{\begin{pmatrix} 1 & x & 1 \\ x-2 & 3 & -3 \\ 4 & -4 & x+5 \end{pmatrix}} \sim \overset{\quad C_2-xC_1 \quad C_3-C_1}{\begin{pmatrix} 1 & 0 & 0 \\ x-2 & -x^2+2x+3 & -x-1 \\ 4 & -4x-4 & x+1 \end{pmatrix}}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -x^2+2x+3 & -x-1 \\ 0 & -4x-4 & x+1 \end{pmatrix} \begin{matrix} \\ L_2-(x-2)L_1 \\ L_3-4L_1 \end{matrix}$$

$$\sim \overset{\quad C_3 \quad C_2}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -x-1 & -x^2+2x+3 \\ 0 & x+1 & -4x-4 \end{pmatrix}} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & x^2-2x-3 \\ 0 & 0 & -x^2-2x-1 \end{pmatrix} \begin{matrix} \\ -L_2 \\ L_3+L_2 \end{matrix}$$

$$\sim \overset{\quad C_3-(x-3)C_2}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & 0 \\ 0 & 0 & -(x+1)^2 \end{pmatrix}} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & 0 \\ 0 & 0 & (x+1)^2 \end{pmatrix} \begin{matrix} \\ \\ -L_3 \end{matrix}$$

Donc les diviseurs élémentaires de U sont 1, $x+1$ et $(x+1)^2$.

On conclut cette section avec la preuve du théorème des diviseurs élémentaires :

Preuve du théorème 8.1.1. Soit N un sous-module de A^n. En vertu du théorème 8.3.1, N possède une base, donc en particulier il possède un système de générateurs finis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ et par suite, on peut écrire

$$N = \mathrm{Col}_A(U) \quad \text{où} \quad U = (\mathbf{u}_1 \cdots \mathbf{u}_m) \in \mathrm{Mat}_{n\times m}(A).$$

D'après le théorème 8.5.5, on peut, par une suite d'opérations élémentaires de lignes et de colonnes, amener la matrice U sous la forme

$$U' = \left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right)$$

pour un entier $r \geq 0$ et des éléments non nul $d_1, \ldots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$. D'après le théorème 8.5.4, il existe un automorphisme φ de A^n tel que

$$\mathrm{Col}_A(U') = \varphi(\mathrm{Col}_A(U) = \varphi(N).$$

Désignons par $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base canonique de A^n, et posons $C_j = \varphi^{-1}(\mathbf{e}_j)$ pour chaque $j = 1, \ldots, n$. Puisque φ^{-1} est un automorphisme de A^n, la proposition 6.5.13 montre que $\{C_1, \ldots, C_n\}$ est une base de A^n. Par ailleurs, on sait, grâce au lemme 8.4.5, que $\{d_1\mathbf{e}_1, \ldots, d_r\mathbf{e}_r\}$ est une base de $\mathrm{Col}_A(U')$. Alors le lemme 8.5.2 appliqué à φ^{-1} nous apprend que

$$\{\varphi^{-1}(d_1\mathbf{e}_1), \ldots, \varphi^{-1}(d_r\mathbf{e}_r)\} = \{d_1C_1, \ldots, d_rC_r\}$$

est une base de $\varphi^{-1}(\mathrm{Col}_A(U')) = N$. □

Exercices.

8.5.1. Démontrer le lemme 8.5.2.

8.5.2. Déterminer les diviseurs élémentaires de chacune des matrices suivantes :

(i) $\begin{pmatrix} 6 & 4 & 0 \\ 2 & 0 & 3 \\ 2 & 2 & 8 \end{pmatrix} \in \mathrm{Mat}_{3\times3}(\mathbb{Z})$, (ii) $\begin{pmatrix} 9 & 6 & 12 \\ 3 & 0 & 4 \end{pmatrix} \in \mathrm{Mat}_{2\times3}(\mathbb{Z})$,

(iii) $\begin{pmatrix} x^3 & 0 & x^2 \\ x^2 & x & 0 \\ 0 & x & x^3 \end{pmatrix} \in \mathrm{Mat}_{3\times3}(\mathbb{Q}[x])$, (iv) $\begin{pmatrix} x^3-1 & 0 & x^4-1 \\ x^2-1 & x+1 & 0 \end{pmatrix} \in \mathrm{Mat}_{2\times3}(\mathbb{Q}[x])$.

8.5.3. Soit A un anneau euclidien, et soit $U \in \mathrm{Mat}_{n\times m}(A)$. Montrer que l'idéal I de A engendré par tous les éléments de U n'est pas affecté par les opérations élémentaires de lignes et de colonnes. En déduire que le premier diviseur élémentaire de U est un générateur de I et que, par suite, ce diviseur est uniquement déterminé par U à multiplication près par une unité.

8.5.4. Soit A et U comme dans l'exercice 8.5.3, et soit k un entier avec $1 \leq k \leq \min(n, m)$. Pour tout choix d'entiers $i_1, \ldots, i_k$ et $j_1, \ldots, j_k$ avec

$$1 \leq i_1 < i_2 < \cdots < i_k \leq n \quad \text{et} \quad 1 \leq j_1 < j_2 < \cdots < j_k \leq m,$$

on désigne par $U_{i_1,\ldots,i_k}^{j_1,\ldots,j_k}$ la sous-matrice de U de format $k \times k$ formée par les éléments de U appartenant aux lignes d'indices $i_1, \ldots, i_k$ et aux colonnes d'indices $j_1, \ldots, j_k$. Soit I_k l'idéal de A engendré par les déterminants de toutes ces sous-matrices (appelés les *mineurs* $k \times k$ de U).

(i) Montrer que cet idéal n'est pas affecté par les opérations élémentaires de lignes et de colonnes.

(ii) Soit $d_1, \ldots, d_r$ les diviseurs élémentaires de U. Déduire de (i) que $I_k = (d_1 \cdots d_k)$ si $k \leq r$ et que $I_k = (0)$ si $k > r$.

(iii) Conclure que l'entier r est le plus grand entier pour lequel U contient une sous-matrice de format $r \times r$ de déterminant non nul, et que les diviseurs élémentaires $d_1, \ldots, d_r$ de U sont uniquement déterminés par U à multiplication près par des unités dans A^*.

Note. Cette construction généralise celle de l'exercice 8.5.3 puisque I_1 est simplement l'idéal de A engendré par tous les éléments de U.

8.6 Algorithmes

La démonstration du théorème 8.1.1 donnée à la section précédente fournit en principe le moyen de calculer la base $\{C_1, \ldots, C_n\}$ de A^n qui apparaît dans l'énoncé de ce théorème. Le but de cette section est de donner un algorithme pratique pour arriver à cette fin, puis d'en déduire un algorithme pour le calcul de la forme rationnelle d'un endomorphisme d'un espace vectoriel de dimension finie. On commence par l'observation suivante :

Lemme 8.6.1. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$, soit U' une matrice obtenue en appliquant à U une opération élémentaire de colonnes, et soit E la matrice obtenue en appliquant la même opération à I_m (la matrice identité $m \times m$). Alors on a :*

$$U' = U\,E.$$

Preuve. Désignons par $C_1, \ldots, C_m$ les colonnes de U. On distingue trois cas :

(I) La matrice U' est obtenue en permutant les colonnes i et j de U, où $i < j$. On a alors

$$U' = (C_1 \cdots \overset{i}{\underset{\vee}{C_j}} \cdots \overset{j}{\underset{\vee}{C_i}} \cdots C_m)$$

$$= (C_1 \cdots C_i \cdots C_j \cdots C_m) \begin{pmatrix} & & & i & & j & & & \\ & & & \downarrow & & \downarrow & & & \\ 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & \cdots & 1 & & & \leftarrow i \\ & & & \vdots & \ddots & \vdots & & & \\ & & & 1 & \cdots & 0 & & & \leftarrow j \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}$$

$$= U\,E.$$

(II) La matrice U' est obtenue en ajoutant à la colonne i de U le produit de la colonne j de U par un élément a de A, où $j \neq i$. Dans ce cas

$$U' = (C_1 \cdots \overset{i}{\vee} C_i + aC_j \cdots \overset{j}{\vee} C_j \cdots C_m)$$

$$= (C_1 \cdots C_i \cdots C_j \cdots C_m) \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \leftarrow i \\ & & \vdots & \ddots & & & \\ & & a & \cdots & 1 & & \leftarrow j \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

$$= U\,E.$$

(Cette présentation suppose $i < j$. Le cas $i > j$ est semblable.)

(III) La matrice U' est obtenue en multipliant la i-ième colonne de U par une unité $a \in A^*$. Alors

$$U' = (C_1 \cdots \overset{i}{\vee} aC_i \cdots C_m) = (C_1 \cdots \overset{i}{\vee} C_i \cdots C_m) \begin{pmatrix} & & i & & \\ & & \downarrow & & \\ 1 & & & & \\ & \ddots & & & \\ & & a & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = U\,E.$$

□

Un résultat semblable s'applique aux opérations élémentaires de lignes. L'énoncé suivant est laissé en exercice :

Lemme 8.6.2. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$, soit U' une matrice obtenue en appliquant à U une opération élémentaire de lignes, et soit E la matrice obtenue en appliquant à I_n la même opération. Alors on a :*

$$U' = E\,U.$$

On note aussi :

Lemme 8.6.3. *Soit k un entier positif. Toute matrice obtenue en appliquant à I_k une opération élémentaire de colonnes s'obtient aussi en appliquant à I_k une opération élémentaire de lignes, et vice versa.*

Preuve. (I) Permuter les colonnes i et j de I_k revient à permuter ses lignes i et j.

(II) Ajouter à la colonne i de I_k sa colonne j multipliée par un élément a de A, pour i et j distincts, équivaut à ajouter à la ligne j de I_k sa ligne i multipliée par a :

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \leftarrow i \\ & & \vdots & \ddots & & & \\ & & a & \cdots\ 1 & & & \leftarrow j \\ & & & & \ddots & & \\ & & & & & 1 & \end{pmatrix}$$

(III) Multiplier la colonne i de I_k par une unité a revient à multiplier sa ligne i par a. □

Ces résultats suggèrent la définition suivante.

Définition 8.6.4. Une *matrice élémentaire* $k \times k$ est toute matrice de $\mathrm{Mat}_{k\times k}(A)$ obtenue en appliquant à I_k une opération élémentaire de lignes ou de colonnes.

On note aussitôt :

Proposition 8.6.5. *Une matrice élémentaire est inversible et son inverse est encore une matrice élémentaire.*

Preuve. Soit $E \in \mathrm{Mat}_{k\times k}(A)$ une matrice élémentaire. Elle est obtenue en appliquant à I_k une opération élémentaire de colonnes. Soit F la matrice obtenue en appliquant à I_k l'opération élémentaire inverse. Alors le lemme 8.6.1 donne $I_k = EF$ et $I_k = FE$. Donc E est inversible et $E^{-1} = F$. □

Note. Le fait qu'une matrice élémentaire est inversible se vérifie aussi en observant que son déterminant est une unité de A (voir l'appendice B)

Grâce à la notion de matrice élémentaire, on peut reformuler le théorème 8.5.5 de la manière suivante :

Théorème 8.6.6. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$. Il existe des matrices élémentaires $E_1, \dots, E_s$ de format $n \times n$ et des matrices élémentaires $F_1, \dots, F_t$ de format $m \times m$ telles que*

$$E_s \cdots E_1 \, U \, F_1 \cdots F_t = \left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right) \tag{8.14}$$

pour un entier $r \geq 0$ et des éléments non nuls $d_1, \dots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$.

Preuve. Le théorème 8.5.5 nous apprend qu'on peut transformer U en une matrice de la forme du membre de droite de (8.14) par une suite d'opérations élémentaires de lignes et de colonnes. L'égalité (8.14) s'ensuit en désignant par E_i la matrice élémentaire qui encode la i-ième opération élémentaire de lignes et par F_j celle qui encode la j-ième opération élémentaire de colonnes. □

Comme toute matrice élémentaire est inversible et qu'un produit de matrices inversibles est inversible, on en déduit :

Corollaire 8.6.7. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$. Il existe des matrices inversibles $P \in \mathrm{Mat}_{n\times n}(A)$ et $Q \in \mathrm{Mat}_{m\times m}(A)$ telles que PUQ soit sous forme normale de Smith.*

Pour calculer P et Q, on peut procéder ainsi :

Algorithme 8.6.8. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$.*

- *On construit le tableau "en coin"* $\left[\begin{array}{c|c} U & I_n \\ \hline I_m & \end{array}\right]$.
- *En appliquant une suite d'opérations élémentaires à ses n premières lignes et à ses m premières colonnes, on l'amène sous la forme* $\left[\begin{array}{c|c} S & P \\ \hline Q & \end{array}\right]$ *où S désigne la forme normale de Smith de U.*
- *Alors P et Q sont des matrices inversibles telles que $PUQ = S$.*

Preuve. Soit s le nombre d'opérations élémentaires requises et, pour $i = 0, 1, \dots, s$, soit

$$\left[\begin{array}{c|c} U_i & P_i \\ \hline Q_i & \end{array}\right]$$

le tableau obtenu après i opérations élémentaires, de sorte que la chaîne de tableaux intermédiaires se lise

$$\left[\begin{array}{c|c} U & I_n \\ \hline I_m & \end{array}\right] = \left[\begin{array}{c|c} U_0 & P_0 \\ \hline Q_0 & \end{array}\right] \sim \left[\begin{array}{c|c} U_1 & P_1 \\ \hline Q_1 & \end{array}\right] \sim \cdots \sim \left[\begin{array}{c|c} U_s & P_s \\ \hline Q_s & \end{array}\right] = \left[\begin{array}{c|c} S & P \\ \hline Q & \end{array}\right].$$

On va montrer par récurrence que, pour $i = 0, 1, \ldots, s$, les matrices P_i et Q_i sont inversibles et que $P_iUQ_i = U_i$. Pour $i = s$, cela montrera que P et Q sont inversibles avec $PUQ = S$ comme annoncé.

Pour $i = 0$, l'affirmation à démontrer est claire car I_n et I_m sont inversibles et $I_nUI_m = U$. Supposons maintenant qu'on ait $1 \leq i < s$, que P_{i-1} et Q_{i-1} soient inversibles et que $P_{i-1}UQ_{i-1} = U_{i-1}$. On distingue deux cas.

Cas 1 : La i-ième opération est une opération élémentaire de lignes (sur les n premières lignes). Dans ce cas, si E est la matrice élémentaire correspondante, on a

$$U_i = EU_{i-1}, \quad P_i = EP_{i-1} \quad \text{et} \quad Q_i = Q_{i-1}.$$

Comme E, P_{i-1} et Q_{i-1} sont inversibles, on en déduit que P_i et Q_i sont inversibles et que

$$P_iUQ_i = EP_{i-1}UQ_{i-1} = EU_{i-1} = U_i.$$

Cas 2 : La i-ième opération est une opération élémentaire de colonnes (sur les m premières colonnes). Dans ce cas, si F est la matrice élémentaire correspondante, on a

$$U_i = U_{i-1}F, \quad P_i = P_{i-1} \quad \text{et} \quad Q_i = Q_{i-1}F.$$

Comme F, P_{i-1} et Q_{i-1} sont inversibles, il en va de même de P_i et Q_i et on obtient de nouveau

$$P_iUQ_i = P_{i-1}UQ_{i-1}F = U_{i-1}F = U_i.$$

□

Exemple 8.6.9. Soit $U = \begin{pmatrix} 4 & 8 & 4 \\ 4 & 10 & 8 \end{pmatrix} \in \mathrm{Mat}_{2\times 3}(\mathbb{Z})$. On cherche des matrices inversibles P et Q telles que PUQ soit sous forme normale de Smith.

Solution: On reprend les calculs de l'exemple 8.5.6 en les appliquant au tableau $\left[\begin{array}{c|c} U & I_2 \\ \hline I_3 & \end{array}\right]$. On trouve de cette manière

$$\left[\begin{array}{ccc|cc} 4 & 8 & 4 & 1 & 0 \\ 4 & 10 & 8 & 0 & 1 \\ \hline 1 & 0 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \end{array}\right] \sim \overset{\begin{matrix} C_2 - 2C_1 & C_3 - C_1 \end{matrix}}{\left[\begin{array}{ccc|cc} 4 & 0 & 0 & 1 & 0 \\ 4 & 2 & 4 & 0 & 1 \\ \hline 1 & -2 & -1 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \end{array}\right]}$$

$$\sim \left[\begin{array}{ccc|cc} 4 & 0 & 0 & 1 & 0 \\ 0 & 2 & 4 & -1 & 1 \\ \hline 1 & -2 & -1 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \end{array}\right] \begin{matrix} \\ \\ \\ \\ \\ L_2 - L_1 \end{matrix} \sim \left[\begin{array}{ccc|cc} 4 & 2 & 4 & 0 & 1 \\ 0 & 2 & 4 & -1 & 1 \\ \hline 1 & -2 & -1 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \end{array}\right] \begin{matrix} \\ \\ \\ \\ \\ L_1 + L_2 \end{matrix}$$

$$\begin{array}{c} \\ \sim \\ \\ \\ \\ \\ \sim \\ \\ \\ \\ \\ \sim \\ \\ \\ \\ \\ \sim \\ \\ \\ \\ \\ \end{array}
\left[\begin{array}{ccc|cc} \multicolumn{3}{c}{\scriptstyle C_2\quad C_1} & & \\ 2 & 4 & 4 & 0 & 1 \\ 2 & 0 & 4 & -1 & 1 \\ \hline -2 & 1 & -1 & & \\ 1 & 0 & 0 & & \\ 0 & 0 & 1 & & \end{array}\right]
\left[\begin{array}{ccc|cc} \multicolumn{3}{c}{\scriptstyle C_2-2C_1\quad C_3-2C_1} & & \\ 2 & 0 & 0 & 0 & 1 \\ 2 & -4 & 0 & -1 & 1 \\ \hline -2 & 5 & 3 & & \\ 1 & -2 & -2 & & \\ 0 & 0 & 1 & & \end{array}\right]$$

$$\begin{array}{c} \\ \sim \\ \\ \\ \\ \\ \sim \\ \\ \\ \\ \\ \end{array}
\left[\begin{array}{ccc|cc} 2 & 0 & 0 & 0 & 1 \\ 0 & -4 & 0 & -1 & 0 \\ \hline -2 & 5 & 3 & & \\ 1 & -2 & -2 & & \\ 0 & 0 & 1 & & \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \\ L_2-L_1 \\ \\ \\ \\ \\ \end{array}
\left[\begin{array}{ccc|cc} 2 & 0 & 0 & 0 & 1 \\ 0 & 4 & 0 & 1 & 0 \\ \hline -2 & 5 & 3 & & \\ 1 & -2 & -2 & & \\ 0 & 0 & 1 & & \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \\ -L_2 \\ \\ \\ \\ \\ \end{array}$$

Donc $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{Z})$ et $Q = \begin{pmatrix} -2 & 5 & 3 \\ 1 & -2 & -2 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{Z})$ sont des matrices inversibles (sur $\mathbb{Z}$) telles que

$$PUQ = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \end{pmatrix}$$

soit la forme normale de Smith de U.

Le résultat suivant montre en particulier que toute matrice inversible est un produit de matrices élémentaires. Par suite, le corollaire 8.6.7 est équivalent au théorème 8.6.6 (chacun des deux énoncés implique l'autre).

Théorème 8.6.10. *Soit $P \in \mathrm{Mat}_{n\times n}(A)$. Les conditions suivantes sont équivalentes :*

(i) P est inversible,

(ii) P peut être transformée en I_n par une suite d'opérations élémentaires de lignes,

(iii) P s'écrit comme un produit de matrices élémentaires.

Si ces conditions équivalentes sont remplies, on peut, par une suite d'opérations élémentaires de lignes, transformer la matrice $(P \,|\, I_n) \in \mathrm{Mat}_{n\times 2n}(A)$ en une matrice de la forme $(I_n \,|\, Q)$ et alors on a $P^{-1} = Q$.

Dans les exercices 8.6.2 et 8.6.3, on propose une esquisse de preuve de ce résultat ainsi qu'une méthode pour déterminer si P est inversible et, lorsque c'est le cas, pour transformer P en I_n par des opérations élémentaires de lignes.

Exemple 8.6.11. Soit $P = \begin{pmatrix} x+1 & x^2+x+1 \\ x & x^2+1 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{Q}[x])$. Montrer que P est inversible et calculer P^{-1}.

Solution: On trouve

$$(P \,|\, I) = \left(\begin{array}{cc|cc} x+1 & x^2+x+1 & 1 & 0 \\ x & x^2+1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & x & 1 & -1 \\ x & x^2+1 & 0 & 1 \end{array}\right) \; L_1 - L_2$$

$$\sim \left(\begin{array}{cc|cc} 1 & x & 1 & -1 \\ 0 & 1 & -x & 1+x \end{array}\right) \begin{array}{l} \\ L_2 - xL_1 \end{array} \sim \left(\begin{array}{cc|cc} 1 & 0 & x^2+1 & -x^2-x-1 \\ 0 & 1 & -x & x+1 \end{array}\right) \begin{array}{l} \\ L_1 - xL_2 \end{array}$$

Donc P est inversible (on peut la transformer en I_2 par une suite d'opérations élémentaires de lignes) et

$$P^{-1} = \begin{pmatrix} x^2+1 & -x^2-x-1 \\ -x & x+1 \end{pmatrix}.$$

Le théorème des diviseurs élémentaires (théorème 8.1.1) fournit une base de A^n adaptée à chaque sous-module N de A^n. L'algorithme suivant procure un moyen pratique pour calculer une telle base lorsqu'on connaît un système de générateurs du module N.

Algorithme 8.6.12. *Soit $U \in \mathrm{Mat}_{n\times m}(A)$ et soit $N = \mathrm{Col}_A(U)$.*

- *En utilisant l'algorithme 8.6.8, on construit des matrices inversibles P et Q telles que*

$$PUQ = \left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right) \tag{8.15}$$

  pour un entier $r \geq 0$ et des éléments non nuls $d_1, \ldots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$.
- *Alors les colonnes de P^{-1} forment une base $\{C_1, \ldots, C_n\}$ de A^n telle que $\{d_1C_1, \ldots, d_rC_r\}$ soit une base de N.*

Preuve. Soit $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base canonique de A^n et soient $C_1, \ldots, C_n$ les colonnes de P^{-1}. L'application

$$\begin{array}{rccc} \varphi : & A^n & \longrightarrow & A^n \\ & X & \longmapsto & P^{-1}X \end{array}$$

est un isomorphisme de A-modules (exercice) et par suite la proposition 6.5.13 nous apprend que

$$\{P^{-1}(\mathbf{e}_1), \ldots, P^{-1}(\mathbf{e}_n)\} = \{C_1, \ldots, C_n\}$$

est une base de A^n. Par ailleurs l'égalité 8.15 se réécrit

$$UQ = P^{-1}\left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right) = \Big(d_1C_1 \ \cdots \ d_rC_r \ \mathbf{0} \ \cdots \ \mathbf{0}\Big).$$

Comme Q est inversible, elle est un produit de matrices élémentaires (théorème 8.6.10) et par conséquent UQ s'obtient de U par une suite de transformations élémentaires de colonnes. En vertu du lemme 8.4.3, on en déduit que

$$N = \mathrm{Col}_A(U) = \mathrm{Col}_A(UQ) = \langle d_1C_1, \ldots, d_rC_r\rangle_A.$$

Donc $\{d_1C_1, \dots, d_rC_r\}$ est un système de générateurs de N. Si $a_1, \dots, a_r \in A$ satisfont

$$a_1(d_1C_1) + \cdots + a_r(d_rC_r) = \mathbf{0},$$

alors $(a_1d_1)C_1 + \cdots + (a_rd_r)C_r = \mathbf{0}$, donc $a_1d_1 = \cdots = a_rd_r = 0$ car $C_1, \dots, C_n$ sont linéairement indépendants sur A, et par suite $a_1 = \cdots = a_r = 0$. Cela montre que $d_1C_1, \dots, d_rC_r$ sont linéairement indépendants et en conséquence $\{d_1C_1, \dots, d_rC_r\}$ est une base de N. □

En fait, comme la seconde partie de l'algorithme 8.6.12 ne requiert que la matrice P, on peut se passer de calculer la matrice Q : il suffit d'appliquer l'algorithme 8.6.8 au tableau $(U \mid I_n)$.

Exemple 8.6.13. Soit $N = \left\langle \begin{pmatrix} x \\ x^2 \end{pmatrix}, \begin{pmatrix} x^2 \\ x \end{pmatrix}, \begin{pmatrix} x \\ x^3 \end{pmatrix} \right\rangle_{\mathbb{Q}[x]} \subseteq \mathbb{Q}[x]^2$. Alors $N = \mathrm{Col}_{\mathbb{Q}[x]}(U)$ où $U = \begin{pmatrix} x & x^2 & x \\ x^2 & x & x^3 \end{pmatrix}$. On trouve

$$(U \mid I) = \left(\begin{array}{ccc|cc} x & x^2 & x & 1 & 0 \\ x^2 & x & x^3 & 0 & 1 \end{array}\right) \sim \begin{array}{c} \\ \left(\begin{array}{ccc|cc} x & 0 & 0 & 1 & 0 \\ x^2 & x-x^3 & x^3-x^2 & 0 & 1 \end{array}\right) \end{array} \begin{array}{c} \\ C_2 - xC_1 \quad C_3 - C_1 \end{array}$$

$$\sim \begin{array}{c} \\ \left(\begin{array}{ccc|cc} x & 0 & 0 & 1 & 0 \\ 0 & x-x^3 & x^3-x^2 & -x & 1 \end{array}\right) \end{array} \begin{array}{c} \\ L_2 - xL_1 \end{array} \sim \begin{array}{c} \\ \left(\begin{array}{ccc|cc} x & 0 & 0 & 1 & 0 \\ 0 & x-x^2 & x^3-x^2 & -x & 1 \end{array}\right) \end{array} \begin{array}{c} \\ C_2 + C_3 \end{array}$$

$$\sim \begin{array}{c} \\ \left(\begin{array}{ccc|cc} x & 0 & 0 & 1 & 0 \\ 0 & x-x^2 & 0 & -x & 1 \end{array}\right) \end{array} \begin{array}{c} \\ C_3 + xC_2 \end{array} \sim \begin{array}{c} \\ \left(\begin{array}{ccc|cc} x & 0 & 0 & 1 & 0 \\ 0 & x^2-x & 0 & x & -1 \end{array}\right) \end{array} \begin{array}{c} \\ -L_2 \end{array}$$

Donc les diviseurs élémentaires de N sont x et $x^2 - x = x(x-1)$. En posant $P = \begin{pmatrix} 1 & 0 \\ x & -1 \end{pmatrix}$, on trouve

$$(P \mid I) = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ x & -1 & 0 & 1 \end{array}\right) \sim \begin{array}{c} \\ \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & -1 & -x & 1 \end{array}\right) \end{array} \begin{array}{c} \\ L_2 - xL_1 \end{array} \sim \begin{array}{c} \\ \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & x & -1 \end{array}\right) \end{array} \begin{array}{c} \\ -L_2 \end{array}$$

donc $P^{-1} = \begin{pmatrix} 1 & 0 \\ x & -1 \end{pmatrix}$. Suivant l'algorithme 8.6.12, on conclut que

$$\left\{ C_1 = \begin{pmatrix} 1 \\ x \end{pmatrix},\ C_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}$$

est une base de $\mathbb{Q}[x]^2$ telle que

$$\{xC_1,\ (x^2 - x)C_2\} = \left\{ \begin{pmatrix} x \\ x^2 \end{pmatrix}, \begin{pmatrix} 0 \\ -x^2 + x \end{pmatrix} \right\}$$

soit une base de N.

On conclut ce chapitre avec le résultat suivant qui précise le théorème 7.2.4 et qui, combiné au théorème 7.2.5, permet de construire une base relative à laquelle la matrice d'un opérateur linéaire est sous forme rationnelle.

Théorème 8.6.14. *Soit V un espace vectoriel de dimension finie $n \geq 1$ sur un corps K, soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, et soit $T\colon V \to V$ un opérateur linéaire. Alors il existe des matrices inversibles $P, Q \in \mathrm{Mat}_{n\times n}(K[x])$ et des polynômes unitaires $d_1(x), \ldots, d_n(x) \in K[x]$ avec $d_1(x) \,|\, d_2(x) \,|\, \cdots \,|\, d_n(x)$ tels que*

$$P\,(xI - [\,T\,]_{\mathcal{B}})\,Q \;=\; \begin{pmatrix} d_1(x) & & \\ & \ddots & \\ 0 & & d_n(x) \end{pmatrix}. \tag{8.16}$$

Écrivons

$$P^{-1} = \begin{pmatrix} p_{11}(x) & \cdots & p_{1n}(x) \\ \vdots & & \vdots \\ p_{n1}(x) & \cdots & p_{nn}(x) \end{pmatrix},$$

et, pour la structure de $K[x]$-module sur V associée à T, posons

$$\mathbf{u}_j \;=\; p_{1j}(x)\mathbf{v}_1 + \cdots + p_{nj}(x)\mathbf{v}_n$$

pour $j = 1, \ldots, n$. Alors on a

$$V = K[x]\mathbf{u}_k \oplus \cdots \oplus K[x]\mathbf{u}_n$$

où k désigne le plus petit entier tel que $d_k(x) \neq 1$. De plus l'annulateur de $\mathbf{u}_j$ est engendré par $d_j(x)$ pour $j = k, \ldots, n$.

Preuve. La proposition 8.1.2 nous apprend que l'application

$$\begin{array}{rccl} \psi\colon & K[x]^n & \longrightarrow & V \\ & \begin{pmatrix} p_1(x) \\ \vdots \\ p_n(x) \end{pmatrix} & \longmapsto & p_1(x)\mathbf{v}_1 + \cdots + p_n(x)\mathbf{v}_n \end{array}$$

est un homomorphisme surjectif de $K[x]$-modules, dont le noyau est

$$\ker(\psi) = \mathrm{Col}_{K[x]}(xI - [\,T\,]_{\mathcal{B}}).$$

L'algorithme 8.6.8 permet de construire des matrices inversibles $P, Q \in \mathrm{Mat}_{n\times n}(K[x])$ telles que

$$P(x\mathrm{I} - [\,T\,]_{\mathcal{B}})Q = \begin{pmatrix} d_1(x) & & & & & 0 \\ & \ddots & & & & \\ & & d_r(x) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

pour un entier $r \geq 0$ et des polynômes non nuls $d_1(x), \ldots, d_r(x) \in K[x]$ avec $d_1(x) \mid \cdots \mid d_r(x)$. Comme le déterminant du membre de gauche de cette égalité n'est pas nul, on doit avoir $r = n$. Quitte à modifier P ou Q, on peut faire en sorte que $d_1(x), \ldots, d_n(x)$ soient unitaires. Cela démontre la première assertion du théorème.

L'algorithme 8.6.12 nous apprend que les colonnes de P^{-1} forment une base de $\{C_1, \ldots, C_n\}$ de $K[x]^n$ telle que $\{d_1(x)C_1, \ldots, d_n(x)C_n\}$ soit une base de $\ker(\psi)$. Par construction, on a aussi

$$\mathbf{u}_j = \psi(C_j)$$

pour $j = 1, \ldots, n$. La conclusion suit en reprenant la preuve du théorème 7.2.1 donnée à la section 8.1 (avec $A = K[x]$). □

Ce résultat fournit une nouvelle preuve du théorème de Cayley-Hamilton. On en déduit en effet :

Corollaire 8.6.15. *Dans les notations du théorème 8.6.14, on a*

$$\mathrm{car}_T(x) = d_1(x)d_2(x)\cdots d_n(x)$$

et $d_n(x)$ est le polynôme minimal de T. De plus, on a $\mathrm{car}_T(T) = 0$.

Preuve. En prenant le déterminant des deux membres de (8.16), on trouve

$$\det(P)\,\mathrm{car}_T(x)\,\det(Q) = d_1(x)\cdots d_n(x).$$

Comme P et Q sont des éléments inversibles de $\mathrm{Mat}_{n\times n}(K[x])$, leurs déterminants sont des éléments de $K[x]^* = K^*$, donc l'égalité ci-dessus se réécrit

$$\mathrm{car}_T(x) = a\, d_1(x)\cdots d_n(x)$$

pour une constante $a \in K^*$. Comme les polynômes $\mathrm{car}_T(x)$ et $d_1(x), \ldots, d_n(x)$ sont tous unitaires, on doit avoir $a = 1$. Par ailleurs, le théorème 7.2.4 nous apprend que $d_n(x)$ est le polynôme minimal de T. On en déduit que

$$\mathrm{car}_T(T) = d_1(T) \circ \cdots \circ d_n(T) = 0.$$

□

En appliquant le théorème 7.2.5, on obtient aussi :

Corollaire 8.6.16. *Avec les notations du théorème 8.6.14, posons $m_i = \deg(d_i(x))$ pour $i = k, \ldots, n$. Alors*

$$\mathcal{C} = \left\{\mathbf{u}_k, T(\mathbf{u}_k), \ldots, T^{m_k-1}(\mathbf{u}_k), \ldots\ldots, u_n, T(\mathbf{u}_n), \ldots, T^{m_n-1}(\mathbf{u}_n)\right\}$$

est une base de V telle que

$$[T]_{\mathcal{C}} = \begin{pmatrix} D_k & & 0 \\ & \ddots & \\ 0 & & D_n \end{pmatrix}$$

où D_i désigne la matrice compagnon de $d_i(x)$ pour $i = k, \ldots, n$.

Exemple 8.6.17. Soit V un espace vectoriel sur $\mathbb{Q}$ de dimension 3, soit $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ une base de V, et soit $T\colon V \to V$ l'opérateur linéaire pour lequel

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & -1 & -1 \\ -3 & 2 & 3 \\ 4 & -4 & -5 \end{pmatrix}.$$

Déterminer une base $\mathcal{C}$ de V telle que $[T]_{\mathcal{C}}$ soit sous forme rationnelle.

Solution: On applique d'abord l'algorithme 8.6.8 à la matrice

$$U = xI - [T]_{\mathcal{B}} = \begin{pmatrix} x & 1 & 1 \\ 3 & x-2 & -3 \\ -4 & 4 & x+5 \end{pmatrix}.$$

Comme on a seulement besoin de la matrice P, on travaille avec le tableau $(U \mid I)$. En reprenant les calculs de l'exemple 8.5.7, on trouve

$$\begin{aligned}
(U \mid I) &\sim \left(\begin{array}{ccc|ccc} & C_2 & C_1 & & & \\ 1 & x & 1 & 1 & 0 & 0 \\ x-2 & 3 & -3 & 0 & 1 & 0 \\ 4 & -4 & x+5 & 0 & 0 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} & C_2 - xC_1 & C_3 - C_1 & & & \\ 1 & 0 & 0 & 1 & 0 & 0 \\ x-2 & -x^2+2x+3 & -x-1 & 0 & 1 & 0 \\ 4 & -4x-4 & x+1 & 0 & 0 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -x^2+2x+3 & -x-1 & -x+2 & 1 & 0 \\ 0 & -4x-4 & x+1 & -4 & 0 & 1 \end{array}\right) \begin{array}{c} \\ L_2 - (x-2)L_1 \\ L_3 - 4L_1 \end{array} \\
&\sim \left(\begin{array}{ccc|ccc} & C_3 & C_2 & & & \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -x-1 & -x^2+2x+3 & -x+2 & 1 & 0 \\ 0 & x+1 & -4x-4 & -4 & 0 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & x+1 & x^2-2x-3 & x-2 & -1 & 0 \\ 0 & 0 & -x^2-2x-1 & -x-2 & 1 & 1 \end{array}\right) \begin{array}{c} \\ -L_2 \\ L_3 + L_2 \end{array} \\
&\sim \left(\begin{array}{ccc|ccc} & & C_3 - (x-3)C_2 & & & \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & x+1 & 0 & x-2 & -1 & 0 \\ 0 & 0 & -(x+1)^2 & -x-2 & 1 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & x+1 & 0 & x-2 & -1 & 0 \\ 0 & 0 & (x+1)^2 & x+2 & -1 & -1 \end{array}\right) \begin{array}{c} \\ \\ -L_3 \end{array}
\end{aligned}$$

Donc les diviseurs élémentaires sont 1, $x+1$ et $(x+1)^2$. On trouve que

$$P^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ x-2 & -1 & 0 \\ x+2 & -1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ x-2 & -1 & 0 \\ 4 & 1 & -1 \end{pmatrix}$$

(la vérification de ce calcul est laissée en exercice). On en déduit que

$$V = \mathbb{Q}[x]\mathbf{u}_2 \oplus \mathbb{Q}[x]\mathbf{u}_3 \quad \text{où} \quad \mathbf{u}_2 = -\mathbf{v}_2 + \mathbf{v}_3 \quad \text{et} \quad \mathbf{u}_3 = -\mathbf{v}_3.$$

De plus l'annulateur de $\mathbf{u}_2$ est engendré par $x+1$ et celui de $\mathbf{u}_3$ par $(x+1)^2 = x^2+2x+1$. On conclut que

$$\mathcal{C} = \{\mathbf{u}_2, \mathbf{u}_3, T(\mathbf{u}_3)\} = \{-\mathbf{v}_2 + \mathbf{v}_3,\ -\mathbf{v}_3,\ \mathbf{v}_1 - 3\mathbf{v}_2 + 5\mathbf{v}_3\}$$

est une base de V telle que

$$[T]_{\mathcal{C}} = \begin{pmatrix} \boxed{-1} & 0 & 0 \\ 0 & \boxed{\begin{matrix} 0 & -1 \\ 1 & -2 \end{matrix}} \end{pmatrix}$$

où, sur la diagonale, on reconnaît les matrices compagnon de $x+1$ et de x^2+2x+1.

Note. La première colonne de P^{-1} fournit

$$\mathbf{u}_1 = 1\mathbf{v}_1 + (x-1)\mathbf{v}_2 + 4\mathbf{v}_3 = \mathbf{v}_1 + T(\mathbf{v}_2) - 2\mathbf{v}_2 + 4\mathbf{v}_3 = \mathbf{0}$$

comme il se doit, puisque l'annulateur de $\mathbf{u}_1$ est engendré par $d_1(x) = 1$.

Exercices.

Dans tous les exercices ci-dessous, A désigne un anneau euclidien.

8.6.1. Démontrer le lemme 8.6.2.

8.6.2. Soit $P \in \mathrm{Mat}_{n\times n}(A)$.

(i) Donner un algorithme pour transformer P en une matrice triangulaire supérieure T par une suite d'opérations élémentaires de lignes.

(ii) Montrer que P est inversible si et seulement si les éléments de la diagonale de T sont des unités de A.

(iii) Si les éléments de la diagonale de T sont des unités, donner un algorithme pour transformer T en I_n par une suite d'opérations élémentaires de lignes.

(iv) Conclure que, si P est inversible, alors P peut être transformée en la matrice I_n par une suite d'opérations élémentaires de lignes.

Indice. Pour la partie (ii), montrer P est inversible si et seulement si T est inversible et utiliser le fait qu'une matrice de $\mathrm{Mat}_{n\times n}(A)$ est inversible si et seulement si son déterminant est une unité de A (voir l'appendice B).

8.6.3. Le problème 8.6.2 montre que, dans l'énoncé du théorème 8.6.10, la condition (i) implique la condition (ii). Compléter la preuve de ce théorème en montrant les implications $(ii) \Rightarrow (iii)$ et $(iii) \Rightarrow (i)$, puis en démontrant la dernière affirmation du théorème.

8.6.4. Pour chacune des matrices U de l'exercice 8.5.2, déterminer des matrices inversibles P et Q telles que PUQ soit sous forme normale de Smith.

8.6.5. Dans chaque cas, déterminer une base $\{C_1, \ldots, C_n\}$ de A^n et des éléments non nuls $d_1, \ldots, d_r$ de A avec $d_1 \,|\, d_2 \,|\, \cdots \,|\, d_r$ tels que $\{d_1C_1, \ldots, d_rC_r\}$ soit une base du sous-module donné N de A^n.

(i) $A = \mathbb{Q}[x]$, $\quad N = \left\langle \begin{pmatrix} x \\ x+1 \\ x \end{pmatrix}, \begin{pmatrix} 0 \\ x \\ x^2 \end{pmatrix} \right\rangle_{\mathbb{Q}[x]}$

(ii) $A = \mathbb{R}[x]$, $\quad N = \left\langle \begin{pmatrix} x+1 \\ -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -4 \\ x+1 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \\ x-5 \end{pmatrix} \right\rangle_{\mathbb{R}[x]}$

(iii) $A = \mathbb{Z}$, $\quad N = \left\langle \begin{pmatrix} 4 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 \\ 10 \end{pmatrix}, \begin{pmatrix} 2 \\ 8 \end{pmatrix} \right\rangle_{\mathbb{Z}}$

(iv) $A = \mathbb{Z}$, $\quad N = \left\langle \begin{pmatrix} 4 \\ 12 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 8 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \\ -2 \end{pmatrix} \right\rangle_{\mathbb{Z}}$

8.6.6. Soit V un espace vectoriel de dimension finie sur un corps K, et soit $T\colon V \to V$ une application linéaire. Montrer que $\min_T(x)$ et $\mathrm{car}_T(x)$ ont les mêmes facteurs irréductibles dans $K[x]$.

Indice. Utiliser le Corollaire 8.6.15.

8.6.7. Soit $T\colon \mathbb{Q}^3 \to \mathbb{Q}^3$ une application linéaire. Supposons que $\mathrm{car}_T(x)$ admette un facteur irréductible de degré ≥ 2 dans $\mathbb{Q}[x]$. Montrer que $\mathbb{Q}^3$ est un $\mathbb{Q}[x]$-module cyclique.

8.6.8. Soit V un espace vectoriel de dimension 2 sur $\mathbb{Q}$ avec base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, et soit $T\colon V \to V$ une application linéaire avec $[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$.

(i) Déterminer $\mathrm{car}_T(x)$ et $\min_T(x)$.

(ii) Déterminer une base $\mathcal{C}$ de V telle que $[T]_{\mathcal{C}}$ soit diagonale par blocs avec des matrices compagnon de polynômes sur la diagonale.

8.6.9. Soit $A = \begin{pmatrix} 4 & 2 & 2 \\ 1 & 3 & 1 \\ -3 & -3 & -1 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{Q})$.

(i) Déterminer $\mathrm{car}_A(x)$ et $\min_A(x)$.

(ii) Déterminer une matrice D semblable à A qui soit diagonale par blocs avec des matrices compagnon de polynômes sur la diagonale.

Chapitre 9

Dualité et produit tensoriel

Dans ce chapitre, on présente des constructions d'algèbre linéaire qui jouent un rôle important dans de nombreux domaines des mathématiques, aussi bien en analyse qu'en algèbre.

9.1 Dualité

Définition 9.1.1. Soit V un espace vectoriel sur un corps K. Le *dual* de V est l'espace vectoriel $\mathcal{L}_K(V,K)$ des applications linéaires de V dans K. On le note V^*. Un élément de V^* s'appelle une *forme linéaire* sur V.

Exemple 9.1.2. Soient X un ensemble et $\mathcal{F}(X,K)$ l'espace vectoriel des fonctions $g\colon X \to K$. Pour tout $x \in X$, l'application

$$\begin{array}{rccc} E_x \ : & \mathcal{F}(X,K) & \longrightarrow & K \\ & g & \longmapsto & g(x) \end{array}$$

dite d'*évaluation au point* x est linéaire, donc $E_x \in \mathcal{F}(X,K)^*$.

Proposition 9.1.3. *Supposons que V soit de dimension finie sur K et que $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ soit une base de V. Pour chaque $i = 1, \ldots, n$, on désigne par $f_i\colon V \to K$ l'application linéaire telle que*

$$f_i(\mathbf{v}_j) = \delta_{ij} := \begin{cases} 0 & \text{si } j \neq i, \\ 1 & \text{si } j = i. \end{cases}$$

Alors $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ est une base de V^*.

On dit que $\mathcal{B}^*$ est la base de V^* *duale* à $\mathcal{B}$.

Preuve. On sait que $\mathcal{E} = \{1\}$ est une base de K vu comme espace vectoriel sur lui-même. Alors le théorème 2.4.3 fournit un isomorphisme

$$\begin{array}{ccc} \mathcal{L}_K(V,K) & \xrightarrow{\sim} & \mathrm{Mat}_{1\times n}(K)\,. \\ f & \longmapsto & [\,f\,]_{\mathcal{E}}^{\mathcal{B}} \end{array}$$

Pour $i = 1, \ldots, n$, on trouve que

$$[\, f_i \,]_{\mathcal{E}}^{\mathcal{B}} = \Big([f_i(\mathbf{v}_1)]_{\mathcal{E}} \ \cdots \ [f_i(\mathbf{v}_n)]_{\mathcal{E}} \Big) = (0, \ldots, \overset{i}{\overset{\vee}{1}}, \ldots, 0) = \mathbf{e}_i^t.$$

Comme $\{\mathbf{e}_1^t, \ldots, \mathbf{e}_n^t\}$ est une base de $\mathrm{Mat}_{1\times n}(K)$, on en déduit que $\{f_1, \ldots, f_n\}$ est une base de V^*. □

Proposition 9.1.4. *Soient V et W des espaces vectoriels sur K et soit $T\colon V \to W$ une application linéaire. Pour toute forme linéaire $g\colon W \to K$, la composée $g \circ T\colon V \to K$ est une forme linéaire sur V. L'application*

$$\begin{array}{rccc} T^* : & W^* & \longrightarrow & V^* \\ & g & \longmapsto & g \circ T \end{array}$$

ainsi obtenue est linéaire.

On dit que T^* est la *transposée* de T.

Preuve. La première affirmation découle du fait que la composée de deux applications linéaires est linéaire. Le fait que T^* est linéaire découle de la proposition 2.3.1. □

Exemple 9.1.5. Soient X et Y deux ensembles, et soit $h\colon X \to Y$ une fonction. On vérifie que l'application

$$\begin{array}{rccc} T : & \mathcal{F}(Y, K) & \longrightarrow & \mathcal{F}(X, K) \\ & g & \longmapsto & g \circ h \end{array}$$

est linéaire (exercice). Dans ce contexte, on se propose de déterminer $T^*(E_x)$ où l'expression $E_x \in \mathcal{F}(X, K)^*$ désigne l'évaluation en un point x de X (voir l'exemple 9.1.2).

Pour tout $g \in \mathcal{F}(Y, K)$, on trouve

$$(T^*(E_x))(g) = (E_x \circ T)(g) = E_x(T(g)) = E_x(g \circ h) = (g \circ h)(x) = g(h(x)) = E_{h(x)}(g),$$

donc $T^*(E_x) = E_{h(x)}$.

Proposition 9.1.6. *Soient U, V et W des espaces vectoriels sur K.*

(i) Si $S, T \in \mathcal{L}_K(V, W)$ et $c \in K$, alors

$$(S + T)^* = S^* + T^* \quad \textit{et} \quad (cT)^* = c\,T^*.$$

(ii) Si $S \in \mathcal{L}_K(U, V)$ et $T \in \mathcal{L}_K(V, W)$, alors

$$(T \circ S)^* = S^* \circ T^*.$$

La preuve de ce résultat est laissée en exercice. La partie (i) implique :

Corollaire 9.1.7. *Soient V et W des espaces vectoriels sur K. L'application*

$$\begin{array}{ccc} \mathcal{L}_K(V,W) & \longrightarrow & \mathcal{L}_K(W^*,V^*) \\ T & \longmapsto & T^* \end{array}$$

est linéaire.

Théorème 9.1.8. *Soient V et W des espaces vectoriels de dimension finie sur K, soient*

$$\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \quad et \quad \mathcal{D} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$$

des bases respectives de V et W, et soient

$$\mathcal{B}^* = \{f_1, \ldots, f_n\} \quad et \quad \mathcal{D}^* = \{g_1, \ldots, g_m\}$$

les bases de V^* et W^* duales respectivement à $\mathcal{B}$ et $\mathcal{D}$. Pour tout $T \in \mathcal{L}_K(V,W)$, on a

$$[T^*]_{\mathcal{B}^*}^{\mathcal{D}^*} = \big([T]_{\mathcal{D}}^{\mathcal{B}}\big)^t .$$

Autrement dit, la matrice de la transposée T^* de T relatives aux bases duales $\mathcal{D}^*$ et $\mathcal{B}^*$ est la transposée de celle de T relatives aux bases $\mathcal{B}$ et $\mathcal{D}$.

Preuve. Soit $T \in \mathcal{L}_K(V,W)$. Écrivons $[T]_{\mathcal{D}}^{\mathcal{B}} = (a_{ij})$. Par définition, on a

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m \quad (j = 1, \ldots, n).$$

On sait que $T^* \in \mathcal{L}_K(W^*, V^*)$ et on cherche

$$[T^*]_{\mathcal{B}^*}^{\mathcal{D}^*} = \Big([T^*(g_1)]_{\mathcal{B}^*} \ \cdots \ [T^*(g_m)]_{\mathcal{B}^*}\Big).$$

Fixons un indice $j \in \{1, \ldots, m\}$. Pour tout $i = 1, \ldots, n$, on a

$$(T^*(g_j))(\mathbf{v}_i) = (g_j \circ T)(\mathbf{v}_i) = g_j(T(\mathbf{v}_i)) = g_j\Big(\sum_{k=1}^{m} a_{ki} w_k\Big) = a_{ji}.$$

On en déduit que

$$T^*(g_j) = \sum_{i=1}^{n} a_{ji} f_i$$

(voir l'exercice 9.1.1). Par conséquent, $[T^*(g_j)]_{\mathcal{B}^*}$ est la transposée de la ligne j de $[T]_{\mathcal{D}}^{\mathcal{B}}$. Cela donne bien

$$[T^*]_{\mathcal{B}^*}^{\mathcal{D}^*} = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} = \big([T]_{\mathcal{D}}^{\mathcal{B}}\big)^t$$

□

Définition 9.1.9. Le *bidual* d'un espace vectoriel V est le dual de V^*. On le note V^{**}.

On a donc $V^{**} = \mathcal{L}_K(V^*, K)$. Le résultat suivant est particulièrement important.

Théorème 9.1.10. *Soit V un espace vectoriel sur K.*

(i) Pour tout $\mathbf{v} \in V$, l'application

$$\begin{array}{rccc} E_{\mathbf{v}} : & V^* & \longrightarrow & K \\ & f & \longmapsto & f(\mathbf{v}) \end{array}$$

est linéaire et par suite $E_{\mathbf{v}} \in V^{**}$.

(ii) L'application

$$\begin{array}{rccc} \varphi : & V & \longrightarrow & V^{**} \\ & \mathbf{v} & \longmapsto & E_{\mathbf{v}} \end{array}$$

est elle-aussi linéaire.

(iii) Si V est de dimension finie, alors φ est un isomorphisme.

Preuve. (i) Soit $\mathbf{v} \in V$. Pour tout choix de $f, g \in V^*$ et $c \in K$, on trouve

$$\begin{aligned} E_{\mathbf{v}}(f+g) &= (f+g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}) = E_{\mathbf{v}}(f) + E_{\mathbf{v}}(g)\,, \\ E_{\mathbf{v}}(c\,f) &= (c\,f)(\mathbf{v}) = c\,f(\mathbf{v}) = c\,E_{\mathbf{v}}(f)\,. \end{aligned}$$

(ii) Soient $\mathbf{u}, \mathbf{v} \in V$ et $c \in K$. Pour tout $f \in V^*$, on trouve

$$\begin{aligned} E_{\mathbf{u}+\mathbf{v}}(f) &= f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}) = E_{\mathbf{u}}(f) + E_{\mathbf{v}}(f) = (E_{\mathbf{u}} + E_{\mathbf{v}})(f)\,, \\ E_{c\,\mathbf{v}}(f) &= f(c\,\mathbf{v}) = c\,f(\mathbf{v}) = c\,E_{\mathbf{v}}(f) = (c\,E_{\mathbf{v}})(f)\,, \end{aligned}$$

donc $E_{\mathbf{u}+\mathbf{v}} = E_{\mathbf{u}} + E_{\mathbf{v}}$ et $E_{c\,\mathbf{v}} = c\,E_{\mathbf{v}}$. Cela montre que φ est linéaire.

(iii) Supposons que V soit de dimension finie. Soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V et soit $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ la base de V^* duale à $\mathcal{B}$. On a

$$E_{\mathbf{v}_j}(f_i) = f_i(\mathbf{v}_j) = \delta_{ij} \quad (1 \leq i, j \leq n).$$

Donc $\{E_{\mathbf{v}_1}, \ldots, E_{\mathbf{v}_n}\}$ est la base de $(V^*)^* = V^{**}$ duale à $\mathcal{B}^*$. Comme φ applique $\mathcal{B}$ sur cette base, l'application linéaire φ est un isomorphisme. □

Exercices.

9.1.1. Soit V un espace vectoriel de dimension finie sur un corps K, soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, et soit $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ la base de V^* duale à $\mathcal{B}$. Montrer que, pour tout $f \in V^*$, on a

$$f = f(\mathbf{v}_1)f_1 + \cdots + f(\mathbf{v}_n)f_n.$$

9.1.2. Démontrer la proposition 9.1.6.

9.1.3. On considère les vecteurs $\mathbf{v}_1 = (1,0,1)^t$, $\mathbf{v}_2 = (0,1,-2)^t$ et $\mathbf{v}_3 = (-1,-1,0)^t$ de $\mathbb{R}^3$.

(i) Si f est une forme linéaire sur $\mathbb{R}^3$ telle que

$$f(\mathbf{v}_1) = 1, \quad f(\mathbf{v}_2) = -1, \quad f(\mathbf{v}_3) = 3,$$

et si $\mathbf{v} = (a,b,c)^t$, déterminer $f(\mathbf{v})$.

(ii) Donner explicitement une forme linéaire f sur $\mathbb{R}^3$ telle que $f(\mathbf{v}_1) = f(\mathbf{v}_2) = 0$ mais $f(\mathbf{v}_3) \neq 0$.

(iii) Soit f une forme linéaire sur $\mathbb{R}^3$ telle que $f(\mathbf{v}_1) = f(\mathbf{v}_2) = 0$ et $f(\mathbf{v}_3) \neq 0$. Montrer que, pour $\mathbf{v} = (2,3,-1)^t$, on a $f(\mathbf{v}) \neq 0$.

9.1.4. On considère les vecteurs $\mathbf{v}_1 = (1,0,-i)^t$, $\mathbf{v}_2 = (1,1,1)^t$ et $\mathbf{v}_3 = (i,i,0)^t$ de $\mathbb{C}^3$. Montrer que $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ est une base de $\mathbb{C}^3$ et déterminer la base $\mathcal{B}^*$ de $(\mathbb{C}^3)^*$ qui est duale à $\mathcal{B}$.

9.1.5. Soit $V = \langle\, 1, x, x^2 \,\rangle_{\mathbb{R}}$ l'espace vectoriel des fonctions polynomiales $p\colon \mathbb{R} \to \mathbb{R}$ de degré au plus 2. On définit trois formes linéaires f_1, f_2, f_3 sur V en posant

$$f_1(p) = \int_0^1 p(x)dx, \quad f_2(p) = \int_0^2 p(x)dx, \quad f_3(p) = \int_{-1}^0 p(x)dx.$$

Montrer que $\mathcal{B} = \{f_1, f_2, f_3\}$ est une base de V^* et déterminer sa base duale dans V.

9.1.6. Soient m, n des entiers positifs, soit K un corps, et soient $f_1, \ldots, f_m$ des formes linéaires sur K^n.

(i) Vérifier que la fonction $T\colon K^n \to K^m$ donnée par

$$T(\mathbf{v}) = \begin{pmatrix} f_1(\mathbf{v}) \\ \vdots \\ f_m(\mathbf{v}) \end{pmatrix} \quad \text{pour tout } \mathbf{v} \in K^n,$$

est une application linéaire.

(ii) Montrer que toute application linéaire de K^n dans K^m s'écrit ainsi pour un choix de $f_1, \ldots, f_m \in (K^n)^*$.

9.1.7. Soient n un entier positif et K un corps. Posons $V = \mathrm{Mat}_{n\times n}(K)$. On rappelle que la *trace* d'une matrice $A = (a_{i,j}) \in V$ est la somme des éléments de sa diagonale : $\mathrm{trace}(A) = \sum_{i=1}^n a_{i,i}$.

(i) Vérifier que l'application trace$\colon V \to K$ est une forme linéaire sur V.

(ii) Soit $B \in V$. Montrer que l'application $\varphi_B\colon V \to K$ donnée par $\varphi_B(A) = \mathrm{trace}(AB)$ pour tout $A \in V$ est linéaire.

(iii) Montrer que l'application $\Phi\colon V \to V^*$ donnée par $\Phi(B) = \varphi_B$ pour tout $B \in V$ est linéaire.

(iv) Établir que $\ker \Phi = \{0\}$ et conclure que Φ est un isomorphisme.

9.2 Applications bilinéaires

Soient U, V et W des espaces vectoriels sur un corps K.

Définition 9.2.1. Une *application bilinéaire* de $U \times V$ dans W est une fonction $B\colon U \times V \to W$ qui satisfait

(i) $B(\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2) = B(\mathbf{u}, \mathbf{v}_1) + B(\mathbf{u}, \mathbf{v}_2)$,

(ii) $B(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) = B(\mathbf{u}_1, \mathbf{v}) + B(\mathbf{u}_2, \mathbf{v})$,

(iii) $B(c\,\mathbf{u}, \mathbf{v}) = B(\mathbf{u}, c\,\mathbf{v}) = c\,B(\mathbf{u}, \mathbf{v})$

pour tout choix de $\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2 \in U$, de $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$ et de $c \in K$. Une *forme bilinéaire* sur $U \times V$ est une application bilinéaire de $U \times V$ dans K.

Ainsi, une application bilinéaire de $U \times V$ dans W est une fonction $B\colon U \times V \to W$ telle que,

1) pour tout $\mathbf{u} \in U$, l'application $\begin{array}{rcl} V & \to & W \\ \mathbf{v} & \mapsto & B(\mathbf{u}, \mathbf{v}) \end{array}$ est linéaire ;

2) pour tout $\mathbf{v} \in V$, l'application $\begin{array}{rcl} U & \to & W \\ \mathbf{u} & \mapsto & B(\mathbf{u}, \mathbf{v}) \end{array}$ est linéaire.

On exprime la condition 1) en disant que $\mathcal{B}$ est *linéaire à droite* et la condition 2) en disant que $\mathcal{B}$ est *linéaire à gauche*.

Exemple 9.2.2. L'application

$$\begin{array}{rccc} B\colon & V \times V^* & \longrightarrow & K \\ & (\mathbf{v}, f) & \longmapsto & f(\mathbf{v}) \end{array}$$

est une forme bilinéaire sur $V \times V^*$ car, pour tout choix de $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$, de $f, f_1, f_2 \in V^*$ et de $c \in K$, on a

(i) $(f_1 + f_2)(\mathbf{v}) = f_1(\mathbf{v}) + f_2(\mathbf{v})$,

(ii) $f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$,

(iii) $f(c\,\mathbf{v}) = c\,f(\mathbf{v}) = (c\,f)(\mathbf{v})$.

Exemple 9.2.3. La proposition 2.3.1 montre que la fonction

$$\begin{array}{rcl} \mathcal{L}_K(U,V) \times \mathcal{L}_K(V,W) & \longrightarrow & \mathcal{L}_K(U,W) \\ (S,T) & \longmapsto & T \circ S \end{array}$$

est une application bilinéaire.

Définition 9.2.4. On désigne par $\mathcal{L}^2_K(U, V; W)$ l'ensemble des applications bilinéaires de $U \times V$ dans W.

Le résultat suivant fournit un analogue du théorème 2.2.3 pour les applications bilinéaires.

Théorème 9.2.5. *Pour tout choix de $B_1, B_2 \in \mathcal{L}^2_K(U,V;W)$ et de $c \in K$, les applications*

$$\begin{array}{rcl} B_1 + B_2 : U \times V & \longrightarrow & W \\ (\mathbf{u},\mathbf{v}) & \longmapsto & B_1(\mathbf{u},\mathbf{v}) + B_2(\mathbf{u},\mathbf{v}) \end{array} \quad \text{et} \quad \begin{array}{rcl} c\,B_1 : U \times V & \longrightarrow & W \\ (\mathbf{u},\mathbf{v}) & \longmapsto & c\,B_1(\mathbf{u},\mathbf{v}) \end{array}$$

sont bilinéaires. L'ensemble $\mathcal{L}^2_K(U,V;W)$ muni de l'addition et de la multiplication scalaire ainsi définies est un espace vectoriel sur K.

La preuve de la première affirmation est semblable à celle de la proposition 2.2.1 tandis que la seconde affirmation se démontre sur le modèle de la preuve du théorème 2.2.3. Ces démonstrations sont laissées en exercice.

Dans l'exercice 9.2.3, on propose de démontrer qu'on a un isomorphisme naturel

$$\begin{array}{rcl} \mathcal{L}^2_K(U,V;W) & \xrightarrow{\sim} & \mathcal{L}_K\big(U, \mathcal{L}_K(V,W)\big) \\ B & \longmapsto & (\mathbf{u} \mapsto (\mathbf{v} \mapsto (B(\mathbf{u},\mathbf{v})))) \end{array}$$

On conclut cette section avec les résultats suivants :

Théorème 9.2.6. *Supposons que U et V admettent des bases*

$$\mathcal{A} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\} \quad \text{et} \quad \mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

respectivement. Alors, pour tout choix d'éléments $\mathbf{w}_{ij} \in W$ $(1 \le i \le m, 1 \le j \le n)$, il existe une et une seule application bilinéaire $B\colon U \times V \to W$ qui satisfait

$$B(\mathbf{u}_i, \mathbf{v}_j) = \mathbf{w}_{ij} \quad (1 \le i \le m, 1 \le j \le n). \tag{9.1}$$

Preuve. Soit $\mathbf{w}_{ij} \in W$ $(1 \le i \le m, 1 \le j \le n)$.

1° Unicité. S'il existe une application bilinéaire $B\colon U \times V \to W$ qui satisfait (9.1), alors

$$\begin{aligned} B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n b_j\mathbf{v}_j\Big) &= \sum_{i=1}^m a_i B\Big(\mathbf{u}_i, \sum_{j=1}^n b_j\mathbf{v}_j\Big) \\ &= \sum_{i=1}^m a_i\Big(\sum_{j=1}^n b_j B(\mathbf{u}_i, \mathbf{v}_j)\Big) \\ &= \sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbf{w}_{ij} \end{aligned}$$

pour tout choix de $a_1, \ldots, a_m, b_1, \ldots, b_n \in K$. Cela montre qu'il existe au plus une telle application bilinéaire.

2° Existence. Considérons la fonction $B\colon U \times V \to W$ définie par

$$B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n b_j\mathbf{v}_j\Big) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbf{w}_{ij}.$$

Elle vérifie les conditions (9.1). Il reste à montrer qu'elle est bilinéaire. On trouve

$$\begin{aligned} B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n b_j\mathbf{v}_j + \sum_{j=1}^n b'_j\mathbf{v}_j\Big) &= B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n (b_j+b'_j)\mathbf{v}_j\Big) \\ &= \sum_{i=1}^m\sum_{j=1}^n a_i(b_j+b'_j)\mathbf{w}_{ij} \\ &= \sum_{i=1}^m\sum_{j=1}^n a_ib_j\mathbf{w}_{ij} + \sum_{i=1}^m\sum_{j=1}^n a_ib'_j\mathbf{w}_{ij} \\ &= B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n b_j\mathbf{v}_j\Big) + B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n b'_j\mathbf{v}_j\Big), \end{aligned}$$

pour tout choix de $a_1, \ldots, a_m, b_1, \ldots, b_n, b'_1, \ldots, b'_n \in K$. On vérifie aussi que

$$B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, c\sum_{j=1}^n b_j\mathbf{v}_j\Big) = c\,B\Big(\sum_{i=1}^m a_i\mathbf{u}_i, \sum_{j=1}^n b_j\mathbf{v}_j\Big),$$

pour tout $c \in K$. Donc B est linéaire à droite. On vérifie de même que B est linéaire à gauche. Donc, B est bilinéaire. □

Théorème 9.2.7. *Supposons que U et V admettent des bases*

$$\mathcal{A} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\} \quad \textit{et} \quad \mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

respectivement. Soit $B\colon U \times V \to K$ une forme bilinéaire sur $U \times V$. Alors il existe une et une seule matrice $M \in \mathrm{Mat}_{m\times n}(K)$ telle que

$$B(\mathbf{u}, \mathbf{v}) = [\mathbf{u}]_{\mathcal{A}}^t M [\mathbf{v}]_{\mathcal{B}} \tag{9.2}$$

pour tout $\mathbf{u} \in U$ et tout $\mathbf{v} \in V$. L'élément de la ligne i et de la colonne j de M est $B(\mathbf{u}_i, \mathbf{v}_j)$.

La matrice M s'appelle la *matrice de la forme bilinéaire B relative aux bases $\mathcal{A}$ de U et $\mathcal{B}$ de V*.

Preuve. **1° Existence.** Soient $\mathbf{u} \in U$ et $\mathbf{v} \in V$. Ils s'écrivent

$$\mathbf{u} = \sum_{i=1}^m a_i\mathbf{u}_i \quad \text{et} \quad \mathbf{v} = \sum_{j=1}^n b_j\mathbf{v}_j$$

avec $a_1, \ldots, a_m, b_1, \ldots, b_n \in K$. On trouve

$$\begin{aligned} B(\mathbf{u}, \mathbf{v}) &= \sum_{i=1}^m\sum_{j=1}^n a_ib_jB(\mathbf{u}_i, \mathbf{v}_j) \\ &= (a_1, \ldots, a_m) \begin{pmatrix} B(\mathbf{u}_1, \mathbf{v}_1) & \cdots & B(\mathbf{u}_1, \mathbf{v}_n) \\ \vdots & & \vdots \\ B(\mathbf{u}_m, \mathbf{v}_1) & \cdots & B(\mathbf{u}_m, \mathbf{v}_n) \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \\ &= [\mathbf{u}]_{\mathcal{A}}^t \Big(B(\mathbf{u}_i, \mathbf{v}_j)\Big) [\mathbf{v}]_{\mathcal{B}}\,. \end{aligned}$$

2° Unicité. Supposons que $M \in \mathrm{Mat}_{m\times n}(K)$ satisfasse la condition (9.2) pour tout $\mathbf{u} \in U$ et tout $\mathbf{v} \in V$. Pour $i = 1, \ldots, m$ et $j = 1, \ldots, n$, on trouve

$$\begin{aligned} B(\mathbf{u}_i, \mathbf{v}_j) &= [\mathbf{u}_i]_{\mathcal{A}}^t M [\mathbf{v}_j]_{\mathcal{B}} = (0, \ldots, \overset{\overset{i}{\vee}}{1}, \ldots, 0) M \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j \\ &= \text{élément d'indice } (i,j) \text{ de } M. \end{aligned}$$

□

Exemple 9.2.8. Supposons que $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ soit une base de V. Soit $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ la base duale de V^*. On se propose de calculer la matrice de la forme bilinéaire

$$\begin{array}{rccc} B: & V \times V^* & \longrightarrow & K \\ & (\mathbf{v}, f) & \longmapsto & f(\mathbf{v}) \end{array}$$

relativement à ces bases.

Solution: On a $B(\mathbf{v}_i, f_j) = f_j(\mathbf{v}_i) = \delta_{ij}$ pour $1 \leq i, j \leq n$, donc la matrice de B est $(\delta_{ij}) = I_n$, la matrice identité $n \times n$.

Exercices.

9.2.1. Soient V et W des espaces vectoriels sur un corps K et soit $\mathbf{v} \in V$. Montrer que l'application

$$\begin{array}{rccc} B: & V^* \times W & \longrightarrow & W \\ & (f, \mathbf{w}) & \longmapsto & f(\mathbf{v})\mathbf{w} \end{array}$$

est bilinéaire.

9.2.2. Soient T, U, V et W des espaces vectoriels sur un corps K, soit $B\colon U \times V \to T$ une application bilinéaire et soit $L\colon T \to W$ une application linéaire. Montrer que la composée

$$L \circ B\colon U \times V \longrightarrow W$$

est une application bilinéaire.

9.2.3. Soient U, V, W des espaces vectoriels sur un corps K. Pour tout $B \in \mathcal{L}_K^2(U, V; W)$ et tout $\mathbf{u} \in U$, on désigne par $B_{\mathbf{u}}\colon V \to W$ l'application linéaire donnée par

$$B_{\mathbf{u}}(\mathbf{v}) = B(\mathbf{u}, \mathbf{v}) \quad (\mathbf{v} \in V).$$

(i) Soit $B \in \mathcal{L}^2_K(U,V;W)$. Vérifier que l'application

$$\begin{array}{rccc} \varphi_B : & U & \longrightarrow & \mathcal{L}_K(V;W) \\ & \mathbf{u} & \longmapsto & B_{\mathbf{u}} \end{array}$$

est linéaire.

(ii) Montrer l'application

$$\begin{array}{rccc} \varphi : & \mathcal{L}^2_K(U,V;W) & \longrightarrow & \mathcal{L}_K(U,\mathcal{L}_K(V;W)) \\ & B & \longmapsto & \varphi_B \end{array}$$

est un isomorphisme d'espaces vectoriels sur K.

9.2.4. Soient $\{\mathbf{e}_1,\mathbf{e}_2\}$ la base canonique de $\mathbb{R}^2$ et $\{\mathbf{e}'_1,\mathbf{e}'_2,\mathbf{e}'_3\}$ celle de $\mathbb{R}^3$. Donner une formule pour $f((x_1,x_2,x_3),(y_1,y_2))$ où $f\colon \mathbb{R}^3\times\mathbb{R}^2\to\mathbb{R}$ désigne l'application bilinéaire déterminée par les conditions

$$\begin{array}{lll} f(\mathbf{e}'_1,\mathbf{e}_1)=1, & f(\mathbf{e}'_1,\mathbf{e}_2)=0, & f(\mathbf{e}'_2,\mathbf{e}_1)=3, \\ f(\mathbf{e}'_2,\mathbf{e}_2)=-1, & f(\mathbf{e}'_3,\mathbf{e}_1)=4, & f(\mathbf{e}'_3,\mathbf{e}_2)=-2. \end{array}$$

En déduire la valeur de $f(\mathbf{u},\mathbf{v})$ si $\mathbf{u}=(1,2,1)^t$ et $\mathbf{v}=(1,0)^t$.

9.2.5. Soient U, V et W des espaces vectoriels sur un corps K, soient $B_i\colon U\times V\to W$ $(i=1,2)$ des applications bilinéaires, et soit $c\in K$. Montrer que les applications

$$\begin{array}{rccc} B_1+B_2\colon U\times V & \longrightarrow & W \\ (\mathbf{u},\mathbf{v}) & \longmapsto & B_1(\mathbf{u},\mathbf{v})+B_2(\mathbf{u},\mathbf{v}) \end{array} \qquad \text{et} \qquad \begin{array}{rccc} cB_1\colon U\times V & \longrightarrow & W \\ (\mathbf{u},\mathbf{v}) & \longmapsto & cB_1(\mathbf{u},\mathbf{v}) \end{array}$$

sont bilinéaires. Montrer ensuite que $\mathcal{L}^2_K(U,V\,;\,W)$ est un espace vectoriel pour ces opérations (comme il s'agit d'une vérification assez longue, vous pouvez vous contenter simplement de de montrer l'existence d'un élément neutre pour l'addition, de montrer que tout élément possède un inverse additif, et de démontrer un des deux axiomes de distributivité).

9.2.6. Soit V un espace vectoriel de dimension finie n sur un corps K, et soit $\mathcal{B}=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ une base de V. Pour toute paire d'entiers $i,j\in\{1,\ldots,n\}$, on désigne par $g_{i,j}\colon V\times V\to K$ la forme bilinéaire déterminée par les conditions

$$g_{i,j}(\mathbf{v}_k,\mathbf{v}_\ell)=\begin{cases}1 & \text{si } (k,\ell)=(i,j),\\ 0 & \text{sinon.}\end{cases}$$

Montrer que $\mathcal{B}=\{\,g_{i,j}\,;\,1\le i,j\le n\,\}$ est une base de l'espace vectoriel $\mathrm{Bil}(V,K):=\mathcal{L}^2_K(V,V\,;\,K)$ de toutes les applications bilinéaires de $V\times V$ dans K.

9.2.7. Soit V un espace vectoriel sur un corps K de caractéristique différente de 2. On dit qu'une forme bilinéaire $f\colon V\times V\to K$ est *symmétrique* si $f(\mathbf{u},\mathbf{v})=f(\mathbf{v},\mathbf{u})$ pour tout choix de $\mathbf{u}\in V$ et $\mathbf{v}\in V$. On dit qu'elle est *anti-symmétrique* si $f(\mathbf{u},\mathbf{v})=-f(\mathbf{v},\mathbf{u})$ quels que soient $\mathbf{u}\in V$ et $\mathbf{v}\in V$.

(i) Montrer que l'ensemble S des formes bilinéaires symmétriques sur $V \times V$ est un sous-espace de $\mathrm{Bil}(V,K) := \mathcal{L}^2_K(V,V\,;\,K)$.
(ii) Montrer que l'ensemble A des formes bilinéaires anti-symmétriques sur $V \times V$ est aussi un sous-espace de $\mathrm{Bil}(V,K)$.
(iii) Montrer que $\mathrm{Bil}(V,K) = S \oplus A$.

9.2.8. Déterminer une base de l'espace des formes bilinéaires symmétriques de $\mathbb{R}^2 \times \mathbb{R}^2$ dans $\mathbb{R}$ et une base de l'espace des formes bilinéaires anti-symmétriques de $\mathbb{R}^2 \times \mathbb{R}^2$ dans $\mathbb{R}$.

9.2.9. Soit $\mathcal{C}(\mathbb{R})$ l'espace vectoriel des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$, et soit V le sous-espace de $\mathcal{C}(\mathbb{R})$ engendré (sur $\mathbb{R}$) par les fonctions 1, $\sin(x)$ et $\cos(x)$. Montrer que l'application

$$\begin{array}{rl} B\colon V \times V & \longrightarrow \mathbb{R} \\ (f,g) & \longmapsto \int_0^\pi f(t)g(\pi - t)dt \end{array}$$

est une forme bilinéaire et déterminer sa matrice par rapport à la base $\mathcal{A} = \{1, \sin(x), \cos(x)\}$ de V.

9.2.10. Soit $V = \langle\, 1, x, \ldots, x^n \,\rangle_{\mathbb{R}}$ l'espace vectoriel des fonctions polynomiales $p\colon \mathbb{R} \to \mathbb{R}$ de degré au plus n. Déterminer la matrice de l'application bilinéaire

$$\begin{array}{rl} B\colon V \times V & \longrightarrow \mathbb{R} \\ (f,g) & \longmapsto \int_0^1 f(t)g(t)dt \end{array}$$

par rapport à la base $\mathcal{A} = \{1, x, \ldots, x^n\}$ de V.

9.3 Produit tensoriel

Tout au long de cette section, on fixe deux espaces vectoriels U et V de dimension finie sur un corps K.

Si T est un espace vectoriel, on peut considérer une application bilinéaire de $U \times V$ dans T comme un produit

$$\begin{array}{rcl} U \times V & \longrightarrow & T \\ (\mathbf{u},\mathbf{v}) & \longmapsto & \mathbf{u} \otimes \mathbf{v} \end{array} \tag{9.3}$$

à valeurs dans T. Dire que cette application est bilinéaire signifie que le produit en question satisfait :

$$\begin{aligned} (\mathbf{u}_1 + \mathbf{u}_2) \otimes \mathbf{v} &= \mathbf{u}_1 \otimes \mathbf{v} + \mathbf{u}_2 \otimes \mathbf{v}, \\ \mathbf{u} \otimes (\mathbf{v}_1 + \mathbf{v}_2) &= \mathbf{u} \otimes \mathbf{v}_1 + \mathbf{u} \otimes \mathbf{v}_2, \\ (c\,\mathbf{u}) \otimes \mathbf{v} &= \mathbf{u} \otimes (c\,\mathbf{v}) = c\,\mathbf{u} \otimes \mathbf{v}, \end{aligned}$$

pour tout choix de $\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2 \in U$, de $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$ et de $c \in K$.

Étant donné une application bilinéaire comme ci-dessus (9.3), on vérifie sans peine que, pour toute application linéaire $L\colon T \to W$, la composée $B\colon U \times V \to W$ donnée par

$$B(\mathbf{u}, \mathbf{v}) = L(\mathbf{u} \otimes \mathbf{v}) \quad \text{pour tout } (\mathbf{u}, \mathbf{v}) \in U \times V \tag{9.4}$$

est une application bilinéaire (voir l'exercice 9.2.2).

Notre but est de construire une application bilinéaire (9.3) qui soit *universelle* au sens où toute application bilinéaire $B\colon U \times V \to W$ s'écrive sous la forme (9.4) pour un et un seul choix d'application linéaire $L\colon T \to W$. Une application bilinéaire (9.3) qui possède cette propriété s'appelle un *produit tensoriel* de U et de V. La définition suivante résume cette discussion.

Définition 9.3.1. Un *produit tensoriel* de U et de V est la donnée d'un espace vectoriel T sur K et d'une application bilinéaire

$$\begin{array}{ccc} U \times V & \longrightarrow & T \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \mathbf{u} \otimes \mathbf{v} \end{array}$$

qui possède la propriété suivante. Pour tout espace vectoriel W et toute application bilinéaire $B\colon U \times V \to W$, il existe une et une seule application linéaire $L\colon T \to W$ telle que $B(\mathbf{u}, \mathbf{v}) = L(\mathbf{u} \otimes \mathbf{v})$ pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$.

Cette condition est schématisée par le diagramme suivant :

$$\begin{array}{ccc} U \times V & \xrightarrow{\quad B \quad} & W \\ {\scriptstyle\otimes}\big\downarrow & \ddots^{\urcorner} & \\ T & \ddots^{L} & \end{array} \tag{9.5}$$

Dans ce schéma, la flèche verticale représente le produit tensoriel, la flèche horizontale représente une application bilinéaire quelconque, et la flèche diagonale en pointillé représente l'unique application linéaire L qui, par composition avec le produit tensoriel, redonne B. Pour un tel choix de L, la flèche horizontale B est la composée des deux autres flèches et on dit que le diagramme (9.5) est *commutatif*.

La proposition suivante démontre l'existence d'un produit tensoriel pour U et V.

Proposition 9.3.2. *Soit $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ une base de U, soit $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, et soit $\{\mathbf{t}_{ij}\,;\, 1 \leq i \leq m, 1 \leq j \leq n\}$ une base d'un espace vectoriel T de dimension mn sur K. L'application bilinéaire (9.3) déterminée par les conditions*

$$\mathbf{u}_i \otimes \mathbf{v}_j = \mathbf{t}_{ij} \quad (1 \leq i \leq m, 1 \leq j \leq n) \tag{9.6}$$

est un produit tensoriel de U et de V.

Preuve. Le théorème 9.2.6 montre qu'il existe une application bilinéaire $(\mathbf{u},\mathbf{v}) \longmapsto \mathbf{u}\otimes\mathbf{v}$ de $U\times V$ dans T qui satisfait (9.6). Il reste à voir qu'elle possède la propriété universelle requise par la définition 9.3.1.

Pour ce faire, on fixe une application bilinéaire quelconque $B\colon U\times V\to W$. Puisque l'ensemble $\{\mathbf{t}_{ij}\,;\,1\le i\le m, 1\le j\le n\}$ est une base de T, il existe une et une seule application linéaire $L\colon T\to W$ telle que

$$L(\mathbf{t}_{ij}) = B(\mathbf{u}_i,\mathbf{v}_j) \quad (1\le i\le m, 1\le j\le n) \tag{9.7}$$

Alors l'application $B'\colon U\times V\to W$ donnée par $B'(\mathbf{u},\mathbf{v}) = L(\mathbf{u}\otimes\mathbf{v})$ pour tout $(\mathbf{u},\mathbf{v})\in U\times V$ est bilinéaire et satisfait

$$B'(\mathbf{u}_i,\mathbf{v}_j) = L(\mathbf{u}_i\otimes\mathbf{v}_j) = L(\mathbf{t}_{ij}) = B(\mathbf{u}_i,\mathbf{v}_j) \quad (1\le i\le m,\ 1\le j\le n).$$

En vertu du théorème 9.2.6, cela implique que $B' = B$. Donc l'application L remplie la condition

$$B(\mathbf{u},\mathbf{v}) = L(\mathbf{u}\otimes\mathbf{v}) \quad \text{pour tout } (\mathbf{u},\mathbf{v})\in U\times V.$$

Enfin, L est la seule application linéaire de T dans W qui remplisse cette condition car cette dernière implique (9.7) laquelle détermine L de manière unique. □

La preuve donnée ci-dessus utilise de manière cruciale l'hypothèse que U et V sont de dimension finie sur K. On peut toutefois démontrer l'existence du produit tensoriel de manière générale sans faire cette hypothèse, mais nous ne le ferons pas ici.

On observe ensuite que le produit tensoriel est unique à isomorphisme près.

Proposition 9.3.3. *Supposons que*

$$\begin{array}{ccc} U\times V & \longrightarrow & T \\ (\mathbf{u},\mathbf{v}) & \longmapsto & \mathbf{u}\otimes\mathbf{v} \end{array} \qquad \textit{et} \qquad \begin{array}{ccc} U\times V & \longrightarrow & T' \\ (\mathbf{u},\mathbf{v}) & \longmapsto & \mathbf{u}\otimes'\mathbf{v} \end{array}$$

soient deux produits tensoriels de U et de V. Alors il existe un isomorphisme $L\colon T\to T'$ tel que

$$L(\mathbf{u}\otimes\mathbf{v}) = \mathbf{u}\otimes'\mathbf{v}$$

pour tout $(\mathbf{u},\mathbf{v})\in U\times V$.

Preuve. Puisque $\otimes'$ est bilineéaire et que $\otimes$ est un produit tensoriel, il existe une application linéaire $L\colon T\to T'$ telle que

$$\mathbf{u}\otimes'\mathbf{v} = L(\mathbf{u}\otimes\mathbf{v}) \tag{9.8}$$

pour tout $(\mathbf{u},\mathbf{v})\in U\times V$. Il reste à montrer que L est un isomorphisme. Pour cela, on note d'abord que, symétriquement, puisque $\otimes$ est bilinéaire et que $\otimes'$ est un produit tensoriel, il existe une application linéaire $L'\colon T'\to T$ telle que

$$\mathbf{u}\otimes\mathbf{v} = L'(\mathbf{u}\otimes'\mathbf{v}) \tag{9.9}$$

pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$. En combinant (9.8) et (9.9), on obtient

$$(L' \circ L)(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \otimes \mathbf{v} \quad \text{et} \quad (L \circ L')(\mathbf{u} \otimes' \mathbf{v}) = \mathbf{u} \otimes' \mathbf{v}$$

pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$. Or les applications identités $I\colon T \to T$ et $I'\colon T' \to T'$ satisfont aussi

$$I(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \otimes \mathbf{v} \quad \text{et} \quad I'(\mathbf{u} \otimes' \mathbf{v}) = \mathbf{u} \otimes' \mathbf{v}$$

pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$. Comme $\otimes$ et $\otimes'$ sont des produits tensoriels, cela implique que $L' \circ L = I$ et que $L \circ L' = I'$. Donc L est une application linéaire inversible d'inverse L', et par suite c'est un isomorphisme. □

En combinant les propositions 9.3.2 et 9.3.3, on obtient aussitôt.

Proposition 9.3.4. *Soit $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ une base de U, soit $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ une base de V, et soit $\otimes\colon U \times V \to T$ un produit tensoriel de U et de V. Alors*

$$\{\mathbf{u}_i \otimes \mathbf{v}_j \,;\, 1 \leq i \leq m, 1 \leq j \leq n\} \tag{9.10}$$

est une base de T.

Preuve. La proposition 9.3.2 montre qu'il existe un produit tensoriel $\otimes\colon U \times V \to T$ tel que (9.10) soit une base de T. Si $\otimes'\colon U \times V \to T'$ est un autre produit tensoriel, alors, selon la proposition 9.3.3, il existe un isomorphisme $L\colon T \to T'$ tel que $\mathbf{u} \otimes' \mathbf{v} = L(\mathbf{u} \otimes \mathbf{v})$ pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$. Comme (9.10) est une base de T, on en déduit que les produits

$$\mathbf{u}_i \otimes' \mathbf{v}_j = L(\mathbf{u}_i \otimes \mathbf{v}_j) \quad (1 \leq i \leq m,\ 1 \leq j \leq n)$$

constituent une base de T'. □

Le théorème suivant résume l'essentiel des résultats obtenus jusqu'à présent :

Théorème 9.3.5. *Il existe un produit tensoriel $\otimes\colon U \times V \to T$ de U et de V. Si $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ est une base de U et $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ une base de V, alors $\{\mathbf{u}_i \otimes \mathbf{v}_j \,;\, 1 \leq i \leq m, 1 \leq j \leq n\}$ est une base de T.*

On dit souvent que T lui-même est le produit tensoriel de U et de V et on le note $U \otimes V$ ou plus précisément $U \otimes_K V$ lorsqu'on veut insister sur le fait que U et V sont considérés comme espaces vectoriels sur K. Le théorème 9.3.5 donne aussitôt

$$\dim_K(U \otimes_K V) = \dim_K(U) \dim_K(V).$$

Le fait que le produit tensoriel ne soit pas unique, mais unique à isomorphisme près, ne pose pas de problèmes en pratique, dans la mesure où les relations de dépendance linéaire entre les produits $\mathbf{u} \otimes \mathbf{v}$ avec $\mathbf{u} \in U$ et $\mathbf{v} \in V$ ne dépendent pas du produit tensoriel $\otimes$. En effet si, pour un produit tensoriel donné $\otimes\colon U \times V \to T$, on a :

$$c_1\mathbf{u}_1 \otimes \mathbf{v}_1 + c_2\mathbf{u}_2 \otimes \mathbf{v}_2 + \cdots + c_s\mathbf{u}_s \otimes \mathbf{v}_s = \mathbf{0} \tag{9.11}$$

avec $c_1, \ldots, c_s \in K$, $\mathbf{u}_1, \ldots, \mathbf{u}_s \in U$ et $\mathbf{v}_1, \ldots, \mathbf{v}_s \in V$, et si $\otimes' \colon U \times V \to T'$ est un autre produit tensoriel, alors en désignant par $L \colon T \to T'$ l'isomorphisme tel que $L(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \otimes' \mathbf{v}$ pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$ et en appliquant L aux deux membres de (9.11), on obtient

$$c_1\mathbf{u}_1 \otimes' \mathbf{v}_1 + c_2\mathbf{u}_2 \otimes' \mathbf{v}_2 + \cdots + c_s\mathbf{u}_s \otimes' \mathbf{v}_s = \mathbf{0}.$$

Exemple 9.3.6. Soit $\{\mathbf{e}_1, \mathbf{e}_2\}$ la base canonique de K^2. Une base de $K^2 \otimes_K K^2$ est

$$\{\mathbf{e}_1 \otimes \mathbf{e}_1,\ \mathbf{e}_1 \otimes \mathbf{e}_2,\ \mathbf{e}_2 \otimes \mathbf{e}_1,\ \mathbf{e}_2 \otimes \mathbf{e}_2\}.$$

En particulier, on a $\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1 \neq \mathbf{0}$, donc $\mathbf{e}_1 \otimes \mathbf{e}_2 \neq \mathbf{e}_2 \otimes \mathbf{e}_1$. On trouve aussi

$$\begin{aligned}(\mathbf{e}_1 + \mathbf{e}_2) \otimes (\mathbf{e}_1 - \mathbf{e}_2) &= \mathbf{e}_1 \otimes (\mathbf{e}_1 - \mathbf{e}_2) + \mathbf{e}_2 \otimes (\mathbf{e}_1 - \mathbf{e}_2) \\ &= \mathbf{e}_1 \otimes \mathbf{e}_1 - \mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1 - \mathbf{e}_2 \otimes \mathbf{e}_2.\end{aligned}$$

Plus généralement, on a

$$\begin{aligned}(a_1\mathbf{e}_1 + a_2\mathbf{e}_2) \otimes (b_1\mathbf{e}_1 + b_2\mathbf{e}_2) &= \Big(\sum_{i=1}^{2} a_i\mathbf{e}_i\Big) \otimes \Big(\sum_{j=1}^{2} b_j\mathbf{e}_j\Big) \\ &= \sum_{i=1}^{2}\sum_{j=1}^{2} a_ib_j\mathbf{e}_i \otimes \mathbf{e}_j \\ &= a_1b_1\mathbf{e}_1 \otimes \mathbf{e}_1 + a_1b_2\mathbf{e}_1 \otimes \mathbf{e}_2 + a_2b_1\mathbf{e}_2 \otimes \mathbf{e}_1 + a_2b_2\mathbf{e}_2 \otimes \mathbf{e}_2.\end{aligned}$$

Exemple 9.3.7. On sait que le produit scalaire dans $\mathbb{R}^n$ est une forme bilinéaire

$$\begin{array}{ccc}\mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R}\ . \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \langle \mathbf{u}, \mathbf{v} \rangle\end{array}$$

Donc il existe une et une seule application linéaire $L \colon \mathbb{R}^n \otimes_{\mathbb{R}} \mathbb{R}^n \to \mathbb{R}$ telle que

$$\langle \mathbf{u}, \mathbf{v} \rangle = L(\mathbf{u} \otimes \mathbf{v})$$

pour tout choix de $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Soit $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ la base canonique de $\mathbb{R}^n$. Alors l'ensemble $\{\mathbf{e}_i \otimes \mathbf{e}_j\,;\, 1 \leq i \leq n, 1 \leq j \leq n\}$ est une base de $\mathbb{R}^n \otimes_{\mathbb{R}} \mathbb{R}^n$ et par suite L est déterminée par les conditions

$$L(\mathbf{e}_i \otimes \mathbf{e}_j) = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} \quad (1 \leq i \leq n, 1 \leq j \leq n).$$

Exercices.

9.3.1. Soit $\{\mathbf{e}_1, \mathbf{e}_2\}$ la base canonique de $\mathbb{R}^2$. Montrer que $\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1$ ne peut pas s'écrire sous la forme $\mathbf{u} \otimes \mathbf{v}$ avec $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$.

9.3.2. Soit $\{\mathbf{e}_1, \mathbf{e}_2\}$ la base canonique de $\mathbb{R}^2$, soit $B\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ l'application bilinéaire déterminée par les conditions

$$B(\mathbf{e}_1, \mathbf{e}_1) = (1, 0), \quad B(\mathbf{e}_1, \mathbf{e}_2) = (2, 1), \quad B(\mathbf{e}_2, \mathbf{e}_1) = (-1, 3) \quad \text{et} \quad B(\mathbf{e}_2, \mathbf{e}_2) = (0, -1),$$

et soit $L\colon \mathbb{R}^2 \otimes_{\mathbb{R}} \mathbb{R}^2 \to \mathbb{R}^2$ l'application linéaire telle que $B(\mathbf{u}, \mathbf{v}) = L(\mathbf{u} \otimes \mathbf{v})$ pour tout choix de $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$. Calculer

(i) $L(\mathbf{e}_1 \otimes \mathbf{e}_2 - \mathbf{e}_2 \otimes \mathbf{e}_1)$,

(ii) $L((\mathbf{e}_1 + 2\mathbf{e}_2) \otimes (-\mathbf{e}_1 + \mathbf{e}_2))$,

(iii) $L(\mathbf{e}_1 \otimes \mathbf{e}_1 + (\mathbf{e}_1 + \mathbf{e}_2) \otimes \mathbf{e}_2)$.

9.3.3. Soit V un espace vectoriel de dimension finie sur un corps K. Montrer qu'il existe une et une seule forme linéaire $L\colon V^* \otimes_K V \to K$ qui satisfait la condition $L(f \otimes \mathbf{v}) = f(\mathbf{v})$ pour tout $f \in V^*$ et tout $\mathbf{v} \in V$.

9.3.4. Soient U et V des espaces vectoriels de dimension finie sur un corps K et soit $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ une base de U. Montrer que tout élément $\mathbf{t}$ de $U \otimes_K V$ s'écrit sous la forme

$$\mathbf{t} = \mathbf{u}_1 \otimes \mathbf{v}_1 + \mathbf{u}_2 \otimes \mathbf{v}_2 + \cdots + \mathbf{u}_m \otimes \mathbf{v}_m$$

pour un et un seul choix de vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in V$.

9.3.5. Soient U, V des espaces vectoriels de dimension finie sur un corps K.

(i) Soient U_1 et V_1 des sous-espaces de U et V respectivement et soit T_1 le sous-espace de $U \otimes_K V$ engendré par les produits $\mathbf{u} \otimes \mathbf{v}$ avec $\mathbf{u} \in U_1$ et $\mathbf{v} \in V_1$. Montrer que l'application

$$\begin{array}{rcl} U_1 \times V_1 & \longrightarrow & T_1 \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \mathbf{u} \otimes \mathbf{v} \end{array}$$

est un produit tensoriel de U_1 et de V_1. On peut donc identifier T_1 à $U_1 \otimes_K V_1$.

(ii) Supposons que $U = U_1 \oplus U_2$ et que $V = V_1 \oplus V_2$. Montrer que

$$U \otimes_K V = (U_1 \otimes_K V_1) \oplus (U_2 \otimes_K V_1) \oplus (U_1 \otimes_K V_2) \oplus (U_2 \otimes_K V_2).$$

9.3.6. Soit V un espace vectoriel de dimension finie sur $\mathbb{R}$. En considérant $\mathbb{C}$ comme un espace vectoriel sur $\mathbb{R}$, on forme le produit tensoriel $\mathbb{C} \otimes_{\mathbb{R}} V$.

(i) Montrer que, pour tout $\alpha \in \mathbb{C}$, il existe une application linéaire $L_\alpha\colon \mathbb{C} \otimes_{\mathbb{R}} V \to \mathbb{C} \otimes_{\mathbb{R}} V$ unique qui satisfait $L_\alpha(\beta \otimes \mathbf{v}) = (\alpha\beta) \otimes \mathbf{v}$ quels que soient $\beta \in \mathbb{C}$ et $\mathbf{v} \in V$.

(ii) Montrer que, pour tout choix de $\alpha, \alpha' \in \mathbb{C}$, on a $L_\alpha \circ L_{\alpha'} = L_{\alpha\alpha'}$ et $L_\alpha + L_{\alpha'} = L_{\alpha+\alpha'}$. En déduire que $\mathbb{C} \otimes_{\mathbb{R}} V$ est un espace vectoriel sur $\mathbb{C}$ pour la multiplication externe donnée par $\alpha\mathbf{t} = L_\alpha(\mathbf{t})$ pour tout $\alpha \in \mathbb{C}$ et tout $\mathbf{v} \in V$.

(iii) Montrer que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est une base de V en tant qu'espace vectoriel sur $\mathbb{R}$, alors $\{1 \otimes \mathbf{v}_1, \ldots, 1 \otimes \mathbf{v}_n\}$ est une base de $\mathbb{C} \otimes_{\mathbb{R}} V$ en tant qu'espace vectoriel sur $\mathbb{C}$.

On dit que $\mathbb{C} \otimes_{\mathbb{R}} V$ est le *complexifié* de V.

9.3.7. Avec les notations du problème précédent, montrer que tout élément $\mathbf{t}$ de $\mathbb{C} \otimes_{\mathbb{R}} V$ s'écrit sous la forme $\mathbf{t} = 1 \otimes \mathbf{u} + i \otimes \mathbf{v}$ pour un et un seul choix de $\mathbf{u}, \mathbf{v} \in V$, de même que tout nombre complexe $\alpha \in \mathbb{C}$ s'écrit sous la forme $\alpha = a + ib$ pour un et un seul choix de $a, b \in \mathbb{R}$. Étant donné α et $\mathbf{t}$ comme ci-dessus, déterminer la décomposition du produit $\alpha\mathbf{t}$.

9.3.8. Soient V et W des espaces vectoriels de dimension finie sur un corps K.

(i) Vérifier que, pour tout $f \in V^*$ et tout $\mathbf{w} \in W$, l'application

$$\begin{array}{rccc} L_{f,\mathbf{w}} \colon & V & \longrightarrow & W \\ & \mathbf{v} & \longmapsto & f(\mathbf{v})\mathbf{w} \end{array}$$

est linéaire.

(ii) Vérifier aussi que l'application

$$\begin{array}{rccc} & V^* \times W & \longrightarrow & \mathcal{L}_K(V,W) \\ & (f, \mathbf{w}) & \longmapsto & L_{f,\mathbf{w}} \end{array}$$

est bilinéaire.

(iii) En déduire qu'il existe un et un seul isomorphisme $L \colon V^* \otimes_K W \xrightarrow{\sim} \mathcal{L}_K(V,W)$ satisfaisant $L(f \otimes \mathbf{w}) = L_{f,\mathbf{w}}$ quels que soient $f \in V^*$ et $\mathbf{w} \in W$.

9.3.9. Avec les notations du problème précédent, soit $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ une base de V, soit $\{f_1, \ldots, f_n\}$ la base de V^* duale à $\mathcal{B}$, et soit $\mathcal{D} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ une base de W. Fixons $T \in \mathcal{L}_K(V,W)$ et posons $[T]_{\mathcal{D}}^{\mathcal{B}} = (a_{ij})$. Montrer que, sous l'isomorphisme L de la partie (iii) du problème précédent, on a

$$T \,=\, L\Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} f_j \otimes \mathbf{w}_i\Big).$$

9.3.10. Soit K un corps de caractéristique différente de 2, et soit V un espace vectoriel sur K de dimension n avec base $\mathcal{A} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Pour chaque paire d'éléments $\mathbf{u}$, $\mathbf{v}$ de V, on définit

$$\mathbf{u} \wedge \mathbf{v} := \frac{1}{2}(\mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u}) \quad \text{et} \quad \mathbf{u} \cdot \mathbf{v} := \frac{1}{2}(\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}).$$

Soient $\bigwedge^2 V$ le sous-espace de $V \otimes_K V$ engendré par les vecteurs $\mathbf{u} \wedge \mathbf{v}$, et $\mathrm{Sym}^2(V)$ celui engendré par les vecteurs $\mathbf{u} \cdot \mathbf{v}$.

(i) Montrer que l'ensemble $\mathcal{B}_1 = \{\mathbf{v}_i \wedge \mathbf{v}_j \,;\, 1 \leq i < j \leq n\}$ constitue une base de $\bigwedge^2 V$ tandis que $\mathcal{B}_2 = \{\mathbf{v}_i \cdot \mathbf{v}_j \,;\, 1 \leq i \leq j \leq n\}$ est une base de $\mathrm{Sym}^2(V)$. En déduire que $V \otimes_K V = \bigwedge^2 V \oplus \mathrm{Sym}^2(V)$.

(ii) Montrer que l'application $\varphi_1 \colon V \times V \to \bigwedge^2 V$ donnée par $\varphi_1(\mathbf{u}, \mathbf{v}) = \mathbf{u} \wedge \mathbf{v}$ est bilinéaire anti-symétrique et que l'application $\varphi_2 \colon V \times V \to \mathrm{Sym}^2(V)$ donnée par $\varphi_2(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ est bilinéaire symétrique (voir l'exercice 9.2.7 pour les définitions).

(iii) Soit $B \colon V \times V \to W$ une application bilinéaire anti-symétrique quelconque. Montrer qu'il existe une et une seule application linéaire $L_1 \colon \bigwedge^2 V \to W$ telle que $B = L_1 \circ \varphi_1$.

(iv) Énoncer et démontrer l'analogue de (iii) pour une application bilinéaire symétrique.

Indice. Pour (iii), soit $L \colon V \otimes V \to W$ l'application linéaire qui satisfait $B(\mathbf{u}, \mathbf{v}) = L(\mathbf{u} \otimes \mathbf{v})$ quels que soient $\mathbf{u}, \mathbf{v} \in V$. Montrer que $\mathrm{Sym}^2(V) \subseteq \ker(L)$ et que $B(\mathbf{u}, \mathbf{v}) = L(\mathbf{u} \wedge \mathbf{v})$ pour tout choix de $\mathbf{u}, \mathbf{v} \in V$.

9.4 Produit de Kronecker

Dans cette section, on fixe des espaces vectoriels U et V de dimension finie sur K et on fixe un produit tensoriel

$$\begin{array}{rcl} U \times V & \longrightarrow & U \otimes_K V\,. \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \mathbf{u} \otimes \mathbf{v} \end{array}$$

On note d'abord :

Proposition 9.4.1. *Soient $S \in \mathrm{End}_K(U)$ et $T \in \mathrm{End}_K(V)$. Alors il existe un et seul élément de $\mathrm{End}_K(U \otimes_K V)$, noté $S \otimes T$, tel que*

$$(S \otimes T)(\mathbf{u} \otimes \mathbf{v}) = S(\mathbf{u}) \otimes T(\mathbf{v})$$

pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$.

Preuve. L'application

$$\begin{array}{rccl} B\,: & U \times V & \longrightarrow & U \otimes_K V \\ & (\mathbf{u}, \mathbf{v}) & \longmapsto & S(\mathbf{u}) \otimes T(\mathbf{v}) \end{array}$$

est linéaire à gauche car

$$\begin{aligned} B(\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}) &= S(\mathbf{u}_2 + \mathbf{u}_2) \otimes T(\mathbf{v}) \\ &= (S(\mathbf{u}_1) + S(\mathbf{u}_2)) \otimes T(\mathbf{v}) \\ &= S(\mathbf{u}_1) \otimes T(\mathbf{v}) + S(\mathbf{u}_2) \otimes T(\mathbf{v}) \\ &= B(\mathbf{u}_1, \mathbf{v}) + B(\mathbf{u}_2, \mathbf{v}) \end{aligned} \qquad \text{et} \qquad \begin{aligned} B(c\,\mathbf{u}, \mathbf{v}) &= S(c\,\mathbf{u}) \otimes T(\mathbf{v}) \\ &= (c\,S(\mathbf{u})) \otimes T(\mathbf{v}) \\ &= c\,(S(\mathbf{u}) \otimes T(\mathbf{v})) \\ &= c\,B(\mathbf{u}, \mathbf{v}) \end{aligned}$$

pour tout choix de $\mathbf{u}, \mathbf{u}_1, \mathbf{u}_2 \in U$, $\mathbf{v} \in V$ et $c \in K$. De même, on vérifie que B est linéaire à droite. Donc c'est une application bilinéaire. En vertu des propriétés du produit tensoriel, on conclut qu'il existe une et une seule application linéaire $L\colon U \otimes_K V \to U \otimes_K V$ telle que

$$L(\mathbf{u} \otimes \mathbf{v}) = S(\mathbf{u}) \otimes T(\mathbf{v})$$

pour tout $\mathbf{u} \in U$ et tout $\mathbf{v} \in V$. □

Le produit tensoriel des opérateurs défini par la proposition 9.4.1 possède de nombreuses propriétés (voir les exercices). On se contente ici de montrer

Proposition 9.4.2. *Soient $S_1, S_2 \in \mathrm{End}_K(U)$ et $T_1, T_2 \in \mathrm{End}_K(V)$. Alors on a*

$$(S_1 \otimes T_1) \circ (S_2 \otimes T_2) = (S_1 \circ S_2) \otimes (T_1 \circ T_2).$$

Preuve. Pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$, on trouve

$$\begin{aligned} ((S_1 \otimes T_1) \circ (S_2 \otimes T_2))(\mathbf{u} \otimes \mathbf{v}) &= (S_1 \otimes T_1)((S_2 \otimes T_2)(\mathbf{u} \otimes \mathbf{v})) \\ &= (S_1 \otimes T_1)(S_2(\mathbf{u}) \otimes T_2(\mathbf{v})) \\ &= S_1(S_2(\mathbf{u})) \otimes T_1(T_2(\mathbf{v})) \\ &= (S_1 \circ S_2)(\mathbf{u}) \otimes (T_1 \circ T_2)(\mathbf{v}) \end{aligned}$$

En vertu de la proposition 9.4.1, cela signifie que $(S_1 \otimes T_1) \circ (S_2 \otimes T_2) = (S_1 \circ S_2) \otimes (T_1 \circ T_2)$. □

Définition 9.4.3. Le produit de Kronecker d'une matrice $A = (a_{ij}) \in \mathrm{Mat}_{m\times m}(K)$ par une matrice $B = (b_{ij}) \in \mathrm{Mat}_{n\times n}(K)$ est la matrice

$$A\dot{\times}B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1m}B \\ a_{21}B & a_{22}B & \cdots & a_{2m}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mm}B \end{pmatrix} \in \mathrm{Mat}_{mn\times mn}(K).$$

L'intérêt de cette notion tient dans le résultat suivant :

Proposition 9.4.4. *Soient $S \in \mathrm{End}_K(U)$ et $T \in \mathrm{End}_K(V)$, et soient*

$$\mathcal{A} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\} \quad \textit{et} \quad \mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$$

des bases respectives de U et de V. Alors

$$\mathcal{C} = \{\mathbf{u}_1 \otimes \mathbf{v}_1, \ldots, \mathbf{u}_1 \otimes \mathbf{v}_n; \mathbf{u}_2 \otimes \mathbf{v}_1, \ldots, \mathbf{u}_2 \otimes \mathbf{v}_n; \ldots\ldots; \mathbf{u}_m \otimes \mathbf{v}_1, \ldots, \mathbf{u}_m \otimes \mathbf{v}_n\}$$

est une base de $U \otimes_K V$ et on a

$$[\,S \otimes T\,]_{\mathcal{C}} = [\,S\,]_{\mathcal{A}}\dot{\times}[\,T\,]_{\mathcal{B}}.$$

Preuve. La matrice $[\,S \otimes T\,]_{\mathcal{C}}$ se décompose naturellement en blocs

$$\begin{array}{c} \\ \left\{\begin{array}{c} \mathbf{u}_1 \otimes \mathbf{v}_1 \\ \vdots \\ \mathbf{u}_1 \otimes \mathbf{v}_n \end{array}\right. \\ \vdots \\ \left\{\begin{array}{c} \mathbf{u}_m \otimes \mathbf{v}_1 \\ \vdots \\ \mathbf{u}_m \otimes \mathbf{v}_n \end{array}\right. \end{array} \begin{array}{c} \overbrace{\mathbf{u}_1 \otimes \mathbf{v}_1, \ldots, \mathbf{u}_1 \otimes \mathbf{v}_n} \quad \cdots \quad \overbrace{\mathbf{u}_m \otimes \mathbf{v}_1, \ldots, \mathbf{u}_m \otimes \mathbf{v}_n} \\ \left(\begin{array}{c|c|c} C_{11} & \cdots & C_{1m} \\ \hline \vdots & & \vdots \\ \hline C_{m1} & \cdots & C_{mm} \end{array}\right) \end{array}.$$

Pour $k, l \in \{1, \ldots, n\}$ fixés, le bloc C_{kl} est le morceau suivant

$$\begin{array}{c} \\ \vdots \\ \mathbf{u}_k \otimes \mathbf{v}_1 \\ \vdots \\ \mathbf{u}_k \otimes \mathbf{v}_n \\ \vdots \end{array} \begin{array}{c} \cdots \quad \mathbf{u}_l \otimes \mathbf{v}_1, \ldots, \mathbf{u}_l \otimes \mathbf{v}_n \quad \cdots \\ \left(\begin{array}{c|c|c} & & \\ \hline & C_{kl} & \\ \hline & & \end{array}\right) \end{array}.$$

L'élément (i,j) de C_{kl} est donc égal au coefficient de $\mathbf{u}_k \otimes \mathbf{v}_i$ dans l'écriture de $(S \times T)(\mathbf{u}_l \otimes \mathbf{v}_j)$ comme conbinaison linéaire des éléments de la base $\mathcal{C}$. Pour le calculer, écrivons

$$A = [\,S\,]_{\mathcal{A}} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} \quad \text{et} \quad B = [\,T\,]_{\mathcal{B}} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

On trouve

$$\begin{aligned} (S \otimes T)(\mathbf{u}_l \otimes \mathbf{v}_j) &= S(\mathbf{u}_l) \otimes T(\mathbf{v}_j) \\ &= (a_{1l}\mathbf{u}_1 + \cdots + a_{ml}\mathbf{u}_l) \otimes (b_{1j}\mathbf{v}_1 + \cdots + b_{nj}\mathbf{v}_n) \\ &= \cdots + (a_{kl}b_{1j}\mathbf{u}_k \otimes \mathbf{v}_1 + \cdots + a_{kl}b_{nj}\mathbf{u}_k \otimes \mathbf{v}_n) + \cdots \end{aligned}$$

donc le coefficient (i,j) de C_{kl} est $a_{kl}b_{ij}$, égal à celui de $a_{kl}B$. Cela montre que $C_{kl} = a_{kl}B$ et par suite

$$[\,S \otimes T\,]_{\mathcal{C}} = A \dot{\times} B = [\,S\,]_{\mathcal{A}} \dot{\times} [\,T\,]_{\mathcal{B}}.$$

□

Exercices.

9.4.1. Désignons par $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$ la base canonique de $\mathbb{R}^2$ et par $\mathcal{E}' = \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ celle de $\mathbb{R}^3$. Soient $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ et $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ les opérateurs linéaires pour lesquels

$$[\,S\,]_{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} \quad \text{et} \quad [\,T\,]_{\mathcal{E}'} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Calculer

(i) $(S \otimes T)(3\mathbf{e}_1 \otimes \mathbf{e}'_1)$,

(ii) $(S \otimes T)(2\mathbf{e}_1 \otimes \mathbf{e}'_1 - \mathbf{e}_2 \otimes \mathbf{e}'_2)$,

(iii) $(S \otimes T)(2(\mathbf{e}_1 + \mathbf{e}_2) \otimes \mathbf{e}'_3 - \mathbf{e}_2 \otimes \mathbf{e}'_3)$.

9.4.2. Soient U et V des espaces vectoriels de dimension finie sur un corps K, et soient $S\colon U \to U$ et $T\colon V \to V$ des applications linéaires inversibles. Montrer que l'application linéaire $S \otimes T\colon U \otimes_K V \to U \otimes_K V$ est aussi inversible, et donner une formule liant S^{-1}, T^{-1} et $(S \otimes T)^{-1}$.

9.4.3. Soient $S \in \mathrm{End}_K(U)$ et $T \in \mathrm{End}_K(V)$, où U et V désignent des espaces vectoriels de dimension finie sur un corps K. Soit U_1 un sous-espace S-invariant de U et soit V_1 un sous-espace T-invariant de V. On considère $U_1 \otimes_K V_1$ comme un sous-espace de $U \otimes_K V$ comme dans l'exercice 9.3.5. Sous ces hypothèses, montrer que $U_1 \otimes_K V_1$ est un sous-espace $S \otimes T$-invariant de $U \otimes_K V$ et que $(S \otimes T)|_{U_1 \otimes_K V_1} = S|_{U_1} \otimes T|_{V_1}$.

9.4.4. Soient $S \in \mathrm{End}_{\mathbb{C}}(U)$ et $T \in \mathrm{End}_{\mathbb{C}}(V)$, où U et V désignent des espaces vectoriels de dimensions respectives m et n sur $\mathbb{C}$. On choisit une base $\mathcal{A}$ de U et une base $\mathcal{B}$ de V telles que les matrices $[S]_{\mathcal{A}}$ et $[T]_{\mathcal{B}}$ soient sous forme canonique de Jordan. Montrer que, pour la base correspondante $\mathcal{C}$ de $U \otimes_{\mathbb{C}} V$ donnée par la proposition 9.4.4, la matrice $[S \otimes T]_{\mathcal{C}}$ est triangulaire supérieure. En calculant son déterminant, montrer que $\det(S \otimes T) = \det(S)^n \det(T)^m$.

9.4.5. Soient U et V des espaces vectoriels de dimension finie sur un corps K.

(i) Montrer que l'application

$$\begin{array}{rcl} B : \mathrm{End}_K(U) \times \mathrm{End}_K(V) & \longrightarrow & \mathrm{End}_K(U \otimes_K V) \\ (S, T) & \longmapsto & S \otimes T \end{array}$$

est bilinéaire.

(ii) Vérifier que l'image de B engendre $\mathrm{End}_K(U \otimes_K V)$ en tant qu'espace vectoriel sur K.

(iii) Conclure que l'application linéaire de $\mathrm{End}_K(U) \otimes_K \mathrm{End}_K(V)$ dans $\mathrm{End}_K(U \otimes_K V)$ associée à B est un isomorphisme.

9.4.6. Soient U et V des espaces vectoriels de dimension finie sur un corps K.

(i) Soient $f \in U^*$ et $g \in V^*$. Montrer qu'il existe une forme linéaire $f \otimes g \colon U \otimes_K V \to K$ unique telle que $(f \otimes g)(\mathbf{u} \otimes \mathbf{v}) = f(\mathbf{u})g(\mathbf{v})$ pour tout $(\mathbf{u}, \mathbf{v}) \in U \times V$.

(ii) Vérifier que l'application $B \colon U^* \times V^* \to (U \otimes_K V)^*$ donnée par $B(f, g) = f \otimes g$ pour tout $(f, g) \in U^* \times V^*$ est bilinéaire et que son image engendre $(U \otimes_K V)^*$ en tant qu'espace vectoriel sur K. Conclure que l'application linéaire de $U^* \otimes_K V^*$ dans $(U \otimes_K V)^*$ associée à B est un isomorphisme.

(iii) Soient $S \in \mathrm{End}_K(U)$ et $T \in \mathrm{End}_K(V)$. Montrer que, si on identifie $U^* \otimes_K V^*$ à $(U \otimes_K V)^*$ grâce à l'isomorphisme établi en (ii), alors on a $S^* \otimes T^* = (S \otimes T)^*$.

9.5 Produits tensoriels multiples

Théorème 9.5.1. *Soient U, V, W des espaces vectoriels de dimension finie sur un corps K. Il existe un et un seul isomorphisme $L \colon (U \otimes_K V) \otimes_K W \to U \otimes_K (V \otimes_K W)$ tel que*

$$L((\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w}) = \mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w})$$

pour tout $(\mathbf{u}, \mathbf{v}, \mathbf{w}) \in U \times V \times W$.

Preuve. Soient $\{\mathbf{u}_1, \ldots, \mathbf{u}_l\}$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ et $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ des bases respectives de U, V et W. Alors

$$\{(\mathbf{u}_i \otimes \mathbf{v}_j) \otimes \mathbf{w}_k \,;\, 1 \leq i \leq l, 1 \leq j \leq m, 1 \leq k \leq n\},$$
$$\{\mathbf{u}_i \otimes (\mathbf{v}_j \otimes \mathbf{w}_k) \,;\, 1 \leq i \leq l, 1 \leq j \leq m, 1 \leq k \leq n\}$$

sont des bases respectives de $(U \otimes_K V) \otimes_K W$ et $U \otimes_K (V \otimes_K W)$. Il donc existe une et une seule application linéaire L de $(U \otimes_K V) \otimes_K W$ dans $U \otimes_K (V \otimes_K W)$ telle que

$$L((\mathbf{u}_i \otimes \mathbf{v}_j) \otimes \mathbf{w}_k) = \mathbf{u}_i \otimes (\mathbf{v}_j \otimes \mathbf{w}_k)$$

pour $i = 1, \ldots, l$, $j = 1, \ldots, m$ et $k = 1, \ldots, n$. Par construction, L est un isomorphisme. Enfin, pour des vecteurs quelconques

$$\mathbf{u} = \sum_{i=1}^{l} a_i \mathbf{u}_i \in U, \quad \mathbf{v} = \sum_{j=1}^{m} b_j \mathbf{v}_j \in V \quad \text{et} \quad \mathbf{w} = \sum_{k=1}^{n} c_k \mathbf{w}_k \in W,$$

on trouve

$$\begin{aligned}
L((\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w}) &= L\Big(\Big(\Big(\sum_{i=1}^{l} a_i \mathbf{u}_i\Big) \otimes \Big(\sum_{j=1}^{m} b_j \mathbf{v}_j\Big)\Big) \otimes \Big(\sum_{k=1}^{n} c_k \mathbf{w}_k\Big)\Big) \\
&= L\Big(\sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} a_i b_j c_k (\mathbf{u}_i \otimes \mathbf{v}_j) \otimes \mathbf{w}_k\Big) \\
&= \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} a_i b_j c_k \, \mathbf{u}_i \otimes (\mathbf{v}_j \otimes \mathbf{w}_k) \\
&= \Big(\sum_{i=1}^{l} a_i \mathbf{u}_i\Big) \otimes \Big(\Big(\sum_{j=1}^{m} b_j \mathbf{v}_j\Big) \otimes \Big(\sum_{k=1}^{n} c_k \mathbf{w}_k\Big)\Big) \\
&= \mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w}),
\end{aligned}$$

comme requis. □

L'isomorphisme $(U \otimes_K V) \otimes_K W \simeq U \otimes_K (V \otimes_K W)$ donné par le théorème 9.5.1 permet d'identifier ces deux espaces, et de les désigner par $U \otimes_K V \otimes_K W$. On écrit alors $\mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$ au lieu de $(\mathbf{u} \otimes \mathbf{v}) \otimes \mathbf{w}$ ou de $\mathbf{u} \otimes (\mathbf{v} \otimes \mathbf{w})$.

En général, le théorème 9.5.1 permet d'omettre les parenthèses dans un produit tensoriel d'un nombre quelconque d'espaces vectoriels $V_1, \ldots, V_r$ de dimension finie sur K. Ainsi, on écrit

$$V_1 \otimes_K V_2 \otimes_K \cdots \otimes_K V_r \quad \text{au lieu de} \quad V_1 \otimes_K (V_2 \otimes_K (\cdots \otimes_K V_r))$$

et

$$\mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_r \quad \text{au lieu de} \quad \mathbf{v}_1 \otimes (\mathbf{v}_2 \times (\cdots \otimes \mathbf{v}_r)).$$

On a alors

$$\dim_K(V_1 \otimes_K V_2 \otimes_K \cdots \otimes_K V_r) = \dim_K(V_1) \dim_K(V_2) \cdots \dim_K(V_r).$$

Exercices.

9.5.1. Soient V et W des espaces vectoriels de dimension finie sur un corps K. Montrer qu'il existe un et un seul isomorphisme $L\colon V \otimes_K W \to W \otimes_K V$ tel que $L(\mathbf{v} \otimes \mathbf{w}) = \mathbf{w} \otimes \mathbf{v}$ pour tout $\mathbf{v} \in V$ et tout $\mathbf{w} \in W$.

9.5.2. Soient r un entier positif et $V_1, \ldots, V_r, W$ des espaces vectoriels sur K. On dit qu'une fonction

$$\varphi : V_1 \times \cdots \times V_r \longrightarrow W$$

est *multilinéaire* ou encore *r-linéaire* si elle est séparément linéaire en chacun de ses arguments, c'est-à-dire si elle sastifait

ML1. $\varphi(\mathbf{v}_1, \ldots, \mathbf{v}_i + \mathbf{v}'_i, \ldots, \mathbf{v}_r) = \varphi(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_r) + \varphi(\mathbf{v}_1, \ldots, \mathbf{v}'_i, \ldots, \mathbf{v}_r)$

ML2. $\varphi(\mathbf{v}_1, \ldots, c\,\mathbf{v}_i, \ldots, \mathbf{v}_r) = c\,\varphi(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_r)$

quels que soient $i \in \{1, \ldots, r\}, \mathbf{v}_1 \in V_1, \ldots, \mathbf{v}_i, \mathbf{v}'_i \in V_i, \ldots, \mathbf{v}_r \in V_r$ et $c \in K$.

Montrer que la fonction

$$\begin{array}{rccc} \psi : & V_1 \times V_2 \times \cdots \times V_r & \longrightarrow & V_1 \otimes_K V_2 \otimes_K \cdots \otimes_K V_r \\ & (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r) & \longmapsto & \mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_r \end{array}$$

est multilinéaire.

9.5.3. Soient $V_1, \ldots, V_r$ et W comme dans l'exercice précédent. Le but de ce nouvel exercice est de montrer que, pour toute application multilinéaire $\varphi\colon V_1 \times \cdots \times V_r \to W$, il existe une et une seule application linéaire $L\colon V_1 \otimes_K \cdots \otimes_K V_r \to W$ telle que

$$\varphi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = L(\mathbf{v}_1 \otimes \mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_r)$$

pour tout $(\mathbf{v}_1, \ldots, \mathbf{v}_r) \in V_1 \times \cdots \times V_r$.

On procède par récurrence sur r. Pour $r = 2$, cela découle de la définition du produit tensoriel. Supposons $r \geq 3$ et le résultat vrai pour les application multilinéaires de $V_2 \times \cdots \times V_r$ dans W.

(i) Montrer que, pour chaque choix de $\mathbf{v}_1 \in V_1$, il existe une et une seule application linéaire $L_{\mathbf{v}_1}\colon V_2 \otimes_K \cdots \otimes_K V_r \to W$ telle que

$$\varphi(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r) = L_{\mathbf{v}_1}(\mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_r)$$

pour tout $(\mathbf{v}_2, \ldots, \mathbf{v}_r) \in V_2 \times \cdots \times V_r$.

(ii) Montrer que l'application

$$\begin{array}{rccc} B : & V_1 \times (V_2 \otimes_K \cdots \otimes_K V_r) & \longrightarrow & W \\ & (\mathbf{v}_1, \alpha) & \longmapsto & L_{\mathbf{v}_1}(\alpha) \end{array}$$

est bilinéaire.

(iii) En déduire qu'il existe une application linéaire $L\colon V_1 \otimes_K (V_2 \otimes_K \cdots \otimes_K V_r) \to W$ et une seule telle que $\varphi(\mathbf{v}_1, \ldots, \mathbf{v}_r) = L(\mathbf{v}_1 \otimes (\mathbf{v}_2 \otimes \cdots \otimes \mathbf{v}_r))$ pour tout $(\mathbf{v}_1, \ldots, \mathbf{v}_r) \in V_1 \times \cdots \times V_r$.

9.5.4. Soit V un espace vectoriel de dimension finie sur un corps K. En utilisant le résultat de l'exercice précédent, montrer qu'il existe une application linéaire $L\colon V^* \otimes_K V \otimes_K V \to V$ unique telle que $L(f \otimes \mathbf{v}_1 \otimes \mathbf{v}_2) = f(\mathbf{v}_1)\mathbf{v}_2$ pour tout choix de $(f, \mathbf{v}_1, \mathbf{v}_2) \in V^* \times V \times V$.

9.5.5. Soient V_1, V_2, V_3 et W des espaces vectoriels de dimension finie sur un corps K, et soit $\varphi : V_1 \times V_2 \times V_3 \to W$ une application trilinéaire. Montrer qu'il existe une et une seule application bilinéaire $B\colon V_1 \times (V_2 \otimes_K V_3) \to W$ telle que $B(\mathbf{v}_1, \mathbf{v}_2 \otimes \mathbf{v}_3) = \varphi(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ pour tout choix de $\mathbf{v}_j \in V_j$ $(j = 1, 2, 3)$.

Chapitre 10

Espaces euclidiens

Après des rappels généraux sur les espaces euclidiens, on définit les notions d'opérateur orthogonal et d'opérateur symétrique sur un tel espace, puis on démontre des théorèmes de structure pour ces opérateurs (théorèmes spectraux). On conclut en montrant que tout opérateur linéaire sur un espace euclidien de dimension finie est la composée d'un opérateur symétrique défini positif suivi par un opérateur orthogonal. Cette écriture est appelée *décomposition polaire* d'un opérateur. Par exemple dans $\mathbb{R}^2$, un opérateur symétrique défini positif est un produit de dilatations suivant des axes orthogonaux tandis qu'un opérateur orthogonal est simplement une rotation autour de l'origine ou encore une symétrie par rapport à une droite passant par l'origine. Donc tout opérateur linéaire sur $\mathbb{R}^2$ est la composée de dilatations suivant des axes orthogonaux suivies par une rotation ou une symétrie.

10.1 Rappels

Définition 10.1.1. Soit V un espace vectoriel réel. Un *produit scalaire* sur V est une forme bilinéaire

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \langle \mathbf{u}, \mathbf{v} \rangle \end{array}$$

qui est *symétrique* :

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \quad \text{quels que soient } \mathbf{u}, \mathbf{v} \in V$$

et *définie positive* :

$$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \quad \text{pour tout } \mathbf{v} \in V, \quad \text{avec} \quad \langle \mathbf{v}, \mathbf{v} \rangle = 0 \iff \mathbf{v} = \mathbf{0}\,.$$

Définition 10.1.2. Un *espace euclidien* est un espace vectoriel réel V muni d'un produit scalaire.

Exemple 10.1.3. Soit n un entier positif. On définit un produit scalaire sur $\mathbb{R}^n$ en posant

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1 y_1 + \cdots + x_n y_n.$$

Muni de ce produit scalaire, $\mathbb{R}^n$ devient un espace euclidien. Quand on parle de $\mathbb{R}^n$ en tant qu'espace euclidien sans spécifier de produit scalaire, on sous-entend qu'il s'agit du produit scalaire défini ci-dessus.

Exemple 10.1.4. Soit V l'espace vectoriel des fonctions continues $f\colon \mathbb{R} \to \mathbb{R}$ qui sont 2π-périodiques, c'est-à-dire qui satisfont $f(x+2\pi) = f(x)$ pour tout $x \in \mathbb{R}$. Alors V est un espace euclidien pour le produit scalaire

$$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\mathrm{d}x\,.$$

Exemple 10.1.5. Soit n un entier positif. Alors $\mathrm{Mat}_{n\times n}(\mathbb{R})$ est un espace euclidien pour le produit scalaire $\langle A, B\rangle = \mathrm{trace}(AB^t)$.

Pour le reste de la section, on fixe un espace euclidien V avec son produit scalaire $\langle\ ,\ \rangle$. On note d'abord :

Proposition 10.1.6. *Tout sous-espace W de V est un espace euclidien pour le produit scalaire de V restreint à W.*

Preuve. En effet, le produit scalaire de V fournit par restriction une application bilinéaire

$$\begin{array}{ccc} W \times W & \longrightarrow & \mathbb{R} \\ (\mathbf{w}_1, \mathbf{w}_2) & \longmapsto & \langle \mathbf{w}_1, \mathbf{w}_2\rangle \end{array}$$

et on vérifie sans peine que celle-ci constitue un produit scalaire au sens de la définition 10.1.1. □

Quand on parle d'un sous-espace de V comme d'un espace euclidien, c'est à ce produit scalaire que l'on fait référence.

Définition 10.1.7. La *norme* d'un vecteur $\mathbf{v} \in V$ est

$$\|\mathbf{v}\| = \langle \mathbf{v}, \mathbf{v}\rangle^{1/2}.$$

On dit qu'un vecteur $\mathbf{v} \in V$ est *unitaire* si $\|\mathbf{v}\| = 1$.

On rappelle que la norme possède les propriétés suivantes, dont les deux premières découlent immédiatement de la définition :

Proposition 10.1.8. *Pour tout choix de $\mathbf{u}, \mathbf{v} \in V$ et de $c \in \mathbb{R}$, on a*

(i) $\|\mathbf{v}\| \geq 0$ *avec égalité si et seulement si* $\mathbf{v} = \mathbf{0}$,

(ii) $\|c\,\mathbf{v}\| = |c|\,\|\mathbf{v}\|$,

(iii) $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$ *(inégalité de Cauchy-Schwarz),*

(iv) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ *(inégalité du triangle).*

Les relations (i), (ii) et (iv) montrent que $\|.\|$ est bien une norme au sens propre du terme. Ainsi la norme permet de mesurer la "longueur" des vecteurs de V. Elle permet aussi de définir la "distance" $\mathrm{d}(\mathbf{u},\mathbf{v})$ entre deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ de V en posant

$$\mathrm{d}(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|.$$

Grâce à l'inégalité du triangle, on vérifie bien que cela définit une métrique sur V.

L'inégalité (iii) dite de Cauchy-Schwarz permet de définir l'angle $\theta \in [0,\pi]$ entre deux vecteurs non nuls $\mathbf{u}$ et $\mathbf{v}$ de V en posant

$$\theta = \arccos\left(\frac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}\right). \tag{10.1}$$

En particulier, cet angle θ vaut $\pi/2$ si et seulement si $\langle \mathbf{u},\mathbf{v}\rangle = 0$. Pour cette raison, on dit que des vecteurs $\mathbf{u},\mathbf{v}$ de V sont *perpendiculaires* ou *orthogonaux* si $\langle \mathbf{u},\mathbf{v}\rangle = 0$.

Enfin, si $\mathbf{v} \in V$ est non nul, il découle de (ii) que le produit

$$\|\mathbf{v}\|^{-1}\mathbf{v}$$

est un vecteur unitaire parallèle à $\mathbf{v}$ (il fait avec $\mathbf{v}$ un angle de 0 radian en vertu de (10.1)).

Définition 10.1.9. Soient $\mathbf{v}_1,\ldots,\mathbf{v}_n \in V$.

1) On dit que $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ est un *sous-ensemble orthogonal* de V si $\mathbf{v}_i \neq \mathbf{0}$ pour $i = 1,\ldots,n$ et si $\langle \mathbf{v}_i,\mathbf{v}_j\rangle = 0$ pour toute paire d'entiers (i,j) avec $1 \leq i,j \leq n$ et $i \neq j$.
2) On dit que $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ est une *base orthogonale* de V, si c'est à la fois une base de V et un sous-ensemble orthogonal de V.
3) On dit que $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ est une *base orthonormée* de V si c'est une base orthogonale de V formée de vecteurs unitaires, c'est-à-dire tels que $\|\mathbf{v}_i\| = \langle \mathbf{v}_i,\mathbf{v}_i\rangle^{1/2} = 1$ pour $i = 1,\ldots,n$.

La proposition suivante rappelle l'importance de ces notions.

Proposition 10.1.10.

(i) Tout sous-ensemble orthogonal de V est linéairement indépendant.

(ii) Supposons que $\mathcal{B} = \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ soit une base orthogonale de V. Alors, pour tout $\mathbf{v} \in V$, on a

$$\mathbf{v} = \frac{\langle \mathbf{v},\mathbf{v}_1\rangle}{\langle \mathbf{v}_1,\mathbf{v}_1\rangle}\mathbf{v}_1 + \cdots + \frac{\langle \mathbf{v},\mathbf{v}_n\rangle}{\langle \mathbf{v}_n,\mathbf{v}_n\rangle}\mathbf{v}_n.$$

(iii) Si $\mathcal{B} = \{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ est une base orthonormée de V, alors, pour tout $\mathbf{v} \in V$, on a

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \langle \mathbf{v},\mathbf{v}_1\rangle \\ \vdots \\ \langle \mathbf{v},\mathbf{v}_n\rangle \end{pmatrix}.$$

Preuve. Supposons d'abord $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ soit un sous-ensemble orthogonal de V, et soit $\mathbf{v} \in \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_{\mathbb{R}}$. On peut écrire

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

avec $a_1, \ldots, a_n \in \mathbb{R}$. En prenant le produit scalaire des deux membres de cette égalité avec $\mathbf{v}_i$, on trouve

$$\langle \mathbf{v}, \mathbf{v}_i \rangle = \Big\langle \sum_{j=1}^{n} a_j\mathbf{v}_j, \mathbf{v}_i \Big\rangle = \sum_{j=1}^{n} a_j \langle \mathbf{v}_j, \mathbf{v}_i \rangle = a_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle$$

et par suite

$$a_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \quad (i = 1, \ldots, n).$$

Si on choisit $\mathbf{v} = \mathbf{0}$, cette formule donne $a_1 = \cdots = a_n = 0$. Donc l'ensemble $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est linéairement indépendant et par suite, c'est une base de $\langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle_{\mathbb{R}}$. Cela démontre à la fois (i) et (ii). L'énoncé (iii) découle immédiatement de (ii). □

Le résultat suivant montre que, si V est de dimension finie, il possède une base orthogonale. En particulier, cela vaut pour tout sous-espace de dimension finie de V.

Théorème 10.1.11 (Procédé d'orthogonalisation de Gram-Schmidt)**.**
Supposons que $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ soit une base de V. On définit récursivement :

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{w}_1, \\
\mathbf{v}_2 &= \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1, \\
\mathbf{v}_3 &= \mathbf{w}_3 - \frac{\langle \mathbf{w}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{w}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2, \\
&\vdots \\
\mathbf{v}_n &= \mathbf{w}_n - \frac{\langle \mathbf{w}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{w}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle}\mathbf{v}_{n-1}.
\end{aligned}$$

Alors, pour chaque $i = 1, \ldots, n$, l'ensemble $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ est une base orthogonale du sous-espace $\langle \mathbf{w}_1, \ldots, \mathbf{w}_i \rangle_{\mathbb{R}}$. En particulier, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est une base orthogonale de V.

Preuve. Il suffit de montrer la première assertion. Pour cela, on procède par récurrence sur i. Pour $i = 1$, il est clair que $\{\mathbf{v}_1\}$ est une base orthogonale de $\langle \mathbf{w}_1 \rangle_{\mathbb{R}}$ car $\mathbf{v}_1 = \mathbf{w}_1 \neq \mathbf{0}$. Supposons que $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$ soit une base orthogonale de $\langle \mathbf{w}_1, \ldots, \mathbf{w}_{i-1} \rangle_{\mathbb{R}}$ pour un entier i avec $2 \leq i \leq n$. Par définition, on a

$$\mathbf{v}_i = \mathbf{w}_i - \frac{\langle \mathbf{w}_i, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{w}_i, \mathbf{v}_{i-1} \rangle}{\langle \mathbf{v}_{i-1}, \mathbf{v}_{i-1} \rangle}\mathbf{v}_{i-1}, \tag{10.2}$$

et on note que les dénominateurs $\langle \mathbf{v}_1, \mathbf{v}_1 \rangle, \ldots, \langle \mathbf{v}_{i-1}, \mathbf{v}_{i-1} \rangle$ des coefficients de $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ ne sont pas nuls car chacun des vecteurs $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ est non nul. On déduit de cette formule que

$$\langle \mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_i \rangle_{\mathbb{R}} = \langle \mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{w}_i \rangle_{\mathbb{R}} = \langle \mathbf{w}_1, \ldots, \mathbf{w}_{i-1}, \mathbf{w}_i \rangle_{\mathbb{R}}.$$

Ainsi, pour conclure que $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ est une base orthogonale de $\langle \mathbf{w}_1, \ldots, \mathbf{w}_i \rangle_{\mathbb{R}}$, il reste seulement à montrer que c'est un ensemble orthogonal (en vertu de la proposition 10.1.10). Pour cela, il suffit de montrer que $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ pour $j = 1, \ldots, i-1$. Grâce à (10.2), on trouve en effet

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \langle \mathbf{w}_i, \mathbf{v}_j \rangle - \sum_{k=1}^{i-1} \frac{\langle \mathbf{w}_i, \mathbf{v}_k \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle} \langle \mathbf{v}_k, \mathbf{v}_j \rangle = 0$$

pour $j = 1, \ldots, i-1$. Cela complète la démonstration. □

On termine cette section avec le rappel de la notion de complément orthogonal et de deux résultats qui s'y rapportent.

Définition 10.1.12. Le *complément orthogonal* d'un sous-espace W de V est

$$W^{\perp} := \{\mathbf{v} \in V \,;\, \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ pour tout } \mathbf{w} \in W\}.$$

Proposition 10.1.13. *Soit W un sous-espace de V. Alors $W^{\perp}$ est un sous-espace de V qui satisfait $W \cap W^{\perp} = \{\mathbf{0}\}$. De plus, si $W = \langle \mathbf{w}_1, \ldots, \mathbf{w}_m \rangle_{\mathbb{R}}$, alors*

$$W^{\perp} = \{\mathbf{v} \in V \,;\, \langle \mathbf{v}, \mathbf{w}_i \rangle = 0 \text{ pour } i = 1, \ldots, m\}.$$

Preuve. On laisse en exercice le soin de montrer que $W^{\perp}$ est un sous-espace de V. On note que, si $\mathbf{w} \in W \cap W^{\perp}$, alors, on a $\langle \mathbf{w}, \mathbf{w} \rangle = 0$ et par suite $\mathbf{w} = 0$. Cela montre que $W \cap W^{\perp} = \{\mathbf{0}\}$. Enfin, si $W = \langle \mathbf{w}_1, \ldots, \mathbf{w}_m \rangle$, on trouve

$$\begin{aligned} \mathbf{v} \in W^{\perp} &\iff \langle \mathbf{v}, a_1\mathbf{w}_1 + \cdots + a_m\mathbf{w}_m \rangle = 0 \quad \forall\, a_1, \ldots, a_m \in \mathbb{R} \\ &\iff a_1\langle \mathbf{v}, \mathbf{w}_1 \rangle + \cdots + a_m\langle \mathbf{v}, \mathbf{w}_m \rangle = 0 \quad \forall\, a_1, \ldots, a_m \in \mathbb{R} \\ &\iff \langle \mathbf{v}, \mathbf{w}_1 \rangle = \cdots = \langle \mathbf{v}, \mathbf{w}_m \rangle = 0, \end{aligned}$$

d'où la dernière assertion de la proposition. □

Théorème 10.1.14. *Soit W un sous-espace de V de dimension finie. Alors, on a :*

$$V = W \oplus W^{\perp}.$$

Preuve. Puisque, selon la proposition 10.1.13, on a $W \cap W^{\perp} = \{\mathbf{0}\}$, il reste à monter que $V = W + W^{\perp}$. Soit $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ une base orthogonale de W (il en existe une car W est de dimension finie). Étant donné $\mathbf{v} \in V$, on pose

$$\mathbf{w} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_m \rangle}{\langle \mathbf{w}_m, \mathbf{w}_m \rangle}\mathbf{w}_m. \tag{10.3}$$

Pour $i = 1, \ldots, m$, on trouve que

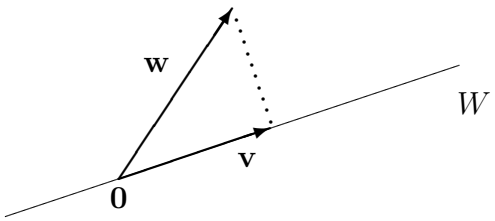
$$\langle \mathbf{v} - \mathbf{w}, \mathbf{w}_i \rangle = \langle \mathbf{v}, \mathbf{w}_i \rangle - \langle \mathbf{w}, \mathbf{w}_i \rangle = 0,$$

donc $\mathbf{v} - \mathbf{w} \in W^{\perp}$, et par suite

$$\mathbf{v} = \mathbf{w} + (\mathbf{v} - \mathbf{w}) \in W + W^{\perp}.$$

Le choix de $\mathbf{v} \in V$ étant arbitraire, cela montre que $V = W + W^{\perp}$, comme annoncé. □

Si W est un sous-espace de dimension finie de V, le théorème 10.1.14 nous apprend que, pour tout $\mathbf{v} \in V$, il existe un et un seul vecteur $\mathbf{w} \in W$ tel que $\mathbf{v} - \mathbf{w} \in W^\perp$.



On dit que ce vecteur $\mathbf{w}$ est la *projection orthogonale* de $\mathbf{v}$ sur W. Si $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ est une base orthogonale de W, alors cette projection $\mathbf{w}$ est donnée par (10.3).

Exemple 10.1.15. Calculer la projection orthogonale de $\mathbf{v} = (0,1,1,0)^t$ sur le sous-espace W de $\mathbb{R}^4$ engendré par $\mathbf{w}_1 = (1,1,0,0)^t$ et $\mathbf{w}_2 = (1,1,1,1)^t$.

Solution: On applique d'abord l'algorithme de Gram-Schmidt à la base $\{\mathbf{w}_1, \mathbf{w}_2\}$ de W. On trouve

$$\mathbf{v}_1 = \mathbf{w}_1 = (1,1,0,0)^t$$
$$\mathbf{v}_2 = \mathbf{w}_2 - \frac{\langle \mathbf{w}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 = (1,1,1,1)^t - \frac{2}{2}(1,1,0,0)^t = (0,0,1,1)^t.$$

Donc $\{\mathbf{v}_1, \mathbf{v}_2\}$ est une base orthogonale de W. La projection orthogonale de $\mathbf{v}$ sur W est donc

$$\frac{\langle \mathbf{v}, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1 + \frac{\langle \mathbf{v}, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} \mathbf{v}_2 = \frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2 = \frac{1}{2}(1,1,1,1)^t.$$

Exercices.

10.1.1. Montrer que $\mathcal{B} = \{(1,-1,3)^t, (-2,1,1)^t, (4,7,1)^t\}$ est une base orthogonale de $\mathbb{R}^3$ et calculer $[(a,b,c)^t]_{\mathcal{B}}$.

10.1.2. Déterminer $a, b, c \in \mathbb{R}$ de telle sorte que l'ensemble

$$\{(1,2,1,0)^t, (1,-1,1,3)^t, (2,-1,0,-1)^t, (a,b,c,1)^t\}$$

devienne une base orthogonale de $\mathbb{R}^4$.

10.1.3. Soient U et V des espaces euclidiens munis de produits scalaires $\langle\ ,\ \rangle_U$ et $\langle\ ,\ \rangle_V$ respectivement. Montrer que la formule

$$\big\langle (\mathbf{u}_1, \mathbf{v}_1), (\mathbf{u}_2, \mathbf{v}_2) \big\rangle = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle_U + \langle \mathbf{v}_1, \mathbf{v}_2 \rangle_V$$

définit un produit scalaire sur $U \times V$ pour lequel $(U \times \{0\})^\perp = \{0\} \times V$.

10.1.4. Soit V un espace euclidien de dimension n muni d'un produit scalaire $\langle\ ,\ \rangle$.

(i) Vérifier que, pour chaque $\mathbf{u} \in V$, la fonction $f_{\mathbf{u}} \colon V \to \mathbb{R}$ donnée par

$$f_{\mathbf{u}}(\mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle \quad \text{pour tout } \mathbf{v} \in V$$

appartient à V^*. Vérifier ensuite que l'application $\varphi \colon V \to V^*$ donnée par $\varphi(\mathbf{u}) = f_{\mathbf{u}}$ pour tout $\mathbf{u} \in V$ est un isomorphisme d'espaces vectoriels sur $\mathbb{R}$.

(ii) Soit $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ une base orthonormée de V. Montrer que la base de V^* duale à $\mathcal{B}$ est $\mathcal{B}^* = \{f_1, \ldots, f_n\}$ où $f_i := \varphi(\mathbf{u}_i) = f_{\mathbf{u}_i}$ pour $i = 1, \ldots, n$.

10.1.5. Soit V l'espace vectoriel des fonctions continues $f \colon \mathbb{R} \to \mathbb{R}$ qui sont 2π-périodiques.

(i) Montrer que la formule

$$\langle f, g \rangle = \int_0^{2\pi} f(x) g(x) \mathrm{d}x$$

définit un produit scalaire sur V.

(ii) Montrer que $\{1, \sin(x), \cos(x)\}$ est un sous-ensemble orthogonal de V pour ce produit scalaire.

(iii) Soit $f \colon \mathbb{R} \to \mathbb{R}$ la fonction périodique en "dents de scie" donnée sur $[0, 2\pi]$ par

$$f(x) = \begin{cases} x & \text{si } 0 \leq x \leq \pi, \\ 2\pi - x & \text{si } \pi \leq x \leq 2\pi, \end{cases}$$

et étendue par périodicité à tout $\mathbb{R}$. Calculer la projection de f sur le sous-espace W de V engendré par $\{1, \sin(x), \cos(x)\}$.

10.1.6. Soit $V = \mathrm{Mat}_{n \times n}(\mathbb{R})$.

(i) Montrer que la formule $\langle A\,,\, B \rangle = \mathrm{trace}(AB^t)$ définit un produit scalaire sur V.

(ii) Pour $n = 3$, déterminer une base orthonormée du sous-espace $S = \{A \in V\,;\, A^t = A\}$ des matrices symétriques.

10.1.7. Sur l'espace vectoriel $\mathcal{C}[-1, 1]$ des fonctions continues de $[-1, 1]$ dans $\mathbb{R}$, on se donne le produit scalaire

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x)\, dx\,.$$

Déterminer une base orthogonale du sous-espace $V = \langle\, 1, x, x^2 \,\rangle_{\mathbb{R}}$.

10.1.8. Soit W le sous-espace de $\mathbb{R}^4$ engendré par $(1, -1, 0, 1)^t$ et $(1, 1, 0, 0)^t$. Déterminer une base de $W^\perp$.

10.1.9. Soient W_1 et W_2 des sous-espaces d'un espace euclidien V de dimension finie. Montrer que $W_1 \subseteq W_2$ si et seulement si $W_2^\perp \subseteq W_1^\perp$.

10.2 Opérateurs orthogonaux

Dans cette section, on fixe un espace euclidien V de dimension finie n, et une base orthonormée $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V. On note d'abord :

Lemme 10.2.1. *Pour tout choix de $\mathbf{u}, \mathbf{v} \in V$, on a*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \big([\mathbf{u}]_{\mathcal{B}}\big)^t [\mathbf{v}]_{\mathcal{B}}.$$

Preuve. Soient $\mathbf{u}, \mathbf{v} \in V$. Écrivons

$$\mathbf{u} = \sum_{i=1}^{n} a_i \mathbf{u}_i \quad \text{et} \quad \mathbf{v} = \sum_{i=1}^{n} b_i \mathbf{u}_i$$

avec $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$. On trouve

$$\langle \mathbf{u}, \mathbf{v} \rangle = \Big\langle \sum_{i=1}^{n} a_i \mathbf{u}_i, \sum_{j=1}^{n} b_j \mathbf{v}_j \Big\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle = \sum_{i=1}^{n} a_i b_i$$

car $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$. Sous forme matricielle, cette égalité se réécrit

$$\langle \mathbf{u}, \mathbf{v} \rangle = (a_1, \ldots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = [\mathbf{u}]_{\mathcal{B}}^t [\mathbf{v}]_{\mathcal{B}}.$$

□

Grâce à ce lemme, nous pouvons maintenant montrer :

Proposition 10.2.2. *Soit $T \colon V \to V$ un opérateur linéaire. Les conditions suivantes sont équivalentes :*

(i) $\|T(\mathbf{u})\| = 1$ *pour tout* $\mathbf{u} \in V$ *avec* $\|\mathbf{u}\| = 1$,

(ii) $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ *pour tout* $\mathbf{v} \in V$,

(iii) $\langle T(\mathbf{u}), T(\mathbf{v}) \rangle = \langle \mathbf{u}, \mathbf{v} \rangle$ *pour tout choix de* $\mathbf{u}, \mathbf{v} \in V$,

(iv) $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ *est une base orthonormée de* V,

(v) $[T]_{\mathcal{B}}^t [T]_{\mathcal{B}} = I_V$,

(vi) $[T]_{\mathcal{B}}$ *est inversible et* $[T]_{\mathcal{B}}^{-1} = [T]_{\mathcal{B}}^t$.

Preuve. $(i) \Longrightarrow (ii)$: Supposons d'abord que $T(\mathbf{u})$ soit unitaire pour tout vecteur unitaire $\mathbf{u}$ de V. Soit $\mathbf{v}$ un vecteur quelconque de V. Si $\mathbf{v} \neq \mathbf{0}$, alors $\mathbf{u} := \|\mathbf{v}\|^{-1}\mathbf{v}$ est unitaire, donc $T(\mathbf{u}) = \|\mathbf{v}\|^{-1} T(\mathbf{v})$ est unitaire et par suite $\|T(\mathbf{v})\| = \|\mathbf{v}\|$. Si $\mathbf{v} = \mathbf{0}$, on a $T(\mathbf{v}) = \mathbf{0}$ et alors l'égalité $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ est encore vérifiée.

$(ii) \Longrightarrow (iii)$: Supposons que $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ pour tout $\mathbf{v} \in V$. Soient $\mathbf{u}, \mathbf{v}$ des vecteurs quelconques de V. On trouve

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle \\ &= \langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u},\mathbf{v}\rangle + \|\mathbf{v}\|^2. \end{aligned} \tag{10.4}$$

En appliquant cette formule à $T(\mathbf{u}) + T(\mathbf{v})$, on obtient

$$\|T(\mathbf{u}) + T(\mathbf{v})\|^2 = \|T(\mathbf{u})\|^2 + 2\langle T(\mathbf{u}), T(\mathbf{v})\rangle + \|T(\mathbf{v})\|^2. \tag{10.5}$$

Comme $\|T(\mathbf{u}) + T(\mathbf{v})\| = \|T(\mathbf{u}+\mathbf{v})\| = \|\mathbf{u}+\mathbf{v}\|$, et que $\|T(\mathbf{u})\| = \|\mathbf{u}\|$ et $\|T(\mathbf{v})\| = \|\mathbf{v}\|$, la comparaison de (10.4) et de (10.5) livre $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$.

$(iii) \Longrightarrow (iv)$: Supposons que $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ pour tout choix de $\mathbf{u}, \mathbf{v} \in V$. Comme $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ est une base orthonormée de V, on en déduit que

$$\langle T(\mathbf{u}_i), T(\mathbf{u}_j)\rangle = \langle \mathbf{u}_i, \mathbf{u}_j\rangle = \delta_{ij} \quad (1 \leq i, j \leq n).$$

Donc $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ est un sous-ensemble orthogonal de V formé de vecteurs unitaires. D'après la proposition 10.1.10 (i), cet ensemble est linéairement indépendant. Comme il consiste de n vecteurs et que V est de dimension n, c'est une base de V, donc une base orthonormée de V.

$(iv) \Longrightarrow (v)$: Supposons que $\{T(\mathbf{u}_1), \ldots, T(\mathbf{u}_n)\}$ soit une base orthonormée de V. Soient $i, j \in \{1, \ldots, n\}$. Grâce au lemme 10.2.1, on obtient

$$\delta_{ij} = \langle T(\mathbf{u}_i), T(\mathbf{u}_j)\rangle = [T(\mathbf{u}_i)]_{\mathcal{B}}^t \, [T(\mathbf{u}_j)]_{\mathcal{B}}.$$

Or $[T(\mathbf{u}_i)]_{\mathcal{B}}^t$ est la ligne i de la matrice $[T]_{\mathcal{B}}^t$, et $[T(\mathbf{u}_j)]_{\mathcal{B}}$ est la colonne j de $[T]_{\mathcal{B}}$. Donc $[T(\mathbf{u}_i)]_{\mathcal{B}}^t \, [T(\mathbf{u}_j)]_{\mathcal{B}} = \delta_{ij}$ est l'élément de la ligne i et de la colonne j de $[T]_{\mathcal{B}}^t \, [T]_{\mathcal{B}}$. Le choix de $i, j \in \{1, \ldots, n\}$ étant arbitraire, cela signifie que $[T]_{\mathcal{B}}^t \, [T]_{\mathcal{B}} = I$.

$(v) \Longrightarrow (vi)$: Supposons que $[T]_{\mathcal{B}}^t \, [T]_{\mathcal{B}} = I$. Alors $[T]_{\mathcal{B}}$ est inversible d'inverse $[T]_{\mathcal{B}}^t$.

$(vi) \Longrightarrow (i)$: Enfin, supposons que $[T]_{\mathcal{B}}$ soit inversible et que $[T]_{\mathcal{B}}^{-1} = [T]_{\mathcal{B}}^t$. Si $\mathbf{u}$ est un vecteur unitaire quelconque de V, on trouve, grâce au lemme 10.2.1,

$$\begin{aligned} \|T(\mathbf{u})\|^2 &= \langle T(\mathbf{u}), T(\mathbf{u})\rangle = [T(\mathbf{u})]_{\mathcal{B}}^t \, [T(\mathbf{u})]_{\mathcal{B}} \\ &= \left([T]_{\mathcal{B}}[\mathbf{u}]_{\mathcal{B}}\right)^t \left([T]_{\mathcal{B}}[\mathbf{u}]_{\mathcal{B}}\right) = [\mathbf{u}]_{\mathcal{B}}^t \, [T]_{\mathcal{B}}^t \, [T]_{\mathcal{B}}[\mathbf{u}]_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}}^t \, [\mathbf{u}]_{\mathcal{B}} = \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{u}\|^2 = 1. \end{aligned}$$

Donc $T(\mathbf{u})$ est aussi un vecteur unitaire.

□

Définition 10.2.3.

- Un opérateur linéaire $T\colon V \to V$ qui remplit les conditions équivalentes de la proposition 10.2.2 est appelé un *opérateur orthogonal* ou encore *opérateur unitaire réel*.

- Une matrice inversible $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ telle que $A^{-1} = A^t$ est appelée une *matrice orthogonale*.

La condition (i) de la proposition 10.2.2 demande que T applique l'ensemble des vecteurs unitaires de V sur lui-même. Cela explique la terminologie d'opérateur unitaire.

La condition (iii) de la proposition 10.2.2 signifie que T préserve le produit scalaire. En particulier, cela requiert que T applique tout couple de vecteurs orthogonaux $\mathbf{u}, \mathbf{v}$ de V sur un autre couple de vecteurs orthogonaux. Pour cette raison, on dit qu'un tel opérateur est orthogonal.

Enfin la condition (vi) signifie :

Corollaire 10.2.4. *Un opérateur linéaire $T\colon V \to V$ est orthogonal si et seulement si sa matrice relative à une base orthonormée de V est une matrice orthogonale.*

Exemple 10.2.5. Les opérateurs linéaires suivants sont orthogonaux :

- une rotation autour de l'origine dans $\mathbb{R}^2$,
- la symétrie par rapport à une droite passant par l'origine dans $\mathbb{R}^2$,
- une rotation autour d'une droite passant par l'origine dans $\mathbb{R}^3$,
- la symétrie par rapport à un plan passant par l'origine dans $\mathbb{R}^3$.

Définition 10.2.6. On désigne par $\mathcal{O}(V)$ l'ensemble des opérateurs orthogonaux de V et par $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices orthogonales $n \times n$.

On laisse en exercice le soin de vérifier que $\mathcal{O}(V)$ est un sous-groupe de $\mathrm{GL}(V) = \mathrm{End}_K(V)^*$, que $\mathcal{O}_n(\mathbb{R})$ est un sous-groupe de $\mathrm{GL}_n(\mathbb{R}) = \mathrm{Mat}_{n\times n}(\mathbb{R})^*$, et que ces deux groupes sont isomorphes.

Exemple 10.2.7. Soit $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'opérateur linéaire dont la matrice relative à la base canonique $\mathcal{E}$ de $\mathbb{R}^3$ est

$$[T]_{\mathcal{E}} = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}.$$

Montrer que T est un opérateur orthogonal.

Solution: Comme $\mathcal{E}$ est une base orthonormée de $\mathbb{R}^3$, il suffit de montrer que $[T]_{\mathcal{E}}$ est une matrice orthogonale. On trouve

$$[T]_{\mathcal{E}}^t\,[T]_{\mathcal{E}} = \frac{1}{9}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix} = I,$$

donc $[T]_{\mathcal{E}}^{-1} = [T]_{\mathcal{E}}^t$, comme requis.

Exercices.

10.2.1. Soit $T\colon V \to V$ un opérateur orthogonal. Montrer que T préserve les angles, c'est-à-dire que si $\mathbf{u}$ et $\mathbf{v}$ sont des vecteurs non nuls de V, l'angle entre $T(\mathbf{u})$ et $T(\mathbf{v})$ est égal à celui entre $\mathbf{u}$ et $\mathbf{v}$.

10.2.2. Soit $T\colon V \to V$ un opérateur linéaire inversible. Montrer que les conditions suivantes sont équivalentes :

(i) $\langle T(\mathbf{u}), T(\mathbf{v})\rangle = 0$ pour tout choix de $\mathbf{u}, \mathbf{v} \in V$ vérifiant $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

(ii) $T = c\,S$ pour un opérateur unitaire $S\colon V \to V$ et un nombre réel $c \neq 0$.

Note. On ne peut donc pas définir un opérateur orthogonal comme un opérateur linéaire qui conserve l'orthogonalité. En ce sens, la terminologie d'opérateur unitaire est plus appropriée.

10.2.3. Soit $\mathcal{E}$ une base orthonormée de V et soit $\mathcal{B}$ une base quelconque de V. Montrer que la matrice de changement de coordonnées $P = [\,I\,]_{\mathcal{E}}^{\mathcal{B}}$ est orthogonale si et seulement si $\mathcal{B}$ est aussi une base orthonormée.

10.2.4. Soit $T\colon V \to V$ un opérateur orthogonal. Supposons que W soit un sous-espace T-invariant de V. Montrer que $W^{\perp}$ est aussi T-invariant.

10.2.5. Soit $T\colon V \to V$ un opérateur orthogonal. Montrer que $\det(T) = \pm 1$.

10.2.6. Soit V un espace euclidien de dimension n. Montrer que $\mathcal{O}(V)$ est un sous-groupe de $\mathrm{GL}(V) = \mathrm{End}_K(V)^*$, que $\mathcal{O}_n(\mathbb{R})$ est un sous-groupe de $\mathrm{GL}_n(\mathbb{R}) = \mathrm{Mat}_{n\times n}(\mathbb{R})^*$, et définir un isomorphisme $\varphi\colon \mathcal{O}(V) \xrightarrow{\sim} \mathcal{O}_n(\mathbb{R})$.

10.2.7. Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ une matrice orthogonale. Montrer que les colonnes de A forment une base orthonormée de $\mathbb{R}^n$.

10.2.8. Soient V et W des espaces euclidiens quelconques. Supposons qu'il existe une fonction $T\colon V \to W$ telle que $T(\mathbf{0}) = \mathbf{0}$ et $\|T(\mathbf{v}_1) - T(\mathbf{v}_2)\| = \|\mathbf{v}_1 - \mathbf{v}_2\|$ pour tout choix de $\mathbf{v}_1, \mathbf{v}_2 \in V$.

(i) Montrer que $\langle T(\mathbf{v}_1), T(\mathbf{v}_2)\rangle = \langle \mathbf{v}_1, \mathbf{v}_2\rangle$ pour tout choix de $\mathbf{v}_1, \mathbf{v}_2 \in V$.

(ii) En déduire que $\|c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2)\| = \|c_1\mathbf{v}_1 + c_2\mathbf{v}_2\|$ quels que soient $\mathbf{v}_1, \mathbf{v}_2 \in V$ et $c_1, c_2 \in \mathbb{R}$.

(iii) Conclure que, pour tout choix de $\mathbf{v} \in V$ et de $c \in \mathbb{R}$, on a $\|T(c\,\mathbf{v}) - c\,T(\mathbf{v})\| = 0$, donc $T(c\,\mathbf{v}) = c\,T(\mathbf{v})$.

(iv) Par un raisonnement semblable, montrer que, quels que soient les vecteurs $\mathbf{v}_1, \mathbf{v}_2 \in V$, on a $\|T(\mathbf{v}_1 + \mathbf{v}_2) - T(\mathbf{v}_1) - T(\mathbf{v}_2)\| = 0$, donc $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$.

(v) Conclure que T est une application linéaire. En particulier, montrer que si $V = W$ est de dimension finie, alors T est un opérateur orthogonal au sens de la définition 10.2.3.

Une fonction $T\colon V \to W$ qui satisfait les conditions énoncées en début d'exercice s'appelle une *isométrie*.

Indice. Pour (i), utiliser la formule (10.4) avec $\mathbf{u} = T(\mathbf{v}_1)$ et $\mathbf{v} = -T(\mathbf{v}_2)$ puis avec $\mathbf{u} = \mathbf{v}_1$ et $\mathbf{v} = -\mathbf{v}_2$. Pour (ii), utiliser cette même formule avec $\mathbf{u} = c_1 T(\mathbf{v}_1)$ et $\mathbf{v} = c_2 T(\mathbf{v}_2)$ puis avec $\mathbf{u} = c_1\mathbf{v}_1$ et $\mathbf{v} = c_2\mathbf{v}_2$.

10.3 Opérateur adjoint

Les notations étant les mêmes que dans la section précédente, on montre d'abord :

Théorème 10.3.1. *Soit $B\colon V \times V \to \mathbb{R}$ une forme bilinéaire sur V. Il existe un et un seul couple d'opérateurs linéaires S et T sur V tels que*

$$B(\mathbf{u}, \mathbf{v}) = \langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, S(\mathbf{v})\rangle \tag{10.6}$$

quels que soient $\mathbf{u}, \mathbf{v} \in V$. De plus, soit $\mathcal{E}$ une base orthonormée de V et soit M la matrice de B relative à la base $\mathcal{E}$. Alors on a

$$[\,S\,]_{\mathcal{E}} = M \quad \text{et} \quad [\,T\,]_{\mathcal{E}} = M^t. \tag{10.7}$$

Preuve. **1° Existence.** Soient S et T les opérateurs linéaires sur V déterminés par les conditions (10.7), et soient $\mathbf{u}, \mathbf{v} \in V$. Puisque M est la matrice de B dans la base $\mathcal{E}$, on a

$$B(\mathbf{u}, \mathbf{v}) = [\mathbf{u}]_{\mathcal{E}}^t\, M [\mathbf{v}]_{\mathcal{E}}.$$

Comme $[T(\mathbf{u})]_{\mathcal{E}} = M^t\,[\mathbf{u}]_{\mathcal{E}}$ et que $[S(\mathbf{v})]_{\mathcal{E}} = M[\mathbf{v}]_{\mathcal{E}}$, le lemme 10.2.1 livre

$$\begin{aligned} &\langle T(\mathbf{u}), \mathbf{v}\rangle = \left(M^t\,[\mathbf{u}]_{\mathcal{E}}\right)^t [\mathbf{v}]_{\mathcal{E}} = [\mathbf{u}]_{\mathcal{E}}^t\, M[\mathbf{v}]_{\mathcal{E}} = B(\mathbf{u}, \mathbf{v}) \\ \textit{et}\quad &\langle \mathbf{u}, S(\mathbf{v})\rangle = [\mathbf{u}]_{\mathcal{E}}^t\, (M[\mathbf{v}]_{\mathcal{E}}) = B(\mathbf{u}, \mathbf{v}), \end{aligned}$$

comme requis.

2° Unicité. Supposons que S', T' soient des opérateurs linéaires quelconques sur V tels que

$$B(\mathbf{u}, \mathbf{v}) = \langle T'(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, S'(\mathbf{v})\rangle$$

pour tout choix de $\mathbf{u}, \mathbf{v} \in V$. Par soustraction de ces égalités avec (10.6), on obtient

$$0 = \langle T'(\mathbf{u}) - T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, S'(\mathbf{v}) - S(\mathbf{v})\rangle$$

quels que soient $\mathbf{u}, \mathbf{v} \in V$. On en déduit que, pour tout $\mathbf{u} \in V$, on a

$$T'(\mathbf{u}) - T(\mathbf{u}) \in V^{\perp} = \{\mathbf{0}\},$$

donc $T'(\mathbf{u}) = T(\mathbf{u})$ et par suite $T' = T$. De même, on trouve que $S' = S$. □

Corollaire 10.3.2. *Pour tout opérateur linéaire T sur V, il existe un et seul opérateur linéaire sur V noté T^*, tel que*

$$\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T^*(\mathbf{v})\rangle$$

quels que soient $\mathbf{u}, \mathbf{v} \in V$. Si $\mathcal{E}$ est une base orthonormée de V, on a $[T^*]_{\mathcal{E}} = [\,T\,]_{\mathcal{E}}^t$.

Preuve. La fonction $B\colon V \times V \to \mathbb{R}$ donnée par $B(\mathbf{u}, \mathbf{v}) = \langle T(\mathbf{u}), \mathbf{v}\rangle$ pour tout $(\mathbf{u}, \mathbf{v}) \in V \times V$ est bilinéaire (exercice). Pour conclure, il suffit d'appliquer le théorème 10.3.1 à cette forme bilinéaire. □

Définition 10.3.3. Soit T un opérateur linéaire sur V. L'opérateur $T^*\colon V \to T$ défini par le corollaire 10.3.2 s'appelle l'*adjoint* de T. On dit que T est un opérateur *symétrique* ou *auto-adjoint* si $T^* = T$.

En vertu du corollaire 10.3.2, on obtient aussitôt :

Proposition 10.3.4. *Soit $T\colon V \to V$ un opérateur linéaire. Les conditions suivantes sont équivalentes :*

(i) *T est symétrique,*

(ii) *$\langle T(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, T(\mathbf{v})\rangle$ quels que soient $\mathbf{u}, \mathbf{v} \in V$,*

(iii) *il existe une base orthonormée $\mathcal{E}$ de V telle que $[T]_{\mathcal{E}}$ soit une matrice symétrique,*

(iv) *$[T]_{\mathcal{E}}$ est une matrice symétrique pour toute base orthonormée $\mathcal{E}$ de V.*

Les propriétés suivantes sont laissées en exercice.

Proposition 10.3.5. *Soient S et T des opérateurs linéaires sur V, et soit $c \in \mathbb{R}$. On a :*

(i) $(cT)^* = cT^*$,

(ii) $(S+T)^* = S^* + T^*$,

(iii) $(S \circ T)^* = T^* \circ S^*$,

(iv) $(T^*)^* = T$.

(v) *De plus, si T est inversible, alors T^* l'est aussi et $(T^*)^{-1} = (T^{-1})^*$.*

Les propriétés (i) et (ii) signifient que l'application

$$\begin{array}{rcl} \mathrm{End}_{\mathbb{R}}(V) & \longrightarrow & \mathrm{End}_{\mathbb{R}}(V) \\ T & \longmapsto & T^* \end{array}$$

est *linéaire*. Les propriétés (ii) et (iii) s'expriment en disant que c'est un *anti-homomorphisme d'anneaux*. Plus précisément, à cause de (iv), on dit que c'est une *anti-involution*.

Exemple 10.3.6. Toute matrice symétrique $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ définit un opérateur symétrique de $\mathbb{R}^n$

$$\begin{array}{rccc} T_A : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & X & \longmapsto & AX \end{array}.$$

On peut le vérifier directement à partir de la définition en notant que, pour tout choix de $X, Y \in \mathbb{R}^n$, on a

$$\langle T_A(X), Y\rangle = (AX)^t Y = X^t A^t Y = X^t AY = \langle X, T_A(Y)\rangle.$$

Cela découle aussi de proposition 10.3.4. En effet, comme la base canonique $\mathcal{E}$ de $\mathbb{R}^n$ est orthonormée et que $[T_A]_{\mathcal{E}} = A$ est une matrice symétrique, l'opérateur linéaire T_A est symétrique.

Exemple 10.3.7. Soit W un sous-espace de V, et soit $P\colon V \to V$ la projection orthogonale sur W (c'est à dire l'application linéaire qui envoie $\mathbf{v} \in V$ sur sa projection orthogonale sur W). Alors P est un opérateur symétrique.

En effet, le théorème 10.1.14 montre que $V = W \oplus W^\perp$. L'application P est simplement la projection sur le premier facteur de cette somme suivie de l'inclusion de W dans V. Soient $\mathbf{u}, \mathbf{v}$ des vecteurs quelconques de V. On trouve

$$\langle P(\mathbf{u}), \mathbf{v}\rangle = \langle P(\mathbf{u})\,,\, (\mathbf{v} - P(\mathbf{v})) + P(\mathbf{v})\rangle = \langle P(\mathbf{u})\,,\, \mathbf{v} - P(\mathbf{v})\rangle + \langle P(\mathbf{u}), P(\mathbf{v})\rangle.$$

Comme $P(\mathbf{u}) \in W$ et que $\mathbf{v} - P(\mathbf{v}) \in W^\perp$, on a $\langle P(\mathbf{u}), \mathbf{v} - P(\mathbf{v})\rangle = 0$ et la dernière égalité se réécrit

$$\langle P(\mathbf{u}), \mathbf{v}\rangle = \langle P(\mathbf{u}), P(\mathbf{v})\rangle.$$

Un calcul tout à fait semblable donne

$$\langle \mathbf{u}, P(\mathbf{v})\rangle = \langle P(\mathbf{u}), P(\mathbf{v})\rangle.$$

Donc, on a $\langle \mathbf{u}, P(\mathbf{v})\rangle = \langle P(\mathbf{u}), \mathbf{v}\rangle$, ce qui montre que P est symétrique.

Exercices.

10.3.1. Démontrer la proposition 10.3.5.

10.3.2. Soit $T\colon V \to V$ un opérateur linéaire. Montrer que $\ker(T^*) = (\mathrm{Im}(T))^\perp$ et que $\mathrm{Im}(T^*) = (\ker(T))^\perp$.

10.3.3. Soit $T\colon V \to V$ un opérateur linéaire. Montrer que T est un opérateur orthogonal si et seulement si $T^* \circ T = I$.

10.3.4. On sait que $V = \mathrm{Mat}_{n\times n}(\mathbb{R})$ est un espace euclidien pour le produit scalaire $\langle A, B\rangle = \mathrm{trace}(AB^t)$ (voir l'exercice 10.1.6). On fixe $C \in V$ et on considère l'opérateur linéaire $T\colon V \to V$ donné par $T(A) = CA$ pour tout $A \in V$. Montrer que l'adjoint T^* de T est donné par $T^*(A) = C^t A$ pour tout $A \in V$.

Indice. Utiliser le fait que $\mathrm{trace}(AB) = \mathrm{trace}(BA)$ quels que soient $A, B \in V$.

10.4 Théorèmes spectraux

On fixe encore un espace euclidien V de dimension finie $n \geq 1$. On montre d'abord :

Lemme 10.4.1. *Soit $T\colon V \to V$ un opérateur linéaire sur V. Il existe un sous-espace T-invariant de V de dimension 1 ou 2.*

Preuve. On considère V comme un $\mathbb{R}[x]$-module pour le produit $p(x)\mathbf{v} = p(T)(\mathbf{v})$. Soit $\mathbf{v}$ un élément non nul de V. D'après la proposition 7.1.7, l'annulateur de $\mathbf{v}$ est engendré par un polynôme unitaire $p(x) \in \mathbb{R}[x]$ de degré ≥ 1. Écrivons $p(x) = q(x)a(x)$ où $q(x)$ est un facteur irréductible unitaire de $p(x)$, et posons $\mathbf{w} = a(x)\mathbf{v}$. Alors l'annulateur de $\mathbf{w}$ est l'idéal de $\mathbb{R}[x]$ engendré par $q(x)$. En vertu de la proposition 7.1.7, le sous-module $W := \mathbb{R}[x]\mathbf{w}$ est un sous-espace T-invariant de V de dimension égale au degré de $q(x)$. Or le théorème 5.6.5 montre que tous les polynômes irréductibles de $\mathbb{R}[x]$ sont de degré 1 ou 2. Donc W est de dimension 1 ou 2. □

Ce lemme suggère la définition suivante :

Définition 10.4.2. Soit $T\colon V \to V$ un opérateur linéaire sur V. On dit qu'un sous-espace T-invariant W de V est *minimal* ou *irréductible* si $W \neq \{\mathbf{0}\}$ et si les seuls sous-espaces T-invariants de V contenus dans W sont $\{\mathbf{0}\}$ et W.

Il découle de cette définition et du lemme 10.4.1 que tout sous-espace T-invariant non nul de V contient un sous-espace T-invariant minimal W, et qu'un tel sous-espace W est de dimension 1 ou 2.

Pour un opérateur orthogonal ou symétrique, on a en plus :

Lemme 10.4.3. *Soit $T\colon V \to V$ un opérateur linéaire orthogonal ou symétrique, et soit W un sous-espace T-invariant de V. Alors $W^\perp$ est aussi T-invariant.*

Preuve. Supposons d'abord que T soit orthogonal. Alors $T|_W\colon W \to W$ est un opérateur orthogonal sur W (il préserve la norme des éléments de W). Par suite, $T|_W$ est inversible et en particulier surjectif. Donc on a $T(W) = W$.

Soit $\mathbf{v} \in W^\perp$. Pour tout $\mathbf{w} \in W$, on trouve que

$$\langle T(\mathbf{v}), T(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle = 0,$$

donc $T(\mathbf{v}) \in (T(W))^\perp = W^\perp$. Cela montre que $W^\perp$ est T-invariant.

Supposons maintenant que T soit symétrique. Soit $\mathbf{v} \in W^\perp$. Pour tout $\mathbf{w} \in W$, on trouve

$$\langle T(\mathbf{v}), \mathbf{w}\rangle = \langle \mathbf{v}, T(\mathbf{w})\rangle = 0,$$

car $T(\mathbf{w}) \in W$. Par suite, on a $T(\mathbf{v}) \in W^\perp$. Le choix de $\mathbf{v} \in W^\perp$ étant arbitraire, cela montre que $W^\perp$ est T-invariant. □

Le lemme 10.4.3 implique aussitôt.

Proposition 10.4.4. *Soit $T\colon V \to V$ un opérateur linéaire orthogonal ou symétrique. Alors V est une somme directe de sous-espaces T-invariants minimaux*

$$V = W_1 \oplus \cdots \oplus W_s$$

qui sont orthogonaux deux à deux, c'est-à-dire tels que $W_i \subseteq W_j^\perp$ pour toute paire d'entiers distincts i et j avec $1 \leq i, j \leq s$.

Pour compléter notre analyse des opérateurs linéaires $T\colon V \to V$ qui sont orthogonaux ou symétriques, on est ainsi ramené à étudier la restriction de T à un sous-espace T-invariant minimal. Grâce au lemme 10.4.1, on sait qu'un tel sous-espace est de dimension 1 ou 2. On commence avec le cas d'un opérateur orthogonal.

Proposition 10.4.5. *Soit $T\colon V \to V$ un opérateur orthogonal, soit W un sous-espace T-invariant minimal de V, et soit $\mathcal{B}$ une base orthonormée de W. Alors ou bien on a*

$$\dim_{\mathbb{R}}(W) = 1 \quad \textit{et} \quad [T|_W]_{\mathcal{B}} = (\pm 1)$$

pour un choix de signe $\pm$, ou bien on a

$$\dim_{\mathbb{R}}(W) = 2 \quad \textit{et} \quad [T|_W]_{\mathcal{B}} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

pour un nombre réel θ.

Preuve. D'après le lemme 10.4.1, le sous-espace W est de dimension 1 ou 2. De plus la restriction $T|_W\colon W \to W$ de T à W est opérateur orthogonal car $T|_W$ préserve la norme (voir la définition 10.2.3). Donc $[T|_W]_{\mathcal{B}}$ est une matrice orthogonale, en vertu du corollaire 10.2.4.

1° Si $\dim_{\mathbb{R}}(W) = 1$, alors $[T|_W]_{\mathcal{B}} = (a)$ avec $a \in \mathbb{R}$. Puisque $[T|_W]_{\mathcal{B}}$ est une matrice orthogonale, on doit avoir $(a)^t(a) = 1$, donc $a^2 = 1$ et par suite $a = \pm 1$.

2° Si $\dim_{\mathbb{R}}(W) = 2$, alors les colonnes de $[T|_W]_{\mathcal{B}}$ forment une base orthonormée de $\mathbb{R}^2$ (voir l'exercice 10.2.7), et par suite cette matrice s'écrit

$$[T|_W]_{\mathcal{B}} = \begin{pmatrix} \cos(\theta) & -\epsilon\sin(\theta) \\ \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix}$$

avec $\theta \in \mathbb{R}$ et $\epsilon = \pm 1$. On en déduit que

$$\mathrm{car}_{T|_W}(x) = \begin{vmatrix} x - \cos(\theta) & \epsilon\sin(\theta) \\ -\sin(\theta) & x - \epsilon\cos(\theta) \end{vmatrix} = x^2 - (1+\epsilon)\cos(\theta)x + \epsilon.$$

Si $\epsilon = -1$, ce polynôme devient $\mathrm{car}_{T|_W}(x) = x^2 - 1 = (x-1)(x+1)$. Alors 1 et -1 sont des valeurs propres de $T|_W$, et pour chacune de ces valeurs propres l'opérateur $T|_W\colon W \to W$ possède un vecteur propre. Mais si $\mathbf{w} \in W$ est un vecteur propre de $T|_W$ alors $\mathbb{R}\mathbf{w}$ est un sous-espace T-invariant de V de dimension 1 contenu dans W, en contradiction avec l'hypothèse que W est un sous-espace T-invariant minimal. Donc, on doit avoir $\epsilon = 1$ et la preuve est complète. □

Dans le cas d'un opérateur symétrique, la situation est encore plus simple.

Proposition 10.4.6. *Soit $T\colon V \to V$ un opérateur symétrique, et soit W un sous-espace T-invariant minimal de V. Alors W est de dimension 1.*

Preuve. On sait que W est de dimension 1 ou 2. Supposons qu'il soit de dimension 2, et désignons par $\mathcal{B}$ une base orthonormée de W. La restriction $T|_W : W \to W$ est un opérateur symétrique sur W car pour tout choix de $\mathbf{w}_1, \mathbf{w}_2 \in W$, on a

$$\langle T|_W(\mathbf{w}_1), \mathbf{w}_2\rangle = \langle T(\mathbf{w}_1), \mathbf{w}_2\rangle = \langle \mathbf{w}_1, T(\mathbf{w}_2)\rangle = \langle \mathbf{w}_1, T|_W(\mathbf{w}_2)\rangle.$$

Par suite, suivant la proposition 10.3.4, la matrice $[T|_W]_{\mathcal{B}}$ est symétrique. Elle s'écrit donc

$$[T|_W]_{\mathcal{B}} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

avec $a, b, c \in \mathbb{R}$. On trouve ainsi

$$\mathrm{car}_{T|_W}(x) = \begin{vmatrix} x-a & -b \\ -b & x-c \end{vmatrix} = x^2 - (a+c)x + (ac - b^2).$$

Le discriminant de ce polynôme est

$$(a+c)^2 - 4(ac-b^2) = (a-c)^2 + 4b^2 \geq 0.$$

Donc $\mathrm{car}_{T|_W}(x)$ possède au moins une racine réelle. Cette dernière étant une valeur propre de $T|_W$, il lui correspond un vecteur propre $\mathbf{w}$. Alors $\mathbb{R}\mathbf{w}$ est un sous-espace T-invariant de V contenu dans W, en contradiction avec l'hypothèse de minimalité de W. Ainsi W est nécessairement de dimension 1. □

En combinant les trois dernières propositions, on obtient une description des opérateurs orthogonaux ou symétriques sur V.

Théorème 10.4.7. *Soit $T : V \to V$ un opérateur linéaire.*

(i) Si T est orthogonal, il existe une base orthonormée $\mathcal{B}$ de V telle que $[T]_{\mathcal{B}}$ soit diagonale par blocs avec des blocs de la forme

$$(1), \quad (-1) \quad ou \quad \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{10.8}$$

(avec $\theta \in \mathbb{R}$) sur la diagonale.

(ii) Si T est symétrique, il existe une base orthonormée $\mathcal{B}$ de V telle que $[T]_{\mathcal{B}}$ soit une matrice diagonale.

Le résultat (ii) s'exprime encore en disant qu'un opérateur symétrique est diagonalisable dans une base orthonormée.

Preuve. Selon la proposition 10.4.4, l'espace V se décompose en somme directe

$$V = W_1 \oplus \cdots \oplus W_s$$

de sous-espaces T-invariants minimaux $W_1, \dots, W_s$ qui sont orthogonaux deux à deux. Choisissons une base orthonormée $\mathcal{B}_i$ de W_i pour chaque $i = 1, \dots, s$. Alors

$$\mathcal{B} = \mathcal{B}_1 \amalg \cdots \amalg \mathcal{B}_s$$

est une base orthonormée de V telle que

$$[T]_{\mathcal{B}} = \begin{pmatrix} [T|_{W_1}]_{\mathcal{B}_1} & & 0 \\ & \ddots & \\ 0 & & [T|_{W_s}]_{\mathcal{B}_s} \end{pmatrix}.$$

Si T est un opérateur orthogonal, la proposition 10.4.5 montre que chacun des blocs $[T|_{W_i}]_{\mathcal{B}_i}$ sur la diagonale est de la forme (10.8). D'autre part, si T est symétrique, la proposition 10.4.6 montre que chaque W_i est de dimension 1. Donc, dans ce cas, $[T|_W]_{\mathcal{B}}$ est diagonale. □

Corollaire 10.4.8. *Soit $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ une matrice symétrique. Il existe une matrice orthogonale $P \in \mathcal{O}_n(\mathbb{R})$ telle que $P^{-1}AP = P^tAP$ soit diagonale.*

En particulier, ce corollaire nous apprend que toute matrice symétrique réelle est diagonalisable sur $\mathbb{R}$.

Preuve. Soit $\mathcal{E}$ la base canonique de $\mathbb{R}^n$ et soit $T_A\colon \mathbb{R}^n \to \mathbb{R}^n$ l'opérateur linéaire associé à A, de matrice $[T_A]_{\mathcal{E}} = A$. Comme A est symétrique et que $\mathcal{E}$ est une base orthonormée de $\mathbb{R}^n$, l'opérateur T_A est symétrique (proposition 10.3.4). En vertu du théorème 10.4.7, il existe donc une base orthonormée de $\mathcal{B}$ de $\mathbb{R}^n$ telle que $[T_A]_{\mathcal{B}}$ soit diagonale. En posant $P = [I]_{\mathcal{E}}^{\mathcal{B}}$, on obtient

$$[T_A]_{\mathcal{B}} = P^{-1}[T_A]_{\mathcal{E}}P = P^{-1}AP$$

(voir la section 2.6). Enfin, comme $\mathcal{B}$ et $\mathcal{E}$ sont deux bases orthonormées de $\mathbb{R}^n$, l'exercice 10.2.3 nous apprend que P est une matrice orthogonale, c'est-à-dire inversible avec $P^{-1} = P^t$. □

Exemple 10.4.9. Soit $\mathcal{E}$ la base canonique de $\mathbb{R}^3$ et soit $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'opérateur linéaire dont la matrice en base $\mathcal{E}$ est

$$[T]_{\mathcal{E}} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

On vérifie sans peine que $[T]_{\mathcal{E}}^t [T]_{\mathcal{E}} = I$. Comme $\mathcal{E}$ est une base orthonormée de $\mathbb{R}^3$, cela signifie que T est un opérateur orthogonal. On trouve

$$\mathrm{car}_T(x) = \begin{vmatrix} x & 0 & -1 \\ -1 & x & 0 \\ 0 & -1 & x \end{vmatrix} = x^3 - 1 = (x-1)(x^2+x+1),$$

donc 1 est la seule valeur propre réelle de T. Comme

$$[T]_{\mathcal{E}} - I = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

est de rang 2, on trouve que l'espace propre de T pour la valeur propre 1 est de dimension 1 engendré par le vecteur unitaire

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Alors $W_1 = \langle \mathbf{u}_1 \rangle_{\mathbb{R}}$ est un sous-espace T-invariant de $\mathbb{R}^3$ et par suite on obtient la décomposition

$$\mathbb{R}^3 = W_1 \oplus W_1^{\perp}$$

où

$$W_1^{\perp} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \; ; \; x + y + z = 0 \right\}$$

est aussi un sous-espace T-invariant de $\mathbb{R}^3$. On trouve que

$$\left\{ \mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{u}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\}$$

est une base orthonormée de $W_1^{\perp}$. Le calcul donne

$$T(\mathbf{u}_2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = -\frac{1}{2}\mathbf{u}_2 + \frac{\sqrt{3}}{2}\mathbf{u}_3, \quad T(\mathbf{u}_3) = \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = -\frac{\sqrt{3}}{2}\mathbf{u}_2 - \frac{1}{2}\mathbf{u}_3.$$

On a aussi $T(\mathbf{u}_1) = \mathbf{u}_1$ puisque $\mathbf{u}_1$ est un vecteur propre de T pour la valeur propre 1. On conclut que $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ est une base orthonormée de $\mathbb{R}^3$ telle que

$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & -1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi/3) & -\sin(2\pi/3) \\ 0 & \sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix}$$

On lit sur cette matrice que T fixe chacun des vecteurs de la droite W_1 tandis qu'il effectue une rotation d'angle $2\pi/3$ dans le plan $W_1^{\perp}$. Donc T est une rotation d'angle $2\pi/3$ autour de la droite W_1.

Note. Cette dernière description de T ne la détermine pas entièrement car T^{-1} est aussi une rotation d'angle $2\pi/3$ autour de W_1. Pour distinguer T de T^{-1}, il faudrait encore préciser le sens de cette rotation.

Exercices.

10.4.1. Soit $R\colon \mathbb{R}^4 \to \mathbb{R}^4$ l'application linéaire dont la matrice relative à la base canonique $\mathcal{E}$ de $\mathbb{R}^4$ est

$$[R]_{\mathcal{E}} = \frac{1}{2}\begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}.$$

(i) Montrer que R est un opérateur orthogonal.

(ii) Déterminer une base orthonormée $\mathcal{B}$ de $\mathbb{R}^3$ dans laquelle la matrice de R est diagonale par blocs avec des blocs de taille 1 ou 2 sur la diagonale, et donner $[R]_{\mathcal{B}}$.

10.4.2. Soit $A = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}$.

(i) Montrer que A est une matrice orthogonale.

(ii) Déterminer une matrice orthogonale U pour laquelle le produit $U^t\,A\,U$ est diagonal par blocs avec des blocs de taille 1 ou 2 sur la diagonale, et donner ce produit.

10.4.3. Pour chacune des matrices A suivantes, déterminer une matrice orthogonale U telle que $U^t\,A\,U$ soit une matrice diagonale et donner cette matrice diagonale.

$$\text{(i) } A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad \text{(ii) } A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{(iii) } A = \begin{pmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{pmatrix}.$$

10.4.4. Soit T un opérateur symétrique sur un espace euclidien V de dimension n. Supposons que T admette n valeurs propres distinctes $\lambda_1, \ldots, \lambda_n$ et que $\mathbf{v}_1, \ldots, \mathbf{v}_n$ soient des vecteurs propres respectifs pour ces valeurs propres. Montrer que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ est une base orthogonale de V.

10.4.5. Soit T un opérateur symétrique sur un espace euclidien V de dimension finie. Montrer que les conditions suivantes sont équivalentes :

(i) Toutes les valeurs propres de T sont positives.

(ii) On a $\langle T(\mathbf{v}), \mathbf{v}\rangle > 0$ pour tout $\mathbf{v} \in V$ avec $\mathbf{v} \neq 0$.

10.4.6. Soit V un espace euclidien de dimension finie, et soient S et T des opérateurs symétriques sur V qui commutent entre eux (i.e. $S \circ T = T \circ S$). Montrer que chacun des espaces propres de T est S-invariant. Déduire de cette observation qu'il existe une base orthogonale $\mathcal{B}$ de V telle que $[S]_{\mathcal{B}}$ et $[T]_{\mathcal{B}}$ soient toutes deux diagonales. (On dit que S et T sont *simultanément* diagonalisables.)

10.5 Décomposition polaire

Soit V un espace euclidien de dimension finie.

Définition 10.5.1. On dit qu'un opérateur symétrique $S\colon V \to V$ est *défini positif* si la forme bilinéaire symétrique

$$\begin{array}{ccc} V \times V & \longrightarrow & \mathbb{R} \\ (\mathbf{u}, \mathbf{v}) & \longmapsto & \langle S(\mathbf{u}), \mathbf{v}\rangle = \langle \mathbf{u}, S(\mathbf{v})\rangle \end{array} \tag{10.9}$$

qui lui est associée est définie positive, c'est-à-dire satisfait

$$\langle S(\mathbf{v}), \mathbf{v}\rangle > 0 \quad \text{pour tout} \quad \mathbf{v} \in V \setminus \{\mathbf{0}\}.$$

En vertu de la définition 10.1.1, cela revient à demander que (10.9) constitue un produit scalaire sur V.

Exemple 10.5.2. Soit $T\colon V \to V$ un opérateur inversible quelconque. Alors $T^* \circ T\colon V \to V$ est un opérateur symétrique défini positif.

En effet, les propriétés de l'adjointe données par la proposition 10.3.5 montrent que

$$(T^* \circ T)^* = T^* \circ (T^*)^* = T^* \circ T,$$

donc $T^* \circ T$ est un opérateur symétrique. Par ailleurs, pour tout $\mathbf{v} \in V \setminus \{\mathbf{0}\}$, on a $T(\mathbf{v}) \neq \mathbf{0}$, donc

$$\langle (T^* \circ T)(\mathbf{v}), \mathbf{v}\rangle = \langle T^*(T(\mathbf{v})), \mathbf{v}\rangle = \langle T(\mathbf{v}), T(\mathbf{v})\rangle > 0.$$

Cela signifie que $T^* \circ T$ est aussi défini positif.

Proposition 10.5.3. *Soit $S\colon V \to V$ un opérateur linéaire. Les conditions suivantes sont équivalentes :*

(i) S est un opérateur symétrique défini positif;

(ii) Il existe une base orthonormée $\mathcal{B}$ de V telle que $[S]_{\mathcal{B}}$ soit une matrice diagonale avec des nombres positifs sur la diagonale.

Preuve. $(i) \Longrightarrow (ii)$: Supposons d'abord que S soit symétrique défini positif. Puisque S est symétrique, le théorème 10.4.7 nous apprend qu'il existe une base orthonormée $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V telle que $[S]_{\mathcal{B}}$ s'écrive

$$[S]_{\mathcal{B}} = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix} \tag{10.10}$$

avec $\sigma_1, \ldots, \sigma_n \in \mathbb{R}$. Alors, pour $j = 1, \ldots, n$, on a $S(\mathbf{u}_j) = \sigma_j \mathbf{u}_j$ et par suite

$$\langle S(\mathbf{u}_j), \mathbf{u}_j\rangle = \sigma_j \langle \mathbf{u}_j, \mathbf{u}_j\rangle = \sigma_j \|\mathbf{u}_j\|^2 = \sigma_j.$$

Comme S est défini positif, on conclut que les éléments $\sigma_1, \ldots, \sigma_n$ de la diagonale de $[S]_{\mathcal{B}}$ sont positifs.

$(ii) \Longrightarrow (i)$: Réciproquement, supposons que $[S]_{\mathcal{B}}$ s'écrive sous la forme (10.10) pour une base orthonormée $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ et des nombres réels positifs $\sigma_1, \ldots, \sigma_n$. Comme $[S]_{\mathcal{B}}$ est symétrique, la proposition 10.3.4 montre que S est un opérateur symétrique. De plus, si $\mathbf{v}$ est un vecteur non nul de V, il s'écrit $\mathbf{v} = a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n$ avec $a_1, \ldots, a_n \in \mathbb{R}$ non tous nuls, et on trouve

$$\langle S(\mathbf{v}), \mathbf{v}\rangle = \Big\langle \sum_{j=1}^{n} a_j S(\mathbf{u}_j), \sum_{j=1}^{n} a_j \mathbf{u}_j \Big\rangle = \Big\langle \sum_{j=1}^{n} a_j \sigma_j \mathbf{u}_j, \sum_{j=1}^{n} a_j \mathbf{u}_j \Big\rangle = \sum_{j=1}^{n} a_j^2 \sigma_j > 0,$$

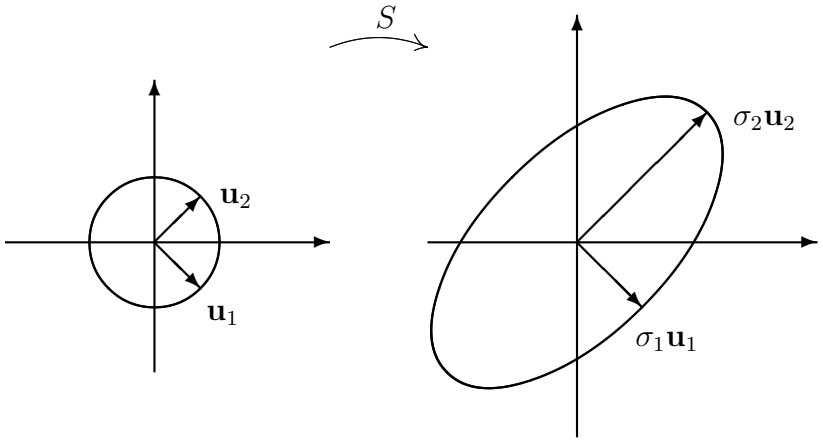
donc S est défini positif. □

Signification géométrique dans $\mathbb{R}^2$

Soit $S\colon \mathbb{R}^2 \to \mathbb{R}^2$ un opérateur symétrique sur $\mathbb{R}^2$, et soit $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$ une base orthonormée de $\mathbb{R}^2$ telle que

$$[S]_{\mathcal{B}} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

avec $\sigma_1 > 0$ et $\sigma_2 > 0$. Alors S peut être vue comme la composée de deux "dilations" selon des axes orthogonaux l'une par un facteur σ_1 dans la direction de $\mathbf{u}_1$ et l'autre par un facteur σ_2 dans la direction de $\mathbf{u}_2$. L'effet global est que S applique le cercle de rayon 1 centré à l'origine sur une ellipse de demi-axes $\sigma_1\mathbf{u}_1$ et $\sigma_2\mathbf{u}_2$.



Proposition 10.5.4. *Soit $S\colon V \to V$ un opérateur symétrique défini positif. Il existe un et un seul opérateur symétrique défini positif $P\colon V \to V$ tel que $S = P^2$.*

En d'autres termes, un opérateur symétrique défini positif admet une et une seule racine carrée du même type.

Preuve. D'après la proposition 10.5.3, il existe une base orthonormée $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ de V telle que

$$[S]_{\mathcal{B}} = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix}$$

avec $\sigma_1, \ldots, \sigma_n > 0$. Soit $P\colon V \to V$ l'opérateur linéaire pour lequel

$$[\,P\,]_{\mathcal{B}} = \begin{pmatrix} \sqrt{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_n} \end{pmatrix}.$$

Comme $[\,S\,]_{\mathcal{B}} = ([\,P\,]_{\mathcal{B}})^2 = [\,P^2\,]_{\mathcal{B}}$, on a $S = P^2$. Comme en plus $[\,P\,]_{\mathcal{B}}$ est diagonale avec des nombres positifs sur sa diagonale, la proposition 10.5.3 nous apprend que c'est un opérateur symétrique défini positif. Pour l'unicité de P, voir l'exercice 10.5.1. □

On conclut ce chapitre avec le résultat suivant qui fournit une décomposition des opérateurs linéaires inversibles $T\colon V \to V$ semblable à la décomposition polaire d'un nombre complexe $z \neq 0$ sous la forme $z = r\,u$ où r est un nombre réel positif et u un nombre complexe de norme 1 (voir l'exercice 10.5.3).

Théorème 10.5.5. *Soit $T\colon V \to V$ un opérateur inversible. Il existe un et un seul opérateur symétrique défini positif $P\colon V \to V$ et un et un seul opérateur orthogonal $U\colon V \to V$ tels que $T = U \circ P$.*

Preuve. L'exemple 10.5.2 nous apprend que $T^* \circ T\colon V \to V$ est un opérateur symétrique défini positif. Selon la proposition 10.5.4, il existe un opérateur symétrique $P\colon V \to V$ défini positif tel que

$$T^* \circ T = P^2.$$

Comme P est défini positif, c'est un opérateur inversible. On pose

$$U = T \circ P^{-1}.$$

Alors, on a

$$U^* = (P^{-1})^* \circ T^* = (P^*)^{-1} \circ T^* = P^{-1} \circ T^*$$

et par suite

$$U^* \circ U = P^{-1} \circ (T^* \circ T) \circ P^{-1} = P^{-1} \circ P^2 \circ P^{-1} = I.$$

Cela montre que U est un opérateur orthogonal (voir l'exercice 10.3.3). Ainsi on a $T = U \circ P$ comme requis. Pour l'unicité de P et U, voir l'exercice 10.5.2. □

Exemple 10.5.6. Soit $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'opérateur linéaire donné par

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x + 2y \end{pmatrix}.$$

Sa matrice relative à la base canonique $\mathcal{E}$ de $\mathbb{R}^2$ est

$$[\,T\,]_{\mathcal{E}} = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}.$$

Le calcul donne

$$[\,T^* \circ T\,]_{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix},$$

puis que le polynôme caractéristique de $T^* \circ T$ est $x^2 - 10x + 16 = (x-2)(x-8)$. Donc les valeurs propres de cet opérateur sont 2 et 8. On trouve que $(1,-1)^t$ et $(1,1)^t$ sont des vecteurs propres de $T^* \circ T$ pour ces valeurs propres. En vertu de l'exercice 10.4.4, ces derniers forment une base orthogonale de $\mathbb{R}^2$. Alors,

$$\mathcal{B} = \left\{ \mathbf{u}_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} , \mathbf{u}_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

est une base orthonormée de $\mathbb{R}^2$ telle que

$$[\, T^* \circ T \,]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}.$$

Par suite on a $T^* \circ T = P^2$ où $P \colon \mathbb{R}^2 \to \mathbb{R}^2$ est l'opérateur symétrique défini positif pour lequel

$$[\, P \,]_{\mathcal{B}} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{8} \end{pmatrix}.$$

On trouve que

$$[\, I \,]_{\mathcal{E}}^{\mathcal{B}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{et} \quad [\, I \,]_{\mathcal{B}}^{\mathcal{E}} = ([\, I \,]_{\mathcal{E}}^{\mathcal{B}})^{-1} = ([\, I \,]_{\mathcal{E}}^{\mathcal{B}})^t = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

où le dernier calcul utilise le fait que, $\mathcal{E}$ et $\mathcal{B}$ étant des bases orthonormées de $\mathbb{R}^2$, la matrice $[\, I \,]_{\mathcal{E}}^{\mathcal{B}}$ est orthogonale (voir l'exercice 10.2.3). La matrice de P dans la base canonique est donc

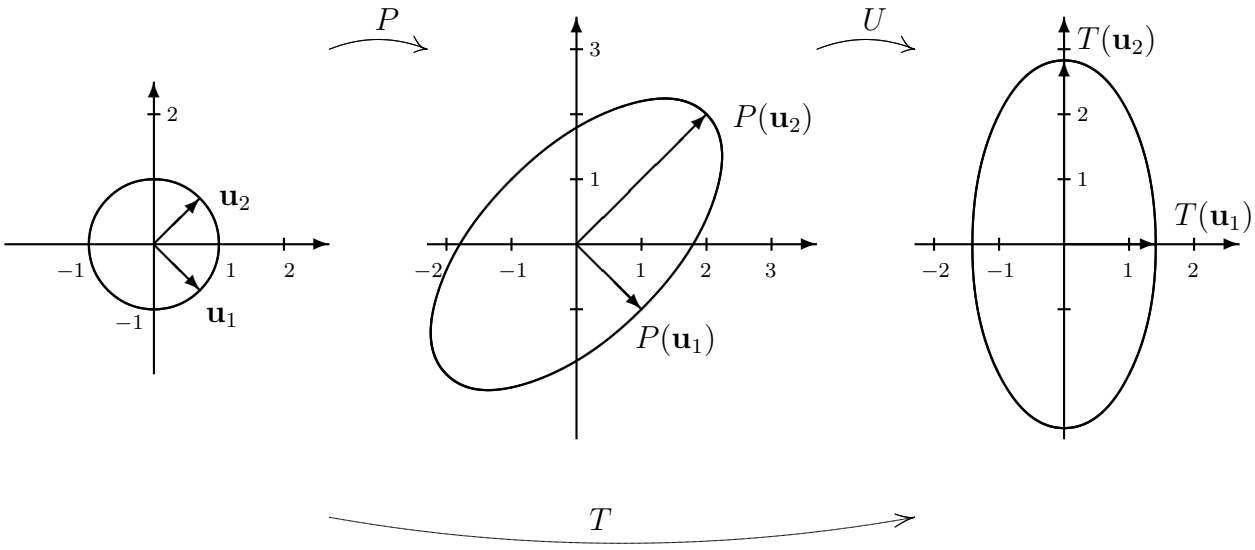
$$[\, P \,]_{\mathcal{E}} = [\, I \,]_{\mathcal{E}}^{\mathcal{B}} \, [\, P \,]_{\mathcal{B}} \, [\, I \,]_{\mathcal{B}}^{\mathcal{E}} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

On note en passant que cette matrice est bien symétrique, comme il se doit. Enfin, en poursuivant avec les étapes de la preuve du théorème 10.5.5, on pose $U = T \circ P^{-1}$. On trouve

$$[\, U \,]_{\mathcal{E}} = \begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix} \frac{1}{4\sqrt{2}} \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix}$$

Sous cette forme, on reconnaît que U est la rotation d'angle $\pi/4$ autour de l'origine dans $\mathbb{R}^2$. En particulier, U est bien un opérateur orthogonal et on a obtenu la décomposition requise $T = U \circ P$, avec P opérateur symétrique qui dilate les vecteurs par un facteur $\sqrt{2}$ dans la

direction de $\mathbf{u}_1$ et par un facteur $\sqrt{8}$ dans la direction de $\mathbf{u}_2$.



Exercices.

10.5.1. Soient $P\colon V \to V$ et $S\colon V \to V$ des opérateurs symétriques définis positifs tels que $P^2 = S$. Montrer que, pour toute valeur propre λ de P, l'espace propre de P pour la valeur propre λ coïncide avec celui de S pour la valeur propre λ^2. Conclure que P est le seul opérateur symétrique défini positif sur V tel que $P^2 = S$.

10.5.2. Soit $P\colon V \to V$ un opérateur symétrique défini positif, soit $U\colon V \to V$ un opérateur orthogonal, et soit $T = U \circ P$. Montrer que $T^* \circ T = P^2$. A l'aide de l'exercice précédent, en déduire que P est le seul opérateur symétrique défini positif et U le seul opérateur orthogonal tels que $T = U \circ P$.

10.5.3. On considère $\mathbb{C}$ comme un espace vectoriel réel et on le munit du produit scalaire

$$\langle a + ib, c + id\rangle = ac + bd.$$

On fixe un nombre complexe $z \neq 0$ et on considère l'application $\mathbb{R}$-linéaire $T\colon \mathbb{C} \to \mathbb{C}$ donnée par $T(w) = zw$ pour tout $w \in \mathbb{C}$.

(i) Vérifier que $\langle w_1, w_2\rangle = \mathrm{Re}(w_1\overline{w_2})$ quels que soient $w_1, w_2 \in \mathbb{C}$.

(ii) Montrer que l'adjoint T^* de T est donné par $T^*(w) = \overline{z}w$ pour tout $w \in \mathbb{C}$.

(iii) Soit $r = |z|$ et soit $u = r^{-1}z$. Montrer que l'opérateur $P\colon \mathbb{C} \to \mathbb{C}$ de multiplication par r est symétrique défini positif, que l'opérateur $U\colon \mathbb{C} \to \mathbb{C}$ de multiplication par u est orthogonal et que $T = U \circ P$ est la décomposition polaire de T.

Appendices

Annexe A

Rappels : groupes, anneaux et corps

Tout au long de la session, nous ferons appel à des notions d'algèbre acquises dans le cours MAT 2543. En particulier, nous ferons beaucoup appel à la notion d'anneau qui est généralement seulement effleurée en MAT 2543. Cet appendice présente un rappel des notions de groupe, anneau et corps avec leurs propriétés les plus simples. Nous omettons la plupart des preuves déjà couvertes en MAT 2543.

A.1 Monoïdes

Une *opération binaire* sur un ensemble E est une fonction $*\colon E \times E \to E$ qui, à chaque couple d'éléments (a, b) de E associe un nouvel élément de E noté $a * b$. Le symbole utilisé pour désigner l'opération peut varier. Le plus souvent, lorsque cela ne cause pas de confusion, on écrit simplement ab pour désigner le résultat d'une opération appliquée à un couple $(a, b) \in E \times E$.

Définition A.1.1. On dit qu'une *opération binaire* $*$ sur un ensemble E est :

- *associative* si $(a * b) * c = a * (b * c)$ quels que soient $a, b, c \in E$
- *commutative* si $a * b = b * a$ quels que soient $a, b \in E$.

Exemples A.1.2. L'addition et la multiplication dans $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des opérations qui sont à la fois associatives et commutatives. Par contre, pour un entier $n \geq 2$, la multiplication dans l'ensemble $\mathrm{Mat}_{n\times n}(\mathbb{R})$ des matrices carrées de dimension n à coefficients dans $\mathbb{R}$ est associative mais non commutative. De même la composition dans l'ensemble $\mathrm{End}(S)$ des applications d'un ensemble S dans lui-même est associative mais elle n'est pas commutative si $|S| \geq 3$. Le produit vectoriel sur $\mathbb{R}^3$ est une opération binaire non commutative et non associative.

Un fait important est que si une opération $*$ sur un ensemble E est associative, alors la manière de grouper les termes dans un produit d'un nombre fini d'éléments $a_1, \ldots, a_n$ de E

n'affecte pas le résultat, pourvu que l'ordre soit préservé. Par exemple, il y a cinq manières de faire le produit de quatre éléments $a_1, \ldots, a_4$ de E tout en respectant leur ordre. Ce sont :

$((a_1 * a_2) * a_3) * a_4$, $(a_1 * (a_2 * a_3)) * a_4$, $(a_1 * a_2) * (a_3 * a_4)$, $a_1 * (a_2 * (a_3 * a_4))$, $a_1 * ((a_2 * a_3) * a_4)$.

On peut vérifier que si $*$ est associative, alors tous ces produits sont égaux. Pour cette raison, si $*$ est une opération associative, on écrit simplement

$$a_1 * a_2 * \cdots * a_n$$

pour désigner le produit d'éléments $a_1, \ldots, a_n$ de E, dans cet ordre (sans mettre de parenthèses).

Si l'opération $*$ est à la fois associative et commutative, alors l'ordre dans lequel on multiplie les éléments de E n'affecte pas leur produit. On a par exemple

$$a_1 * a_2 * a_3 = a_1 * a_3 * a_2 = a_3 * a_1 * a_2 = \cdots .$$

Ici, le verbe "multiplier" et le terme "produit" sont pris au sens générique : "multiplier" signifie "appliquer l'opération" et le mot "produit" signifie le "résultat de l'opération". Dans un contexte précis, on utilisera les termes qui s'appliquent.

Cas particulier : Le symbole $+$ est réservé pour désigner une opération associative et commutative. On dit alors que l'opération est une *addition*, on parle d'*additionner* des éléments. Le résultat d'une addition s'appelle une *somme*.

Définition A.1.3. Soit $*$ une opération binaire sur un ensemble E. On dit qu'un élément e de E est *élément neutre* pour $*$ si

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

pour tout $a \in E$.

On montre que si E possède un élément neutre pour une opération $*$, alors il n'en possède qu'un seul. Généralement, on désigne cet élément par 1 mais si l'opération est notée $+$, on le désigne par 0.

Exemples A.1.4. Soit n un entier positif. L'élément neutre pour l'addition dans $\mathrm{Mat}_{n\times n}(\mathbb{R})$ est la matrice O_n de format $n \times n$ dont tous les coefficients sont nuls. L'élément neutre pour la multiplication dans $\mathrm{Mat}_{n\times n}(\mathbb{R})$ est la matrice I_n de format $n \times n$ dont les éléments de la diagonale sont égaux à 1 et les autres égaux à 0.

Définition A.1.5. Un *monoïde* est un ensemble E muni d'une opération associative pour laquelle E possède un élément neutre.

Un monoïde est donc une paire $(E, *)$ formée d'un ensemble E et d'une opération $*$ sur E avec les propriétés ci-dessus. Quand on parle d'un monoïde E sans préciser l'opération, c'est que celle-ci est sous-entendue. Dans ce cas, on écrit généralement ab pour désigner le produit de deux éléments a et b de E.

Exemple A.1.6. Les paires $(\mathbb{Z},+)$, $(\mathbb{Z},\cdot)$, $(\mathbb{N},+)$ et $(\mathbb{N},\cdot)$ sont des monoïdes (rappelons que $\mathbb{N}=\{0,1,2,\ldots\}$ désigne l'ensemble des entiers positifs ou nul). Pour tout entier $n\geq 1$, les paires $(\mathrm{Mat}_{n\times n}(\mathbb{R}),+)$ et $(\mathrm{Mat}_{n\times n}(\mathbb{R}),\cdot)$ sont aussi des monoïdes. Enfin, si S est un ensemble, alors $(\mathrm{End}(S),\circ)$ est un monoïde.

Définition A.1.7. Soit $(E,*)$ un monoïde. Pour tout entier $n\geq 1$ et tout $a\in E$, on désigne par

$$a^n=\underbrace{a*a*\cdots*a}_{n\text{ fois}}$$

le produit de n copies de a. Si l'opération est notée $+$ (donc commutative), on écrit plutôt

$$n\,a=\underbrace{a+a+\cdots+a}_{n\text{ fois}}$$

pour désigner la somme de n copies de a.

Dans ce contexte, on a la règle suivante dite :

Règle des exposants. *Si $(E,*)$ est un monoïde, on a*

$$a^m*a^n=a^{m+n}\quad\text{et}\quad(a^m)^n=a^{mn}$$

pour tout $a\in E$ et toute paire d'entiers $m,n\geq 1$. De plus, si $(E,*)$ est un monoïde commutatif*, c'est-à-dire si $*$ est commutative (en plus d'être associative), on a aussi*

$$a^m*b^m=(a*b)^m$$

pour tout entier $m\geq 1$ et tout choix de $a,b\in E$.

Enfin, si $(E,+)$ est un monoïde commutatif noté additivement, ces formules deviennent

$$ma+na=(m+n)a,\quad n(ma)=(nm)a\quad\text{et}\quad ma+mb=m(a+b)$$

quels que soient $a,b\in E$ et les entiers $m,n\geq 1$.

A.2 Groupes

La notion de groupe est centrale en algèbre. Dans ce cours, les groupes qui interviendront sont surtout des groupes abéliens mais comme on rencontrera aussi des groupes non abéliens, il convient de rappeler la notion plus générale. On rappelle d'abord :

Définition A.2.1. Soit $(E,*)$ un monoïde. On dit qu'un élément a de E est *inversible* (pour l'opération $*$) s'il existe $b\in E$ tel que

$$a*b=b*a=1 \tag{A.1}$$

On montre que si a est un élément inversible d'un monoïde $(E, *)$, alors il existe un seul élément b de E qui remplit la condition (A.1). On l'appelle l'*inverse* de a et on le note a^{-1}. On a donc

$$a * a^{-1} = a^{-1} * a = 1 \tag{A.2}$$

pour tout élément inversible a de $(E, *)$. On montre aussi :

Lemme A.2.2. *Soient a et b des éléments inversibles d'un monoïde $(E, *)$. Alors*

(i) a^{-1} est inversible et $(a^{-1})^{-1} = a$,

(ii) ab est inversible et $(ab)^{-1} = b^{-1}a^{-1}$.

Cas particulier important : Si l'opération du monoïde est commutative notée $+$, on écrit $-a$ pour désigner l'inverse d'un élément inversible a et on a :

$$a + (-a) = 0.$$

Si a, b sont deux éléments inversibles de $(E, +)$, les énoncés (i) et (ii) du lemme A.2.2 deviennent :

(i') $-a$ est inversible et $-(-a) = a$.

(ii') $a + b$ est inversible et $-(a + b) = (-a) + (-b)$.

Exemples A.2.3. - Tous les éléments de $(\mathbb{Z}, +)$ sont inversibles.

- L'ensemble des éléments inversibles de $(\mathbb{Z}, \cdot)$ est $\{1, -1\}$.
- Si S est un ensemble, les éléments inversibles de $(\mathrm{End}(S), \circ)$ sont les fonctions bijectives $f : S \to S$.

Définition A.2.4. Un *groupe* est un monoïde $(G, *)$ dont tous les éléments sont inversibles.

Si on explicite cette définition en tenant compte des définitions précédentes, un groupe est un ensemble G muni d'une opération binaire $*$ qui possède les propriétés suivantes :

G1. $(a * b) * c = a * (b * c)$ quels que soient $a, b, c \in G$;

G2. il existe $1 \in G$ tel que $1 * a = a * 1 = a$ quel que soit $a \in G$;

G3. pour tout $a \in G$, il existe $b \in G$ tel que $a * b = b * a = 1$.

Définition A.2.5. On dit qu'un groupe $(G, *)$ est *abélien* ou *commutatif* si l'opération est commutative.

Exemples A.2.6. - Les paires $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +)$ constituent des groupes abéliens.

- Si S est un ensemble, l'ensemble $\mathrm{Aut}(S)$ des bijections de S dans lui-même est un groupe sous la composition.

Le deuxième exemple est un cas particulier du résultat suivant :

Proposition A.2.7. *L'ensemble E^* des éléments inversibles d'un monoïde E est un groupe pour l'opération de E restreinte à E^*.*

Preuve. Le lemme A.2.2 montre que si $a, b \in E^*$, alors leur produit ab demeure dans E^*. Donc par restriction, l'opération de E fournit une opération binaire sur E^*. Elle est associative et, comme $1 \in E^*$, elle possède un élément neutre dans E^*. Enfin, pour tout $a \in E^*$, le lemme A.2.2 montre que son inverse a^{-1} (dans E) appartient à E^*. Comme il satisfait $a\,a^{-1} = a^{-1}\,a = 1$, c'est aussi un inverse de a dans E^*. □

Exemple A.2.8. Si S est un ensemble, les éléments inversibles de $(\mathrm{End}(S), \circ)$ sont les fonctions bijectives de S dans lui-même. Donc l'ensemble $\mathrm{Aut}(S)$ des bijections $f\colon S \to S$ est un groupe sous la composition.

Exemple A.2.9. Soit un entier $n \geq 1$. On sait que l'ensemble $\mathrm{Mat}_{n\times n}(\mathbb{R})$ est un monoïde sous la multiplication. Les éléments inversibles de ce monoïde sont les matrices $n \times n$ de déterminant non nul. Ceux-ci forment donc un groupe sous la multiplication. On note ce groupe $\mathrm{GL}_n(\mathbb{R})$:

$$\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathrm{Mat}_{n\times n}(\mathbb{R})\,;\, \det(A) \neq 0\}.$$

On l'appelle le *groupe linéaire général d'ordre n sur $\mathbb{R}$*.

Définition A.2.10. Soit a un élément d'un groupe G. Pour tout entier $n \in \mathbb{Z}$, on définit

$$a^n := \begin{cases} \underbrace{a\,a\,\cdots\,a}_{n \text{ fois}} & \text{si } n \geq 1, \\ 1 & \text{si } n = 0, \\ \underbrace{a^{-1}\cdots a^{-1}}_{-n \text{ fois}} & \text{si } n \leq -1. \end{cases}$$

Alors la règle des exposants devient :

Proposition A.2.11. *Soit a un élément d'un groupe G. Pour tout choix d'entiers $m, n \in \mathbb{Z}$, on a :*

(i) $a^m a^n = a^{m+n}$

(ii) $(a^m)^n = a^{mn}$.

Si G est abélien et si $a, b \in G$, on a aussi

(iii) $(ab)^m = a^m b^m$ *pour tout* $m \in \mathbb{Z}$.

La règle des exposants implique comme cas particulier $(a^{-1})^{-1} = a^{(-1)(-1)} = a^1 = a$.

Dans le cas d'un groupe abélien $(G, +)$ noté additivement, on définit plutôt, pour tout $n \in \mathbb{Z}$,

$$na = \begin{cases} \underbrace{a + a + \cdots + a}_{n \text{ fois}} & \text{si } n \geq 1, \\ 0 & \text{si } n = 0, \\ \underbrace{(-a) + \cdots + (-a)}_{-n \text{ fois}} & \text{si } n \leq -1, \end{cases}$$

et la règle devient

(i) $(ma) + (na) = (m+n)a$,
(ii) $m(na) = (mn)a$,
(iii) $m(a+b) = ma + mb$,
pour tout choix d'entiers $m, n \in \mathbb{Z}$ et d'éléments $a, b \in G$.

A.3 Sous-groupes

Définition A.3.1. Un *sous-groupe* d'un groupe G est un sous-ensemble H de G tel que
SG1. $1 \in H$,
SG2. $ab \in H$ pour tout choix de $a, b \in H$,
SG3. $a^{-1} \in H$ pour tout $a \in H$.

On montre sans peine qu'un sous-groupe H d'un groupe G est lui-même un groupe pour l'opération de G restreinte à H (grï£¡ce à **SG2**, le produit dans G induit par restriction un produit sur H).

Exemples A.3.2. - L'ensemble $2\mathbb{Z}$ des entiers pairs est un sous-groupe de $(\mathbb{Z}, +)$.
- L'ensemble $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ des nombres réels non nuls est un groupe pour la multiplication et l'ensemble $\mathbb{R}_+^* = \{x \in \mathbb{R}\,;\, x > 0\}$ des nombres réels positifs est un sous-groupe de $(\mathbb{R}^*, \cdot)$.

On rappelle aussi que l'ensemble des sous-groupes d'un groupe G est stable sous l'intersection.

Proposition A.3.3. *Soit $(G_i)_{i\in I}$ une collection de sous-groupes d'un groupe G. Alors leur intersection $\cap_{i\in I} G_i$ est encore un sous-groupe de G.*

Par exemple, si X est un sous-ensemble d'un groupe G, l'intersection de tous les sous-groupes de G qui contiennent X est un sous-groupe de G et, comme il contient X, c'est le plus petit sous-groupe de G qui contienne X. On l'appelle le *sous-groupe de G engendré par X*, et on le note $\langle X \rangle$. Dans le cas où X est réduit à un élément, on trouve

Proposition A.3.4. *Soit a un élément d'un groupe G. Le sous-groupe de G engendré par a est*

$$\langle a \rangle = \{a^n\,;\, n \in \mathbb{Z}\}.$$

Preuve. Posons $H = \{a^n\,;\, n \in \mathbb{Z}\}$. On a

$$1 = a^0 \in H.$$

De plus pour tout choix de $m, n \in \mathbb{Z}$, la règle des exposants (proposition A.2.11) donne

$$a^m a^n = a^{m+n} \in H \quad \text{et} \quad (a^m)^{-1} = a^{-m} \in H.$$

Donc H est un sous-groupe de G. Il contient $a^1 = a$. Comme tout sous-groupe de G qui contient a contient aussi H tout entier, H est le plus petit sous-groupe de G qui contienne a, et par suite $H = \langle a \rangle$. □

Que G soit abélien ou non, le sous-groupe $\langle a \rangle$ engendré par un élément est toujours abélien. Si G est abélien avec sa loi de groupe notée additivement, alors le sous-groupe engendré par un élément a de G s'écrit plutôt :

$$\langle a \rangle = \{n\, a\,;\, n \in \mathbb{N}\}.$$

Définition A.3.5. On dit qu'un groupe G est *cyclique* si G est engendré par un élément, c'est-à-dire s'il existe $a \in G$ tel que $G = \langle a \rangle$.

Exemple A.3.6. Le groupe $(\mathbb{Z}, +)$ est cyclique engendré par 1.

Exemple A.3.7. Soit n un entier positif. L'ensemble $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ des nombres complexes non nuls est un groupe sous la multiplication et l'ensemble

$$C_n = \{z \in \mathbb{C}^*\,;\, z^n = 1\}$$

des racines n-ièmes de l'unité est le sous-groupe de $\mathbb{C}^*$ engendré par $e^{2\pi i/n} = \cos(2\pi/n) + i\sin(2\pi/n)$.

A.4 Homorphismes de groupes

Définition A.4.1. Soient G et H deux groupes. Un *homomorphisme* de G dans H est une fonction $\varphi : G \to H$ telle que

$$\varphi(a\, b) = \varphi(a)\, \varphi(b)$$

pour toute paire d'éléments a, b de G.

On rappelle que si $\varphi : G \to H$ est un homomorphisme de groupes, alors :

$$\varphi(1_G) = 1_H \quad \text{et} \quad \varphi(a^{-1}) = \varphi(a)^{-1} \text{ pour tout } a \in G,$$

où 1_G et 1_H désignent respectivement les éléments neutres de G et de H. Plus généralement, on a

$$\varphi(a^n) = \varphi(a)^n$$

pour tout $a \in G$ et tout $n \in \mathbb{Z}$.

Dans le cas où G et H sont des groupes abéliens notés additivement, il faut interpréter la définition A.4.1 de la manière suivante : un *homomorphisme* de G dans H est une fonction $\varphi : G \to H$ telle que

$$\varphi(a + b) = \varphi(a) + \varphi(b)$$

pour toute paire d'éléments a, b de G. Une telle fonction satisfait $\varphi(0_G) = 0_H$ et $\varphi(-a) = -\varphi(a)$ pour tout $a \in G$, et plus généralement $\varphi(na) = n\varphi(a)$ pour tout $n \in \mathbb{Z}$ et tout $a \in G$.

On rappelle aussi les résultats suivants :

Proposition A.4.2. *Soit $\varphi : G \to H$ et $\psi : H \to K$ des homomorphismes de groupes.*

(i) La composée $\psi \circ \varphi : G \to K$ est un homomorphisme.

(ii) Si φ est bijectif, la fonction inverse $\varphi^{-1} : H \to G$ est aussi un homomorphisme bijectif.

Un homomorphisme bijectif s'appelle un *isomorphisme*. On dit qu'un groupe G est *isomorphe* à un groupe H s'il existe un isomorphisme de G dans H. Cela définit une relation d'équivalence sur la classe des groupes.

Exemple A.4.3. L'application

$$\begin{array}{rccc} \exp : & \mathbb{R} & \to & \mathbb{R}_+^* \\ & x & \mapsto & e^x \end{array}$$

est un homomorphisme de groupes car $e^{x+y} = e^x e^y$ pour tout choix de $x, y \in \mathbb{R}$. Comme il est bijectif, c'est un isomorphisme. Donc les groupes $(\mathbb{R}, +)$ et $(\mathbb{R}_+^*, \cdot)$ sont isomorphes.

Proposition A.4.4. *Soit $\varphi : G \to H$ un homomorphisme de groupes.*

(i) L'ensemble

$$\ker(\varphi) := \{a \in G \,;\, \varphi(a) = 1_H\}$$

est un sous-groupe de G appelé le noyau *de φ.*

(ii) L'ensemble

$$\mathrm{Im}(\varphi) := \{\varphi(a) \,;\, a \in G\}$$

*est un sous-groupe de H appelé l'*image *de φ.*

(iii) L'application φ est injective si et seulement si $\ker(\varphi) = \{1_G\}$. elle surjective si et seulement si $\mathrm{Im}(\varphi) = H$.

Exemple A.4.5. L'application

$$\begin{array}{rccc} \varphi : & \mathbb{R}^* & \longrightarrow & \mathbb{R}^* \\ & x & \longmapsto & x^2 \end{array}$$

est un homomorphisme de groupes de noyau $\ker(\varphi) = \{-1, 1\}$ et d'image $\mathrm{Im}(\varphi) = \mathbb{R}_+^*$.

Exemple A.4.6. Soit $n \in \mathbb{N}^*$. L'application

$$\begin{array}{rccc} \varphi : & \mathrm{GL}_n(\mathbb{R}) & \longrightarrow & \mathbb{R}^* \\ & A & \longmapsto & \det(A) \end{array}$$

est homomorphisme de groupes car $\det(AB) = \det(A)\det(B)$ pour toute paire de matrices $A, B \in \mathrm{GL}_n(\mathbb{R})$ (voir l'exemple A.2.9 pour la définition de $\mathrm{GL}_n(\mathbb{R})$). Son noyau est

$$\mathrm{SL}_n(\mathbb{R}) := \{A \in GL_n(\mathbb{R}) \,;\, \det(A) = 1\}$$

On l'appelle le *groupe linéaire spécial* d'ordre n sur $\mathbb{R}$. C'est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$. Par ailleurs, on vérifie aisément que l'image de φ est $\mathbb{R}^*$, c'est-à-dire que φ est un homomorphisme surjectif.

Exemple A.4.7. Soient G et H deux groupes. Leur produit cartésien

$$G \times H = \{(g,h)\,;\, g \in G, h \in H\}$$

est un groupe pour le produit composantes à composantes :

$$(g,h)\cdot(g',h') = (gg',hh').$$

Pour cette structure de groupe, la projection

$$\begin{array}{rccc} \pi\ : & G\times H & \longrightarrow & G \\ & (g,h) & \longmapsto & g \end{array}$$

est un homomorphisme surjectif de noyau

$$\ker(\pi) = \{(1_G,h)\,;\, h \in H\} = \{1_G\} \times H.$$

A.5 Anneaux

Définition A.5.1. Un *anneau* est un ensemble A muni de deux opérations binaires

$$\begin{array}{rccc} A\times A & \longrightarrow & A \\ (a,b) & \longmapsto & a+b \end{array} \qquad \text{et} \qquad \begin{array}{rccc} A\times A & \longrightarrow & A\,, \\ (a,b) & \longmapsto & ab \end{array}$$

appelées respectivement *addition* et *multiplication*, qui remplissent les conditions suivantes :

(i) sous l'addition A est un groupe abélien,

(ii) sous la multiplication A est un monoïde,

(iii) la multiplication est distributive sur l'addition : on a

$$a(b+c) = ab+ac \quad \text{et} \quad (a+b)c = ac+bc$$

quels que soient $a,b,c \in A$.

Explicitement les trois conditions de la définition se traduisent par les huit axiomes suivants :

A1. $a+(b+c) = (a+b)+c \quad \forall a,b,c \in A.$

A2. $a+b = b+a \quad \forall a,b \in A.$

A3. Il existe un élément 0 de A tel que $a+0 = a$ pour tout $a \in A$.

A4. Pour tout $a \in A$, il existe $-a \in A$ tel que $a+(-a) = 0$.

A5. $a(bc) = (ab)c \quad \forall a,b,c \in A.$

A6. Il existe un élément 1 de A tel que $a\,1 = 1\,a = a$.

A7. $a(b+c) = ab+ac \quad \forall a,b,c \in A.$

A8. $(a+b)c = ac+bc \quad \forall a,b,c \in A.$

On dit qu'un anneau est *commutatif* si sa multiplication est commutative auquel cas les conditions **A7** et **A8** sont équivalentes.

Puisqu'un anneau est un groupe abélien sous l'addition, la somme de n éléments $a_1, \cdots, a_n$ d'un anneau A ne dépend ni de la manière de grouper les termes ni de l'ordre dans lequel on les additionne. La somme de ces éléments est simplement notée

$$a_1 + \cdots + a_n \quad \text{ou} \quad \sum_{i=1}^{n} a_i.$$

L'élément neutre pour l'addition, caractérisé par la condition **A3**, s'appelle le *zéro* de l'anneau A et on le note 0. Chaque élément a de A possède un unique inverse additif noté $-a$ et caractérisé par la condition **A4**. Pour tout $n \in \mathbb{Z}$, on définit

$$na = \begin{cases} \underbrace{a + +a + \cdots + a}_{n \text{ fois}} & \text{si } n \geq 1, \\ 0 & \text{si } n = 0, \\ \underbrace{(-a) + \cdots + (-a)}_{-n \text{ fois}} & \text{si } n \leq -1, \end{cases}$$

comme il est coutume dans tout groupe abélien. Alors on a

$$(m+n)a = ma + na, \quad m(na) = (mn)a \quad \text{et} \quad m(a+b) = ma + mb,$$

pour tout choix de $m, n \in \mathbb{Z}$ et de $a, b \in A$.

De même, puisqu'un anneau est un monoïde sous la multiplication, le produit de n éléments $a_1, a_2, \ldots, a_n$ d'un anneau A ne dépend pas de la manière de grouper les termes dans le produit pourvu qu'on en respecte l'ordre. Le produit de ces éléments est simplement noté

$$a_1\, a_2\, \cdots\, a_n.$$

Si l'anneau A est commutatif, le produit d'éléments de A ne dépend pas non plus de l'ordre dans lequel on les multiplie. Enfin, l'élément neutre de A pour la multiplication, noté 1 et caractérisé par la condition **A5**, s'appelle l'*unité* de A. Pour tout entier $n \geq 1$ et tout $a \in A$, on définit comme il se doit

$$a^n = \underbrace{a\, a\, \cdots\, a}_{n \text{ fois}}.$$

On obtient alors les règles des exposants :

$$a^m a^n = a^{m+n}, \quad (a^m)^n = a^{mn} \tag{A.3}$$

pour tout choix d'entiers $m, n \geq 1$ et de $a \in A$.

On dit que deux éléments a, b de A *commutent* si $ab = ba$. Dans ce cas, on a aussi

$$(ab)^m = a^m b^m$$

pour tout $m \in \mathbb{N}^*$. En particulier, cette formule est satisfaite quels que soient $a, b \in A$ si A est un anneau commutatif.

On dit qu'un élément a d'un anneau A est *inversible* s'il possède un inverse multiplicatif, c'est-à-dire s'il existe $b \in A$ tel que $ab = ba = 1$. Comme on l'a vu à la section A.2, cet élément b est alors caractérisé par cette condition. On le note a^{-1} et on l'appelle l'*inverse* de a. Pour un élément inversible a de A, on peut étendre la définition de a^m à tout entier $m \in \mathbb{Z}$ de sorte que les formules (A.3) demeurent valables quels que soient $m, n \in \mathbb{Z}$ (voir la section A.2).

Dans un anneau A, les lois d'adition et de multiplication sont liées par les conditions de distributivité à deux termes **A7** et **A8**. Par récurrence sur n, on en déduit des propriétés de distributivité à n termes :

$$\begin{aligned} a(b_1 + \cdots + b_n) &= ab_1 + \cdots + ab_n\,, \\ (a_1 + \cdots + a_n)b &= a_1 b + \cdots + a_n b\,, \end{aligned}$$

quels que soient $a, a_1, \ldots, a_n, b, b_1, \ldots, b_n \in A$. Plus généralement encore, en combinant ces deux propriétés, on trouve

$$\left(\sum_{i=1}^{m} a_i\right)\left(\sum_{j=1}^{n} b_j\right) = \sum_{i=1}^{m} a_i\left(\sum_{j=1}^{n} b_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j$$

quels que soient $a_1, \ldots, a_n, b_1, \ldots, b_n \in A$. De plus, si $a, b \in A$ commutent, on obtient la "formule du binôme"

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

Enfin, les axiomes de distributivité impliquent

$$(ma)b = a(mb) = m(ab)$$

quels que soient $a, b \in A$ et $m \in \mathbb{Z}$. En particulier, en prenant $m = 0$ puis $m = -1$, on retrouve les propriétés bien connues

$$0\,a = a\,0 = 0 \quad \text{et} \quad (-a)b = a(-b) = -(ab)$$

où, dans la première série d'égalités, le symbole 0 représente le zéro de l'anneau.

Rappelons aussi que, puisqu'un anneau A est monoïde sous la multiplication, l'ensemble des éléments inversibles de A forme un groupe sous la multiplication. On note ce groupe A^*.

Exemple A.5.2. Pour tout entier $n \geq 1$, l'ensemble $\mathrm{Mat}_{n\times n}(\mathbb{R})$ est un anneau sous l'addition et la multiplication matricielles. Le groupe $(\mathrm{Mat}_{n\times n}(\mathbb{R}))^*$ des éléments inversibles de cet anneau est noté $\mathrm{GL}_n(\mathbb{R})$ (voir l'exemple A.2.9).

Exemple A.5.3. Soit $n \geq 1$ un entier et soit A un anneau quelconque. L'ensemble

$$\mathrm{Mat}_{n\times n}(A) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} ; a_{11}, \ldots, a_{nn} \in A \right\}$$

des matrices $n \times n$ à coefficients dans A est un anneau pour les opérations

$$\begin{aligned} (a_{ij}) + (b_{ij}) &= (a_{ij} + b_{ij}) \\ (a_{ij}) \cdot (b_{ij}) &= (c_{ij}) \quad \text{où } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \end{aligned}$$

Cela ne requiert pas que l'anneau A soit commutatif.

Exemple A.5.4. Soient X un ensemble et A un anneau. On désigne par $\mathcal{F}(X, A)$ l'ensemble des fonctions de X dans A. Étant donné $f, g \in \mathcal{F}(X, A)$, on définit $f + g$ et fg comme étant les fonctions de X dans A données par

$$(f + g)(x) = f(x) + g(x) \quad \text{et} \quad (fg)(x) = f(x)g(x)$$

pour tout $x \in X$. Alors $\mathcal{F}(X, A)$ est un anneau pour ces opérations.

Exemple A.5.5. Les triplets $(\mathbb{Z}, +, \cdot), (\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ sont des anneaux commutatifs.

Exemple A.5.6. Soit n un entier positif. On définit une relation d'équivalence sur $\mathbb{Z}$ en posant

$$a \equiv b \mod n \iff n \text{ divise } a - b$$

Alors $\mathbb{Z}$ se partitionne en exactement n classes d'équivalence représentées par les entiers $0, 1, \ldots, n-1$. On désigne par $\bar{a}$ la classe d'équivalence de $a \in \mathbb{Z}$ et on l'appelle plus précisément la *classe de congruence* de a modulo n. L'ensemble des classes de congruence modulo n est noté $\mathbb{Z}/n\mathbb{Z}$. On a donc

$$\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\}.$$

On définit la somme et le produit de deux classes de congruence modulo n par

$$\overline{a} + \overline{b} = \overline{a + b} \quad \text{et} \quad \overline{a}\,\overline{b} = \overline{ab}\,,$$

les classes résultantes $\overline{a + b}$ et $\overline{ab}$ étant indépendantes des choix des représentants a de $\overline{a}$ et b de $\overline{b}$. Pour ces opérations $\mathbb{Z}/n\mathbb{Z}$ est un anneau commutatif avec n éléments. On l'appelle l'*anneau des entiers modulo* n.

A.6 Sous-anneaux

Définition A.6.1. Un *sous-anneau* d'un anneau A est un sous-ensemble B de A tel que

SA1. $0, 1 \in B$.
SA2. $a + b \in B$ pour tout choix de $a, b \in B$.
SA3. $-a \in B$ pour tout $a \in A$.
SA4. $ab \in B$ pour tout choix de $a, b \in B$.

Un tel sous-ensemble B est lui-même un anneau pour les opérations de A restreintes à B. La notion de sous-anneau est transitive : si B est un sous-anneau d'un anneau A et si C est un sous-anneau de B, alors C est un sous-anneau de A.

Exemple A.6.2. Dans la chaîne d'inclusions $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ chaque anneau est un sous-anneau de son successeur.

Exemple A.6.3. Soit A un anneau quelconque. L'ensemble

$$P := \{m\, 1_A\,;\, m \in \mathbb{Z}\}$$

des multiples entiers de l'unité 1_A de A est un sous-anneau de A. En effet, on a

$$O_A = 0 \cdot 1_A \in P, \qquad 1_A = 1 \cdot 1_A \in P,$$

et, quels que soient $m, n \in \mathbb{Z}$,

$$\begin{aligned}(m\, 1_A) + (n\, 1_A) &= (m+n)\, 1_A \in P,\\ -(m\, 1_A) &= (-m)\, 1_A \in P\\ (m\, 1_A)(n\, 1_A) &= (mn)\, 1_A \in P.\end{aligned}$$

Ce sous-anneau P est le plus petit sous-anneau de A.

Exemple A.6.4. Soit A un sous-anneau d'un anneau *commutatif* C et soit $c \in C$. L'ensemble

$$A[c] := \{a_0 + a_1\, c + \cdots + a_n\, c^n\,;\, n \in \mathbb{N}^*, a_0, \ldots, a_n \in A\}$$

est le plus petit sous-anneau de C qui contienne A et c (exercice).

A.7 Homomorphismes d'anneaux

Définition A.7.1. Soient A et B deux anneaux. Un *homomorphisme* de A dans B est une fonction $\varphi : A \to B$ telle que

1) $\varphi(1_A) = 1_B$,
2) $\varphi(a + a') = \varphi(a) + \varphi(a')$ quels que soient $a, a' \in A$,
3) $\varphi(a\, a') = \varphi(a)\varphi(a')$ quels que soient $a, a' \in A$.

La condition 2) signifie que $\varphi : A \to B$ est un homomorphisme de groupes abéliens sous l'addition. Elle entraîne en particulier $\varphi(O_A) = O_B$. Par contre, la condition 3) n'entraîne pas nécessairement que $\varphi(1_A) = 1_B$. C'est pourquoi on ajoute la condition 1).

Si $\varphi\colon A \to B$ est un homomorphisme d'anneaux et si $a \in A$, on a en particulier

$$\begin{aligned} \varphi(na) &= n\varphi(a) \quad \text{pour tout } n \in \mathbb{Z} \\ \varphi(a^n) &= \varphi(a)^n \quad \text{pour tout } n \in \mathbb{N}^* . \end{aligned}$$

Exemple A.7.2. Soit A un anneau. On vérifie sans peine que l'ensemble

$$U = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} ; a, b, c \in A \right\}$$

est un sous-anneau de $\mathrm{Mat}_{2\times 2}(A)$ et que l'application $\varphi : U \to A$ donnée par

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

est un homomorphisme d'anneaux.

Proposition A.7.3. *Soient $\varphi\colon A \to B$ et $\psi\colon B \to C$ des homomorphismes d'anneaux*

(i) La composée $\psi \circ \varphi\colon A \to C$ est un homomorphisme d'anneaux.

(ii) Si φ est bijective, alors la fonction inverse $\varphi^{-1} : B \to A$ est aussi un homomorphisme bijectif.

Comme dans le cas des groupes, un homomorphisme d'anneaux bijectif s'appelle un *isomorphisme d'anneaux*. On dit qu'un anneau A est *isomorphe* à un anneau B, s'il existe un isomorphisme d'anneaux de A dans B. Cela définit une relation d'équivalence sur la classe des anneaux.

Définition A.7.4. Soit $\varphi\colon A \to B$ un homomorphisme d'anneaux. On définit le *noyau* de φ par

$$\ker(\varphi) := \{a \in A\,;\, \varphi(a) = 0\}$$

et l'*image* de φ par

$$\mathrm{Im}(\varphi) := \{\varphi(a)\,;\, a \in A\}.$$

Ce sont donc respectivement le noyau et l'image de φ en tant qu'homomorphisme de groupes abéliens. Par suite, $\ker(\varphi)$ est un sous-groupe de A pour l'addition, et $\mathrm{Im}(\varphi)$ est un sous-groupe de B pour l'addition. En vertu de la partie (iii) de la proposition A.4.4, on en déduit qu'un homomorphisme d'anneaux $\varphi : A \to B$ est

- injectif $\iff$ $\ker(\varphi) = \{0\}$,
- surjectif $\iff$ $\mathrm{Im}(\varphi) = B$.

On vérifie plus précisément que $\mathrm{Im}(\varphi)$ est un sous-anneau de B. Par contre, $\ker(\varphi)$ n'est généralement pas un sous-anneau de A. En effet, on trouve que

$$1_A \in \ker(\varphi) \quad \Longleftrightarrow \quad 1_B = O_B \quad \Longleftrightarrow \quad B = \{O_B\}.$$

Par suite $\ker(\varphi)$ est un sous-anneau de A si et seulement si B est l'anneau trivial réduit à un seul élément $O = 1$. On note cependant que si $a \in A$ et si $x \in \ker(\varphi)$, alors on a

$$\varphi(a\,x) = \varphi(a)\varphi(x) = \varphi(a)O = O$$

$$\varphi(x\,a) = \varphi(x)\varphi(a) = O\varphi(a) = O$$

donc ax et xa appartiennent à $\ker(\varphi)$. Cela motive la définition suivante :

Définition A.7.5. Un *idéal* d'un anneau A est un sous-ensemble I de A tel que

(i) $0 \in I$

(ii) $x + y \in I$ quels que soient $x, y \in I$,

(iii) $ax, xa \in I$ pour tout $a \in A$ et tout $x \in I$.

La condition (iii) implique que $-x = (-1)x \in I$ pour tout $x \in I$. Donc un idéal de A est un sous-groupe de A sous l'addition qui remplit la condition (iii). Lorsque A est abélien, cette dernière condition devient simplement

$$ax \in I \quad \text{pour tout } a \in A \text{ et tout } x \in I.$$

En vertu des observations qui précèdent la définition A.7.5, on peut encore énoncer :

Proposition A.7.6. *Le noyau d'un homomorphisme d'anneaux $\varphi : A \to B$ est un idéal de A.*

Exemple A.7.7. Pour l'homomorphisme $\varphi : U \to A$ de l'exemple A.7.2, on a

$$\ker(\varphi) = \left\{ \begin{pmatrix} 0 & b \\ 0 & c \end{pmatrix} ;\, b, c \in A \right\}.$$

Ce dernier est un idéal de U.

A.8 Corps

Définition A.8.1. Un *corps* est un anneau commutatif K avec $0 \neq 1$, dont tous les éléments non nuls sont inversibles.

Si K est un corps, chaque élément non nul a de K possède un inverse multiplicatif a^{-1}. En plus de l'addition et de la multiplication, on peut donc définir dans K une opération de soustraction

$$a - b := a + (-b)$$

et une autre de division

$$\frac{a}{b} := a\,b^{-1} \quad \text{si } b \neq 0.$$

Exemple A.8.2. Les ensembles $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps pour les opérations usuelles.

On rappelle enfin le résultat suivant :

Proposition A.8.3. *Soit $n \in \mathbb{N}^*$. L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est un nombre premier.*

Pour chaque nombre premier p, on désigne par $\mathbb{F}_p$ le corps à p éléments.

Pour $p = 2$, les tables d'addition et de multiplication dans $\mathbb{F}_2 = \{\bar{0}, \bar{1}\}$ sont données par

$$\begin{array}{c|cc} + & \bar{0} & \bar{1} \\ \hline \bar{0} & \bar{0} & \bar{1} \\ \bar{1} & \bar{1} & \bar{0} \end{array} \qquad \text{et} \qquad \begin{array}{c|cc} \cdot & \bar{0} & \bar{1} \\ \hline \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{0} & \bar{1} \end{array}.$$

Dans $\mathbb{F}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$, elles sont données par

$$\begin{array}{c|ccc} + & \bar{0} & \bar{1} & \bar{2} \\ \hline \bar{0} & \bar{0} & \bar{1} & \bar{2} \\ \bar{1} & \bar{1} & \bar{2} & \bar{0} \\ \bar{2} & \bar{2} & \bar{0} & \bar{1} \end{array} \qquad \text{et} \qquad \begin{array}{c|ccc} \cdot & \bar{0} & \bar{1} & \bar{2} \\ \hline \bar{0} & \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{0} & \bar{1} & \bar{2} \\ \bar{2} & \bar{0} & \bar{2} & \bar{1} \end{array}.$$

Annexe B

Le déterminant

Dans tout cet appendice, on fixe un anneau commutatif quelconque A, avec $0 \neq 1$. On étend la notion de déterminant aux matrices carrées à coefficients dans A et on montre que la plupart des propriétés usuelles du déterminant (celles qu'on connaît pour les matrices à coefficients dans un corps) s'appliquent à cette situation générale avec des modifications mineures. Pour y parvenir, on utilise la notion de A-module présentée au chapitre 6.

B.1 Applications multilinéaires

Définition B.1.1. Soient $M_1, \ldots, M_s$ et N des A-modules. On dit qu'une fonction

$$\varphi\colon M_1 \times \cdots \times M_s \to N$$

est *multilinéaire* si elle satisfait

ML1. $\varphi(\mathbf{u}_1, \ldots, \mathbf{u}'_j + \mathbf{u}''_j, \ldots, \mathbf{u}_s) = \varphi(\mathbf{u}_1, \ldots, \mathbf{u}'_j, \ldots, \mathbf{u}_s) + \varphi(\mathbf{u}_1, \ldots, \mathbf{u}''_j, \ldots, \mathbf{u}_s)$,

ML2. $\varphi(\mathbf{u}_1, \ldots, a\mathbf{u}_j, \ldots, \mathbf{u}_s) = a\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_j, \ldots, \mathbf{u}_s)$,

pour tout entier $j = 1, \ldots, s$, tout choix de $\mathbf{u}_1 \in M_1, \ldots, \mathbf{u}_j, \mathbf{u}'_j, \mathbf{u}''_j \in M_j, \ldots, \mathbf{u}_s \in M_s$, et tout $a \in A$.

En d'autres termes, une fonction $\varphi\colon M_1 \times \cdots \times M_s \to N$ est multilinéaire si, pour chaque entier $j = 1, \ldots, s$ et pour chaque choix de $\mathbf{u}_1 \in M_1$, ... , $\mathbf{u}_s \in M_s$, la fonction

$$\begin{array}{rccl} \overline{\varphi} : & M_j & \longrightarrow & N \\ & \mathbf{u} & \longmapsto & \varphi(\mathbf{u}_1, \ldots, \mathbf{u}_{j-1}, \mathbf{u}, \mathbf{u}_{j+1}, \ldots, \mathbf{u}_s) \end{array}$$

est un homomorphisme de A-modules. On montre d'abord :

Proposition B.1.2. *Soient $M_1, \ldots, M_s$ des A-modules libres, et soit $\{\mathbf{v}_1^{(j)}, \ldots, \mathbf{v}_{n_j}^{(j)}\}$ une base de M_j pour $j = 1, \ldots, s$. Pour tout A-module N et tout choix de*

$$\mathbf{w}_{i_1 \ldots i_s} \in N \quad (1 \leq i_1 \leq n_1, \ \ldots, \ 1 \leq i_s \leq n_s),$$

il existe une et une seule application multilinéaire $\varphi\colon M_1 \times \cdots \times M_s \to N$ telle que

$$\varphi(\mathbf{v}_{i_1}^{(1)}, \ldots, \mathbf{v}_{i_s}^{(s)}) = \mathbf{w}_{i_1 \ldots i_s} \quad (1 \leq i_1 \leq n_1, \ \ldots, \ 1 \leq i_s \leq n_s). \tag{B.1}$$

Preuve. Supposons qu'une application multilinéaire $\varphi\colon M_1 \times \cdots \times M_s \to N$ satisfasse les conditions (B.1). Pour tout choix de $\mathbf{u}_1 \in M_1, \ldots, \mathbf{u}_s \in M_s$, on peut écrire

$$\mathbf{u}_1 = \sum_{i_1=1}^{n_1} a_{i_1,1}\mathbf{v}_{i_1}^{(1)}, \quad \mathbf{u}_2 = \sum_{i_2=1}^{n_2} a_{i_2,2}\mathbf{v}_{i_2}^{(2)}, \ \ldots\ , \ \mathbf{u}_s = \sum_{i_s=1}^{n_s} a_{i_s,s}\mathbf{v}_{i_s}^{(s)}, \tag{B.2}$$

pour des éléments $a_{i,j}$ de A. Alors en vertu de la multilinéarité de φ, on obtient

$$\begin{aligned}
\varphi(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_s) &= \sum_{i_1=1}^{n_1} a_{i_1,1}\, \varphi(\mathbf{v}_{i_1}^{(1)}, \mathbf{u}_2, \ldots, \mathbf{u}_s) \\
&= \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2} a_{i_1,1}\, a_{i_2,2}\, \varphi(\mathbf{v}_{i_1}^{(1)}, \mathbf{v}_{i_2}^{(2)}, \mathbf{u}_3, \ldots, \mathbf{u}_s) \\
&\quad \cdots\cdots \\
&= \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_s=1}^{n_s} a_{i_1,1}a_{i_2,2}\ldots a_{i_s,s}\, \varphi(\mathbf{v}_{i_1}^{(1)}, \mathbf{v}_{i_2}^{(2)}, \ldots, \mathbf{v}_{i_s}^{(s)}) \\
&= \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2}\cdots\sum_{i_s=1}^{n_s} a_{i_1,1}a_{i_2,2}\ldots a_{i_s,s}\, \mathbf{w}_{i_1\ldots i_s}.
\end{aligned} \tag{B.3}$$

Cela montre qu'il y a au plus une fonction multilinéaire φ qui satisfait (B.1). Par ailleurs, pour $\mathbf{u}_1 \in M_1, \ldots, \mathbf{u}_s \in M_s$ donnés, il y a un seul choix de $a_{ij} \in A$ qui satisfait (B.2) et par suite la formule (B.3) définit bien une fonction $\varphi\colon M_1 \times \cdots \times M_s \to N$. On laisse en exercice le soin de vérifier que celle-ci est multilinéaire et qu'elle satisfait (B.1). □

Dans cet appendice, on s'intéresse à un type particulier d'application multilinéaire :

Définition B.1.3. Soient M et N des A-modules, et soit n un entier positif. On dit qu'une application multilinéaire

$$\varphi : M^n \longrightarrow N$$

est *alternée* si elle satisfait $\varphi(\mathbf{u}_1, \ldots, \mathbf{u}_n) = \mathbf{0}$ pour tout choix de $\mathbf{u}_1, \ldots, \mathbf{u}_n \in M$ non tous distincts.

Le qualificatif "alternée" est justifié par la proposition suivante :

Proposition B.1.4. *Soit $\varphi\colon M^n \longrightarrow N$ une application multilinéaire alternée, et soient i et j des entiers avec $1 \le i < j \le n$. Pour tout choix de $\mathbf{u}_1, \ldots, \mathbf{u}_n \in M$, on a*

$$\varphi(\mathbf{u}_1, \ldots, \overset{i}{\underset{\vee}{\mathbf{u}_j}}, \ldots, \overset{j}{\underset{\vee}{\mathbf{u}_i}}, \ldots, \mathbf{u}_n) = -\varphi(\mathbf{u}_1, \ldots, \overset{i}{\underset{\vee}{\mathbf{u}_i}}, \ldots, \overset{j}{\underset{\vee}{\mathbf{u}_j}}, \ldots, \mathbf{u}_n).$$

Autrement dit, le signe de φ change lorsqu'on permute deux de ses arguments.

Preuve. Fixons $\mathbf{u}_1, \ldots, \mathbf{u}_n \in M$. Pour tout $\mathbf{u} \in M$, on a :

$$\varphi(\mathbf{u}_1, \ldots, \overset{i}{\underset{\vee}{\mathbf{u}}}, \ldots, \overset{j}{\underset{\vee}{\mathbf{u}}}, \ldots, \mathbf{u}_n) = 0.$$

En choisissant $\mathbf{u} = \mathbf{u}_i + \mathbf{u}_j$ et en utilisant la multilinéarité de φ, on trouve

$$\begin{aligned} 0 &= \varphi(\mathbf{u}_1, \dots, \overset{i}{\underset{\vee}{\mathbf{u}_i + \mathbf{u}_j}}, \dots, \overset{j}{\underset{\vee}{\mathbf{u}_i + \mathbf{u}_j}}, \dots, \mathbf{u}_n) \\ &= \varphi(\mathbf{u}_1, \dots, \mathbf{u}_i, \dots, \mathbf{u}_i, \dots, \mathbf{u}_n) + \varphi(\mathbf{u}_1, \dots, \mathbf{u}_i, \dots, \mathbf{u}_j, \dots, \mathbf{u}_n) \\ &\quad + \varphi(\mathbf{u}_1, \dots, \mathbf{u}_j, \dots, \mathbf{u}_i, \dots, \mathbf{u}_n) + \varphi(\mathbf{u}_1, \dots, \mathbf{u}_j, \dots, \mathbf{u}_j, \dots, \mathbf{u}_n) \\ &= \varphi(\mathbf{u}_1, \dots, \mathbf{u}_i, \dots, \mathbf{u}_j, \dots, \mathbf{u}_n) + \varphi(\mathbf{u}_1, \dots, \mathbf{u}_j, \dots, \mathbf{u}_i, \dots, \mathbf{u}_n). \end{aligned}$$

La conclusion suit. □

Soit S_n l'ensemble des bijections de l'ensemble $\{1, 2, \dots, n\}$ dans lui-même. C'est un groupe sous la composition qu'on appelle le *groupe des permutations* de $\{1, 2, \dots, n\}$. On dit qu'un élément τ de S_n est une *transposition* s'il existe des entiers i et j avec $1 \leq i < j \leq n$ tels que $\tau(i) = j$, $\tau(j) = i$ et $\tau(k) = k$ pour tout $k \neq i, j$. Dans le cours de théorie de groupes, on montre que tout $\sigma \in S_n$ s'écrit comme un produit de transpositions

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_\ell.$$

Cette écriture n'est pas unique, mais pour chacune d'elles la parité de l'entier ℓ est la même. On définit la *signature* de σ par

$$\epsilon(\sigma) = (-1)^\ell$$

de sorte que toutes les transpositions ont pour signature -1. On montre aussi que l'application $\epsilon\colon S_n \to \{1, -1\}$ ainsi définie est un homomorphisme de groupes. À l'aide de ces notions, on peut généraliser la proposition B.1.4 de la manière suivante :

Proposition B.1.5. *Soit $\varphi\colon M^n \longrightarrow N$ une application multilinéaire alternée, et soit $\sigma \in S_n$. Pour tout choix de $\mathbf{u}_1, \dots, \mathbf{u}_n \in M$, on a*

$$\varphi(\mathbf{u}_{\sigma(1)}, \dots, \mathbf{u}_{\sigma(n)}) = \epsilon(\sigma)\varphi(\mathbf{u}_1, \dots, \mathbf{u}_n).$$

Preuve. La proposition B.1.4 montre que, pour toute transposition $\tau \in S_n$, on a

$$\varphi(\mathbf{u}_{\tau(1)}, \dots, \mathbf{u}_{\tau(n)}) = -\varphi(\mathbf{u}_1, \dots, \mathbf{u}_n).$$

indépendamment du choix de $\mathbf{u}_1, \dots, \mathbf{u}_n \in M$. On en déduit

$$\varphi(\mathbf{u}_{\nu(\tau(1))}, \dots, \mathbf{u}_{\nu(\tau(n))}) = -\varphi(\mathbf{u}_{\nu(1)}, \dots, \mathbf{u}_{\nu(n)}) \tag{B.4}$$

pour tout $\nu \in S_n$.

Écrivons $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_\ell$ où $\tau_1, \tau_2, \dots, \tau_\ell$ sont des transpositions. Posons $\sigma_0 = 1$ et

$$\sigma_i = \tau_1 \circ \cdots \circ \tau_i \quad \text{pour} \quad i = 1, \dots, \ell,$$

de sorte que $\sigma_i = \sigma_{i-1} \circ \tau_i$ pour $i = 1, \dots, \ell$. Pour chacun de ces entiers i, la relation (B.4) livre

$$\varphi(\mathbf{u}_{\sigma_i(1)}, \dots, \mathbf{u}_{\sigma_i(n)}) = -\varphi(\mathbf{u}_{\sigma_{i-1}(1)}, \dots, \mathbf{u}_{\sigma_{i-1}(n)}).$$

Comme $\sigma = \sigma_\ell$, ces ℓ relations entraînent

$$\varphi(\mathbf{u}_{\sigma(1)}, \dots, \mathbf{u}_{\sigma(n)}) = (-1)^\ell \varphi(\mathbf{u}_1, \dots, \mathbf{u}_n).$$

La conclusion suit, car $\epsilon(\sigma) = (-1)^\ell$. □

La proposition B.1.5 entraîne la conséquence suivante :

Théorème B.1.6. *Soit M un A-module libre avec base $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, et soit $c \in A$. Il existe une et une seule application multilinéaire alternée $\varphi : M^n \to A$ telle que $\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = c$. Elle est donnée par*

$$\varphi\Big(\sum_{i=1}^n a_{i,1}\mathbf{e}_i\,,\ \dots\ ,\ \sum_{i=1}^n a_{i,n}\mathbf{e}_i\Big) = c\sum_{\sigma\in S_n}\epsilon(\sigma)\,a_{\sigma(1),1}\dots a_{\sigma(n),n} \tag{B.5}$$

pour tout choix de $a_{i,j} \in A$.

Preuve. Supposons d'abord qu'il existe une application multilinéaire alternée $\varphi : M^n \to A$ telle que $\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = c$. Soient

$$\mathbf{u}_1 = \sum_{i=1}^n a_{i,1}\mathbf{e}_i, \quad \mathbf{u}_2 = \sum_{i=1}^n a_{i,2}\mathbf{e}_i,\ \dots\ , \mathbf{u}_n = \sum_{i=1}^n a_{i,n}\mathbf{e}_i \tag{B.6}$$

des éléments de M écrits comme combinaisons linéaires de $\mathbf{e}_1, \dots, \mathbf{e}_n$. Puisque φ est multilinéaire, on a

$$\varphi(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = \sum_{i_1=1}^n\sum_{i_2=1}^n\cdots\sum_{i_n=1}^n a_{i_1,1}a_{i_2,2}\dots a_{i_n,n}\,\varphi(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \dots, \mathbf{e}_{i_n}) \tag{B.7}$$

(c'est un cas particulier de la formule (B.3)). Pour des entiers $i_1, \dots, i_n \in \{1, 2, \dots, n\}$ non tous distincts, on a $\varphi(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) = 0$ en vertu de la définition B.1.3. Par contre si $i_1, \dots, i_n \in \{1, 2, \dots, n\}$ sont distincts, ils déterminent une bijection σ de $\{1, 2, \dots, n\}$ donnée par $\sigma(1) = i_1, \sigma(2) = i_2, \dots, \sigma(n) = i_n$ et, grâce à la proposition B.1.5, on obtient

$$\varphi(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) = \varphi(\mathbf{e}_{\sigma(1)},\ \dots,\ \mathbf{e}_{\sigma(n)}) = \epsilon(\sigma)\,\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = c\,\epsilon(\sigma).$$

Donc la formule (B.7) se réécrit

$$\varphi(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = c\sum_{\sigma\in S_n}\epsilon(\sigma)\,a_{\sigma(1),1}a_{\sigma(2),2}\dots a_{\sigma(n),n}.$$

Cela montre que φ est nécessairement donnée par (B.5).

Pour conclure, il reste seulement à montrer que la fonction $\varphi : M^n \to A$ donnée par (B.5) est multilinéaire alternée. On note d'abord qu'elle est multilinéaire (exercice). Il reste à vérifier que, si $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n \in M$ ne sont pas tous distincts, alors $\varphi(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = 0$.

Supposons donc qu'il existe des indices i et j avec $1 \le i < j \le n$ tels que $\mathbf{u}_i = \mathbf{u}_j$. On désigne par E l'ensemble des permutations σ de S_n telles que $\sigma(i) < \sigma(j)$, par τ la transposition de S_n qui interchange i et j, et par E^c le complément de E dans S_n, c'est-à-dire l'ensemble des permutations σ de S_n telles que $\sigma(i) > \sigma(j)$. On constate qu'un élément de S_n appartient à E^c si et seulement s'il est de la forme $\sigma \circ \tau$ avec $\sigma \in E$. On en déduit que

$$\varphi(\mathbf{u}_1, \dots, \mathbf{u}_n) = c\sum_{\sigma\in E}\Big(\epsilon(\sigma)\,a_{\sigma(1),1}\dots a_{\sigma(n),n} + \epsilon(\sigma\circ\tau)\,a_{\sigma(\tau(1)),1}\dots a_{\sigma(\tau(n)),n}\Big).$$

Or, pour tout $\sigma \in S_n$, on a à la fois $\epsilon(\sigma \circ \tau) = \epsilon(\sigma)\epsilon(\tau) = -\epsilon(\sigma)$ et

$$a_{\sigma(\tau(1)),1} \dots a_{\sigma(\tau(n)),n} = a_{\sigma(1),\tau(1)} \dots a_{\sigma(n),\tau(n)} = a_{\sigma(1),1} \dots a_{\sigma(n),n},$$

la dernière égalité découlant du fait que $\mathbf{u}_i = \mathbf{u}_j$. On conclut que $\varphi(\mathbf{u}_1, \dots, \mathbf{u}_n) = 0$, comme annoncé. □

Corollaire B.1.7. *Soit M un A-module libre. Alors, toutes les bases de M renferment le même nombre d'éléments.*

Preuve. Supposons au contraire que $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ et $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ soient des bases de M de cardinalités différentes. Supposons qu'on ait $m < n$ et désignons par $\varphi\colon M^n \longrightarrow A$ la forme mulitilinéaire alternée telle que $\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$. Puisque $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ est une base de M, on peut écrire

$$\mathbf{e}_1 = \sum_{i=1}^{m} a_{i,1}\mathbf{f}_i \ , \ \dots \ , \ \mathbf{e}_n = \sum_{i=1}^{m} a_{i,n}\mathbf{f}_i$$

pour des éléments $a_{i,j}$ de A. Comme φ est multilinéaire, on en déduit

$$\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} a_{i_1,1} \dots a_{i_n,n}\, \varphi(\mathbf{f}_{i_1}, \dots, \mathbf{f}_{i_n}). \tag{B.8}$$

Or, pour tout choix de $i_1, \dots, i_n \in \{1, \dots, m\}$, au moins deux des entiers $i_1, \dots, i_n$ sont égaux (car $n > m$) et par suite, comme φ est alternée, on obtient $\varphi(\mathbf{f}_{i_1}, \dots, \mathbf{f}_{i_n}) = 0$. Donc, l'égalité (B.8) entraîne $\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = 0$. Comme on a choisi φ telle que $\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$, c'est impossible. □

B.2 Le déterminant

Soit n un entier positif. On sait que

$$A^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \ ; \ a_1, \dots, a_n \in A \right\}$$

est un A-module libre qui admet pour base

$$\left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} , \dots , \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Le théorème B.1.6 appliqué à $M = A^n$ montre que, pour ce choix de base, il existe une et une seule application multilinéaire alternée $\varphi\colon (A^n)^n \to A$ telle que $\varphi(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$, et que

celle-ci est donnée par

$$\varphi\Big(\begin{pmatrix} a_{1,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ \vdots \\ a_{n,2} \end{pmatrix}, \dots, \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{n,n} \end{pmatrix}\Big) = \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{\sigma(1),1} \dots a_{\sigma(n),n}\,.$$

Or, le produit cartésien $(A^n)^n = A^n \times \cdots \times A^n$ de n copies de A^n s'identifie naturellement à l'ensemble $\mathrm{Mat}_{n\times n}(A)$ des matrices $n \times n$ à coefficients dans A. Il suffit d'associer à chaque élément $(C_1, \dots, C_n)$ de $(A^n)^n$ la matrice $(C_1, \dots, C_n) \in \mathrm{Mat}_{n\times n}(A)$ dont les colonnes sont $C_1, \dots, C_n$. De cette façon, le n-uplet $(\mathbf{e}_1, \dots, \mathbf{e}_n) \in (A^n)^n$ correspond à la matrice identité I_n de format $n \times n$. On obtient ainsi :

Théorème B.2.1. *Il existe une et une seule fonction* $\det\colon \mathrm{Mat}_{n\times n}(A) \to A$ *qui satisfait* $\det(I_n) = 1$ *et qui, vue comme fonction des n colonnes d'une matrice, est multilinéaire alternée. Elle est donnée par*

$$\det \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} = \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{\sigma(1),1} \dots a_{\sigma(n),n}\,. \tag{B.9}$$

Cette fonction est appelée le *déterminant*. On en déduit aussitôt l'effet d'une transformation élémentaire de colonnes sur le déterminant :

Proposition B.2.2. *Soit $U \in \mathrm{Mat}_{n\times n}(A)$ et soit U' une matrice obtenue en appliquant à U une opération élémentaire de colonnes. Il y a trois cas :*

(I) Si U' est obtenue en permutant deux colonnes de U, alors $\det(U') = -\det(U)$ *;*

(II) Si U' est obtenue en ajoutant à la colonne i de U le produit de sa colonne j par un élément a de A, pour des entiers i et j distincts, alors $\det(U') = \det(U)$ *;*

(III) Si U' est obtenue en multipliant la colonne i de U par une unité $a \in A^*$, alors $\det(U') = a\,\det(U)$.

Preuve. Désignons par $C_1, \dots, C_n$ les colonnes de U. L'égalité $\det(U') = -\det(U)$ du cas (I) découle de la proposition B.1.4. Dans le cas (II), on trouve

$$\begin{aligned} \det(U') &= \det(C_1, \dots, \overset{i}{\underset{\vee}{C_i + a\,C_j}}, \dots, \overset{j}{\underset{\vee}{C_j}}, \dots, C_n) \\ &= \det(C_1, \dots, \overset{i}{\underset{\vee}{C_i}}, \dots, \overset{j}{\underset{\vee}{C_j}}, \dots, C_n) + a\,\det(C_1, \dots, \overset{i}{\underset{\vee}{C_j}}, \dots, \overset{j}{\underset{\vee}{C_j}}, \dots, C_n) \\ &= \det(U) \end{aligned}$$

car le déterminant est nul pour une matrice qui possède deux colonnes égales. Enfin, dans le cas (III), on obtient

$$\det(U') = \det(C_1, \dots, \overset{i}{\underset{\vee}{a\,C_i}}, \dots, C_n) = a\,\det(C_1, \dots, \overset{i}{\underset{\vee}{C_i}}, \dots, C_n) = a\,\det(U)$$

comme annoncé. □

Lorsque A est un anneau euclidien, on peut amener toute matrice $U \in \mathrm{Mat}_{n\times n}(A)$ sous la forme triangulaire inférieure par une suite d'opérations élémentaires de colonnes (voir le théorème 8.4.6). Grâce à la proposition précédente, on peut suivre l'effet de ces opérations sur le déterminant. Pour en déduire la valeur de $\det(U)$, on utilise alors :

Proposition B.2.3. *Si $U \in \mathrm{Mat}_{n\times n}(A)$ est une matrice triangulaire inférieure, son déterminant est le produit des éléments de sa diagonale.*

Preuve. Écrivons $U = (a_{ij})$. Puisque U est triangulaire inférieure, on a $a_{ij} = 0$ si $i > j$. Pour toute permutation $\sigma \in S_n$ avec $\sigma \neq 1$, il existe un entier $i \in \{1, 2, \ldots, n\}$ tel que $\sigma(i) > i$ et alors on obtient

$$a_{\sigma(1),1} \ldots a_{\sigma(i),i} \ldots a_{\sigma(n),n} = 0.$$

Alors la formule (B.9) donne $\det(U) = a_{1,1}\, a_{2,2} \ldots a_{n,n}$. □

On note aussi que le déterminant de toute matrice carrée est égal à celui de sa transposée :

Proposition B.2.4. *Pour toute matrice $U \in \mathrm{Mat}_{n\times n}(A)$ on a $\det(U^t) = \det(U)$.*

Preuve. Écrivons $U = (a_{ij})$. Pour tout $\sigma \in S_n$, on a

$$a_{\sigma(1),1} \ldots a_{\sigma(n),n} = a_{1,\sigma^{-1}(1)} \ldots a_{n,\sigma^{-1}(n)}.$$

Comme on a aussi $\epsilon(\sigma^{-1}) = \epsilon(\sigma)^{-1} = \epsilon(\sigma)$, la formule (B.9) livre

$$\det(U) = \sum_{\sigma\in S_n} \epsilon(\sigma^{-1})\, a_{1,\sigma^{-1}(1)} \ldots a_{n,\sigma^{-1}(n)} = \sum_{\sigma\in S_n} \epsilon(\sigma)\, a_{1,\sigma(1)} \ldots a_{n,\sigma(n)} = \det(U^t).$$

□

En vertu de ce résultat, on en déduit que la fonction $\det\colon \mathrm{Mat}_{n\times n}(A) \to A$, vue comme fonction des n lignes d'une matrice est multilinéaire alternée. On obtient aussi l'analogue suivant de la proposition B.2.2 pour les opérations élémentaires de lignes en remplaçant partout le mot "colonne" par "ligne".

Proposition B.2.5. *Soit $U \in \mathrm{Mat}_{n\times n}(A)$ et soit U' une matrice obtenue en appliquant à U une opération élémentaire de lignes. Il y a trois cas :*

(I) Si U' est obtenue en permutant deux lignes de U, alors $\det(U') = -\det(U)$;

(II) Si U' est obtenue en ajoutant à la ligne i de U le produit de sa ligne j par un élément a de A, pour des entiers i et j distincts, alors $\det(U') = \det(U)$;

(III) Si U' est obtenue en multipliant la ligne i de U par une unité $a \in A^*$, alors $\det(U') = a\det(U)$.

Preuve. Par exemple, dans le cas (II), la matrice $(U')^t$ s'obtient de U^t en ajoutant à la colonne i de U^t le produit de sa colonne j par a. La proposition B.2.2 donne dans ce cas $\det((U')^t) = \det(U^t)$. Grâce à la proposition B.2.4, on conclut que $\det(U') = \det(U)$. □

De même on obtient :

Proposition B.2.6. *Le déterminant d'une matrice triangulaire $U \in \mathrm{Mat}_{n\times n}(A)$ est égal au produit des éléments de sa diagonale.*

Preuve. Si U est triangulaire inférieure, cela découle de la proposition B.2.3. Si U est triangulaire supérieure, sa transposée U^t est triangulaire inférieure, donc $\det(U) = \det(U^t)$ est le produit des éléments de la diagonale de U^t. Or, ce sont les mêmes éléments que ceux de la diagonale de U. □

On conclut cette section avec le résultat suivant.

Théorème B.2.7. *Pour tout paire de matrices $U, V \in \mathrm{Mat}_{n\times n}(A)$, on a*

$$\det(UV) = \det(U)\det(V).$$

Preuve. Fixons une matrice $U \in \mathrm{Mat}_{n\times n}(A)$, et considérons la fonction $\varphi\colon (A^n)^n \to A$ donnée par

$$\varphi(C_1, \ldots, C_n) = \det(UC_1, \ldots, UC_n)$$

pour tout choix de n colonnes $C_1, \ldots, C_n \in A^n$. On constate que φ est multilinéaire alternée. De plus, si $C_1, \ldots, C_n$ sont les colonnes d'une matrice V, alors $UC_1, \ldots, UC_n$ sont les colonnes de UV. Donc, on a aussi

$$\varphi(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det(UI) = \det(U).$$

Comme le théorème B.1.6 montre qu'il existe une et une seule application multilinéaire alternée de $(A^n)^n$ dans A telle que $\varphi(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det(U)$ et que la fonction $\psi\colon (A^n)^n \to A$ donnée par

$$\psi(C_1, \ldots, C_n) = \det(U)\det(C_1, \ldots, C_n)$$

en est une autre, on conclut que $\det(UC_1, \ldots, UC_n) = \det(U)\det(C_1, \ldots, C_n)$ pour tout $(C_1, \ldots, C_n) \in (A^n)^n$, donc $\det(UV) = \det(U)\det(V)$ pour tout $V \in \mathrm{Mat}_{n\times n}(A)$. □

B.3 L'adjointe d'une matrice

Le but de cette dernière section est de montrer que les formules usuelles du développement d'un déterminant selon une ligne ou une colonne demeurent vraies pour les matrices à coefficients dans un anneau commutatif quelconque A, et que la notion d'adjointe et ses propriétés s'étendent aussi à ce cadre. Pour ce faire, on commence par donner une présentation alternative du déterminant qui est plus commode pour décrire le déterminant d'une sous-matrice.

On rappelle d'abord que le *nombre d'inversions* d'un n-uplet d'entiers distincts $(i_1, \ldots, i_n)$, noté $N(i_1, \ldots, i_n)$, est défini comme le nombre de couples d'indices (k, ℓ) avec $1 \leq k < \ell \leq n$

tels que $i_k > i_\ell$. Par exemple, on a $N(1,5,7,3) = 2$ et $N(4,3,1,2) = 4$. Toute permutation σ de S_n s'identifiant au n-uplet $(\sigma(1), \ldots, \sigma(n))$ de ses valeurs, on montre en algèbre que

$$\epsilon(\sigma) = (-1)^{N(\sigma(1),\ldots,\sigma(n))}.$$

Alors, la formule (B.9) appliquée à une matrice $U = (a_{i,j}) \in \mathrm{Mat}_{n\times n}(A)$ devient

$$\det(U) = \sum_{(i_1,\ldots,i_n)\in S_n} (-1)^{N(i_1,\ldots,i_n)} a_{i_1,1} \cdots a_{i_n,n}.$$

Comme le déterminant de U est égal à celui de sa transposée, on en déduit encore que

$$\det(U) = \sum_{(j_1,\ldots,j_n)\in S_n} (-1)^{N(j_1,\ldots,j_n)} a_{1,j_1} \cdots a_{n,j_n}.$$

Grâce à cette formule, si $1 \leq f_1 < f_2 < \cdots < f_k \leq n$ et $1 \leq g_1 < g_2 < \cdots < g_k \leq n$ sont deux sous-suites croissantes de k éléments parmi les entiers $1, 2, \ldots, n$, alors la sous-matrice $U' = (a_{f(i),g(j)})$ de U de format $k \times k$ formée des éléments de U appartenant aux lignes d'indices $f_1, \ldots, f_k$ et aux colonnes d'indices $g_1, \ldots, g_k$ a pour déterminant

$$\det(U') = \sum (-1)^{N(j_1,\ldots,j_k)} a_{f_1,j_1} \cdots a_{f_k,j_k} \tag{B.10}$$

où la somme porte sur toutes les permutations $(j_1, \ldots, j_k)$ de $(g_1, \ldots, g_k)$.

Définition B.3.1. Soit $U = (a_{i,j})$ une matrice $n \times n$ à coefficients dans A. Pour $i, j = 1, \ldots, n$, le *mineur* d'indice (i,j) de U, noté $M_{i,j}(U)$ est le déterminant de la sous-matrice de format $(n-1) \times (n-1)$ de U obtenue en retirant de U sa ligne i et sa colonne j, tandis que le *cofacteur* d'indice (i,j) de U est

$$C_{i,j}(U) := (-1)^{i+j} M_{i,j}(U).$$

La *matrice des cofacteurs* de U est la matrice de format $n \times n$ dont l'élément (i,j) est le cofacteur d'indice (i,j) de U. Enfin, l'*adjointe* de U, notée $\mathrm{Adj}(U)$ est la transposée de la matrice des cofacteurs de U.

On montre d'abord la formule du dévelopement selon la ligne 1 :

Lemme B.3.2. *Soit $U = (a_{i,j}) \in \mathrm{Mat}_{n\times n}(A)$. On a $\det(U) = \sum_{j=1}^{n} a_{1,j} C_{1,j}(U)$.*

Preuve. La formule (B.10) appliquée au calcul de $M_{1,j}(U)$ livre

$$M_{1,j}(U) = \sum_{(j_1,\ldots,j_{n-1})\in E_j} (-1)^{N(j_1,\ldots,j_{n-1})} a_{2,j_1} \cdots a_{n,j_{n-1}},$$

la somme portant sur l'ensemble E_j des permutations $(j_1, \ldots, j_{n-1})$ de $(1, \ldots, \widehat{j}, \ldots, n)$ c'est-à-dire l'ensemble des $(n-1)$-uplets d'entiers $(j_1, \ldots, j_{n-1})$ tels que $(j, j_1, \ldots, j_{n-1}) \in S_n$.

Comme $N(j, j_1, \ldots, j_{n-1}) = N(j_1, \ldots, j_{n-1}) + j - 1$ pour chacun de ces $(n-1)$-uplets, on en déduit que

$$\begin{aligned}\det(U) &= \sum_{(j_1,\ldots,j_n)\in S_n} (-1)^{N(j_1,\ldots,j_n)} a_{1,j_1} \cdots a_{n,j_n} \\ &= \sum_{j=1}^{n} a_{1,j} \sum_{(j_1,\ldots,j_{n-1})\in E_j} (-1)^{N(j,j_1,\ldots,j_{n-1})} a_{2,j_1} \cdots a_{n,j_{n-1}} \\ &= \sum_{j=1}^{n} a_{1,j} (-1)^{j-1} C_{1,j}(U),\end{aligned}$$

et la conclusion suit. □

Plus généralement, on montre les formules suivantes :

Proposition B.3.3. *Soit $U = (a_{i,j}) \in \mathrm{Mat}_{n\times n}(A)$ et soient $k, \ell \in \{1, 2, \ldots, n\}$. On a*

$$\sum_{j=1}^{n} a_{k,j} C_{\ell,j}(U) = \sum_{j=1}^{n} a_{j,k} C_{j,\ell}(U) = \begin{cases} \det(U) & \text{si } k = \ell, \\ 0 & \text{sinon.} \end{cases}$$

Preuve. Supposons d'abord $k = \ell$ et désignons par U' la matrice obtenue à partir de U en faisant glisser sa ligne k de la position k à la position 1, de sorte que, pour tout $j = 1, \ldots, n$, on ait $M_{1,j}(U') = M_{k,j}(U)$. En appliquant le lemme B.3.2 à la matrice U', on en déduit

$$\det(U) = (-1)^{k-1} \det(U') = (-1)^{k-1} \sum_{j=1}^{n} a_{k,j} C_{1,j}(U') = \sum_{j=1}^{n} a_{k,j} C_{k,j}(U).$$

Si $k \neq \ell$, on déduit de la formule ci-dessus que $\sum_{j=1}^{n} a_{k,j} C_{\ell,j}(U)$ est égal au déterminant de la matrice U'' obtenue à partir de U en remplaçant sa ligne ℓ par sa ligne k. Comme cette matrice possède deux lignes égales (celles d'indices k et ℓ), son déterminant est nul. Ainsi on a $\sum_{j=1}^{n} a_{k,j} C_{\ell,j}(U) = \det(U)\delta_{k,\ell}$ quels que soient k et ℓ. En appliquant cette formule à la matrice U^t et en notant que $C_{j,\ell}(U) = C_{\ell,j}(U^t)$ quels que soient j et ℓ, on en déduit que

$$\det(U)\delta_{k,\ell} = \det(U^t)\delta_{k,\ell} = \sum_{j=1}^{n} a_{j,k} C_{\ell,j}(U^t) = \sum_{j=1}^{n} a_{j,k} C_{j,\ell}(U).$$

□

Théorème B.3.4. *Pour toute matrice $U \in \mathrm{Mat}_{n\times n}(A)$, on a*

$$U\mathrm{Adj}(U) = \mathrm{Adj}(U)U = \det(U) I.$$

Preuve. En effet, l'élément d'indice (ℓ, j) de $\mathrm{Adj}(U)$ étant $C_{j,\ell}(U)$, les formules de la proposition B.3.3 signifient que l'élément (k, ℓ) du produit $U\mathrm{Adj}(U)$ et l'élément (ℓ, k) du produit $\mathrm{Adj}(U)U$ sont tous deux égaux à $\det(U)\delta_{k,\ell}$ quels que soient $k, \ell \in \{1, \ldots, n\}$. □

On conclut avec le critère suivant :

Corollaire B.3.5. *Les éléments inversibles de l'anneau* $\mathrm{Mat}_{n\times n}(A)$ *sont ceux dont le déterminant est un éément inversible de* A.

Preuve. Soit $U \in \mathrm{Mat}_{n\times n}(A)$. Si U est inversible, il existe une matrice $V \in \mathrm{Mat}_{n\times n}(A)$ telle que $UV = I$. En prenant le déterminant des deux membres de cette égalité, on obtient, grâce au théorème B.2.7, $\det(U)\det(V) = \det(I) = 1$. Comme $\det(V) \in A$, on en déduit que $\det(U) \in A^*$. Réciproquement si $\det(U) \in A^*$, il existe $c \in A$ tel que $c\det(U) = 1$ et alors le théorème ci-dessus montre que la matrice $V = c\mathrm{Adj}(U)$ de $\mathrm{Mat}_{n\times n}(A)$ satisfait $UV = VU = c\det(U)I = I$, donc U est inversible. □